Proceedings of the

Basque Conference on Cyber-Physical Systems and Artificial Intelligence

May 18-19, 2022

Palacio Miramar, San Sebastian, Spain

Editor: Manuel Graña, UPV/EHU

DOI 10.5281/zenodo.6562355

MSCA-RISE 2017, CybSPEED
Grant #777720

To those that came before us. To those that will come after us. To those that understand what we say. To those that do not understand us. To those that live in fear, may them be freed from it by enlightenment because fear is always born from ignorance. To those that can look at the times we live and smile at them. We will be thankful if they can share the joke with us.

Foreword

The CybSPEED project, funded in the call MSCA-RISE-2017 was a four year project. The first two years, 2018 and 2019 were very intense, accomplishing an astounding percentage of the planned secondments, with many personal interactions on top of the research work. We were ready for another two years of research and heated discussions, and lots of publications. Then came the pandemic. When it was obvious that the pandemic status of the world has arrived to stay we asked for a project suspension due to the management of the pandemic, and when the suspension period ended the doubts about everything did not vanish in thin air, they stayed as very solid black clouds. In this state of affairs we tried to keep calm and carry on, doing our best in the worst conditions for free creative thinking. So some experimental works have been carried out, some new publications and PhD Thesis have been added to the already long list of results, extraordinary with regard to the project budget according to external reviews. In this frame of mind, shared by 90% of humanity, that of confusion and uncertainty and diminishing trust, we faced a final milestone of our project in the form of a formal conference gathering researchers from diverse countries against all odds. These proceedings are the formal outcome of this gathering of minds that we will wish to signal the beginning of the return of the good times.

San Sebastian, May 2022 *Manuel Grana for the CybSPEED research teams*

Preface

These proceedings are the formal record of the meeting held in the Palacio Miramar of San Sebastian in May 18-19, 2022. The attendance of the meeting was composed of invited speakers from several countries and of researchers from the beneficiaries and third country partners of the CybSPEED project. This is the first version that will be published as open access publication in zenodo.org, that will be superseded by improved versions in the near future, when we correct minor formal errors and include some papers that have not been processed before the conference, or new versions of papers already included. (DOI 10.5281/zenodo.6562355) The conference web page will continue to be updated after the conference with new announcements, including special issues in selected journals.

San Sebastian,
May 2022

Manuel Graña
CybSPEED research teams

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation 352 programme under the Marie Sklodowska-Curie grant agreement No 777720.

Contents

Chapter 1
Open-Source Tools for Human-Robot Interaction

Javier de Lope[1] and Manuel Graña[2]

Abstract This paper reviews the state of art of open-source tools that are used in a Human-Robot Interaction context. They are arranged in two areas. First, libraries and frameworks based on deep learning, approach in which most of the solutions are currently based. Second, we are considering the integration with robots, thus we summarize some architectures and frameworks defined on Robot Operating System.

1.1 Introduction

Human-Robot Interaction (HRI) integrates concepts about how the communication between humans and robots is made. Initially the problem was limited to how to define commands to operate robots but once robotic systems and tasks increased in complexity, it was need to consider a broader communication channel for the interaction.

Modern robotic systems are aware of direct human commands but also of body motion and pose, gesture motion, facial expressions, emotion in utterances, etc. The goal is to analyse the human actions with a complex sensory system [6]. We are not longer interested on simple recognizing, we want to recognize the human intentions in order to predict and immitate actions.

It started under the so-called Social Signal Processing (SSP) that aims at providing computers with the ability to sense and understand human social signals [15]. These social signals include several behavioral cues as for example affective or cognitive states like fear, joy or disagreement and culture-specific interactive signals link wink or thumbs up.

Javier de Lope
Department of Artificial Intelligence, Universidad Politécnica de Madrid (UPM), Spain

M. Graña
Computational Intelligence Group, University of the Basque Country (UPV/EHU), Spain

To detect and recognize those actions the sensory system needs to be more complex as in the architecture, communication and inference as in the own sensors capabilities. Generally sensory system must provide data based on computer vision inputs, speech analysis and biometry. We need new architectures and tools, which were able to generate the appropriate data on time.

The rest of the paper is organized as follows. First, we review the most used open-source toolkits and frameworks. We divide them into deep learning-based and ROS frameworks because of these two technologies are frequently used when robots and artificial intelligence are used. Then, we summarize the domain alternatives by giving some insigths on how the field will advance in the next few years.

1.2 Deep learning-based tools

Thanks to the new deep learning approach, the neural networks are being satisfactorily applied to many tasks where the use was not advised. Convolutional layers to extract image features and new proposal for backpropagating the error have made possible to improve the performance in classification, detection and recognition tasks.

TensorFlow [1] was probably the first effort to create a software library for deep neural networks. As the use was complex and repetitive, it appeared higher level software libraries as for example Keras [3] that made easier the creation of network models by including macro-like routines to add and manage layers, optimizers, and other low level blocks provided by TensorFlow. Keras was posteriorly adopted by the TensorFlow ecosystem but there are many other high level deep learning frameworks as Caffe [7] or PyTorch [11].

Now the tendency is to provide higer level frameworks. The basic block is not, for example, the layer as in the above libraries, it is the whole net model, usually pretrained with large datasets and ready to be used in user applications after some adapatation to the problem (transfer learning). Or even complete computer vision or audio processsing pipelines to model, for example, the human body.

1.2.1 Intel RealSense SDK

Intel RealSense [4] is a family of devices with depth perception capabilities (stereo depth, LiDAR, etc.) to be applied in a wide range of applications, particularly HRI related. The current SDK 2.0 facilitates the connection of the devices to the computer by giving the drivers and low level routines to acquire RGB and depth images and low level image processing.

The SDK 1.x also solved the low level layer but provided a very interesting framework to build complex applications based on the hardware devices. The framework covered computer vision and and automatic speech processing tasks such as face and hand detection via images and facial landmarks position estimation and hand pose modeling. This framework has been discontinued after the next mayor upgrade and it is no longer available for its use. Probably some of the functionalities have been included in newer Intel toolkits.

1.2.2 OpenVINO

OpenVINO [5] is an open source toolkit for optimizing and deploying deep learning models onto Intel hardware. It is based on public models provided by frameworks such as TensorFlow [1] or PyTorch [11]. The toolkit boosts the performance in typical artificial intelligence tasks as computer vision and automatic speech processing.

Since it uses well-known deep neural networks, which are generally available in other frameworks, the main interest consists in its integration with the range of supported hardware and platforms.

1.2.3 MediaPipe

MediaPipe [9] is a framework for building pipelines over sensory data provided by Google. It offers *solutions* to be directly used by final applications or for research prototypes. Some of these solutions imply video processing and include face detection, facial landmarks detection and segmentation, hand and body pose estimation with skeleton-like models, or object detection and recognition for robot navigation.

Conceptually, MediaPipe is at least one level higher than software libraries such as Keras [3] or PyTorch [11]. In fact those libraries are used when solutions require deep learning techniques, although the use is transparent to the user. A kind of transfer learning is performed when deep neural networks are used. The complex applications can developed over these solutions, which add a new abstraction level over the sensory pipeline processing. As an example a sign language recognition system can be developed over the hand detection solution.

The API is available from several programming languages and the solutions can be deployed on computers and mobile devices.

1.2.4 Other deep learning middleware

Similar middleware solutions like the MediaPipe are been provided by specialized companies. Sometimes they are not provided as open source but fully usable via trial versions as the Ailia SDK [2] by the Japanese Ax Inc.

Those middleware generally provide specialized inference systems via transfer learning, which improve the performance in multi-CPU and multi-GPU environments. They use pretrained deep neural networks models almost ready to be integrated in applications.

1.3 The ROS ecosystem

The Robot Operating System (ROS) [13] is currently the *de facto* standard to develop robotic applications. It provides an efficient intercommunication protocol between the elements like sensors, actuators and controllers and a model architecture that allows to create applications quickly.

Using ROS as base layer, companies and research institutions add modules to control proprietary manipulators or wheeled and legged mobile robots, to acquire data from cameras and other sensory hardware for robot navigation or manipulation tasks, to build environment models in which robots operate for localizing them, and many others.

In the last few years robotic applications are starting to consider humans inside the control loop. As software as hardware is reaching a maturity level that makes possible to take seriously this old idea. For example, manipulators are not longer specialized cells in integrated manufacturing environments, they freely collaborate with humans to solve the task [16].

Under this perspective ROS is being extended with frameworks that make easier the human integration in the loop and allow to model part of his external or internal activity.

1.3.1 Ambient assisted living

One of the initial efforts to integrate humans in ROS was the proposal of a multi-robot architecture for ambient assisted living [8]. The aim is to control an autonomous wheelchair and social robots by using a ROS-based architecture to help the user in navigation tasks.

It can not be considered under the current HRI concept because of, for example, it does not include a multimodal nature in sensors, but strictly speaking it serves as interface between the human and the wheelchair. Also, the conclusions project depict how to integrate modules for emotion detection based on electroencephalography (EEG) and controllers based on voice-oriented dialogues and gestures.

1.3.2 ROS4HRI

ROS4HRI [10] is a framework that introduces a set of conventions and standard interfaces for HRI designed to be used with ROS. The lack of a ROS standard for HRI is partially caused by maturity level of some HRI algorithms and pipelines compared to technological problems such as simultaneous localization and mapping (SLAM). Most usual HRI task such as skeleton tracking, face recognition, gaze identification, and speech processing use their own set of interfaces and conventions, which usually contain proprietary solutions.

ROS4HRI specifies a ROS-based representation model for humans and ROS conventions and data types to integrate multi-modal and interoperable social signal processing pipelines. The representation of a person is based on a combination of four unique identifiers: a face, a body, a voice, and a person identifier. These identifiers are not mutually exclusive and only some of them might be available at a given time for a particular person. They are asigned to each person as soon they are detected by the system. Also, an URDF kinematic model with 15 links for humans is defined, which is used to define poses and compatible with ROS ecosystem tools such as RViz.

The current version is focused on ROS1. Most of the messages and topics structure should be straightforwardly transferable to ROS2.

1.3.3 ROS-Neuro

ROS-Neuro [14] is an extension of ROS to use as middleware for human neural interfaces (HNI). The aim is to place the HNIs at the same conceptual and implementation level than robotic systems. Thus, robotic applications such as control of telepresence robots, powered wheelchairs, robotic arms or exoskeletons with brain-machine interface (BMI) can share the same architecture and workflow. The acquisition, processing, and classification of nerual data can be included into a closed-loop controlled robotic system.

ROS-Neuro generalizes the conventional ROS architecture by providing modules to gather neurophysiological signals, to record the acquired data, to process and decode it, and to infer the intention of the user. It provides several standard interfaces for data acquisition from different commercial electroencephalography (EEG) and electromyography (EMG) devices.

ROS-Neuro has been used and tested in different experiments and situations with commercial EEG/EMG devices and various robotic platforms. One of the most extensive experimentation has been carried out in the Cybathlon neurotobotic championship [12], in which several international teams compete in different disciplines such as wheelchairs or exoskeletons races.

The current version of ROS-Neuro is based on ROS1, although the transition to ROS2 is already scheduled.

1.4 Concluding remarks

Human-Robot Interaction and Social Signal Processing are fields that are being modified thanks to the new open source libraries and frameworks that provide computing pipelines to compute highly complex human oriented tasks. These frameworks are built on high level software libraries that add an abstraction level over the hardware and conceptual methods by hidding scientific and technological complexity.

Currently it is easier to build applications based on complex sensory systems, which model pose and movement of humans to determine what activities are carried out, read the neural activity to approximate the state of awareness, or record the voice to recognize emotions in utterances. All these new completely functional applications offer a new challenge to researchers to increase the complexity in the interaction with humans.

The change of paradigm will be possible to add new methods to these frameworks, which will be received by the community and will serve to create even more complex applications.

Acknowledgments

This work has been partially supported by FEDER funds through MINECO project PID2020-116346GB-I00. This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 777720.

References

1. Abadi, M., et al.: TensorFlow: Large-scale machine learning on heterogeneous systems (2015), https://www.tensorflow.org/
2. ax Inc.: ailia SDK (2019), https://axinc.jp/en/
3. Chollet, F., et al.: Keras (2015), https://github.com/fchollet/keras
4. Intel Corp.: Intel RealSense SDK (2015), https://intelrealsense.com/
5. Intel Corp.: OpenVINO (2018), https://docs.openvino.ai/
6. Ji, Y., Yang, Y., Shen, F., Shen, H.T., Li, X.: A survey of human action analysis in hri applications. IEEE Transactions on Circuits and Systems for Video Technology **30**(7), 2114–2128 (2020). https://doi.org/10.1109/TCSVT.2019.2912988
7. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. arXiv preprint arXiv:1408.5093 (2014)
8. Li, R., Oskoei, M.A., Hu, H.: Towards ros based multi-robot architecture for ambient assisted living. In: 2013 IEEE International Conference on Systems, Man, and Cybernetics. pp. 3458–3463 (2013). https://doi.org/10.1109/SMC.2013.590
9. Lugaresi, C., et al.: MediaPipe: A framework for building perception pipelines (2019). https://doi.org/10.48550/ARXIV.1906.08172

10. Mohamed, Y., Lemaignan, S.: ROS for human-robot interaction. In: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 3020–3027 (2021). https://doi.org/10.1109/IROS51168.2021.9636816
11. Paszke, A., et al.: Pytorch: An imperative style, high-performance deep learning library. In: Wallach, H., et al. (eds.) Advances in Neural Information Processing Systems 32, pp. 8024–8035. Curran Associates, Inc. (2019)
12. Perdikis, S., Tonin, L., Saeedi, S., Schneider, C., Millán, J.d.R.: The Cybathlon BCI race: Successful longitudinal mutual learning with two tetraplegic users. PLOS Biology **16**(5), 1–28 (05 2018). https://doi.org/10.1371/journal.pbio.2003787
13. Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R., Ng, A.: ROS: an open-source Robot Operating System. In: IEEE ICRA Workshop on Open Source Software. vol. 3 (2009)
14. Tonin, L., Beraldo, G., Tortora, S., Menegatti, E.: ROS-Neuro: An open-source platform for neurorobotics. Frontiers in Neurorobotics **16** (2022). https://doi.org/10.3389/fnbot.2022.886050
15. Vinciarelli, A., Pantic, M., Bourlard, H.: Social signal processing: Survey of an emerging domain. Image and Vision Computing **27**(12), 1743–1759 (2009). https://doi.org/10.1016/j.imavis.2008.11.007
16. Weiss, A., Wortmeier, A.K., Kubicek, B.: Cobots in industry 4.0: A roadmap for future practice studies on human-robot collaboration. IEEE Transactions on Human-Machine Systems **51**(4), 335–345 (2021). https://doi.org/10.1109/THMS.2021.3092684

Chapter 2
The Alienist*: a journey in neuroscience from simple binary classifications to concentration inequalities

J.M. Górriz[1,2,3] and SIPBA collaborators[1,2]

Abstract Our understanding of the human brain comes primarily from *in vivo* imaging that has evolved from small samples to population studies. Most neurological and psychiatric disorders are now associated with cortical patterns of disease pathogenesis and progression, leading to increasing hope for high quality, reliable prognostic markers of treatment response that are the foundation of personalised medicine. Whole-brain analyses comprising a large number of statistical tests are frequently undertaken to ensure inclusion of relevant structures. These analyses have been traditionally conducted with classical statistics, either hypothesis testing or Bayesian inference, that relies on assumptions that are frequently violated. Consequentially, inflated type I error rates have become problematic, and a key contributor to the replication crisis. Technological advances are increasing spatial and temporal resolutions as well as the range of available measurements of anatomy and physiology; a true exemplar of the *curse of dimensionality*. In this context, analyses of contemporary large image repositories retain the difficulties associated with small sample sizes. One promising solution is machine learning, where high-dimensional relationships between datasets are empirically established. Resampling and cross-validation in combination with statistical classifiers have been proposed for brain decoding. Beyond that, Statistical Learning Theory (SLT) focuses on identifying discriminant functions for pattern recognition that optimise the robustness and complexity of the input-output relationship. Estimating dependencies, unlike in classical statistics, characterizes the actual relationships with a limited dataset. The Alienist project proposes a new view of deriving inferences from brain imaging data by applying SLT in a model-free, i.e. agnostic, methodology based on concentration

J.M. Górriz
DASCI Institute, University of Granada Spain e-mail: gorriz@ugr.es sipba.ugr.es

Dpt. Psychiatry, University of Cambridge e-mail: jg825@cam.ac.uk
· Dpt. Electronics, University of Cádiz, Spain
e-mail: juanmanuel.gorriz@uca.es

* The Statistical Agnostic Mapping: Machine Learning in King Neuroscience's court

inequalities that exploits an accurate and reliable data-driven algorithm: Statistical Agnostic Mapping.

2.1 Introduction

Over the last decades, neuroscience has transitioned from qualitative case reports to quantitative, longitudinal and multivariate population studies in the quest for defining patterns of disease pathogenesis, prognostic indicators and treatment response. Classical statistics has primarily contributed to neuroscience through infernece based on null-hypothesis (H_0) testing [1] or more recently, with Bayesian Inference (BI) based on the conditional probability of the model parameters given the data ($p(\theta, \mathbf{y})$). This methodology has been extensively exploited by the brain mapping community for conducting whole-brain exploratory analyses [2].

There are two existing cultures regarding the use of statistical modeling to extract relevant information from data [3]. One of them, widely used in neuroscience and particularly in neuroimaging, relies on stochastic models behind observed samples, and on the assumption that the hypothesised model structures (e.g. pdf) correspond to the actual structure. The other culture uses an **agnostic theory based on algorithmic models** operating in limited sample sizes, i.e. *structural risk minimisation*, and treats the nature of data as unknown. This ground-breaking approach known as Statistical Agnostic Learning constitutes the foundation of the proposed methodology, and has never before been exploited in neuroscience for statistical inference.

2.1.1 State-of the-art:

In the early 2000s, the use of the general linear models (GLM) and classical statistics was a widely accepted standard in neuroscience for characterising neuroimaging data. The statistical inference of these parametric methods assumes Gaussian distributions for the errors (ϵ) at each level of the model design, i.e.:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\theta} + \epsilon \tag{2.1}$$

where the design matrix ($\mathbf{X}$) contains explanatory variables or constraints of the response (observed) variables ($\mathbf{y}$), and the parameters ($\boldsymbol{\theta}$) are determined using the maximum likelihood (ML) criteria. Inference about these ML estimations $\hat{\boldsymbol{\theta}} = E[\boldsymbol{\theta}]$ i) involves a random/fixed effect analysis, and ii) is performed by the use of a linear compound ($\mathbf{c}$, a contrast weight vector) of the parameter estimates over its covariance in the form of a T statistic:

$$T = \frac{\mathbf{c}^T \hat{\boldsymbol{\theta}}}{\sqrt{\mathbf{c}^T Cov(\hat{\boldsymbol{\theta}})\mathbf{c}}} \tag{2.2}$$

As shown in equation 2.2, to implement this classical procedure we need the covariance of the ML estimator $Cov(\hat{\boldsymbol{\theta}}) = (\mathbf{X}^T \mathbf{C}_\epsilon^{-1} \mathbf{X})^{-1}$, i.e. the error covariance components of the model, that can be recursively estimated via the Expectation-Maximisation (EM) algorithm.

The well-known statistical parametric mapping (SPM) software package [4] calculates a T-statistic independently at every voxel of an image. These statistics are then thresholded by reference to random field theory (RFT) [5] to control the family wise error (i.e type I) rate (FWER). The **main limitation** of this frequentist approach is that the resulting significance maps refer to the probability of observing the absence of any effect; that is (H_0), and does not reflect the likelihood that a particular effect is found in the sample. Thus, significance maps are established by thresholding a p-value map, which, if arbitrarily selected small, can be exploited for reporting **meaningless differences**. On the contrary, the alternative hypothesis (that there is an effect) cannot be rejected. Moreover, the inference process requires other assumptions beyond Gaussianity such as "homogeneity" (i.e. a location is allowed to be either present or absent) and "shift alternative" (the effect is a non-zero average effect over all subjects). Unfortunately, these **assumptions are frequently violated** and advanced approaches are needed that allow regions of the image to contain voxels representing the presence and absence of the effect at some mixing proportion. Alternatively, non-parametric permutation tests can derive asymptotic forms for null distributions, but with a possible loss of statistical power through model assumptions [6].

A major challenge for SPM and similar packages is inference across a very the large number of statistical values that form an image (i.e. multiple comparisons problem). This **mass univariate** scenario forces correction of actual p-values to reduce false positives (FP) using classical methods such as Bonferroni correction, or improved methods like RFT [7]. However, in many cases, the degree of spatial and temporal autocorrelation of the imaging data make these approaches **over-conservative**.

From the end of 2000s, Bayesian inference (BI) started to be exploited either alongside or to replace frequentist inference. Unlike classical statistics, BI uses likelihood to express degrees of belief [8] that a particular effect is present. Bayesian inference is based on the conditional probability of the parameters ($\boldsymbol{\theta}$) given the observations ($\mathbf{y}$), i.e. $p(\boldsymbol{\theta}|\mathbf{y})$, where the sampling model $p(\mathbf{y}|\boldsymbol{\theta})$ and the prior distribution of the parameters $p(\boldsymbol{\theta})$ are specified *a-priori*. From Bayes rule the posterior probability is proportional to both:

$$p(\boldsymbol{\theta}|\mathbf{y}) \sim p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{2.3}$$

Under the GLM model, the conditional and the prior probabilities are again assumed to be Gaussian, and the problem reduces to determining conditional statistical moments, mean and variance, in hierarchical linear observation models. Subsequently, a similar T statistic as that shown in equation 2.2 but defined on conditional estimations, is derived and thresholded to declare the presence of an effect at a voxel, although this final step is unnecessary in BI. In this sense, BI differs in its use of a prior distribution to express the uncertainty before seeing the data, where the posterior distribution (the uncertainty remaining) provides the basis for predicting new values taking into account the uncertainty in estimating $\boldsymbol{\theta}$ [9]. The idea behind BI is to avoid the limitation of frequentist inference on a specific model and giving a mechanism for inferring any model. Thus, models can be addressed for which there is no parametric frequentist solution [8].

The selection and estimation of any specific model under limited sample size scenario **remains an issue in neuroimaging**. Particularly when the model and the interaction between model parameters becomes too complex for an accurate **posterior probability estimation**, or a feasible numerical **computation of the Bayes rule**. In these cases procedures generally exploit a combination of several heuristics that try to solve this intractable high dimensional problem. Examples include several tools in FSL [8], a well-established software package for neuroimaging that has implemented BI. Moreover, at the final stage of classical inference thresholding is necessary to define an effect based on the probability of observing the data under the null-hypothesis (that no effect is present) .

2.2 A new era in neuroscience based on machine learning:

Despite the popularity of machine learning (MLE) as a solution for a wide range of complex problems, its relevance for between-group statistical inference remains an open question. Neuroimaging, in particular, has embraced MLE as a technology to deliver diagnostic and prognostic classification of neurological and psychiatric disorders. Whilst the primary aim of neuroimaging studies is frequently to locate regional between-group differences in brain structure and function, identifying imaging relationships to clinical profiles, genes, and tissue-derived biomarkers is key to establish corroborative evidence. Efforts with MLE in this context are increasing with continuous output variables ([10] and remarked in [11]) rather than the more typical binary classifications.

Several advances for combining p-value maps have been proposed based on the **concept of prevalence** [12] which go beyond fixed and mixed (random) effects models [2]. Common to these approaches is the assumption of a mixing of subject classifications at each voxel that is more realistic than those assumed in classic random effects approaches; for example, homogeneity of a binary activation pattern [12]. However, the latter approach is still based their insight of assuming an ever increasing number of data models. The purpose is to relieve the violation of the parametric assumptions by increasing the complexity of the approach, as an attempt to minimize empirical dependencies. Nevertheless, the concept of prevalence offers the possibility a new framework for modern statistics.

The use of prevalence, for example as a fraction of individuals correctly classified by a MLE algorithm in group comparisons, is not novel in neuroimaging and is indeed the main focus of **predictive inference**; out-of sample generalization approaches, such as Cross-Validation (CV), aim to estimate the accuracy of a classifier on unobserved data in a binary classification problem. Although the methods and goals of predictive CV inference are distinct from classical extrapolation procedures[13], they are exploited within frameworks aimed **at assessing statistical significance** [11]. Bootstrapping, binomial or permutation ("resampling") tests [14] are all examples that have been demonstrated as relevant outside of classical statistics, filling otherwise-unmet inferential needs. In predictive CV inference we assume the existence of classes (H_1) that can be differentiated by classifiers. These classifiers measure their performance in terms of accuracy or prevalence on an independent dataset. In this case, we accept or reject (improperly in a statistical sense) the alternative hypothesis, H_1, using empirical confidence intervals such as standard deviations of the classification accuracy from training folds. In addition, in small sample size scenarios, the most popular K-fold CV method [15] is sub-optimal under unstable conditions [16, 17, 18]. In such circumstances, the predictive power of the trained classifiers can be arguable. Moreover, it has been recently demonstrated that when using only a classifier's empirical accuracy as a test statistic, the probability of detecting differences between two distributions is lower than that of a *bona fide* statistical test [19, 20].

2.3 From explaining observations to explaining conditions

The observation and explanatory variable domains are connected. This statement establishes that the GLM solution and the subsequent statistical inference in general hierarchical models (that is, a set of recursive equations as shown in eq. 2.1) can be obtained by solving a set of multiclass classification/regression problems at each level, in terms of a regression of the (*label*) explanatory matrices.

Proof As a simple proof of concept, consider the following example. Given a set of observations $\{\mathbf{y}_i\}$, for $i = 1, \ldots, N$, we are interested in explaining a set of "explanatory" binary-coded variables $\mathbf{x}_i$ by a matrix $\mathbf{W}$ of parameters. This problem, referred to here as the inverse problem in the *label domain*, is also known as the linear regression of an Indicator Matrix or linear regression model (LRM) [9]. Note that, in this model, explanatory variables instead of observed responses are regressed. This regression can be more accurate depending on the nature of the data to be fitted, in particular for a low number of discrete categories in the specified design matrix $\mathbf{X} = [x_{im}]$. If we have M classes then $\mathbf{X}$ is a $N \times M$ matrix, where each row $i = 1, \ldots, N$ contains a single $x_{im} = 1$, for $m = 1, \ldots, M$, $\mathbf{Y}$ is the $N \times P$ matrix of column responses $\mathbf{y}_i$ and $\mathbf{W}$ is a $P \times M$ coefficient matrix. Thus, we fit a LRM of the form:

$$\mathbf{X} = \mathbf{YW} \tag{2.4}$$

where the P dimension denotes multimodal (e.g. PET, structural or functional MRI) or multiframe (e.g. over time) acquisitions given the same indicator response matrix $\mathbf{X}$. Following the ML methodology, the best estimation (least squares, LS) is given by:

$$\hat{\mathbf{W}} = (\mathbf{Y}^T\mathbf{Y})^{-1}\mathbf{Y}^T\mathbf{X} \tag{2.5}$$

which regresses inputs of observations on to a novel set of labels or constraints, i.e. $\hat{\mathbf{X}} = \mathbf{Y}\hat{\mathbf{W}}$. The novel set $\hat{\mathbf{X}}$ can be seen as an approximation of the constraints for the set of observation vectors $\mathbf{y}_i$, or an approximation of the posterior probability $p(class = m|\mathbf{y})$. Thus, it allows computation of an error model as $\epsilon_{LS} = \mathbf{X} - \hat{\mathbf{X}}$.

For simplicity, and to connect with the GLM, as shown in equation 2.1, let $P = 1$ in the LRM, then $\mathbf{W} = \mathbf{w}$ is a $1 \times M$ row vector and $\mathbf{Y} = \mathbf{y}$ is an $N \times 1$ column vector. A simple relation between the GLM and LRM approximations can be found taking into account:

$$\mathbf{X} = \mathbf{y}\mathbf{w} + \epsilon_{LS} \tag{2.6}$$

Thus, the corresponding (random effects) GLM is:

$$\mathbf{y} = (\mathbf{X} - \epsilon_{LS})\hat{\boldsymbol{\theta}} \tag{2.7}$$

where we define $\hat{\boldsymbol{\theta}} = \mathbf{w}^T(\mathbf{w}\mathbf{w}^T)^{-1}$, and the GLM noise model is derived using $\epsilon = -\epsilon_{LS}\hat{\boldsymbol{\theta}}$. The scalar term of equation 2.7 can be expressed with the LS solution as:

$$(\mathbf{w}\mathbf{w}^T)^{-1} = (\mathbf{y}^T\mathbf{y})^2/((\mathbf{X}^T\mathbf{y})^T\mathbf{X}^T\mathbf{y}) = \frac{(\sum_{i=1}^N y_i^2)^2}{\sum_{m=1}^M \sum_{i,j} y_{im}y_{jm}} \tag{2.8}$$

where y_{im} denotes observation i belonging to class m. Thus, a LS linear regression of the observations can be described by a GLM regression on the observations (i.e. a linear regression on the explanatory variables), and vice versa. □

2.4 The LS solution in terms of the optimum GLM

The above simple connection for a single level (random effects) model can be extrapolated to hierarchical models and non-binary design matrices. Thus, parameter fitting in the observation domain by ML or Bayesian estimations can be performed by solving a complex pattern classification/regression problem. Therefore, this **clearly opens a novel path for solving conventional analyses of neuroimaging data**, beyond classical and Bayesian frameworks, by the use of the most recent advances in MLE within statistical agnostic theory [21]. This theory delivers an optimum estimation of the actual error (instead of using LS in the LRM) with limited samples and complex dependencies, generalising to unseen data.

Let assume the model in Eq. 2.1 with optimum solution $\boldsymbol{\theta}$ and zero-mean covariates between $[-c, c]$ in the design matrix, then the LS solution [2] in terms of the latter can be expressed as:

---

[2] the L_1-norm solution could be expressed in a similar fashion but the proof requires a little more effort [22]

$$\mathbf{w} = \frac{\mathbf{y}^T\mathbf{X}}{||\mathbf{y}||^2} = \frac{\boldsymbol{\theta}^T\mathbf{R}_X + \epsilon^T\mathbf{X}}{\boldsymbol{\theta}^T\mathbf{R}_X\boldsymbol{\theta} + \boldsymbol{\theta}^T\mathbf{X}^T\epsilon + ||\epsilon||^2 + \epsilon^T\mathbf{X}\boldsymbol{\theta}} \tag{2.9}$$

where $\mathbf{R}_X = \mathbf{X}^T\mathbf{X}$. If we suppose that $\epsilon^T\mathbf{X} = \mathbf{0}$, assumption that is more than realistic for low-noise levels and balanced experimental conditions we can write:

$$\mathbf{w} = \frac{\boldsymbol{\theta}^T\mathbf{R}_X}{\boldsymbol{\theta}^T\mathbf{R}_X\boldsymbol{\theta} + ||\epsilon||^2} \tag{2.10}$$

where the jth component is $w_j = \frac{\sum_{m,i} X_{im}X_{mj}\theta_i}{\sum_{m,i,j} X_{im}X_{mj}\theta_i\theta_j + ||\epsilon||^2}$. For only one dimension and disjoint covariates, Eq. 2.9 simplifies to:

$$w = \frac{\theta}{\theta^2 + (1/c^2)\sigma_\epsilon^2} \tag{2.11}$$

that generalizes the analysis shown in [22]. Therefore, for low-noise levels the LS solution of the MLE approach is the inverse of the LS solution of the GLM.

2.5 From the p-value to the analysis of the worst case

Data-driven analysis methods based on MLE [23, 24, 25, 26, 27] have demonstrated their ability for detecting activations in fMRI data, outperforming conventional hypothesis-driven approaches, i.e. the standard GLM inference based on random effect models. The core idea on these agnostic (model-free) approaches is to perform an accurate feature extraction based on a fixed-complexity MLE classifier, such as a linear support vector machine (SVM), between predefined groups. They all share the same characteristic processing pipeline of a data-driven multivariate approach that enhances detection ability. Consequently, the SPM derived from the GLM are replaced by a spatial discriminance map (SDM) [23] based on prevalence [19], or some other specific feature extracted at the training stage; e.g. distance to the separating hyperplane [25]. These data-driven maps are then deployed in conventional pipelines for statistical inference, although some approaches depart from the p-value-based frequentist, Bayesian or permutation analyses, and introduce the concept of the probability of the *worst case* in neuroimaging [27].

Beyond empirical techniques for the estimation of performance, MLE is well supported by data-driven SLT, which is primarily devoted to problems of estimating dependencies with **limited amounts of data** [28]. Although MLE approaches were not originally designed to test hypotheses based on prevalence in neuroimaging [29], they are theoretically grounded to provide confidence intervals (protected inference) in the classification of image patterns formulated as maps of statistical significance [27]. The procedure consists of assessing with high probability (the worst case) the quality of the fitted function and its generalization ability, in terms of in- and out-of-sample predictions. This can be achieved by assessing **the upper bounds** of the actual error in a binary classification problem (a confidence interval) and simple significance tests of a population proportion [27] (a prevalence-based test). This multivariate approach (SAM) results in improvements to the test's statistical power based on accuracy. In hypothesis testing, it can be conceptualized as the inverse problem of "carefully rejecting H_0". In other words, we assume H_1 and reject it **under the worst case**, thus accepting H_0, when there is no effect or it is not significant.

Concentration inequalities provide robust and better estimators for neuroimaging. This statement establishes that, under a low dimensional scenario, with a limited sample size and heterogeneous data, it is preferable to assess concentration inequalities cluster- or region-wise, rather than with voxelwise classical CV methods that are ubiquitously used in neuroimaging for estimating the real error of a classifier and statistical inference [11]. This approach is over conservative with regard to the control of false positives (FP), whilst regional assessment provides a good trade-off between statistical power and FP rates, within a permutation analysis [27]. With these statements in mind, we focus on solving a classification/regression problem in a similar way than for classical inference, i.e. voxelwise, or by the use of a multivarate approaches to better control the FP rate. This methodology boosts the impact of one of the more powerful advances in modern statistics in the last decades, the uniform convergence of frequencies of events to their probabilities [28], as well as the concentration inequalities in its general version. They are expressed using several ideas and concepts such as the Vapnik-Chervonenkis (VC) dimension, and the pseudo or fat-shattering dimensions, etc.

2.6 Conclusions

The possibilities for innovating biologically meaningful metrics from structural and functional neuroimaging represent both opportunities and challenges for modern neurology. One the one hand, new analyses are expanding our view of the most complex organ - the human brain - whilst on the other, the wide variety of approaches has led to **"inconsistences"** in the extant literature. The reasons for this are multi-factorial including heterogeneous datasets, multiple scanning sites, differing acquisition protocols and processing methods, and diagnostic definitions. Even findings from studies that use putatively similar metrics and currently available statistical tools can substantially differ from one another. Clearly, deriving consistent multivariate patterns of structural and functional brain differences would be a major advance in our understanding of the complex and heterogeneous alterations induced by developmental, psychiatric and neurological conditions. Confidence in neurobiological substrates would not only build a rich systems narrative, but also serves as a starting point to run 'reverse" development approaches that disentangle the roles of environmental and genetic risk factors in their etiology.

Beyond empirical techniques of performance estimation, MLE is well-established in **data-driven statistical learning theory**, which is primarily devoted to problems of estimating empirical dependencies with limited amounts of data [28]. In short, the goal of modern SLT is to formalize the algorithms proposed in MLE for studying the problem of inference. With reference to its definition, the connection of the SLT-based inference with classical approaches (HT and BI) becomes clear: **it departs from the use of p-values**, i.e. probability of observing the data given no effect, and rather estimates the likelihood that the effect is present in terms of prevalence, in a similar way to the Bayesian formulation. However, with limited sample sizes **our approach is hypothesized to be more conservative and effective than Bayesian methods** as it evaluates the worst-case scenario in the classification of effects without the need for computationally-demanding parameter estimation.

2.6.0.1 Acknowledgements

I would like to thank all my collaborators in the last two decades who took the journey with me. I am so grateful for that. Specially, I would like to thank at the UCAM: John Suckling and Rafael Romero; at the UGR: Javier Ramírez, Carlos Puntonet, Andrés Ortiz, Fermín Segovia, Diego Salas, Ignacio Álvarez, Francisco J. Martinez, Diego Castillo, Miriam López, Rosa Chaves, Antonio Moreno, etc.; at the UCA: Oscar Pernía, Ignacio Turias; at the US: Ruben Martin Clemente and Sergio Cruces; at the University of Leicester: Yudong Zhang and Shui Hua Wang; and at everywhere: Ana Maria Tomé, Elmar W. Lang, Johannes Levin, Su Li, Deniz Erdogmus, Isabella Castiglioni, Afshin Shoeibi, Roohallah Alizadehsani, etc.

This work was supported by the MCIN/ AEI/10.13039/501100011033/ and FEDER “Una manera de hacer Europa” under the RTI2018-098913-B100 project, by the Consejería de Economía, Innovación, Ciencia y Empleo (Junta de Andalucía) and FEDER under CV20-45250, A-TIC-080-UGR18, B-TIC-586-UGR20 and P20-00525 projects.

References

1. Bzdok, D. Classical Statistics and Statistical Learning in Imaging Neuroscience. Front. Neurosci., 06 October 2017 | https://doi.org/10.3389/fnins.2017.00543
2. Friston, K.J. et al. Classical and Bayesian inference in neuroimaging: theory NeuroImage, 16 (2) (2002), pp. 465-483
3. Breiman, L. Statistical Modeling: The Two Cultures. Statistical Science 2001, Vol. 16, No. 3, 199–231
4. Friston, K. et al. Statistical Parametric Maps in functional imaging: A general linear approach Hum. Brain Mapp. 2:189-210 (1995)
5. Frackowiak, R.S. J. et al. Human Brain Function (Second Edition). Chap. 44. Introduction to Random Field Theory. ISBN 978-0-12-264841-0 Academic Press. 867-879, 2004.
6. Heller, R. et al. 2007. Conjunction group analysis: An alternative to mixed/random effect analysis. NeuroImage 37, 1178-1185.
7. Brett, M. et al. An Introduction to Random Field Theory. Human Brain Function (Second Edition), Academic Press, 2004, pp 867-879, ISBN 9780122648410

8. Woolrich, et al. Bayesian analysis of neuroimaging data in FSL. NeuroImage 45 (2009) S173–S186.
9. Hastie T. et al. The elements of statistical learning theory. Data Mining inference and prediction. Ed Springer. isbn 0-387-95284-5. 2001
10. Cohen, J.R., et al. Decoding continuous behavioral variables from neuroimaging data. Front. Neurosci. 5. 2011.
11. Reiss, P.T. et al. Cross-validation and hypothesis testing in neuroimaging: an irenic comment on the exchange between Friston and Lindquist et al. Neuroimage. 2015 August 1; 116: 248–254
12. Rosenblatt, J. D. et al. Revisiting multi-subject random effects in fMRI: Advocating prevalence estimation. NeuroImage 84 (2014): 113-121.
13. Lindquist, M.A. et al. Ironing out the statistical wrinkles in "ten ironic rules". Neuroimage. 2013 Nov 1;81:499-502.
14. Winkler, A.M., et al. Non-parametric combination and related permutation tests for neuroimaging. Human brain mapping 37.4 (2016): 1486-1511.
15. Kohavi, R. A study of CV and bootstrap for accuracy estimation and model selection. Proc. of the 14th international joint conference on AI - Vol. 2 pp 1137-1143 (1995)
16. Górriz, J.M. et al. A Machine Learning Approach to Reveal the NeuroPhenotypes of Autisms. International journal of neural systems, 1850058. 2019.
17. Górriz, J.M. et al. On the computation of distribution-free performance bounds: Application to small sample sizes in neuroimaging. Pattern Recognition 93, 1-13, 2019
18. Varoquaux, G. Cross-validation failure: Small sample sizes lead to large error bars. NeuroImage 180 (2018) 68–77.
19. Rosenblatt, J.D. et al. Better-than-chance classification for signal detection. Biostatistics (2016).
20. Kim, I. et al. Classification accuracy as a proxy for two sample testing Annals of Stat., 2020
21. Haussler, D. Decision theoretic generalizations of the PAC model for neural net and other learning applications. Information and Computation Volume 100, Issue 1, September 1992, Pages 78-150
22. Gorriz, J. M. et al. A hypothesis-driven method based on machine learning for neuroimaging data analysis. arXiv:2202.04397.
23. Mouro-Miranda, J. et al. Classifying brain states and determining the discriminating activation patterns: Support vector machine on functional MRI data. NeuroImage, 28, 980?995. (2005).
24. Wang, Z. et al. Support vector machine learning-based fMRI data group analysis. NeuroImage 36 (4), 1139-1151
25. Wang, Z. A hybrid SVM–GLM approach for fMRI data analysis. Neuroimage 46 (3), 608-615
26. Smith, D.V. et al. Decoding the anatomical network of spatial attention. Proc Natl Acad Sci USA 110:1518–1523. (2013).
27. Gorriz, J.M. et al. Statistical Agnostic Mapping: A framework in neuroimaging based on concentration inequalities. Information Fusion Volume 66, February 2021, Pages 198-212
28. Vapnik, V. Estimation dependencies based on Empirical Data. Springer-Verlach. 1982 ISBN 0-387-90733-5
29. Friston, K.J. Sample size and the fallacies of classical inference. NeuroImage 81 (2013) 503?504

Chapter 3
Characterization of neural circuits underlying auditory processing in healthy and pathological conditions

Maria Isabel Carreño

Abstract This short paper is an statement of the research interests and planning of activities for the next few years.

3.1 Scientific interest

My long-term scientific interest is to understand the mechanisms by which sensory processing can be modulated by internal and external cues and how this determines the behavioral output and environmental adaptations. It has been shown that different contexts as well as different concentrations of specific neurmodulators or hormones may change perception and consequent behavior in typically developing brains1-3. But, why is that? What makes possible that sensory perception is integrated in the brain differently depending on the environment, awareness level or hormonal conditions? And how is this modulation affected in different neurodevelopmental disorders? In order to answer these questions, we need to study the wiring patterns of the brain and molecular signaling that makes this possible. Furthermore, understanding the mechanisms by which experience, and early-life events refine this wiring both at a functional and structural level in sensory cortices, will allow us to intervene if needed. During late stages of development, as well as in response to injury, neuronal circuits have a remarkable capacity to reorganize themselves in response to the patterns of activity of their inputs. This plasticity in turn has a profound impact on the information processing capability of the circuits. One of our most important goals is to be able to identify these cellular rearrangements and circuitry organization and most importantly, understand their translation to behavior in mice and humans.

MI Carreño
CHU Sainte-Justine Research Center, affiliated with Universite de Montreal, Canada

3.1.1 Sensory perception in neurodevelopmental disorders

My research in Dr Di Cristo's lab focuses on investigating the cellular mechanisms underlying both healthy and pathological development of inhibitory GABAergic circuitry. Recent studies suggest that hypofunction of Parvalbumin-positive (PV) cells, a GABAergic interneuron subtype innervating hundreds of postsynaptic targets with multiple synapses clustered around the cell body and proximal dendrites, is fundamental to the pathogenesis of ASDs and could therefore be an essential key to developing novel targeted therapies. We have recently made progress in understanding the molecular mechanisms controlling the development of different inhibitory interneuron circuits in postnatal cortex of different models of neurodevelopmental disorders: Syngap1-Related Disorders (SYNGAP1-RD), Tuberous sclerosis (TSC1), Fragile X Syndrome (FXS) and perinatal hypoxia. Among those, most of my efforts have been oriented to Syngap1-RD, an epileptic encephalopathy characterized by epilepsy, intellectual disability (ID), and ASD. To date more than 800 individuals withSYNGAP1-RD have been reported worldwide, and despite intense research over the last decade, we still do not have a clear picture of the underlying neuropathology. In a recently published work, we reported that different aspects of auditory-processing are similarly impaired in both SYNGAP1-RD patients and Syngap1 haploinsufficient mouse model. Here, for the first time, we have been able to correlate data from SYNGAP1-RD patients and Syngap1+/- mouse model, adding translational knowledge to SYNGAP1 literature4. More recently, using conditional knockout mice, we have found that Syngap1 haploinsufficiency in GABAergic interneurons causes similar sensory processing dysregulation and cognitive inflexibility (Jadhav et al, submitted), suggesting GABAergic circuits play a critical role to the pathophysiological mechanisms underlying aberrant sensory perception in Syngap1- RD. In addition, most of the alterations in sensory perception that we have recently reported4 have also been shown to rely mainly on GABAergic interneurons; in particular, on parvalbumin-expressing (PV) cells, the larger subgroup of cortical GABAergic cells which target the perisomatic domain of excitatory neurons, and on somatostatin (SST)-expressing, dendritic targeting interneurons5. Indeed, recent work has shown that optogenetic suppression of PV cells led to increased responses to both repetitive (standard) and rare (deviant) tones46., reducing auditory habituation, as we observed in Syngap1+/- mice. On the other hand, SST interneurons selectively reduce excitatory responses to frequent tones in an oddball paradigm6 and optogenetic suppression of SST cells reduced the miss match negativity (MMN) peak7,, suggesting impaired deviant sound detection, similarly to what we observed in Syngap1+/- mice. Currently, we are investigating to what extent, inhibitory function is involved in different aspects of auditory perception and whether acute increase of inhibitory tone by chemogenetic approaches rescues altered EEG patterns in Syngap1−/+ mice.

3.1.2 Sexual hormones modulation of sensory perception

In a parallel project, we study sex differences in sensory perception. We believe this is particularly important in psychiatric research, where hormonal effects on brain activity need to be taken into consideration for efficient drug development and administration. Actually, it is well known that sensory perception varies in function of sex and hormonal concentrations. For example, in humans, changes in auditory sensitivity8, and predictive coding and repetition suppression to sounds9 change over the female reproductive cycle. These changes are mainly driven by the large periodic fluctuations of gonadal hormones such as estrogen (E2) and progesterone. In particular, E2 has been suggested to play both neurotrophic and neuroprotective actions in the brain, including modulation of reproductive behavior, synaptic plasticity, sensory perception, and cognition10. At the cellular level, recent studies have shown that PV cells activity fluctuates over the estrous cycle in females 11,12. One of these studies specifically demonstrates an important effect of estrogen in PV neurons excitability12. We believe that understanding how sex hormones regulate inhibitory circuits in sensory cortices will help explaining sex-differences observed in psychiatric drug efficacy. In our current study, we are investigating the role of sexual hormones in auditory perception. Our preliminary observation in mice suggests the existence of different network strategies in auditory perception, which could be related to different communication abilities in males versus females, pointing to the possibility that different ways of neural synchronization may rely on different circuits properties and that the use in clinic of these two different auditory tasks might reveal a specific alterations in auditory circuitry.

3.2 Scientific approaches

Approach 1. Chemogenetic approach to explore PV cell role in sensory processing alterations in Syngap1 haploinsufficient mice.

Adult Syngap1+/− mice and control littermates of both sexes are implanted with an array of eight microdrives and one optic fiber in primary auditory cortex (A1). This device allows recording local field potentials together with single units from the living cells surrounding the tip of the electrodes. After one week of recovery time, we record EEG activity at baseline and during specific sensory stimulation, as previously described4. Brain rhythms, inter-ictal spikes and sensory responses to auditory paradigms are evaluated during spontaneous movement, immobility and different sleep phases. To evaluate whether increasing inhibitory transmission, in particular PV interneurons activity, rescues neurological alterations in Syngap1−/+ mice, we use a chemogenetic approach. In particular, we employ PV-Cre mice that are also haploinsufficient for Syngap1 (through the breeding of PV-Cre knockin and Syngap1−/+ mice). We inject stereotaxically Cre-dependent hM3Dq-mCherry AAV or control Cre-dependent mCherry AAV in the basal forebrain of PV-Cre;

Syngap1−/+ mice (and control littermates). We then repeat EEG recordings following administration of CNO, which activates hM3Dq and increases the excitability of PV neurons.

Approach 2. Optotagging to evaluate cell-specific activity changes over the reproductive cycle in mice

In collaboration with Dr J Csicsvari (IST, Austria), I have leant to build microdrives, use optogenetics tools and perform spike activity analysis. Here, the use of optotagging allows us to accurately identify different neuronal types. We employ PV and SST reporter mice that are Syngap1 haploinsufficient (through the breeding of PV-Cre and SST-Cre knockin with Syngap1−/+ mice). Similar to our chemogenetic experiments, we then inject stereotaxically Cre-dependent ChR2-mCherry adeno-associated virus (AAV) or control Cre-dependent mCherry AAV in A1. In this case, subsequent light stimulation activates specifically PV or SST neurons allowing their identification during different sensory modality stimulation. Wide-band signals are high-passed (0.8-5kHz) and filtered for spike detection. Extracted spike waveforms are identified by their activation upon light stimulation and clustered on the basis of their spike amplitude and waveshape, using principal component analysis. This approach allows us to accurately identify fast-spiking PV and SST cells and to correlate their firing behavior to different sensory stimuli. By performing spike-local field potential coherence and phase locking analysis, we also study how well these cells are tuned to different frequency bands.

Approach 3. RNAScope to evaluate anatomic distribution of E2 receptors.

In order to evaluate E2 receptor expression in inhibitory and excitatory neurons of the corresponding sensory area, we use RNAscope (Advanced Cell Diagnostics) multichannel fluorescent in situ hybridization. We study the neuron-specific expression pattern of E2 receptors in different subtypes of glutamatergic and GABAergic neurons. Confocal images are acquired using a confocal microscope (LEICA SP8) and analyzed with image custom made scripts.

Approach 4. Recording from human subjects to evaluate the translational value of EEG patterns observed in mice.

We also apply a translational approach by recording EEG activity from human subjects during sensory tasks that are equivalent to those we use in mice. First, in order to identify the cognitive phenotype of our cohort, a neuropsychological assessment is also performed. Intellectual functioning of the subjects is evaluated by the Wechsler Adult Intelligence Scale (WAIS-IV > 16 years old). This IQ measure is taken into consideration in further statistical analysis to assess the correlation between intellectual functioning and different aspects of sensory perception, such as auditory entrainment. Later, EEG recordings are performed by using a 128-electrodes Geodesic headset. In order to reduce to minimum, the presence of electrical noise in the signal, EEG recordings are performed in a Faraday cage. Electrode impedances are kept at 40kOhms, as it has been suggested by Tucker 16. EEG signal is recorded at a frequency sampling of 2000 Hz. The software used for these experiments is NetStatio EEG Software v2.0. Screens and speakers used for auditory and visual stimulation are placed 60 cm from the subjects. Auditory and visual trigger generators assure the timing of the stimuli presentation.

Approach 5. Computational approach to reveal hidden patterns in EEG activity.

In order to decipher the fundamental codes of the brain's circuits underlying auditory perception we apply newly developed computational tools. Using deep learning, artificial neural networks and other classification tools we aim at identifying specific firing properties during different auditory tasks. We also apply neural decoding and complexity analysis of the population vector to identify what are the components of neural activity that carry information about the presented stimuli, and how they change upon the activation of different neuromodulatory systems. Furthermore, we also aim at reproducing a computational model that could explain what are the physiological consequences of Syngap1 haploinsufficiency at the populational levels during different auditory stimulations.

References

1. Carreno-Munoz, M. I. et al. Potential Involvement of Impaired BKCa Channel Function in Sensory Defensiveness and Some Behavioral Disturbances Induced by Unfamiliar Environment in a Mouse Model of Fragile X Syndrome. Neuropsychopharmacology 43, 492-502, doi:10.1038/npp.2017.149 (2018).
2. Weber, L. A. et al. Auditory mismatch responses are differentially sensitive to changes in muscarinic acetylcholine versus dopamine receptor function. bioRxiv, 2021.2003.2018.435979, doi:10.1101/2021.03.18.435979 (2021).
3. Parikh, V., Kozak, R., Martinez, V. & Sarter, M. Prefrontal acetylcholine release controls cue detection on multiple timescales. Neuron 56, 141-154, doi:10.1016/j.neuron.2007.08.025 (2007).
4. Carreno-Munoz, M. I. et al. Sensory processing dysregulations as reliable translational biomarkers in SYNGAP1 haploinsufficiency. Brain, doi:10.1093/brain/awab329 (2021).
5. Rudy, B., Fishell, G., Lee, S. & Hjerling-Leffler, J. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. Dev Neurobiol 71, 45-61, doi:10.1002/dneu.20853 (2011).

6. Natan, R. G. et al. Complementary control of sensory adaptation by two types of cortical interneurons. Elife 4, doi:10.7554/eLife.09868 (2015).
7. Hamm, J. P. & Yuste, R. Somatostatin Interneurons Control a Key Component of Mismatch Negativity in Mouse Visual Cortex. Cell Rep 16, 597-604, doi:10.1016/j.celrep.2016.06.037 (2016).
8. Al-Mana, D., Ceranic, B., Djahanbakhch, O. & Luxon, L. M. Alteration in auditory function during the ovarian cycle. Hear Res 268, 114-122, doi:10.1016/j.heares.2010.05.007 (2010).
9. Sumner, R. L. et al. Neural plasticity is modified over the human menstrual cycle: Combined insight from sensory evoked potential LTP and repetition suppression. Neurobiol Learn Mem 155, 422-434, doi:10.1016/j.nlm.2018.08.016 (2018).
10. Simpkins, J. W., Singh, M., Brock, C. & Etgen, A. M. Neuroprotection and estrogen receptors. Neuroendocrinology 96, 119-130, doi:10.1159/000338409 (2012).
11. Picard, N., Takesian, A. E., Fagiolini, M. & Hensch, T. K. NMDA 2A receptors in parvalbumin cells mediate sex-specific rapid ketamine response on cortical activity. Mol Psychiatry 24, 828-838, doi:10.1038/s41380-018-0341-9 (2019).
12. Clemens, A. M. et al. Estrus-Cycle Regulation of Cortical Inhibition. Curr Biol 29, 605-615 e606, doi:10.1016/j.cub.2019.01.045 (2019).

Chapter 4
Lifelong Learning for Robotics: Current Trends and Challenges

Bogdan Raducanu and Joost van de Weijer

Abstract As the robots move from the factory floor to peoples' homes they need to be endowed with near-human intelligence that support long-term interactions. This aim could be achieved by replicating the evolutionary mechanisms that allowed us to evolve. One of them is represented by the ability to learn and adapt to changing conditions during our lifetime. While for humans this is an innate ability, for machines to learn during their lifespan is still in its infancy and limited progress has been achieved so far. Therefore, the current paper aims to review the current trends in lifelong learning for robots and present the challenges that need to be overcome in the future in order to convert them in true companions.

4.1 Introduction

Creating robots showing human-like intelligence is still a very big challenge for the scientific community. Despite of the efforts devoted in the past decades, only partial, limited success has been achieved. The question that arises naturally at this point is: what did we miss in our strategy that prevented us in reaching the desired goal? The traditional paradigm in robotics (which can be found on a large scale nowadays) is represented by task-specific programming. Within this framework, the human designer is the one in charge to specify the task that must be performed by the robot, how it must be performed, its corresponding internal representation and the data structures associated to it. Being restricted from the very beginning by its creator, lacking the ability to 'understand' what, how or why it is doing, the robot has no possibilities to 'become' something more than the purpose it was originally created for.

Bogdan Raducanu and Joost van de Weijer
Computer Vision Center and UAB, Edifici "O" - Campus UAB, 08193 Bellaterra (Barcelona), Spain, e-mail: {bogdan,joost}@cvc.uab.es

On the other hand, developmental framework has always represented a cherished goal in robotics [39]. In this case, opposed to its predecessor, the robot should be endowed with developmental abilities specific to humans: it should evolve from an ‘infancy’ stage to ‘maturity’. This paradigm implies a completely different approach for robot programming. It can not be task-specific anymore because the tasks the robot will perform are unknown in the design phase. Furthermore, it should be able to generate automatic representations of unknown knowledge or skills. However, the human designer is responsible for providing it some basic, ‘building blocks’ of knowledge (the kind of knowledge similar to what we carry in our genetic material) that contain critical specifications that will allow the robot to learn and evolve during its existence. In terms of theory of evolution, the ‘genetic’ information provided will resemble the phylogenetic process taking part in biological systems. Phylogeny refers to those adaptations and modifications in an organism structure that occurs at large time scales (across several generations). On the other hand, the development taking place in a system during its existence is referred to as ontogenetic process. Ontogeny is driven by the phylogenetic factors, but it is also strongly influenced by the system’s surrounding environment. This aspect has been studied in detail by the psychologist Jean Piaget [27], whose work represents a landmark in the theory of cognitive development.

In a similar manner, the cognitive development of a robot should happen through direct, continuous interaction (making use of its perceptual mechanisms) with the environment in which it will be situated [36]. In consequence, the robot will create its own understanding about things and will build its own reality. The cognitive development is an open-ended, incremental and progressive process: a robot cannot learn complex skills successfully without first learning necessary simpler ones. Cognitive development was the

In this paper, we focus on lifelong learning as a key of how cognitive development could be achieved. In the next section, we review lifelong learning in biological and artificial systems. Section 3 will present some of the challenges we still have to overcome in order to deploy lifelong learning paradigm to robotics.

4.2 Lifelong learning

This section aims to briefly review the most relevant aspects of lifelong learning both in biological and artificial systems. A more in-depth survey could be found in [25].

4.2.1 Lifelong learning in biological Systems

The underlying biological mechanisms that are responsible for continual learning determined the development, at a higher, cognitive level, of two separate, interacting memory systems [21, 22, 23]. These findings were formalized as Complementary

Learning Systems (CLS) theory, according to which the hippocampus region of the brain is responsible for short-term memory and rapid learning of novel information and neocortex is seen as long-term memory storage. The information stored in hippocampus is replayed several times for its consolidation in neocortex [24], a process which occurs during the rapid eye movement (REM) stage of sleep [35, 5]. Although the hippocampus is responsible for replaying recent memories, once they are consolidated in neocortex, it is the role of prefrontal cortex (PFC) to recall them [12]. The hippocampus presents a high learning rate and encodes sparse representations to minimize interference meanwhile the neocortex is characterized by a low learning rate and builds overlapping representations of the learned knowledge. In conclusion, the interplay between hippocampus and neocortex functionality is crucial for continual learning.

As it can be appreciated from the above description of CLS, the brain developed robust mechanisms that allow the consolidation and protection of the stored information. Despite that, we are also prone to forget some of the previously learned information, during our lifetime. In general we rarely witness that novel information is interfering with consolidated knowledge (French, 1999) with the exception, of course, of some cognitive degenerative diseases. In the next section we review how these biological fundaments of continual learning have been implemented in computational models.

4.2.2 Lifelong learning in artificial systems

Although humans developed a robust mechanism for continual learning, AI systems based on deep learning do not present the same level of performance. Traditionally, neural networks are trained in a static way, assuming that all needed data is available in advance. However, in practice, this is seldomly the case. When trained to learn the tasks sequentially, they are prone to ‘catastrophic forgetting’, i.e. the most recent information replaces the old one, with a tremendous effect on network’s performance. Recent work has been devoted in proposing several solutions to overcome this shortcoming. Most of the solutions proposed are biologically-inspired and rely on the mechanisms presented in the previous section. They can be roughly divided in three main categories: regularization-based methods, memory replay-based methods, and architecture-based methods

Regularization-based methods.

A first group of approaches is based on weight regularization. The proposed solution in [11] (EWC - Elastic Weight Consolidation), follows a neuro-biological model according to which mammalian neocortex relies on a process of task-oriented synaptic consolidation. In other words, the knowledge about how to perform a previously acquired task is durably stored in a set of synapses which remain stable for long periods of time. An improved version of this algorithm (R-EWC) has been reported recently in [15], where the authors propose a network re-parametrization scheme which takes the form of a factorized rotation of parameter space. When combined with EWC, this approach leads to significant improvement. In [1], the importance of the network's parameters is estimated online and in an unsupervised manner by looking at the sensitivity of the output function instead of the loss. This way, changes to important parameters can then be penalized, preventing thus the loss of important knowledge related to previous tasks, when a new task is learned. Other regularization approaches focuses on preventing changes in activations by using knowledge distillation, like for instance Learning without Forgetting (LwF) [14]. Thus, when learning a new task, the previously learned tasks are prevented from drifting too much. A variant of LwF has been proposed in [8], where they propose to regularize on the cosine similarity between the L2-normalized logits of the previous and current network. Another work which builds upon LwF is [43]. In this case, the authors noticed that there is an asymmetry between old and new classes while training: new classes have explicit and strong supervision, whereas supervision for old classes is weaker and communicated by means of knowledge distillation.

Memory replay-based methods.

Another group of biologically inspired approaches uses a mechanism known as 'memory replay' or 'pseudo-rehearsal' [30] to prevent the catastrophic forgetting. It is called like this because it resembles the memory replay processes which take place in hippocampus. As part of this strategy, the task being learnt is trained jointly with samples from the previous tasks. These samples could be real ones, like in [28] where they are called 'exemplars', or they could be approximate representations of real ones which are generated using Generative Adversarial Networks (GANs) [6]. A similar approach to [28] has been pursued in [17], but in this case the exemplars are used to constrain gradients so that the approximated loss from previous tasks is not increased. An improved version of this approach, which overcomes some efficiency issues has been proposed in [3].

In [34] it is for the first time that the GANs are used to create synthetic samples for memory replay. The success of this approach demonstrated that there is no need to store real data, but this could be replaced by synthetic one generated online. The same idea has been extended to conditional GANs in [40]. A further extension is reported in [42], where they substitute label-conditioned GANs with image-conditioned GANs. As an alternative to memory replay with synthetic data, some other approaches propose to perform memory replay using features extracted from the intermediate levels of a fixed backbone network, like in [41, 16]. The advantage of feature replay relies in the fact that feature distribution is much simpler than image distribution. Therefore, features can be efficiently generated using smaller generators which require a reduced number of sample for training. Additionally, replaying features requires significantly less memory than images, which makes the approach scalable to larger datasets. Finally, several methods study compressing the features that are replayed, thereby significantly increasing the amount of exemplars that can be stored in a limited memory buffer [7, 37].

Architecture-based methods.

A first category of methods from this family are those known as mask-based methods which prevent catastrophic forgetting by isolating the parameters corresponding to a particular task. In other words, these methods apply a mask on weights or layer activations. In [19], by performing network pruning, they are able to pack in a sequential manner multiple tasks in the same network, without increasing its capacity. They exploit network's redundancies in order to free unused parameters and to allocate them to future tasks. This idea is further extended in [18] where they maintain a set of real-valued weights which are thresholded in order to obtain binary masks, that are then applied to existing weights. This strategy is improved in [20], where they use ternary masks, but which are applied on the activations instead of weights. This reduces considerably the number of mask parameters for each new task. The same strategy, i.e. applying a mask on activations has been pursued in [33]. Simultaneously to learning a new task, they also learn some binary attention vectors which define a mask. The masks from previous tasks constrain the updates of the network's weights on the current task. All the previous methods have the disadvantage that they are applied on fixed network, with a fixed capacity. This means they don't scale well when the number of tasks is high. Furthermore, with each new added task, the network capacity is decreasing, thus the performance of the last tasks is decreasing due to lack of enough capacity. In order to cope with these limitations, a second category of methods which involve dynamic architectures has been proposed. In [2], they sequentially add new lateral connections for each new task, and preserve the already learned tasks. A similar idea has been pursued in [31], where the network retains a pool of pretrained models throughout training, and learn lateral connections from these to extract useful features for the new task. This approach has been extended in [32], where they duplicate the network only once to keep the number of parameters constant and combine it with EWC [11] to prevent forgetting. Finally, in [4], they combine the concept of dynamic architecture with memory replay to learn task-specific and task-invariant features.

4.3 Challenges of lifelong learning for robotics

Although the above references represent an important breakthrough for biologically-inspired computational approaches, sustained effort is still required in order to get closer to human performance. For instance, one of the strong limitations is the fact that previous approaches have been tested on relatively small datasets. More experiments are required in order to confirm if they are indeed scalable to large-scale data. Below is a list of challenges we identified and that need to be overcome in order to endow robots with lifelong learning capabilities. For a more in-depth discussion, please refer to [13].

Scalability

Most of the existing algorithms are still limited to a relatively low number of tasks (order of tens as maximum). Real-world robots, existing in a real environment, should be able to cope perhaps with hundreds or perhaps thousands of tasks. None of the previous reviewed lifelong learning approaches could be applied directly in this case. Therefore, new strategies should be devised.

Hardware limitations

Currently, the most successful lifelong learning approaches need big clusters with several state-of-the-art GPUs for training. Usually, robot's hardware is significantly less powerful than a cluster. Since in the lifelong paradigm, the network requires continuously training it should be adapted and optimized in order to cope with the computational resources available onboard, both in terms of capacity and energy. Recent advances in Edge Computing and the release of deep learning framework for mobile devices like TinyML could pave the way for the deployment of lifelong learning approaches on systems with limited computational resources.

Data type

While still most of the lifelong learning methods require labeled data, this would be not possible in the case of the robots. It is assumed they are going to exist in dynamic environments. The lifelong learning implies a continuous, online learning process: the robot can not be shut down and updated offline. Therefore, it has to cope with unlabeled data and decide what is relevant or not. Recent advances in active learning [29] and self-supervised learning [10] have recently shown promising results.

Self supervision

The lifelong learning paradigm for robotics should not be task-specific, therefore it should not require human supervision. The robot should be able to figure out and to adapt to new tasks for which it was not programmed in the beginning. The lifelong learning for robotics should be also an open-ended cumulative learning process, following the curriculum learning paradigm [38], i.e. the robot should gradually learn from simple to complex tasks. Reinforcement learning-based techniques could tackle these problems, but for lifelong learning it has mainly applied in simulation scenarios, e.g. video games [26, 9].

4.4 Conclusions

Although learning continuously throughout our existence is an innate ability for humans, deploying it to artificial systems, in general, and in robots, in particular, is not straightforward. One of the major barriers to overcome is represented by catastrophic forgetting, i.e. a phenomenon which refers to the loss of previously learned knowledge when new information arrives. In this paper we have postulated that lifelong learning could be a possible paradigm in our quest to develop robots showing human-like intelligence. We have reviewed the most relevant approaches of in the field together with the mechanisms to prevent catastrophic forgetting. Finally, we have identified some of the challenges we need to overcome in order to deploy lifelong learning strategies in robots.

4.4.0.1 Acknowledgements

We acknowledge the support from European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement number 777720 (CybSpeed).

References

1. Aljundi, R., Baniloni, F., Eñhoseiny, M., Rohrbach, M., Tuytelaars, T.: Memory aware synapses: Learning what (not) to forget. In: Proc. of ECCV. pp. 139–154. Munich, Germany (2018)
2. Aljundi, R., Chakravarty, P., Tuytelaars, T.: Expert gate: Lifelong learning with a network of experts. In: Proc. of CVPR. pp. 3366–3375. Honolulu (HI), USA (2017)
3. Chaudhry, A., Ranzato, M., Rohrbach, M., Elhoseiny, M.: Efficient lifelong learning with agem. In: Proc. of ICLR. Vancouver, Canada (2019)
4. Ebrahimi, S., Meier, F., Calandra, R., Darrell, T., Rohrbach, M.: Adversarial continual learning. In: Proc. of ECCV. pp. 386–402 (2020)
5. Gais, S., Albouy, G., Boly, M., Dang-Vu, T., Darsaud, A., et al.: Sleep transforms the cerebral trace of declarative memories. Proceedings of the National Academy of Sciences **104**(47), 18778–18783 (2007)
6. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversial nets. In: Proc. of NIPS. Montreal, Canada (2014)
7. Hayes, T.L., Kafle, K., Shrestha, R., Acharya, M., Kanan, C.: Remind your neural network to prevent catastrophic forgetting. In: European Conference on Computer Vision. pp. 466–483. Springer (2020)
8. Hou, S., Pan, X., Loy, C.C., Wang, Z., Lin, D.: Learning a unified classifier incrementally via rebalancing. In: Proc. of CVPR. pp. 831–839. Long Beach (CA), USA (2019)
9. Isele, D., Cosgun, A.: Selective experience replay for lifelong learning. In: Proc. of AAAI. pp. 3302–3309. New Orleans (LA), USA (2018)
10. Jaiswal, A., Babu, A., Zadeh, M., Banerjee, D., Makedon, F.: A survey on contrastive self-supervised learning. arXiv preprint arXiv:2011.00362v3 (2021)
11. Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A.A., et al.: Overcoming catastrophic forgetting in neural networks. Proceedings of the National Academy of Sciences **114**(13), 3521–3526 (2017)

12. Kitamura, T., Ogawa, S., Roy, D., Okuyama, T., Morrissey, M., Smith, L., Redondo, R., Tonegawa, S.: Engrams and circuits crucial for systems consolidation of a memory. Science **356**(6333), 73–78 (2017)
13. Lesort, T., Lomonaco, V., Stoian, A., Maltoni, D., Filliat, D., Diaz-Rodriguez, N.: Continual learning for robotics: Definition, framework, learning strategies, opportunities and challenges. Information Fusion **58**, 52–68 (2020)
14. Li, Z., Hoiem, D.: Learning without forgetting. IEEE Transactions on PAMI **40**(12), 2935–2947 (2018)
15. Liu, X., Masana, M., Herranz, L., van de Weijer, J., Lopez, A., Bagdanov, A.: Rotate your networks: Better weight consolidation and less catastrophic forgetting. In: Proc. of ICPR. pp. 2262–2268. Beijing, China (2018)
16. Liu, X., Wu, C., Menta, M., Herranz, L., Raducanu, B., Bagdanov, A., Jui, S., van de Weijer, J.: Generative feature replay for class-incremental learning. In: Proc. of CVPR Workshops (2020)
17. Lopez-Paz, D., Ranzato, M.: Gradient episodic memory for continual learning. In: Proc. of NeurIPS. pp. 6467–6476. Long Beach (CA), USA (2017)
18. Mallya, A., Davis, D., Lazebnik, S.: Piggyback: Adapting a single network to multiple tasks by learning to mask weights. In: Proc. of ECCV. pp. 67–82. Munich, Germany (2018)
19. Mallya, A., Lazebnik, S.: Packnet: Adding multiple tasks to a single network by iterative pruning. In: Proc. of CVPR. pp. 7765–7773. Salt Lake City (UT), USA (2018)
20. Masana, M., Tuytelaars, T., van de Weijer, J.: Ternary feature masks: continual learning without any forgetting. In: Proc. of CVPR Workshops (2021)
21. McClelland, J., McNaughton, B., O'Reilly, R.: Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. Psychological Review **102**(3), 419–457 (1995)
22. O'Reilly, R., Norman, K.: Hippocampal and neocortical contributions to memory: Advances in the complementary learning systems framework. Trends in Cognitive Sciences **6**(12), 505–510 (2002)
23. O'Reilly, R., R., B., M.D., H., Ketz, N.: Complementary learning systems. Cognitive Science **38**, 1229—-1248 (2014)
24. O'Neill, J., Pleydell-Bouverie, B., Dupret, D., Csicsvari, J.: Play it again: reactivation of waking experience and memory. Trends in Neurosciences **33**(5), 220–229 (2010)
25. Parisi, G., Kemker, R., Part, J., Kanan, C., Wermter, S.: Continual lifelong learning with neural networks: A review. Neural Networks **113**, 54–71 (2019)
26. Pathak, D., Agrawal, P., Efros, A., Darrell, T.: Curiosity-driven exploration by self-supervised prediction. In: Proc. of ICML. pp. 2778—-2787. Sydney, Australia (2017)
27. Piaget, J.: The Origins of Intelligence in Children. Norton & Co. Publishers, New York, NY, USA (1952)
28. Rebuffi, S.A., Kolesnikov, A., Sperl, G., Lampert, C.H.: icarl: Incremental classifier and representation learning. In: Proc. of CVPR. pp. 2001–2010. Honolulu (HI), USA (2017)
29. Ren, P., Xiao, Y., Chang, X., et al.: A survey of deep active learning. arXiv preprint arXiv:2009.00236v2 (2021)
30. Robins, A.: Catastrophic forgetting, rehearsal and pseudorehearsal. Connection Science **7**(2), 123–146 (1995)
31. Rusu, A., Rabinowitz, N., Desjardins, G., et al.: Progressive neural networks. arXiv preprint arXiv:1606.04671 (2016)
32. Schwarz, J., Luketina, J., Czarnecki, W., Grabska-Barwinska, A., Teh, Y., Pascanu, R., Hadsell, R.: Progress & compress: A scalable framework for continual learning. In: Proc. of ICML. Stockholm, Sweden (2018)
33. Serra, J., Suris, D., Miron, M., Karatzoglou, A.: Overcoming catastrophic forgetting with hard attention to the task. In: Proc. of ICML. Stockholm, Sweden (2018)
34. Shin, H., Lee, J.K., Kim, J., Kim, J.: Continual learning with deep generative replay. In: Proc. of NeurIPS. Long Beach (CA), USA (2017)
35. Taupin, P., Gauge, F.: Adult neurogenesis and neural stem cells of the central nervous system in mammals. Journal of neuroscience research **69**(6), 745–749 (2002)

36. Vygotsky, L.: Thought and Language. MIT Press, Cambridge, MA, USA (1962)
37. Wang, K., van de Weijer, J., Herranz, L.: Acae-remind for online continual learning with compressed feature replay. Pattern Recognition Letters **150**, 122–129 (2021)
38. Wang, X., Chen, Y., Zhu, W.: A survey on curriculum learning. arXiv preprint arXiv:2010.13166v2 (2021)
39. Weng, J., McClelland, J., Pentland, A., Sporns, O., Stockman, I., Sur, M., Thelen, E.: Autonomous mental development by robots and animals. Science **291**(5504), 599–600 (2001)
40. Wu, C., Herranz, L., Liu, X., Wang, Y., van de Weijer, J., Raducanu, B.: Memory replay gans: learning to generate images from new categories without forgetting. In: Proc. of NeurIPS. Montreal, Canada (2018)
41. Xiang, Y., Fu, Y., Ji, P., Huang, H.: Incremental learning using conditional adversarial networks. In: Proc. of ICCV. pp. 6619–6628. Seoul, South Korea (2019)
42. Zhai, M., Chen, L., Tung, F., He, J., Nawhal, M., Mori, G.: Lifelong gan: Continual learning for conditional image generation. In: Proc. of ICCV. pp. 2759–2768. Seoul, South Korea (2019)
43. Zhang, J., Zhang, J., Ghosh, S., Li, D., Tasci, S., Heck, L., Zhang, H., Kuo, C.C.J.: Class-incremental learning via deep model consolidation. In: Proc. of WACV. pp. 1131–1140. Snowmass Village (CO), USA (2020)

Chapter 5
An Overview of Cognitive Architectures from the perspective of Lifelong Open-ended Learning Autonomy

R. J. Duro, A. Romero and F. Bellas

Abstract In this paper we analyze a large set of very well-known cognitive architectures in the literature from the perspective of their capabilities for Lifelong Open-ended Learning Autonomy. In this line we have consider whether they contemplate learning, contextual memory systems, motivations or attention and try to provide an idea of whether they were applied to real robot scenarios. As it turns out, none of them are completely ready to address this challenge, but some of them do provide some indications on the paths to follow in some of the aspects they contemplate.

5.1 Introduction

In this paper we are concerned with the idea of making robots autonomous. Autonomy implies the capability of handling variability successfully and robustly in the situations/domains that robots face and that were not considered at design time. Variability can be described and addressed at different levels: From slight variations in the domains the robot operates in, where it is only necessary to adapt a skill the robot already has; to the designer changing the goal the robot must achieve in a given domain, in which case the robot must master a new skill; to a more difficult situation in which domains change and the robot must find its own goals and learn to master the skills to achieve them consistently. This last problem is generally called the open-ended learning (OEL) problem, that is, a robot must be able to learn to operate in domains that were unknown at design time [1], and thus, it must possess open-ended learning autonomy. Additionally, when the domains change continuously and unpredictably, often effectively preventing the robot from being able to

R. J. Duro, A. Romero and F. Bellas
Integrated Group for Engineering Research, CITIC (Centre for Information and Communications Technology Research), Universidade da Coruña, Spain e-mail: {richard, Alejandro.romero.montero, francisco.bellas}@udc.es

master a domain in a single shot, we would be facing a lifelong open-ended learning problem, requiring lifelong open-ended learning autonomy (LOLA).

Increasing the level of autonomy of robotic systems up to LOLA involves jointly solving the problems of open-ended learning [1] and lifelong learning [2]. Dealing with these problems goes beyond specific learning algorithms. It requires the ability to manage all the knowledge that is learned so that it can be contextually related and reused, thus facilitating further learning and exploitation. Furthermore, for robots to learn in complex and unfamiliar domains, it is also necessary to manage their motivations. As well as considering other processes, such as attention, representation, and learning. Thus, for robots to autonomously learn to operate in domains that were not considered at design time and build on this knowledge to address new domains as they emerge during their lifetime, the capabilities mentioned above (and probably some more) must be integrated and regulated. This is the job of cognitive architectures, and these have been studied for decades. However, the purposes for which each one was created was different and usually not related directly to LOLA. Consequently, it makes sense to carry out a brief overview of the main cognitive architectures found in the literature and characterize them in terms of the level of autonomy they allow and their possible adequation to LOLA.

5.2 Cognitive Architectures and LOLA

Cognitive architectures are structures that artificially implement cognition [3]. They allow learning, storing, using and reusing knowledge and can also contemplate developmental or other integration strategies to produce new high-level knowledge nuggets from the elements stored in their memory. Many types of cognitive architectures have been developed over the past four decades, each one addressing different aspects of cognition. They generally base their operation on their abilities to interpret, index and sort the different knowledge nuggets they require based on their content. However, they follow different approaches to this end. Based on [4], cognitive architectures can be classified into 3 basic groups according to the type of representations they can manipulate:

- Symbolic: Most of the original versions traditional cognitive architectures (ACT-R [5], CLARION [6], 4CAPS [7] or SOAR [8]) belong to this group, although some of them have been later hybridized. These architectures are characterized by representing concepts through symbols and having predefined instruction sets to manipulate them. This makes them excellent systems in terms of planning and reasoning. But for the same reason, they present grounding problems and lack the robustness and flexibility needed to adapt to the changing conditions of real environments. In addition to the fact that the designer assumes a high degree of knowledge about the domains and tasks to be performed and therefore provides a lot of knowledge in the form of specific knowledge or even complete sets of rules in some cases. They are therefore not really ready to tackle the LOLA problem.

- Emergent: They are based on sub-symbolic or connectionist approaches. They often follow developmental principles that seek to progressively build system knowledge from scratch through direct interaction with the world. In these, knowledge is often represented and distributed through neural networks. This approach provides a direct path to the autonomous construction of high-level knowledge without grounding problems. Thus, they aim to solve the problems of adaptation to the environment and learning through the concatenation of multiple models in parallel, where information flows through activation signals. However, this introduces a high level of complexity in development and the need for very long learning and interaction processes. It also causes the system to lose transparency, as knowledge is no longer represented by well understood symbols and is distributed throughout the network. Some examples of these architectures are MDB [9], GRAIL [10] or SASE [11].
- Hybrid: Finally, the group of hybrid architectures consists of those that use symbolic representations at the higher processing levels, but include sub-symbolic representations at the low level, just like the emerging connectionist paradigms. These approaches have become quite popular for addressing low-level grounding and domain adaptation problems, but they still require many adjustments to construct symbolic information. In fact, even though there are researchers trying to provide autonomous approaches to bridge the gap between sub-symbolic and symbolic representations [12], they are still not very common in cognitive architectures. This makes these architectures difficult to adapt to general use cases in robotics and their implementations tend to focus on specific functionalities. Recent examples of such architectures are OpenCogPrime [13] or MLECOG [14], or even extensions of more traditional architectures such as ACT-R [15].

From the perspective of Lifelong Open-ended Learning Autonomy, however, and from a functional viewpoint, we must bear in mind that architectures should contemplate, at least, the following components:

- A motivational system that enables open-ended learning, i.e., that allows the robot to search for new goals and select which ones are active each instant of time.
- A memory system that permits storing the acquired knowledge and relate it contextually to facilitate its reuse it in the right conditions so that lifelong learning is made possible.
- An on-line learning system that facilitates acquiring knowledge on the different goals discovered in the different domains, as well as about how to achieve them during robot operation (skills).
- Some type of attention system that helps to reduce the sensory and processing overload of the system would also be very convenient.

Table I provides a summary of the characteristics of the fifteen most relevant architectures analyzed and the requirements they meet in order to achieve LOLA. The criteria for selecting these architectures were that their development is still ongoing and that they have practical applications. In the following subsections we consider the different components mentioned above.

Table 5.1 Features of cognitive architectures

Architecture	Type (following [4])	Design objective	Motivational system Level	Learning system	Contextual memory	Attention mechanism	Real robot applications
EPIC	Symbolic	Emulate human cognition	0	NO	NO	YES	NO
ICARUS	Symbolic	Robotics	1	YES (rule-based)	NO	NO	NO
ADAPT	Hybrid	Computer vision	0	YES (rule-based)	NO	YES	NO
CLARION	Hybrid	Emulate human cognition	1	YES	YES	NO	NO
LIDA	Hybrid	Emulate human cognition	1	YES	YES	YES	NO
iCub	Hybrid	Robotics	3	YES	NO	YES	YES
SOAR	Hybrid	Robotics	2	YES	SI	NO	YES
OpenCogPrime	Hybrid	Artificial General Int. (AGI)	1	YES	YES	NO	NO
DUAL	Hybrid	Emulate human cognition	0	YES (rule-based)	YES	NO	YES
4CAPS	Hybrid	Psychology Emulate human cognition	0	NO	NO	NO	NO
ACT-R	Hybrid	Emulate human cognition	0	YES	YES	YES	YES
MLECOG	Hybrid	Seek Autonomy	2	YES	YES	YES	NO
IMPACT	Hybrid	Robotics	3	YES	NO	NO	NO
MicroPSI	Emergent	Seek autonomy	2	YES	NO	YES	NO
GRAIL	Emergent	Robotics	3	YES	NO	NO	YES
MDB	Emergent	Robotics	2	YES	NO	NO	YES
SASE	Emergent	Model learning	0	YES	NO	NO	YES

5.2.1 Learning

Starting with the last component, and even though it may seem rather obvious, it is important to note that possessing the ability to learn is fundamental to be able to address LOLA. However not all cognitive architectures presented in the literature show this capability. In fact, many symbolic architectures do not implement learning mechanisms and, therefore, their knowledge must be introduced by the designer when they are built. This implies that the domains in which the robot will operate must be known at design time, contradicting the open-ended learning principle. Examples of these are EPIC [16] or 4CAPS [7]. In addition, others within the symbolic or hybrid group, such as DUAL [17], ADAPT [18], ICARUS [19], incorporate learning capabilities, but mostly through top-level rule modification, without a versatile and unrestricted ability to create new rules for new domains. Only within the group of emerging cognitive architectures, such as iCub [20] or MDB [9], and in a small group of hybrid architectures, such as MicroPSI [21], versatile low-level learning mechanisms can be found. Consequently, only these types of architectures would be candidates for achieving LOLA in robots from the point of view of learning.

5.2.2 Motivational System

On the other hand, when we talk about a cognitive architecture having a motivational system, we are talking about having a mechanism in charge of determining what the robot should strive for in a given domain each moment in time. This mechanism can have different functionalities, from being able to guide the robot towards the achievement of a goal, to allowing for the selection of which goal/goals are active at each moment in time, or even to be able to guide the robot to discover new goals. It is worth remembering that possessing these three qualities is what may allow the robot to be able to perform OEL. Considering these qualities, a classification of motivational systems into different levels can be established. The levels we consider in this paper are the following:

- Level 0: The robot has a specific goal set by the designer and the motivational system is able to guide the robot towards the achievement of that goal.
- Level 1: The robot has a series of goals set in advance by the designer and the motivational system is able to select which goals should be active at any given moment in time and guide the robot towards their achievement.
- Level 2: The robot has a series of goals set in advance by the designer and the motivational system is able to select which goals are active at any given moment in time and is capable of autonomously generating sub-goals to reach those goals.
- Level 3: The goals/domains are not known at the time of design and the motivational system is able to discover goals, select which ones are active and guide the robot towards their achievement.

At level 0 we find all the architectures that do not have an explicit motivational system, since all of them allow the systems they control to achieve the specific objective for which they are designed. Examples are symbolic architectures such as EPIC [16], hybrid architectures such as ADAPT [18] and DUAL [17], or emergent architectures such as SASE [11].

If we go to level 1, we have examples such as CLARION [6], which uses a motivational system based on drives. These drives have goals associated to them beforehand, so that the activations of the goals depend on the value of the drives. In other architectures such as LIDA [22], the motivations of the system are set in the form of artificial sensations and emotions. This allows it to appropriately select its goals and, consequently, the actions with which to act on the environment. Something similar to what happens in OpenCogPrime [13], where human motivations of feelings and beliefs are modeled through a motivational system based on the concepts of magicians and anti-magicians.

At level 2, there are multiple different implementations of motivational systems. Architectures such as SOAR have a motivational system that allows them to generate their own subgoals from goals predefined by the designer [8] as a previous step to be able to address a problem. MDB [9] also allows the intrinsic change of goals or motivations, and the generation of subgoals by introducing a satisfaction model. In MDB [9], the degree of fulfillment of motivations is based on both internal and external perceptions of the agent. This is similar to how MLECOG [14] handles motivations and their action choices based on pain/need (and other factors such as distance and availability). Moreover, in MLECOG only a few motivations are given to the system, with all others being developed internally. On the other hand, MicroPSI [21] also has a motivational system based on needs and drives, so that MicroPsi agents use pleasure/distraction signals related to the satisfaction of those drives. Finally, in the iCub [20] architecture it is the affective state that is in charge of providing the motivational cues. Thus, it has affective factors (motivations) that allow it to acquire knowledge and validate it.

Finally, only two of the architectures found in the literature present a motivational system that could be suitable for carrying out open-ended learning. GRAIL [10] and IMPACT [23], present motivational systems composed of intrinsic motivations based on competence. These systems allow them to autonomously learn new skills based on self-generated goals driven by intrinsic motivations (intrinsic goals).

5.2.3 Contextual Memory

If we look at the type of memory systems presented by the different architectures, we can distinguish two main groups. On the one hand, we have a series of architectures that have a computer-type memory, i.e. they see the memory as a hard disk. In them, all the knowledge generated is stored under a label. In this group we can find architectures such as EPIC [16], 4CAPS [7] or SASE [11]. On the other hand, we have a series of architectures that present an associative memory system, more similar to what would be the memory of a human. These associative memories are characterized by the fact that they are able to relate knowledge through the context in which it can be used. Thus, as discussed in the previous section, they would be the most appropriate to be able to achieve LOLA. This group includes architectures such as ACT-R [15], MLECOG [14], OpenCogPrime [13] or CLARION [6]. However, in these architectures the contextual associations are implemented by hand by the designer and are not created autonomously, thus defeating the purpose of LOLA.

5.2.4 Attention

Finally, another component present in most of the reviewed architectures, and that could help to cope with LOLA, are attention mechanisms. Attention is necessary to reduce the amount of sensory information for real-time operation and select the sensory information most relevant to the current situation. Attention allows the architecture to manage real-time operation, reducing the amount of information processed and, consequently, the reaction time of the system. Examples of architectures with attention mechanisms are MicroPSI [21], LIDA [22], MLECOG [14] or iCub [20].

5.3 Discussion

Most of the architectures discussed abover were not created with the objective of achieving a higher level of autonomy, but with the objective of demonstrating/imitating human behaviors. Moreover, the existing architectures work with intelligent agents, but not always with real robots. So, they are not really prepared to work in real environments and with continuous perceptual spaces. Only some emergent architectures have been tested using real robots in laboratory experiments to verify specific cognitive functionalities [20]. Additionally, the fact that none of the existing cognitive architectures has been explicitly designed to address the LOLA problem, implies that most lack some of the necessary components/functionalities to be able to achieve it. Many implement motivational systems, although most of them are not prepared to deal with open-ended learning and remain at lower levels of autonomy. Also, some of the architectures include attention systems and have low-level learning mechanisms. However, very few of them include an associative memory capable of handling contexts and these are usually constructed by the desinger. Finally, none of them addresses the issue of representation learning.

On the other hand, although there are no architectures explicitly designed to address the problem of achieving LOLA in a general way, there is qute a bit of work on different aspects of this field. Thus, there are examples in the literature of architectures such as GRAIL [10] or IMPACT [23], that have been tested in different OEL problems. However, they are tested in simulations [23] or they only address a specific part of the robotic system and therefore cannot be translated to reality [10]. While other architectures such as ACT-R have faced knowledge reuse problems [24], although without using real robots and starting from knowledge previously introduced by the designer.

Finally, it is interesting to comment that all architectures have implicitly or explicitly assumed that robot cognitive systems are given specific and appropriate state-space representations by their designers. That is, designers decide what is relevant from the robot's sensory flow and how these relevant features are represented. Consequently, the learning mechanisms for architectures have focused on how to learn whatever knowledge components the architectures require (direct or inverse models, utility models, policies, etc.) using these predefined state-space representations. So it seems that it would also be important to start addressing the issue of learning representations within the framework of cognitive architectures.

Although not all these blocks have been tackled under the cognitive architectures approach, they have been addressed to a greater or lesser extent in other fields. In this line, the Intrinsically Motivated Open-Ended Learning (IMOL) framework has made great contributions towards achieving agents capable of operating in an open-ended manner and autonomously acquiring knowledge and skills to solve tasks that are not known at design time. These approaches have been used in a wide variety of applications such as state-space exploration, knowledge gathering, autonomous skill learning or autonomous goal selection. However, despite these advances, IMOL systems are still difficult to use in real world applications. This is because these approaches are designed to acquire the maximum possible knowledge from the interaction of the robot with the environment but without considering the purpose for which the robot was designed. This results in an unbounded and unfocused learning that is not adapted to the specific needs of a service robot. A solution to this problem could consist of a motivational mechanism capable of considering and balancing different topologies of motivations. However, this topic is still under study. Moreover, another problem that is not yet solved is the design of specific motivations to trigger representation/redescription processes. It is important to look for motivational mechanisms that allow for seeking better representations, when necessary, since it is something that, as has been commented before, is critical for facilitating learning.

Regarding lifelong learning, fields such as Transfer Learning or Continuous Learning present very promising approaches to the problems of knowledge reuse and task learning in multiple domains, respectively. These approaches have proven to be effective for Deep Learning or Supervised Learning. However, they are not yet applicable to real robotic problems, since the former are not able to solve the issue of catastrophic forgetting, while in the latter the tasks to be performed and the domains of operation of the robot must be known in advance by the designer. Thus, not being intended for LOLA problems, they do not fully cover their needs.

5.4 Conclusions and Perspective

Most current applications of autonomous robots consider a very limited range of autonomy, usually dealing with a limited number of unexpected disturbances in the domain the robot is designed for. They seldom face the problem of autonomously setting goals in previously unknown domains (open-ended autonomy) nor, consequently, using experience from previous domains (Lifelong open-ended autonomy). Cognition and cognitive architectures have been purported as a way to address problems that require higher levels of autonomy, however the mostly programmed-in symbolic representations of traditional general purpose cognitive architectures are not up to the task due to their grounding and domain adaptation problems. Hybrid approaches, on the other hand, have become quite popular to address grounding and domain adaptation at a low level, but they require a lot of tweaking of the symbolic information in the higher levels, thus generally making them inadequate for open-ended learning situations. Finally, most emergent cognitive approaches have never been completely integrated into full cognitive architectures or tested on real market use-cases. In fact, most develpments are incomplete and only address a specific part of the robotic system and thus require more work to be ported to reality.

In general, there has been a lot of work on several aspects pertaining Lifelong Open-ended Learning Autonomy, but generally within areas outside the cognitive architecture realm and hardly ever considering the problem as a whole. These works go from intrinsically motivated structures to provide for goal discovery, to different approaches to knowledge reuse. Thus, the open question that needs to be addressed now is how to integrate this work within operational cognitive architectures that provide the structural components and the internal operational mechanisms needed to achieve the LOLA objective in a way that does not constrain their performand and possibilities.

Acknowledgments

This work has been partially funded by the Xunta the Galicia (grant EDC431C-2021/39). We wish to acknowledge the support received from the Centro de Investigación "CITIC", funded by Xunta de Galicia and the European Union (European Regional Development Fund-Galicia 2014-2020 Program), by grant ED431G 2019/01.

References

1. S. Doncieux et al., “Open-ended learning: a conceptual framework based on representational redescription,” Front. Neurorobot., vol. 12, p. 59, 2018.
2. S. Thrun and T. M. Mitchell, “Lifelong robot learning,” Rob. Auton. Syst., vol. 15, no. 1–2, pp. 25–46, 1995.
3. D. Vernon, Artificial cognitive systems: A primer. MIT Press, 2014.
4. I. Kotseruba and J. K. Tsotsos, “A Review of 40 Years of Cognitive Architecture Research: Core Cognitive Abilities and Practical Applications,” Artif. Intell. Rev., vol. 53, no. 1, pp. 17–94, 2020.
5. C. Lebiere and J. R. Anderson, “A connectionist implementation of the ACT-R production system,” in Proceedings of the fifteenth annual conference of the Cognitive Science Society, 1993, pp. 635–640.
6. R. Sun, “The importance of cognitive architectures: An analysis based on CLARION,” J. Exp. Theor. Artif. Intell., vol. 19, no. 2, pp. 159–193, 2007.
7. S. Varma and M. A. Just, “4CAPS: An Adaptive Architecture for Human Information Processing.,” in AAAI Spring Symposium: Between a Rock and a Hard Place: Cognitive Science Principles Meet AI-Hard Problems, 2006, pp. 91–96.
8. J. E. Laird, A. Newell, and P. S. Rosenbloom, “SOAR: An architecture for general intelligence,” Artif. Intell., vol. 33, no. 1, pp. 1–64, 1987.

9. F. Bellas, R. J. R. J. Duro, A. Faina, and D. Souto, “Multilevel Darwinist Brain (MDB): Artificial Evolution in a Cognitive Architecture for Real Robots,” IEEE Trans. Auton. Ment. Dev., vol. 2, no. 4, pp. 340–354, 2010.
10. V. G. Santucci, G. Baldassarre, and M. Mirolli, “Grail: a goal-discovering robotic architecture for intrinsically-motivated learning,” IEEE Trans. Cogn. Dev. Syst., vol. 8, no. 3, pp. 214–231, 2016.
11. J. Weng, “Developmental robotics: Theory and experiments,” Int. J. Humanoid Robot., vol. 1, no. 2, pp. 199–236, 2004, [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1
12. J. E. Laird, “Toward Cognitive Robotics,” Proc. SPIE, vol. 7332, pp. 73320Z-73320Z–11, 2009.
13. B. Goertzel, “OpenCogPrime: A cognitive synergy based architecture for artificial general intelligence,” in 8th IEEE International Conference on Cognitive Informatics, 2009, pp. 60–68.
14. J. A. Starzyk, S. Member, and J. Graham, “MLECOG: Motivated Learning Embodied Cognitive Architecture,” pp. 1–12, 2015.
15. J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, “An integrated theory of the mind.,” Psychol. Rev., vol. 111, no. 4, pp. 1036–60, 2004, [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/15482072.
16. D. E. Kieras, G. H. Wakefield, E. R. Thompson, N. Iyer, and B. D. Simpson, “Modeling Two-Channel Speech Processing With the EPIC Cognitive Architecture,” Top. Cogn. Sci., vol. 8, no. 1, 2016.
17. A. Nestor and B. Kokinov, “Towards Active Vision in the DUAL Cognitive Architecture,” BulDML Inst. Math. Informatics, vol. 11, no. 1, 2004.
18. D. P. Benjamin, D. M. Lyons, and D. W. Lonsdale, “ADAPT: A Cognitive Architecture for Robotics.,” in ICCM, 2004, pp. 337–338.
19. P. Langley and D. Choi, “A unified cognitive architecture for physical agents,” in Proceedings of the National Conf.e on Artificial Intelligence, 2006, vol. 2.
20. N. G. Tsagarakis et al., “iCub: the design and realization of an open humanoid platform for cognitive and neuroscience research,” Adv. Robot., vol. 1, pp. 1–25, 2007. [21] J. Bach, “MicroPsi 2: The next generation of the MicroPsi framework,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012, vol. 7716 LNAI.
21. D. Friedlander and S. Franklin, “LIDA and a Theory of Mind,” Proceeding 2008 Conf. Artif. Gen. Intell. 2008 Proc. First AGI Conf., no. Franklin, pp. 137–148, 2008, [Online]. Available: http://portal.acm.org/citation.cfm?id=1566188.
22. A. Oddi et al., “Integrating open-ended learning in the sense-plan-act robot control paradigm,” in ECAI 2020, IOS Press, 2020, pp. 2417–2424.
23. D. D. Salvucci, “Integration and reuse in cognitive skill acquisition,” Cogn. Sci., vol. 37, no. 5, pp. 829–860, 2013.

Chapter 6
An Innovative Water Index and Convolution Network for Water Body Segmentation in Remote Sensing Images

Mateusz Starczyk and Anna Kamińska-Chuchmała and Manuel Graña

6.1 Introduction

Segmentation of water regions in multispectral remote sensing images has been a research topic for a long time. Some water indices combining information from the various spectral bands have been proposed and tested in the literature. Usually, water body segmentation is done by thresholding the water index images. Recently, some machine learning proposals have been reported, namely approaches based on deep learning architectures. Specifically, the combined use of water indices and raw remote sensing data as inputs to the neural networks has show improved segmentation results. Following this trend, we propose a new water index combining visible and infrared bands that improves the segmentation results over the state of the art when validated over a quite general dataset.

In this introductory section we will discuss first the state of the art regarding water indices for the segmentation of water bodies in remote sensing images. Secondly, we comment of the machine learning and deep learning alternatives. Finally, we discuss available datasets for computational experiments and their shortcomings for the intended work on validating new water body segmentation. In following sections we present the materials and methods, the achieved validation results, and, finally, our conclusions and lines of future work.

Mateusz Starczyk and Anna Kamińska-Chuchmała
Department of Computer Science and Systems Engineering, Faculty of Information and Communication Technology, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland, e-mail: anna.kaminska-chuchmala@pwr.edu.pl, starczyk.mateusz@gmail.com

Manuel Graña
Computational Intelligence Group, Computer Science Faculty, University of the Basque Country, UPV/EHU, 00685 San Sebastián, Spain; e-mail: manuel.grana@ehuu.es

6.1.1 Water indices in remote sensing images

McFeeters in his pioneering work [1] presented the first water index, the Normalized Difference Water Index (NDWI) designed to highlight water bodies in satellite images. It uses reflected near-infrared radiation and visible green light to eliminate soil and vegetation features. The index is determined as the quotient of the difference and the sum of green and near-infrared light. It is formally defined as follows:

$$NDWI = \frac{Green - NIR}{Green + NIR} \tag{6.1}$$

The shortcomings of the above indice motivated the development of a new water index, the Augmented Normalized Difference Water Index (ANDWI) [2] defined as follows:

$$ANDWI = \frac{Blue + Green + Red - NIR - SWIR_1 - SWIR_2}{Blue + Gree + Red + NIR + SWIR_1 + SWIR_2}, \tag{6.2}$$

In summary, ANDWI classifies as water pixels that have higher reflectance in the visible bands than in the infrared bands, which, for example, can help to cancel out the false positives generated by metal roofs of buildings in urban areas, which have higher shortwave infrared reflectance than pure water, but still not higher than the reflectance of the RGB bands.

Appropriate selection of thresholds to distinguish water pixels from other pixels can significantly reduce classification errors. Studies by Li et al [3] and Jiang et al. [4] have shown that the traditional zero threshold approach can introduce errors due to spatial-temporal variations in contrast and brightness of satellite images. Therefore, they tried to automate threshold selection by using the Otsu binarization algorithm [5].

6.1.2 Machine learning for water segmentation

Intelligent detection using machine learning is a relatively new development for water body segmentation in satellite images, ranging classification, detection, and segmentation of objects in images.

We consider the DeepLab architecture, which has now reached its fourth version, DeepLabv3+ [6]. It consists of two parts, an encoder and a decoder. The encoder module records contextual information by applying pyramidal convolution (atrous convolution) at multiple scales, while a simple but effective decoder module analyzes the segmentation results along object boundaries. In addition, the authors recommend, supporting their results, that the Xception model (a network model consisting of three flows containing separate depth convolution layers) be used as the basis of the network instead of ResNet-101 (a residual network model containing 101 layers of weights).

Artificial Intelligence methods for identifying water bodies on satellite images give much better results than thresholding standard water indices. However, the authors of NDWI-DeepLabv3+ have innovated the recognition methods by creating a model that combines the advantages of both solutions. The combination of solutions not only gives better results, but sets the stage for the development of intelligent methods for water body recognition in satellite images.

6.1.3 Datasets

One of the main problems of training neural networks is the need for a suitable dataset. For water bodies segmentation, datasets are few and most are not prepared for machine learning, i.e., ground truth labels are not available in any way. Popular collections used for machine learning from satellite imagery are: EuroSAT [7], SAT-6 [8], RIT-18 [9], Dstl Satellite Imagery Feature Detection [10].

Taking into account the need to use the datasets for water body segmentation, it is necessary at first to discard the datasets suitable only for classification - EuroSAT and SAT-6. The RIT-18 collection was taken using a drone, which can have a large impact on the accuracy of the learned model on actual images taken by satellites. Additionally, it does not include the shortwave infrared bands that are needed to process images by the ANDWI index. The remaining dataset (Dstl Satellite Imagery Feature Detection) is not an ideal set in the context of water body recognition due to the large number of classes, which may mean that the water in the images will be a small percentage of the total, in which case the model might show biased performance overestimating overall accuracy while failing to detect minority classes. However, this is the only set that meets the conditions needed to teach the model adequately to the problem, hence it will be used in our computational experiments.

6.2 Materials and Methods

To conduct the research, a Python software was developed to create a neural network model, train it, and then test it. The network architecture is based on the DeepLabv3+ model and the Xception backbone network. From the four datasets for the study, the Dstl Satellite Imagery Feature Detection dataset was selected by elimination, which will be used to train and validate the proposed model. The purpose of this research is to verify whether the modification of the NDWI-DeepLabv3+ architecture will produce the intended results, i. e., an increase in the quality of water body segmentation. In addition, plug-ins for the QGIS application have been implemented, enabling water recognition on selected satellite images using the implemented models - NDWI-DeepLabv3+ and its modifications. The plugin allows you to easily change models by replacing their files. If there are better datasets on which to learn a new model, you will be able to easily improve the plug-in's operation by replacing the models. The plugin is available along with the code in the public GitHub repository at [11].

The modification of the model was to replace the NDWI index values with the new ANDWI index values for each pixel. The Figure 6.1 shows the diagram of the new ANDWI-DeepLabv3+ architecture modified by changing the NDWI index layer to ANDWI. The change should significantly improve the water bodies segmentation results.

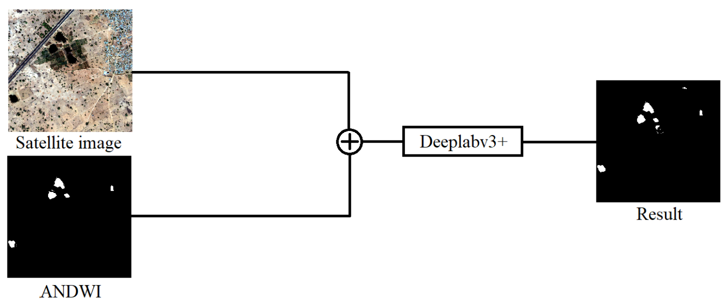


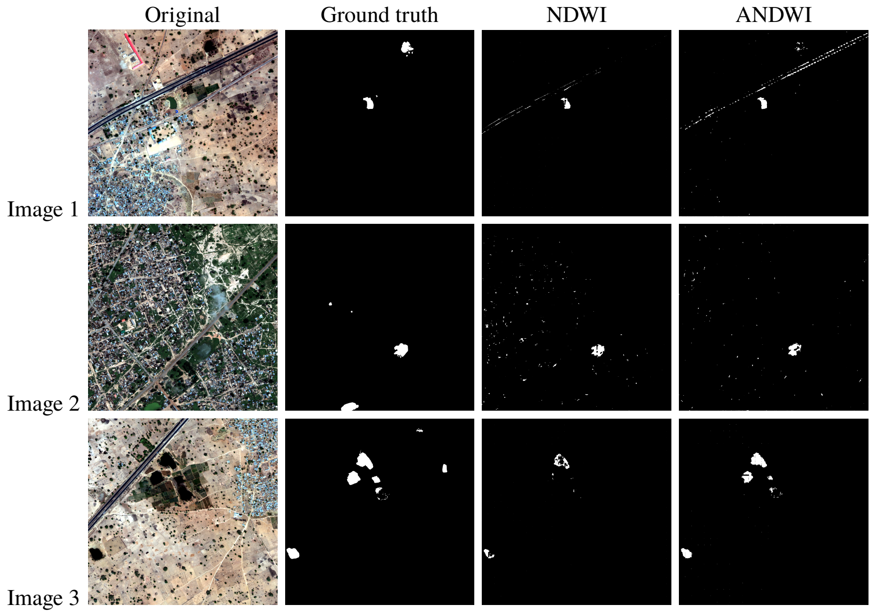
Fig. 6.1 New ANDWI-DeepLabv3+ model created by replacing NDWI index layer with ANDWI index layer.

6.3 Results

Two intelligent algorithms were tested for effectiveness on three different satellite images. The models were learned on the same dataset for 4,000 epochs, which took about 8 hours for each model. The cross entropy function was used as the loss function, which involves measuring the difference between two probability distributions for a random variable or a set of events.

The Table 6.1 presents the results of the experiments performed for the two models. The first model was developed based on the NDWI-DeepLabv3+ architecture, that is, the NDWI index values for each pixel were calculated during data preparation. The second model is a modification of it, by replacing the NDWI index values with a new ANDWI index. Based on the resultsin figure 6.2 was created, from which it is clear that the ANDWI model performs significantly better than NDWI. The new model behaved as expected and significantly outperformed the standard NDWI-DeepLabv3+ model in water body segmentation quality. The models were unexpectedly well learned given that the dataset was not a typical dataset for learning to recognize water bodies, so water was only a small part of its content.

Table 6.1 Algorithms results on three images. In the monochrome images, water is highlighted in white and black indicates the background.



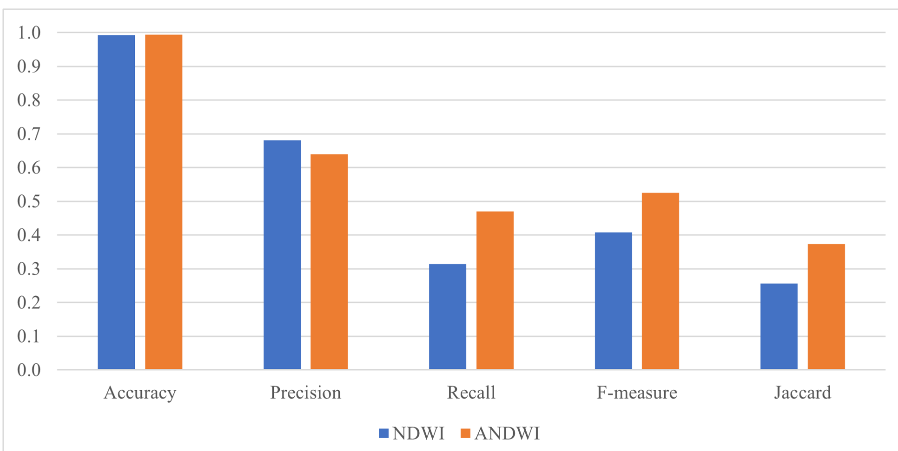


Fig. 6.2 Averaged results of NDWI and ANDWI models.

6.4 Conclusions

The aim of the research was to prepare a modification of the algorithm for recognizing water reservoirs on satellite images, and then to conduct a comparative analysis of the obtained results. In the study, a modification was made in the NDWI-DeepLabv3 + algorithm, which consisted in replacing the layer created on the basis of the NDWI with the layer based on the ANDWI. New ANDWI-DeepLabv3+ algorithm presented in the paper gave promising results during the research, showing a much higher quality of segmentation of water reservoirs than the base version of the algorithm could stand as a basis for further research on intelligent water recognition in satellite images. The new algorithm based on the ANDWI index has great potential for further development, however, before further work on new improvements, new datasets should be created for research that would contain diverse landscapes, along with labels for water segmentation.

References

1. McFeeters, S.K. The use of Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing 17* **1996**, 1425–1432.
2. Rad, A.M., Kreitler, J., Sadegh, M. Augmented Normalized Difference Water Index for improved surface water monitoring. *Environmental Modelling and Software vol. 140 no. 105030* **2021**.
3. Li, W., Du, Z., Ling, F., Zhou, D., Wang, H., Gui, Y., Sun, B., Zhang, X. A comparison of land surface water mapping using the normalized difference water index from TM, ETM+ and ALI. *Remote Sensing vol. 5, Issue 11* **2013**, 5530–5549.
4. Jiang, H., Feng, M., Zhu, Y., Lu, N., Huang, J., Xiao, T. An automated method for extracting rivers and lakes from Landsat imagery. *Remote Sensing vol. 6, Issue 6* **2014**, 5067–5089.
5. Otsu, N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern vol. SMC-9, Issue 1* **1979**, 62–66.
6. Liang-Chieh Chen, A., Zhu, Y., Papandreou, G., Schroff, F., Hartwig, A. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *Proceedings of the ECCV* **2018**, *801–818.*
7. *Helber, P., Bischke, B., Dengel, A., Borth, D. EuroSAT: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 12, no. 7 **2019**, *2217–2226.*
8. *Basu, S., Ganguly, S., Mukhopadhyay, S., DiBiano, R., Karki, M., Nemani, R. DeepSat: a learning framework for satellite imagery.* Association for Computing Machinery **2015**, *1–10.*
9. *Kemker, R., Salvaggio, C., Kanan, C. Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning.* ISPRS Journal of Photogrammetry and Remote Sensing, vol. 145 **2018**, *60–77.*
10. *Dstl Satellite Imagery Feature Detection. Available online: https://www.kaggle.com/c/dstl-satellite-imagery-feature-detection/data*
11. *Plugin for QGIS application: https://github.com/MateuszStarczyk/QGISWaterBodyRecognizerPlugin.*

Chapter 7
A Preliminary Solution for the Supervision of River Basins based on Multispectral Information.

Dora B. Heras[0000−0002−5304−1426] and Francisco Argüello[0000−0001−9279−5426] and Alberto S. Garea[0000−0001−9394−0330], Francisco Javier Cardama, and Álvaro Goldar Dieste

Abstract The sustainable management of forests, land and water is very relevant for different agents with different purposes that range from the health protection of citizens to the conservation of animal and vegetation species. In this paper we summarize the conclusions of the work previously published by the authors regarding the automatic management of river basins. The work focuses on detecting the presence of different vegetation species and non-vegetation structures as a first step to track their evolution. The information used consists of five-band multispectral images taken by unmanned aerial vehicles in the region of Galicia, northern Spain. The proposed solution is a classification chain based on the extraction of spatial, spectral and textural information previous to a hierarchical classification process. We analyze the state and limitations of the work and some future perspectives.

7.1 Introduction

River ecosystems have been increasingly affected by population growth with a consequent degradation of natural resources [1]. Different regulations [3] ensure the protection of river basins understood as the riverbeds and the ecosystems on both sides directly affected by them, establishing the obligation to control the watershed situation. It requires analyzing the specific composition of the watershed related to vegetation, the structure of the riverbanks, and the river connectivity in the riverbed and to a distance of around 100 meters. In this context two processes can be identified in the watershed supervision if we focus on the main characteristics related to

D.B. Heras, AS Garea, FJ Cardama, and A.Goldar Dieste
Departamento de Electrónica e Computación, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

D.B. Heras and F. Argüello
Centro Singular de Investigación en Tecnoloxías Intelixentes, Universidade de Santiago de Compostela, Santiago de Compostela, Spain

vegetation. The first one is to carry out inventories of the river situation from a morphological point view including vegetation and non-vegetation elements. The second one consists in analyzing the situation over time to determine the evolution of the ecosystem, taking corrective measures when necessary and ascertaining their effect. There is currently a lack of tools to carry out these two processes automatically and systematically.

The main objective of this paper is to summarize the work carried out by the authors (mainly [3] and [4]) focusing on producing inventories of the river basins of Galicia, Spain, using five-band multispectral images taken by UAVs and image processing algorithms.

For this work the collaboration of the local agency in charge of the management of the river ecosystems in Galicia, called Augas de Galicia, was critical to establish the main requirements. The software solution discussed here consists of a supervised classification chain based on exploiting spatial, spectral and textural information at the level of segments of the image called superpixels. The tool produces a georeferenced map of the river basin. Execution time and accuracy will be considered to analyze the different alternatives that were proposed.

The rest of the article is organized into four sections. Section 2 presents the description of the proposed solution, a classification chain involving superpixel computation and texture extraction. The experimental results for the evaluation in terms of classification performance and computational cost and the discussion are presented in Section 3. Finally, Section 4 summarizes the main conclusions.

7.2 Watershed mapping solution proposed

The problem of identifying different vegetation species and non-vegetation structures was solved as a supervised classification problem over multispectral images consisting of five different bands. As the number of spectral bands is low, spatial and textural information also needs to be exploited. Information on the spatial structures present in the input image is added to the different spectral features available for each pixel of the image [8]. For identifying different vegetation species, the introduction of additional textural information has been shown to be very useful as the discrimination among different species in the region under study can not be accurately performed based only on spectral signatures [9]. The classification was performed by using SVM, KELM and a CNN. As a result, a spectral-spatial texture-based classification chain was proposed, as shown in Fig. 7.1.

Since the images are large as a consequence of the vast territory to be analyzed, the image processing techniques and algorithms in the classification chain have been selected for fast execution and efficient use of computational resources. The steps involved (See Fig. 7.1) are the following: spatial information extraction, texture codebook generation, texture feature encoding, and classification.

1. **Spatial information extraction**: Although different alternatives were explored for extracting spatial information in our experiment, including the incorporation

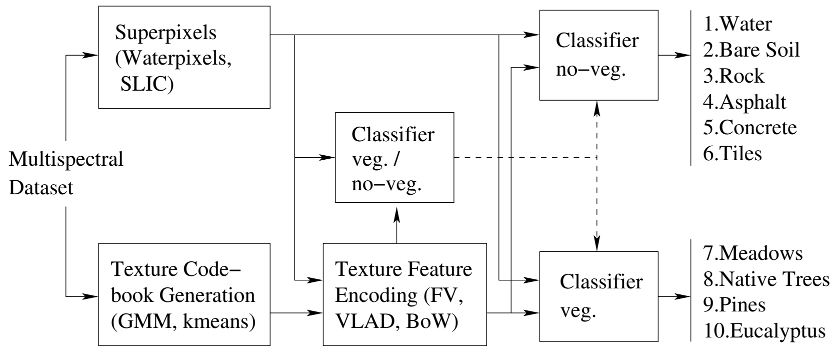


Fig. 7.1 Diagram of the classification chain for riverside mapping.

of different extended morphological profiles [3], the most efficient technique turned out to be the application of a superpixel algorithm.

Superpixel algorithms are segmentation algorithms that divide the image into segments of similar size, whose shape is adapted to the characteristics of the image. Different parameters allow to adjust the segmentation to the image facilitating the the detection of objects and structures. Out of the superpixel algorithms proposed in the literature [7], those showing the best results in our experiments and presenting the lowest execution time are SLIC and waterpixels. Despite their simplicity, that positively impacts the execution time, SLIC and waterpixels adhere well to boundaries and allow easy control of segment size and compactness. Examples of the segmentation maps obtained are shown in Figure 7.2. In the proposed chain, segmentation algorithms play two key roles: a spectral-spatial structure enhancer and a data downsizer, as other further steps in the classification chain are computed over segments instead of individual pixels.

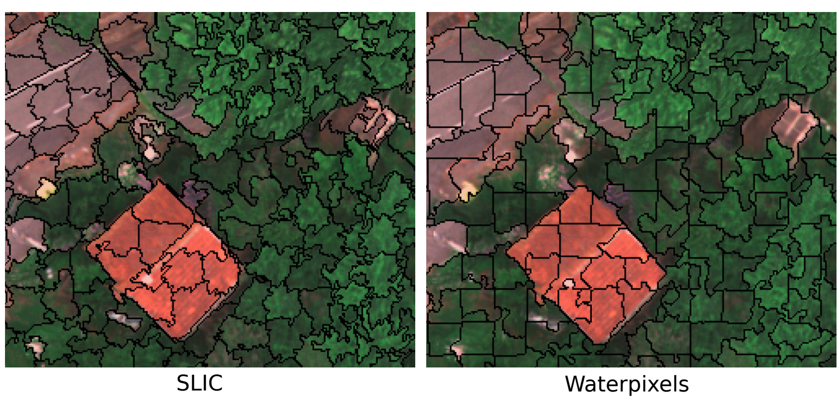


Fig. 7.2 Examples of superpixel computation with segments using SLIC and waterpixel segmentation algorithms.

2. **Texture codebook generation and feature encoding** The reason for using textures in the proposedl solution is to improve the accuracy in the identification of vegetation species that are similar in reflectance properties when only 5-band multispectral information is available. Figure 7.3 represents oak and eucaliptus canopies showing differences in terms of their structures.
The first step in the application of textures is to generate a codebook with a set of basic elements or primitives (called textons) that can be used to represent any texture present in the images. The design principle is to obtain a compact and discriminative codebook that allows classifying the textures [6, 5]. To obtain the dictionary elements, $k$means clustering and Gaussian Mixture Modeling (GMM) together with a CNN were considered. Dictionary elements in $k$means are the cluster centers while GMM uses both mean centers and covariances to describe the spreads of the clusters.
Given the generated texture codebook, feature encoding generates a representation of the image based on the codebook. The three feature encoding methods that were compared in our experiments were Bag of Words (BoW), Vector of Locally Aggregated Descriptors (VLAD) and Fisher Vectors (FV). BoW builds a histogram counting the number of local features assigned to each codeword and, therefore, encodes the zero order statistics of the distribution of local descriptors. VLAD accumulates the differences of local features assigned to each codeword. Finally, FV extends the BoW by encoding higher order statistics (first and second order), including information about the fitting error [6]. In the present study, the texture information is used to distinguish among the different elements present in each superpixel.

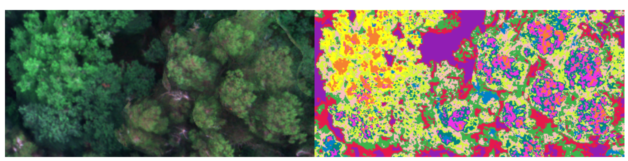


Fig. 7.3 A color composition scene of oak and eucalyptus canopies and its clusterisation with k-means using 10 classes. Left: canopies of oak (on the left) and eucalyptus (on the right). Right: clusterisation map using kmeams.

3. **Classification**. Once the spatial or texture features have been obtained, the final operation is classification. Three supervised classification algorithms are widely used, since they are fast and provide good results: SVM, KELM, and random forest. In the proposed classification chain, the classification is performed at superpixel level, that is, each superpixel provides a single input sample to the classifier, which can be the average of the pixels in the segment or the texture feature. A hierarchical scheme of three classifiers is used, since the separation between vegetation species is especially complicated due to their similar spectra and individual variability. First, a separation is made into two classes: vegetation and the other. A second classifier identifies the different vegetation classes from

each other. Finally, a third classifier performs the separation of the different non-vegetation classes. This is shown in detail in the diagram of Figure 7.1. In the case when a CNN is used for extracting the textures, the same CNN performs the classification process.

7.3 Dataset description

With the objective of monitoring the interaction of the masses of native vegetation with artificial structures and river beds, eight locations in Galicia were studied [2]. They were selected based on the presence of native vegetation, eucalyptus, and pines. The native vegetation, that populates areas near the water streams due to its ability to survive under unstable water conditions, includes oaks, birches, alders, and willows. Different artificial structures are also present and are identified: rooftops covered by tiles, some concrete structures, asphalt roads, stone structures, and bare soil roads. These locations correspond to portions of the rivers Oitaven, Xesta, Eiras, Ermidas, Ferreiras, Das Mestas, Mera, and Ulla.

The datasets were captured by the MicaSense RedEdge multispectral camera mounted on a custom UAV. Its five discrete sensors provide spectral channels at wavelengths of 475 nm (Blue), 560 nm (Green), 668 nm (Red), 717 nm (Edge), and 840 nm (NIR). Each flight captured data over long distances at a height of 120 m, with a spatial resolution of 10 cm/pixel. Each dataset was built as the orthomosaic of the frames captured by the UAV, each of which is of size 1280×960 pixels. The flights were conducted during the summer months of 2018, 2019 and 2020.

The location as well as the relevant classes to be considered in the reference classification maps were selected with the collaboration of the local agency involved in the watershed management of these rivers, Augas de Galicia. The construction of accurate reference data was a long-term process involving these forestry experts and the authors of the paper. Some of the images and the corresponding reference maps can be seen in Figure 7.4 while the characteristics of the eight datasets are detailed in Table 7.1.

7.4 Results and analysis

This section contains information about the experiments, the classification results obtained and their analysis. The experiments were carried out on a PC with a dual-core Intel Pentium G3220 at 3.00 GHz and 8 GB of RAM. The code was written in C and compiled using gcc under Ubuntu 20.04. CNN code was compiled in CUDA and executed on a NVidia GeForce GTX 1050 GPU with 2 GB of memory. Classification results in terms of Overall Accuracy (OA), Average Accuracy (AA), and execution times are presented. The proposed classification chain was run on the eight datasets,

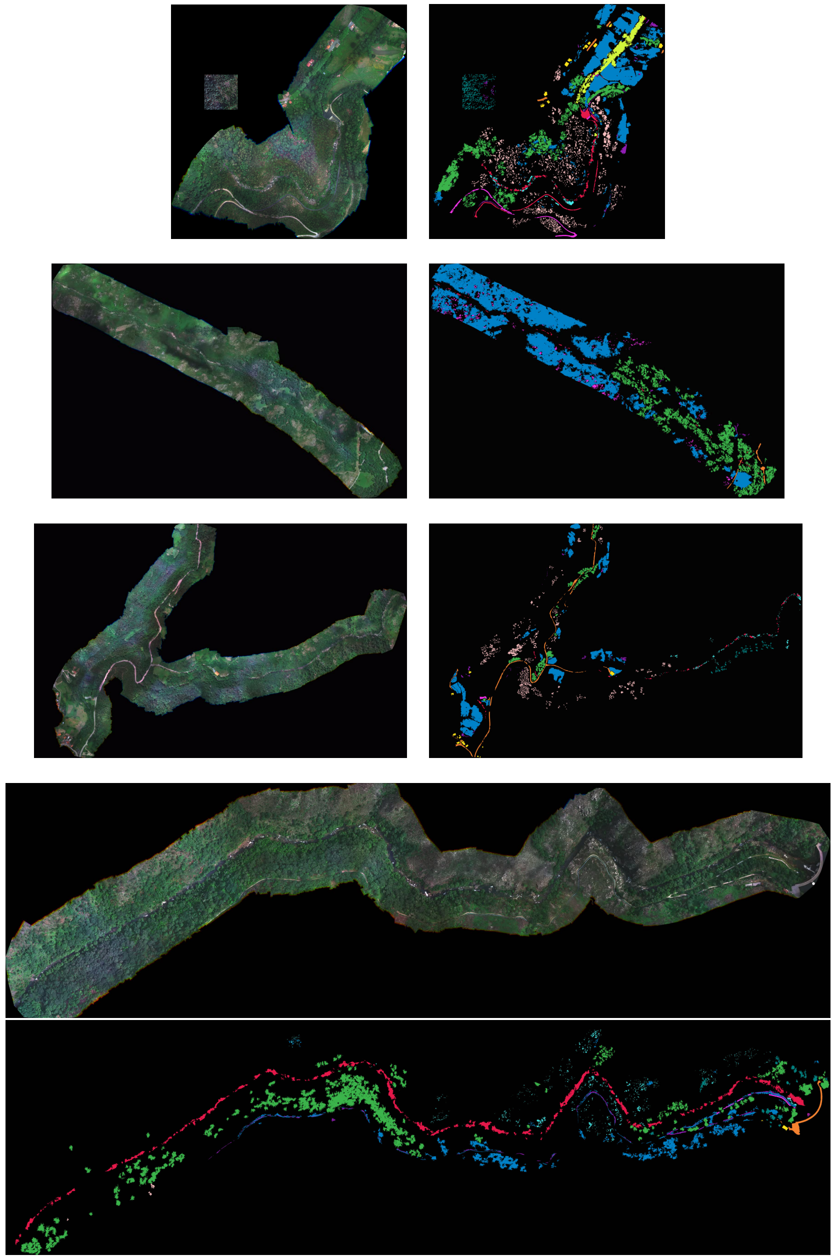


Fig. 7.4 Some of the multispectral datasets used in the experiments. Color composite and reference data for the Oitavén, Xesta, Eiras and Ermida datasets, respectively.

Table 7.1 Datasets used in this study. The "Pixels" column indicates the number of data values in the dataset, while the "Superpixels" column indicates the number of superpixels containing data values, those that are available in the reference data, and those used for training, corresponding to 15% of the total.

Dataset	Size	Pixels	Superpixels
Oitaven	6722×6689 (760×760 m^2)	22.04M	78383/21650/3242
Xesta	17202×3848 (1945×435 m^2)	40.00M	149727/61394/9207
Eiras	18221×5176 (2260×660 m^2)	38.35M	142110/15435/2309
Ermidas	18972×11924 (2190×1390 m^2)	65.56M	248742/23188/3474
Ferreiras	9219×9335 (740×750 m^2)	40.19M	156331/54564/8181
Das Mestas	9040×4915 (920×510 m^2)	27.17M	107979/39229/5882
Mera	22116×10718 (2770×1370 m^2)	99.24M	397799/59989/8995
Ulla	16555×4220 (1420×380 m^2)	46.01M	197094/1461/217

classifying superpixels corresponding to vegetation and non-vegetation elements into ten classes.

7.4.1 Alternatives for the different steps

The aim of the first set of experiments is to analyze how the different alternatives for the different steps of the classification chain influence the classification accuracies and the execution times. Table 7.2 shows the results of the available alternatives for each module using the Oitaven dataset with 15% of training superpixels. The best accuracy results were obtained for the WP (waterpixels), GMM, FV, and KELM algorithms, which are considered the base chain. The rest of the measurements were taken substituting the indicated algorithm in the base chain.

A detailed discussion or the results presented here is avaliable in [4]. Here, we only highlight the most relevant observations. When superpixel algorithms are applied, the parameter indicating the side length of a segment is very relevant as the assignment of pixels to superpixels depends on the structures of the image. For the images in our experiments, a side length of 20 pixels was selected. SLIC and waterpixels were selected due to specially low execution times but they tend to over-segment the image, so a post-processing of the segments is applied in order to achieve segments of a minimum of 100 pixels.

Regarding the texture codebook generation process, the best results were obtained by the GMM clustering algorithm, followed by $k$means. GMM builds the codebook

Table 7.2 Classification accuracies and execution times obtained by the different algorithm alternatives for each module. The Oitaven dataset and 15% training superpixels were used in the experiments.

Step	Algorithm	OA	AA
Superpixel	SLIC	91.75	83.56
	WP	93.03	87.18
Texture Codebook Generation	None	82.71	71.86
	*k*means	91.81	84.72
	GMM	93.03	87.18
Texture Feature Encoding	None	82.71	71.86
	BoW *k*means	82.39	72.16
	BoW GMM	80.12	62.99
	VLAD	91.81	84.72
	FV	93.03	87.18
Classifier	SVM	92.51	86.00
	KELM	93.03	87.18
	RF	86.48	75.62
	CNN	88.93	80.08

from the pixel values using Gaussian distributions that take the mean and deviation as parameters, while in *k*means the codebook is generated taking into account only the cluster centers. Therefore, *k*means only considers spherical clusters, while GMM can be adjusted to elliptic clusters. Regarding computational cost, in all the cases not all the pixels of the dataset were considered for this process.

For the texture feature encoding, four algorithms were considered: BoW *k*means, BoW GMM, VLAD, and FV. FV not only accumulates the differences of local features assigned to each codeword (first order statistics), as VLAD, but also second order statistics. Encoding a larger amount of information increases the accuracy of the classification, but it also increases the execution time. This is the reason for its best results in terms of accuracy but not in terms of time, as shown in the table.

Finally, for the *classification* module, four algorithms were tested: SVM, KELM, random forest, and CNN. The best results were provided by KELM and SVM. These two classifiers are kernel-type algorithms that map the data to a larger feature space. In our experiments, an RBF kernel with $\gamma = 5$ and $C = 16$ was used for both classifiers. CNN does not require the texture generation and encoding modules, but due to its computational requirements the execution was carried out on a GPU. In our case, the CNN is composed of a convolutional layer and two directly connected layers. For the convolutional layer, five filters of size 28×28, max-pooling of factor 2 and sigmoid activation function were used, while the directly connected hidden layer contains 100 neurons and sigmoid activation functions. The precision values obtained by the CNN are not remarkable since this type of algorithms require a fairly large training set to provide good results. Regarding the execution times, the best values were obtained by SVM followed by random forest.

7.4.2 Results for the best configuration

In table 7.3, the results for the proposed classification chain running on the eight datasets are shown. Two types of chains were considered: the one with the best classification accuracies, and a deep-learning one. The chain with better accuracy from the different options discussed in the previous section is WP+FV+KELM, while the deep-learning chain consists of WP+CNN. Table 7.3 shows the classification accuracies and execution times obtained for each chain. The average accuracies for the eight datasets were AA=92.65 and OA=80.39 for the most accurate chain, and AA=89.67 and OA=73.45 for the CNN chain. Regarding execution times, the average values were $t = 735$, and $t = 484$ seconds, for the corresponding three chains.

Table 7.3 Classification accuracies and execution times obtained by algorithm chains: high accuracy (WP+FV+KELM), and deep-learning (WP+CNN). Results are for 15% of training superpixels and execution times displayed in seconds (t^* indicates that the CNN was run on a GPU).

	WP+FV+KELM			WP+CNN		
Dataset	OA	AA	t	OA	AA	t^*
Oitaven	93.03	87.18	217	88.93	80.08	224
Xesta	96.18	81.88	701	94.47	73.17	573
Eiras	94.45	77.58	330	91.45	62.84	236
Ermidas	94.76	87.42	981	92.11	82.27	635
Ferreiras	90.27	74.69	705	86.53	68.75	541
Das Mestas	89.27	82.35	381	86.06	75.39	378
Mera	87.65	63.98	2312	82.19	54.97	1185
Ulla	95.56	88.04	250	95.61	90.10	103
Average	92.65	80.39	735	89.67	73.45	484

7.5 Discussion

The problem analyzed imposes some limitations to be considered. In this case only multispectral information was available and it has not been possible to use the canopy height model. In addition, the spectral ranges of the different tree species are very similar and overlap. In small scenarios, these issues could be compensated with very high resolution, but this is hardly applicable to the large areas that we have in this case.

Taking these limitations into account, a classification based on textures has been used in this study. Although the canopy structure of the different native species (alders, willows, birches, oaks, etc.) is very similar, they are easily distinguishable from pines and eucalyptus. This facilitates the identification of the areas occupied by these invasive species. It was necessary to carry out field visits to identify the tree species and build the reference maps.

In relation to the detection of the river course in the images, it must be taken into account that often the vegetation in the riparian zones of Galicia completely hides the river when viewed from above. Additionally, in summer, some parts of the river dry up and leave rocks or bare soil uncovered. In such cases, the river channel appears fragmented. To solve this issue, algorithms for the repair of partial occlusions in topographic data could be added. The method used in this work has the advantage of ease of implementation and low cost, since medium-sized UAVs and multispectral sensors can be used. The algorithms used have been selected for their favorable relationship between precision and execution time. It should be noted that some parameters could be adjusted to improve accuracy or speed, but not both.

Aproblem that is common to many remote sensing applications such as in this one, is the scarcity of labeled samples. One possible way of mitigating the limitations of remote sensing datasets is the use of deep neural architectures as classifiers [10, 11], in particular, those approaches incorporating data augmentation techniques [12] or those networks specially designed for performing data augmentation such as the adversarial neural networks [13].

Different problems such as the application of transfer learning techniques for extending the applicability of the classification models to other rivers not considered yet, need to be tackled. We also face the next stage, that will require the analysis of new images of temporal series of images of the same areas, change detection analysis for tracking the watershed evolution.

7.5.0.1 Acknowledgements

The authors would like to thank Aguas de Galicia from Xunta de Galicia for the support in the construction of the reference information for the multispectral images. This work was supported in part by the Civil Program UAVs Initiative, promoted by the Xunta de Galicia and developed in partnership with the Babcock company. We also have to acknowledge the support by the Ministerio de Ciencia e Innovación, Government of Spain (grant number PID2019-104834GB-I00), and Consellería de Educación, Universidade e Formación Profesional (accreditation 2019–2022 ED431G-2019/04). All are co-funded by the European Regional Development Fund (ERDF).

References

1. Watershed Conservation Resource Center. Available online: https://www.watershedconservation.org (accessed on 12 January 2022).
2. La vegetación en los Bosques de Rivera Gallegos. Available online: https://www.parqueforestaldesantiago.org/es/bosques-de-ribera/la-vegetacion-de-ribeira-en-galicia/ (accessed on 13 May 2021) (in Spanish).
3. Bascoy, P.G., Garea, A.S.,Heras, D.B., Argüello, F.,Ordóñez, A. Texture-based analysis of hydrographical basins with multispectral imagery. Remote Sensing for Agriculture, Ecosystems,

and Hydrology XXI. SPIE, International Society for Optics and Photonics., vol. 11149, p. 111490Q, (2019).

4. Argüello, F., Heras, D. B., Garea, A. S., Quesada-Barriuso, P. Watershed Monitoring in Galicia from UAV Multispectral Imagery Using Advanced Texture Methods. Remote Sensing, 13(14), 2687 (2021)
5. Rey, S.R.; Blanco, D.B.; Argüello, F. Texture Extraction Techniques for the Classification of Vegetation Species in Hyperspectral Imagery: Bag of Words Approach Based on Superpixels. Remote Sens. 12, 2633 (2020)
6. Liu, L.; Chen, J.; Fieguth, P.; Zhao, G.; Chellappa, R.; Pietikäinen, M. From BoW to CNN: Two decades of texture representation for texture classification. Int. J. Comput. Vis. 127, 74–109 (2019)
7. Stutz, D; Hermans, A.; Leibe, B. Superpixels: An evaluation of the state-of-the-art. *Comput. Vis. Image Underst.*, 166, 1–27 (2018)
8. Fu, Y., Zhao, C., Wang, J., Jia, X., Yang, G., Song, X., Feng, H. An improved combination of spectral and spatial features for vegetation classification in hyperspectral images. Remote Sensing, 9(3), 261 (2017)
9. Szantoi, Z., Escobedo, F., Abd-Elrahman, A., Smith, S., Pearlstine, L. Analyzing fine-scale wetland composition using high resolution imagery and texture features. International Journal of Applied Earth Observation and Geoinformation, 23, 204-212 (2013)
10. Zhong, Z.,Li, J., Luo, Z., Chapman, M. Spectral–Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework, IEEE Transactions on Geoscience and Remote Sensing, 56, (2), pp. 847–858, (2018)
11. Paoletti, N.E., Haut, J.M., Fernandez-Beltran, R., Plaza, J., Plaza, A.J., Pla, F. Deep Pyramidal Residual Networks for Spectral–Spatial Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing, 57 (2), pp. 740–754 (2019)
12. C. Shorten and T. M. Khoshgoftaar, A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6 (1), p. 60 (2019)
13. L. Zhu, L., Chen, Y., Ghamisi, P., Benediktsson, J.A., Generative Adversarial Networks for Hyperspectral Image Classification, IEEE Transactions on Geoscience and Remote Sensing, 56 (9), pp. 5046–5063 (2018)

float

Chapter 8
Capturing Ready-Mix Concrete Operational Workload: A Case Study

Pablo Cleveland

Abstract Ready-Mix Concrete industry, among many others, suffer from overbooking. In this work, we propose a heuristic approach to measure the impact a new order has on the operation workload.

8.1 Introduction

Ready-Mix Concrete is a perishable product, meaning there is a time window within which the product must be delivered. Generally, the orders are structured in dispatch trains which must respect a spacing time between them (established by the client), this further increases the strain on the operation. On top of that the distribution of daily demand presents pronounced peaks and valleys, which affects the size of the required fleet and the periods with idle resources, respectively. All of the aforementioned factors make for a very complex and challenging decision-making environment.

One of the biggest problems they have to face on a daily basis is overbooking, which happens when more dispatches than those that can be processed (without incurring in delay) are scheduled. Overbooking generates operational inefficiencies that companies commonly pass on to the customer through prices. In general, 2 strategies are used, namely, increasing all prices, and raising specific prices. On the one hand, applying a general price increase, although simple to put in practice, this ends up penalizing all customers equally regardless of whether they formed part of the cause of the inefficiency, this grants the competition the opportunity to offer a better price which may cause some customers to flee. On the other hand, applying a penalty to specific prices based on behavior projections using historical data can

Pablo Cleveland
Industrial Engineering Department, Universidad de Chile, Santiago, Chile e-mail: pcleveland@ceine.cl

help to deliver a fairer treatment to customers. Other penalties used that are worthy to take into account are: penalization based on the proximity of purchase to the due date (e.g. concert tickets and airplane tickets) and the application of cancellation fees (e.g. some airlines keep a percentage of the price paid when making a cancellation). Also worthy of note is the tendency in e-commerce environments to adopt a dynamic pricing strategy, i.e. a making the price flexible and responsive to market demands.

In this work, we propose a heuristic approach to measure the increase of strain on the operation caused by the incorporation of a new dispatch train onto the current schedule. Obtaining this is essential to build a dynamic pricing strategy based on operational workload. In consequence, this strategy can then be used to tackle the problem of overbooking by discouraging its occurrence by presenting the customer with prices that take into account the consequences of his choice to the operation. This strategy possesses the following advantages: it naturally rewards clients who make purchases with anticipation, it fairly transfers the extra cost to the actor that generates the disruption, and by presenting several differently priced options to the consumer it accommodates both, clients that have an inflexible willingness to pay but are willing to adapt to another option and clients that don't mind paying extra for maintaining their original plan.

The organization of the rest of this paper is as follows. The next section explains the problem of the study case and its context. Section 8.3 provides the relevant literature with respect to the usual models used to deal with scheduling daily dispatches. Section 8.4 details the proposed heuristic. Finally, in section 8.5, conclusions and future work are presented.

8.2 Problem Description

The data used in this work was provided by a Ready-Mix Concrete Industry in Chile. They manufacture and dispatch of various types of concrete, having to coordinate trucks and depots to deliver the orders of their clients within the requested times (see Fig 14.1). It is worthy of mention that concrete has specific constraints regarding its delivery to clients, due to its perishable nature. For example, the truck cannot unload the product before the requested time, however it is possible (but undesirable) to start the unloading process late (within a given maximum delay). Also, there must be a separation time between two consecutive truck unloads of a given order. The problem that must be addressed consists of obtaining a schedule for every dispatch, where in order to be able to start, both, depot and truck must be available, and time windows of delivery must be respected, meanwhile attempting to minimize the total weighted delay. This is then used to determine the operational strain increase (in operational workload) by simulating the incorporation of a new dispatch train to the pre-computed schedule. This is done for each time block to form a grid that reflects the different impacts depending on the time of insertion of the new order.

8.3 Related Works

Scheduling models are often used for computer and manufacturing processes [6]. A comprehensive explanation of scheduling models in general it is suggested to see the works of Pinedo et al. [21], Leung et al.[15]. Also, it is recommended to check Allahverdi et al. [2] where a comprehensive survey of scheduling problems with setup times is provided and is further expanded in their posterior works in [3] and [1].

There are many applications where scheduling models are useful, so its convenient to classify the problems that are tackled in the literature to better understand the different branches scheduling problems can be grouped in.

First of all, a distinction can be made between deterministic and stochastic scheduling. In the former all attributes are known while in the latter one or more attributes are uncertain. Although, as noted by Liu et al. [16], most stochastic scheduling works focus on the case where job processing times are uncertain, there are also works that deal with random due dates (e.g. Wu and Zhou [23]), random job arrival (e.g. Megow, Uetz and Vredeveld [19]), uncertain release times (e.g. Liu et al. [16]), and stochastic machine breakdowns (Gu et al. [12], Zhou and Cai [25]). However, from here on we focus on deterministic scheduling.

According to Koulamas [14], the key components of deterministic scheduling problems are: machine configuration, job characteristics and the objective function. Moreover, considering the work done by Edis, Oguz and Ozkarahan in [10], we should also add whether the possibility of additional resources is included. On this last subject it is advised to view the work done by Blazewicz, Brawner and Finke in [5] where they deal with discrete resources.

Firstly, considering machine configuration, previous work can be first classified into single machine and multiple machine problems. The multiple machine problems branch can be further subdivided according to the layout of the machines into serial and parallel machines. As an example of parallel machine scheduling problem (PMSP), the work of Balin [4] can be examined. Balin studied the PMSP with fuzzy processing times, presenting a robust genetic algorithm approach within a simulation model that minimizes the makespan and comparing its results with those obtained by the use of the longest processing time dispatching rule.

Furthermore, within the parallel machines layout we can distinguish the cases of identical, uniform, and unrelated parallel machines. The first accounts for the case where the processing time of a job is the same regardless of the machine it is assigned to. The second refers to when the processing time of a job on a given machine is obtained by the machine's speed factor. The last represents the more general case where processing times of jobs on different machines need not be the same. An example of this last case can be found in [24], where Zhang, Wang and Liu take on the unrelated parallel machine scheduling problem (UPMSP) with variable job processing times taking into consideration learning effect, deteriorating jobs and controllable processing times simultaneously. For a more detailed description and classification of PMSP it is suggested to see the review published by Cheng and Sin in [9] and for a review on PMSP with additional resources see the work of Edis, Oguz and Ozkarahan in [10].

Secondly, regarding the objective function, most works minimize the makespan, while some other works use a multiobjective approach like Tamaki et al. in [22] where the performance criteria is a weighted sum of the makespan, the mean lateness, and the maximum lateness. Fewer papers tackle other objective functions like the number of tardy jobs, total tardiness, lateness, or the total absolute deviation from due dates, among others. Hence, some researches have made attempts to consolidate works regarding a particular performance metric to speed up the reviewing step of researches. For instance, Koulamas [14]) presents a survey of the total tardiness problem (while also providing 2 new heuristic for the single machine case and for the parallel machine case respectively). In this research, we choose the total weighted tardiness as the performance measure.

For instance, Tamaki et al. in [22] study a problem of the UPMR class in a plastics forming plant. The additional resource constraints correspond to the use of dies that in combination with the machines determine the production speed. They apply search methods such as simulated annealing (SA), genetic algorithm (GA), local search (LS) and random search (RS) to a binary representation of feasible schedules in order to obtain satisfactory solutions. Making a binary representation of schedules proved to be the key to being able to apply search methods to scheduling problems.

Improving on the work by Tamaki et al. [22], Jeng-Fung [7] proposed a heuristic procedure to obtain schedules that greatly outperforms the best of the search methods, SA, both in solution quality (regarding the performance metric) and runtime. On average, it accomplished a reduction of 11.18% for the performance metric value (makespan) and a 96.03% reduction for runtime.

Chen and Wu in [8] present a heuristic approach based on threshold-accepting method, tabu lists, and improvement procedures to obtain solutions when the performance metric is the total tardiness. They achieve competitive results with this approach capable of obtaining optimal solutions for small instances. Computational results show that this approach also outperforms Tamaki et al. SA method in both solution quality and runtime.

In [11], Fanjul-Peyro et al. study the unspecified dynamic unrelated parallel machine scheduling problem with additional resource and present 2 models UPMR-S and UPMR-P, and also provide three matheuristic strategies that are able to outperform the MIP models on medium sized instances. Their novel model UPMR-P is of particular interest because it is able to dispense of the time index which makes for a more efficient approach.

On a different note, there are some other works that are of interest to our research that do not fall under the category of UPMR scheduling problem. On the one hand, there are some niche models for the ready mixed concrete dispatch problem like the one presented in the works of Maghrebi et al. [20, 17, 18]. The RMC dispatching model depicts depots and customers as nodes and a delivery is depicted as an arc, this formulations resembles ours in this aspect. In [20], they compare a robust genetic algorithm (Robust-GA) and a column generation scheme (CG) as means to obtaining solutions for this model. They found that CG finds better quality solutions but Robust-GA is 40% faster. While in [17] they propose a Benders decomposition scheme and report results from a real instance. Lastly, in [18] they propose another column generation scheme this time for large-scale problems.
On the other hand, Khatami and Salehipour [13] study the problem of coupled task scheduling considering a single machine. They were able to obtain optimal solutions for small instances and also proposed a heuristic that obtains good solutions for instances where optimal solutions can not be obtained.

8.4 Solution Approach

Among the literature the family of models that resembles the problem characteristics the most is the UPMR, although there are some differences. Thus, in our heuristic algorithm we model the problem in a similar fashion. The proposed algorithm goes as follows:

```
Data: DispatchData, NewOrder
Result: Strains_dict
BaseStrain ← UPMR(DispatchData)
for d in Days do
    for hblock in HourlyBlocks do
        NewData ← Joindata(DispatchData, NewOrder, d, hblock)
        NewStrain ← UPMR(NewData)
        Strains_dict(d, hblock) ← NewStrain − BaseStrain
    end
end
```

Algorithm 1: Strain Grid Algorithm

Algorithm 1 recieves as input thge information of a set of dispatches and a new order to be inserted to the schedule. It starts by computing the base operational strain by solving the UPMR model over the set of dispatches without considering the new order. Then for each possible insertion hour the algorithm solves the UPMR model with the extended dispathces set. It ends by returning a grid containing the strain increase for each of the possible insertions hours for the new order.

In Fig. 14.2 we show the results of applying the algorithm to a certain week of data. The results are scaled in order to respect the company's requests with regards to data usage.

8.5 Conclusion

The proposed solution approach returns an approximation of the additional strain generated by the incorporation of a new order to a established set of orders. The resulting grid can be used as guidance for pricing strategies that take the operation conditions into consideration.

Future work should focus on developing a strategy to convert a grid of increased operational workload to a price grid. It is also necessary to generalize the existing scheduling models to adapt to the case where several types of parallel machines must be coordinated. Consequently, efficient solution methods for this new family of models should also be researched.

Acknowledgment

This research was partially supported by the Business Intelligence Research Center (www.ceine.cl) and ANID (CONICYT-PCHA/MagísterNacional/2015-22151700, CONICYT-PFCHA/Doctorado Nacional/2019-21190971)

References

1. Allahverdi, A.: The third comprehensive survey on scheduling problems with setup times/costs. European Journal of Operational Research **377**, 345–378 (2015)
2. Allahverdi, A., Gupta, J., Aldowaisan, T.: A review of scheduling research involving setup considerations. Omega **27**, 219–239 (1999)
3. Allahverdi, A., Ng, C.t., Cheng, T.C.E., Kovalyov, M.: A survey of scheduling problems with setup times or costs. European Journal of Operational Research **187** (2008)
4. Balin, S.: Parallel machine scheduling with fuzzy processing times using a robust genetic algorithm and simulation. Information Sciences **181**, 3551–3569 (2011)
5. Blazewicz, J., Brauner, N., Finke, G.: Scheduling with Discrete Resource Constraints, pp. 23–1, 23 (2004)
6. Blazewicz, J., Ecker, K., Pesch, E., Schmidt, G., Węglarz, J.: Scheduling Computer and Manufacturing Processes (2001)

7. Chen, J.F.: Unrelated parallel machine scheduling with secondary resource constraints. The International Journal of Advanced Manufacturing Technology **26**(3), 285–292 (2005), https://doi.org/10.1007/s00170-003-1622-1, iD: Chen2005
8. Chen, J.F., Wu, T.H.: Total tardiness minimization on unrelated parallel machine scheduling with auxiliary equipment constraints. Omega **34**(1), 81–89 (2006), https://www.sciencedirect.com/science/article/pii/S0305048304001197, iD: 271690
9. Cheng, T., Sin, C.: A state-of-the-art review of parallel-machine scheduling research. European Journal of Operational Research **47**(3), 271–292 (1990). https://doi.org/https://doi.org/10.1016/0377-2217(90)90215-W, https://www.sciencedirect.com/science/article/pii/037722179090215W
10. Edis, E., Oğuz, C., Ozkarahan, I.: Parallel machine scheduling with additional resources: Notation, classification, models and solution methods. European Journal of Operational Research **230**, 449–463 (11 2013). https://doi.org/10.1016/j.ejor.2013.02.042
11. Fanjul-Peyro, L., Perea, F., Ruiz, R.: Mip models and matheuristics for the unrelated parallel machine scheduling problem with additional resources. European Journal of Operational Research **260** (01 2017). https://doi.org/10.1016/j.ejor.2017.01.002
12. Gu, J., Gu, M., Gu, X.: Optimal rules for single machine scheduling with stochastic breakdowns. Mathematical Problems in Engineering **2014**, 260415 (2014), https://doi.org/10.1155/2014/260415
13. Khatami, M., Salehipour, A.: Coupled task scheduling with time-dependent processing times. Journal of Scheduling (2021), https://doi.org/10.1007/s10951-020-00675-2, iD: Khatami2021
14. Koulamas, C.: The total tardiness problem: Review and extensions. Operations research **42**(6), 1025–1041 (1994), https://doi.org/10.1287/opre.42.6.1025, doi: 10.1287/opre.42.6.1025; 28
15. Leung, J.: Handbook of Scheduling: Algorithms, Models, and Performance Analysis. Chapman & Hall/CRC Computer and Information Science Series, CRC Press, USA (2004). https://doi.org/10.5555/983603, https://books.google.cl/books?id=MAY1ZstmGPkC
16. Liu, X., Chu, F., Zheng, F., Chu, C., Liu, M.: Parallel machine scheduling with stochastic release times and processing times. International Journal of Production Research **0**(0), 1–20 (2020). https://doi.org/10.1080/00207543.2020.1812752, https://doi.org/10.1080/00207543.2020.1812752
17. Maghrebi, M., Periaraj, V., Waller, S., Sammut, C.: Using Benders Decomposition for Solving Ready Mixed Concrete Dispatching Problems (2014)
18. Maghrebi, M., Periaraj, V., Waller, S.T., Sammut, C.: Column generation-based approach for solving large-scale ready mixed concrete delivery dispatching problems. Computer-Aided Civil and Infrastructure Engineering **31**(2), 145–159 (2016), https://doi.org/10.1111/mice.12182, https://doi.org/10.1111/mice.12182; 24
19. Megow, N., Uetz, M., Vredeveld, T.: Stochastic online scheduling on parallel machines. In: Proceedings of the Second International Conference on Approximation and Online Algorithms. p. 167–180. WAOA'04, Springer-Verlag, Berlin, Heidelberg (2004). https://doi.org/10.1007/11389811_15, https://doi.org/10.1007/11389811_15
20. Mojtaba, M., Vivek, P., Travis, W.S., Claude, S.: Solving ready-mixed concrete delivery problems: Evolutionary comparison between column generation and robust genetic algorithm, https://doi.org/10.1061/9780784413616.176, 04; doi:10.1061/9780784413616.176
21. Pinedo, M.L.: Scheduling: Theory, Algorithms, and Systems, Fourth Edition. SPRINGER, NEW YORK; 233 SPRING STREET, NEW YORK, NY 10013, UNITED STATES (2012), pT: B; NR: 562; TC: 736; PG: 673; GA: BZI32; UT: WOS:000301677200021
22. Tamaki, H., Hasegawa, Y., Kozasa, J., Araki, M.: Application of search methods to scheduling problem in plastics forming plant: a binary representation approach. In: Proceedings of 32nd IEEE Conference on Decision and Control. pp. 3845–3850 vol.4 (Dec 1993). https://doi.org/10.1109/CDC.1993.325943
23. Wu, X., Zhou, X.: Stochastic scheduling to minimize expected maximum lateness. European Journal of Operational Research **190**(1), 103–115 (2008), https://www.sciencedirect.com/science/article/pii/S0377221707005656, iD: 271700
24. Zhang, G.Q., Wang, J.J., Liu, Y.J.: Scheduling jobs with variable job processing times on unrelated parallel machines. TheScientificWorldJournal **2014**, 242107 (2014)

25. Zhou, X., Cai, X.: General stochastic single-machine scheduling with regular cost functions. Mathematical and Computer Modelling **26**(3), 95–108 (1997), https://www.sciencedirect.com/science/article/pii/S0895717797001349, iD: 271552

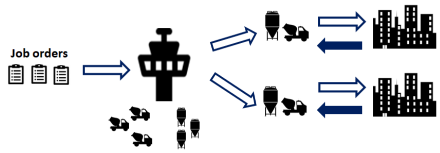


Fig. 8.1 Sketch of the the real-world case study.

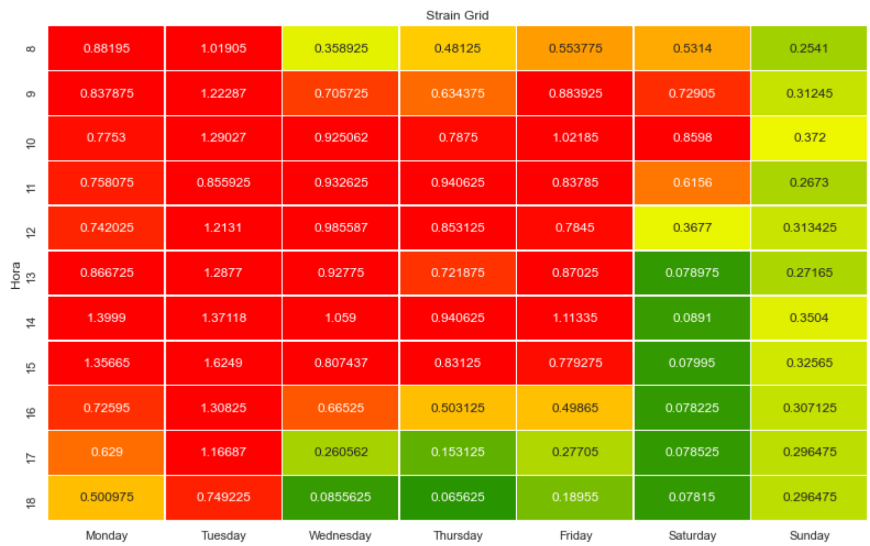


Fig. 8.2 Scaled Operational Strain Grid

Chapter 9
Motivating Artificial Intelligence and Natural Computation Students with NetLogo

P. B. de Moura Oliveira, E. J. Solteiro Pires, and M. S. Fraga

Agents become a mandatory topic in most artificial intelligence courses worldwide. Introducing intelligent agents and multi-agent systems to undergraduate students can be facilitated by an easy to use agent-based modeling and simulation tool. This is the case of NetLogo. Also, in more specific postgraduate courses, such as Natural Computation, NetLogo can be successfully used to model nature and biologically inspired phenomena. This work reports two NetLogo case studies that have been successfully used as laboratory assignments by undergraduates of an artificial intelligent course and master's students of a natural computation course.

9.1 Introduction

Living in the twenties in the 21st century, both teachers and students face new challenges in the teaching/learning methodologies used. This century students are quite different from the 20th century ones. However, it is well known that some lecturers still teach with last century (and older) methodologies. Some of the key questions currently faced by teachers are the following: How to increase students' motivation? Which innovative teaching techniques and technologies to use? How to increase students critical thinking? How to reduce student's passivity in classes? These questions concern teachers of different scientific areas, including computer science. This article addresses the use of an agent-based modelling and simulation software called NetLogo, developed by [1].

P. B. de Moura Oliveira, E. J. Solteiro
Department of Engineering, University of Trás-os-Montes and Alto Douro (UTAD), Vila Real, Portugal

all authors
INESC-TEC Technology and Science, Campus da FEUP, Porto, Portugal {oliveira@utad.pt; epires@utad.pt; milenesofiaa@hotmail.com}

NetLogo is a freely available tool [1] which has been applied in a wide range of different domains [1] to model multi-agent systems. Examples of studies addressing the use of NetLogo within the scope of the above stated questions are the following: motivating students and increasing their participation based on participatory simulation is reported in [2-4] as a form to actively involve students in the learning process, by simulating complex dynamic systems; the encouragement of critical and scientific thinking is addressed both for high-schools and undergraduate students in [5, 6]; within computer science intelligent agents teaching and complex system modelling and simulation [7-10].

In this paper, two case studies are reported: the first concerns a practical assignment proposed to undergraduate students of artificial intelligent. This assignment was proposed the year 2021-2022, relates to one of the first topics addressed in this course: intelligent agents. The second case study addresses the modeling and simulation of covid-19 virus (SARS-Co-V-2) using a multi-agent system. An early version of this case study was proposed as a natural computation course assignment in the 2019-2020 year. Since covid-19 appeared a myriad of techniques have been used to modelled and try to forecast its spread. Examples of studies addressing this problem using multi-agent systems and NetLogo are the following: the modelling of the covid-19 transmission within facilities and evaluate strategies for its reduction is addressed by [11]; the spread of covid-19 in Iran and China is addressed, respectively in [12,13]; a study analyzing the effects of mobilities and travel is reported in [14]; effects of using masks and keeping the social distance to control covid-19 is reported in [15]; the efficacy of measures to control the indoor virus transmission is presented in [16]; the virus transmission in sports facilities is analyzed in [17].

9.2 Case Study I: Grass Planting Agent

The first case study reports an assignment regarding rational agents executed by computer engineering undergraduate students in an artificial intelligence introductory course. The problem addressed is the plantation of grass if an empty field, which represents the world. The machine planting the patches of grass is represented by an agent, designated planter. The planter has two operating modes: planting and maintenance. Initially he must plant all the field patches to obtain a perfect lawn. The world origin and start position for the planter is defined as the left lower corner (see Fig. 1). From that position the grass can be planted by selecting several modes (serpentine, strait ahead, etc.). Students should implement several planting algorithms.The following premises/rules must be fulfilled:

- The field patches can have two states: not planted (represented by brown color) a planted (represented by green color). • The planter heading direction can only be directed to four angles: 0º, 90º, 180º and 270º.
- The planter can only perceive the single patch state located in front of it.
- The planter can only move one patch in each discrete time step (tick).
- In the planting mode, the planter can only plant over a non-planted cell.

When all the field is planted the planter agent changes to a maintenance mode, with the main task to repair damaged patches of grass. At this stage another specie of agent is introduced: the mole. Moles can be introduced in the field in a random location and with a random heading direction, manually by the user or automatically following a probabilistic law. When moles appear in the field, they destroy the grass in the respective patch where they are located. The following additional rules must be fulfilled:

- The activity of moles is determined by a probability value defined by the user using a graphical slider. As this probability value increases the more active should moles be, circulating more in the field and damaging more patches.
- If the planter shocks with a mole in the same patch, the mole dies.
- If two moles move to the same patch a new mole may be born and inserted into the field. Another variable defines the probability of two moles generating a new one.

Besides implementing the proposed predefined protocol, which includes the interface design and respective agent's behavior programming code, students must propose and implement alternative ideas for agent's behavior. They can also include other species of agents with different goals. Part of the grade obtained by students in this assignment depends on the "novelty" associated with the variants proposed. Besides the agent-based learning skills, the aim is to promote critical thinking. In total 90 students executed this assignment both in some laboratory classes and within their individual study component. Students were organized in groups of two, and one week after the deadline to conclude the assignment must deliver a written report. All groups performed a brief oral assignment defense to the teaching staff. As an example, a screenshot is presented in Fig. 1, illustrating the graphical user interface developed a group of students. The feedback received from students both in laboratory classes and in the oral assignment defense was quite positive. Overall students were clearly motivated to develop different agent behaviors.

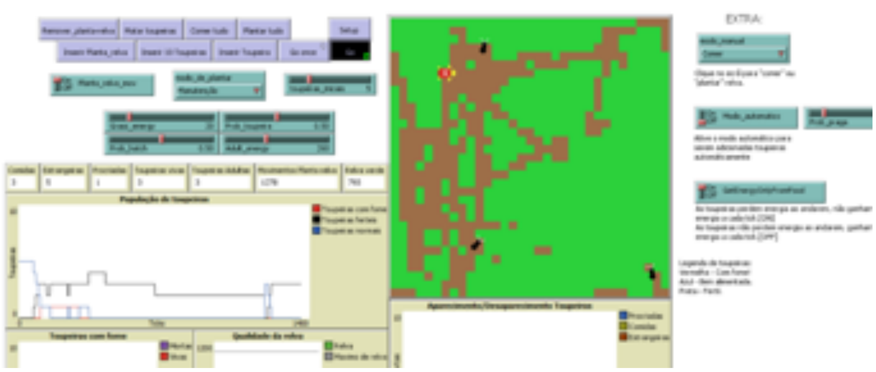
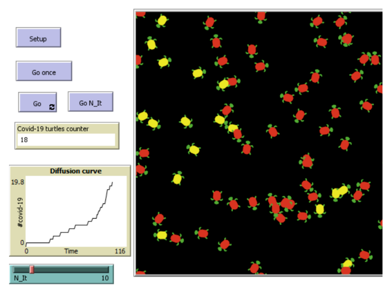


Fig. 9.1 Screenshot of a graphical user interface developed by a group of students.

9.3 Case Study II: Covid-19 Virus Spread Simulation

The second case study idea reported here started as a simple assignment protocol proposed within a natural computation course and evolved into a master's dissertation topic carried out by this paper third author. It is important to state that the students background attending this course ranges from informatic engineering, biomedical engineering and bioengineering, with diverse interests in terms of intelligent agents' applications. The global objective of this work is to explore and develop multi-agent models which can be used to simulate the Corona (SARS-Co-V-2) virus spread, generally referred as covid-19 virus. The scenario can consider a small village or a University population. Initially, there are two basic types of agents: non-contaminated with covid-19 (represented with red color) and contaminated with covid-19 (represented with yellow color). An illustration is presented in Fig. 2.



a) Initial situation. a) Initial situation. b) Evolution after 100 ticks.

Fig. 9.2 Covid-19 contamination evolution: case 1. Agents with no masks.

A discrete time-step (tick) can represent a day in the simulations presented. Different model versions were developed with different complexity levels. The following cases are reported here:

- Case 1: An agent is contaminated if it is in the same patch as a contaminated agent. This case represents the almost total lack of information available to the general population regarding the covid-19 spreading at the end of 2019 and early 2020.
- Case 2: there are three types of agents circulating in the world: non-contaminated without mask (represented by red color); non-contaminated with mask (represented by blue color) and contaminated (represented by yellow color). This models the early 2020 months, with the recommendation adopted in most European countries to use masks to prevent the covid-19 spreading.
- Case 3 (see Fig. 3): there are four types of agents circulating in the world: non-contaminated without mask (represented with red color); non-contaminated with a mask (represented by blue color), vaccinated (represented by green color) and contaminated (represented with yellow color). This case corresponds to the situation in which a user defined population percentage is vaccinated. As it can

be observed from the results presented in Fig. 3, the covid-19 diffusion curves generated by this model are oscillatory with cyclic increases and decreases regarding the number of infected agents. The number of agents can be adjusted by the user. In this case the number of initial agents contaminated was set to 20, the non-mask agents were 400, the agents with masks were 2300 and the vaccinated 300. The probabilities of the different types of agents to be contaminated are the ones specified in the slider presented in Fig. 3. Note, that a variable was defined corresponding to the number of days that, without other complications, an agent is considered seek with covid-19. This value can also be changed, has it has been adjusted over time and it may differ from country to country (e.g. 7 days or 14 days). So, if a contaminated agent respects its quarantine period and keeps itself isolated, the proposed models consider that it does not contaminate other agents during that period. Also, another variable defines a time window corresponding to the average number of days that an agent, after being infected with covid-19, is considered immune. Thus, in this case, even if an agent is in the same patch as another contaminated agent, it is not contaminated.

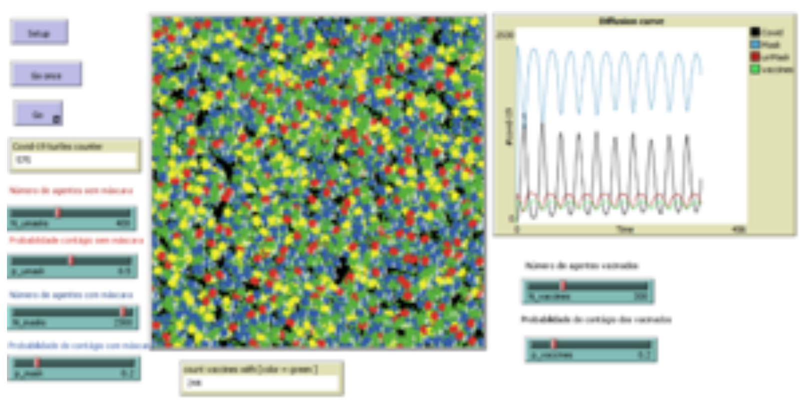


Fig. 9.3 Covid-19 contamination evolution: case 3. Agents with masks, no-masks and vaccines. Evolution after 1 year.

9.4 Conclusion

Two case studies concerning the use in classes of participatory simulations involving multi-agent systems were reported. The agent-based modelling and simulation environment used is the NetLogo. The courses in which these teaching/learning experiment took place are artificial intelligence for informatics engineers undergraduates' students of and natural computation for master's students. The results obtained in terms of student's engagement and active participation in mainly laboratory classes, but also in theoretical ones, are quite positive. The feedback received by students, mostly from oral statements and from statements included in pedagogical enquires, confirms that students really enjoy using NetLogo. Regarding the grass planting agent case study, presenting innovative variants to the original protocol proposed by staff, motivated an interesting competition among students, which naturally motivated them to work hard on the project.

References

1. Wilensky, U., (2022), NetLogo, http://ccl.northwestern.edu/netlogo/, retrieved in 1/5/2022.
2. Wilensky U. and Stroup W., (1999) Learning through participatory simulations: Network-based design for systems learning in classrooms. In Computer- Supported Collaborative Learning (CSCL), Dec. 1999.
3. Kopf S., Scheele N., Winschel L. and Effelsberg W., (2005) Improving Activity and Motivation of Students with Innovative Teaching and Learning Technologies, Proc. of Methods and Technologies for Learning, pp. 551 – 556, Palermo, Italy.
4. Brady C., Orton K., Weintrop D., Anton G., Rodriguez S., and Wilensky U., (2016), All Roads Lead to Computing: Making, Participatory Simulations, and Social Computing as Pathways to Computer Science, in IEEE Transactions on Education, vol. 60, no. 1, pp. 59-66, Feb. 2017, doi: 10.1109/TE.2016.2622680.
5. Gammack D., (2015), Using NetLogo as a Tool to Encourage Scientific Thinking Across Disciplines, Journal of Teaching and Learning with Technology, Vol. 4, No. 1, June, pp. 22-39. doi: 10.14434/jotlt.v4n1.12946
6. 6. Musaeus L. H. and Musaeus P. (2019), Computational Thinking in the Danish High School: Learning Coding, Modeling, and Content Knowledge with NetLogo, SIGCSE '19: Proceedings of the 50th ACM Technical Symposium on Computer Science Education, February, pp. 913–919, https://doi.org/10.1145/3287324.3287452.
7. Dickerson M., (2014), Multi-agent simulation, NetLogo, and the recruitment of computer science majors, Journal of Computing Sciences in Colleges, Vol. 30, Nº 1, pp. 131-139
8. Dickerson M., (2015), Agent-based modeling and NetLogo in the introductory computer science curriculum: tutorial presentation, Journal of Computing Sciences in Colleges, Vol. 30, Nº 5, pp. 174-177
9. Tullis J. G. and Goldstone R.L., (2017), Instruction in Computer Modeling Can Support Broad Application of Complex Systems Knowledge. Front. Educ. 2:4. doi: 10.3389/feduc.2017.00004
10. Sakellariou I., Kefalas P. and Stamatopoulou I., (2008), Teaching Intelligent Agents using NetLogo, Proc. of the ACM-IFIP IEEIII 2008 Informatics Education Europe III Conference, Venice, Italy, December 4-5.
11. Cuevas, E, (2020), An agent-based model to evaluate the COVID-19 transmission risks in facilities, Computers in Biology and Medicine, Vol. 121, https://doi.org/10.1016/j.compbiomed.2020.103827

12. Gharakhanlou N. M. and Hooshangi M, (2020), Spatio-temporal simulation of the novel coronavirus (COVID-19) outbreak using the agent-based modeling approach (case study: Urmia, Iran), Informatics in Medicine Unlocked 20, https://doi.org/10.1016/j.imu.2020.100403
13. Wei Y, Wang J., Song W., Xiu C., Ma L. and Pei T, (2021), Spread of COVID-19 in China: analysis from a city-based epidemic and mobility model, Cities Journal, Vol. 110, Elsevier, https://doi.org/10.1016/j.cities.2020.103010
14. Shannaa M. and Abdallah S., (2020), Agent-based simulation of COVID-19 outbreak within a semi-closed environment, First International Conference of Smart Systems and Emerging Technologies (SMARTTECH), DOI 10.1109/SMARTTECH49988.2020.00059
15. Li K. K. F., Jarvis S. A. and Minhas F., (2021), Elementary Effects Analysis of factors controlling COVID-19 infections in computational simulation reveals the importance of Social Distancing and Mask Usage, Computational Biology in Medicine, Jul;134:104369, doi: 10.1016/j.compbiomed.2021.104369.
16. Farthing T. S. and Lanzas C., (2021), Assessing the efficacy of interventions to control indoor SARS-Cov-2 transmission: An agent-based modeling approach, Epidemics Vol. 37, https://doi.org/10.1016/j.epidem.2021.100524.
17. Qi B., Tan J., Zhang Q., Cao M., Wang X. and Zou Y., (2021) Movement Route Model, Non-Overcrowding and Social Distancing Reduce the Spread of COVID-19 in Sporting Facilities. Int. J. Environ. Res. Public Health, 18, 8212. https://doi.org/10.3390/ijerph18158212.

Chapter 10
Airplane Boarding Simulation using NetLogo

João Costa, Rodrigo Pereira$^{0000-0002-6623-8084}$, E. J. Solteiro Pires$^{0000-0003-3224-4926}$ and P. B. de Moura Oliveira$^{0000-0003-4283-1243}$

Abstract Airplane turnovers are one main worry for airline companies because they can cost airline companies thousands of Euros and impact customer satisfaction due to delays. Airplane boarding is one of the most critical events, and it is necessary to have an appropriate boarding strategy to reduce boarding time. Several strategies are used by companies, such as random boarding, boarding back-to-front, or boarding by blocks. Other more efficient strategies exist but are difficult to implement in real life. In this work, the most used strategies used by companies were simulated using the NetLogo, a multi-agent system, to analyze and clarify the best boarding strategy and compare it to one of many ideal boarding strategies Wilma. A scenario was defined for a narrow-body airplane with one entrance with 24 rows and six seats per row to analyze the boarding performance. Excluding unpractical strategies, the random boarding strategy presented an efficient way of boarding passengers into the airplanes requiring little support from staff.

10.1 Introduction

The airplane has become an integral part of our everyday transport. According to the Federal Aviation Administration (FAA), there are over 20,000 commercial flights per day in the US [4]. The competitiveness among Airline companies requires keeping the airplanes running as long as possible. The boarding is an inevitable process set that can influence the total time on the ground. Also, the passengers' satisfaction is closely associated with the boarding process, which in return, influences the business strategy of airline companies [3]. Airline companies use several boarding strategies

J. Costa, R Pereira, EJ Solteiro, and PB de Moura Oliveira
Departamento de Engenharias, Universidade de Trás-os-Montes e Alto Douro, e-mail: al59259,al68798@alunos.utad.pt, e-mail: epires,oliveira@utad.pt

EJ Solteiro, and PB de Moura Oliveira
INESC-TEC Technology and Science, Porto, Portugal

to board passengers, such as rear-to-front, Wilma (Windows-Middle-Aisle), random, and blocks, among others [11]. In this work, some methods of boarding passengers are studied using multi-agent systems (MAS) and analyzing them to identify the best strategy to board an airplane that takes a minor time.

Intelligent agents are systems capable of automatically understanding and modifying their surrounding environment to find the best results. An agent is anything that can perceive its environment through sensors and act upon the environment through actuators [7]. Intelligent agents are commonly used to solve and analyze optimization problems through MAS.

Some works addresses the boarding problem. Audenaert *et al.* [1] study the boarding problem via multi-agent-based simulation considering several boarding methods were studied, including some disturbances, such as passenger compliance, passengers' groups, and transit passengers. The strategies to board the airplane used in their paper consider a division into blocks of similar size so that passengers assigned to the same block board the plane simultaneously. The order in which a block of passengers can board the plane can be ascending or descending according to their block number. Another strategy used divided the plane into six blocks that start at the bottom of the plane, and then three blocks are skipped, and when there are no more blocks, the ordering restarts again from the back of the plane. Another strategy considered the rows in blocks that can be board mixed and divide the two aisles of one row into two half-blocks. Another strategy used was to perform boarding from windows to the aisle, possibly alternated. Another strategy used was to call every passenger by their seat number individually. Finally, the last strategy used blocks, which will be further refined in sub-blocks using the seats of the same columns, and these sub-blocks are ordered via a pyramidal structure. It is important to note that more strategies were used that are sub-variants of those explained above. Audenaert *et al.* concluded that the passengers' strategies to board based on the seat number and the strategy by pyramid are not very passenger-friendly and costly to implement.

Climer *et al.* [5] experimented six boarding methods using an agent-based computational model and suggested a new method that reduced the boarding time significantly. The boarding methods compared were found in the literature and focused on the following methods: Random, that all the passengers board without order; Wilma, that all passengers seated at the windows are boarding in the first group, followed by the middle seat group and the aisle seat group; Back-to-front, that is boarding from the back to the front with the passengers from the window boarding at first; Blocks, that boarding is in four-row blocks were the back four rows are the first boarding group, followed by the front block and finishing with the center four rows block; Steffen [9] proposed a method where adjacent passengers in line are sitting two rows apart from each other in corresponding seats (*e.g.*, 12A, 10A, 8A, 6A, *etc.*). Kautzka's model combines three methods: Wilma, Back-front, and parallel luggage stowing. The Netlogo tool was used to evaluate the different methods, where different variables were used to improve the simulation, such as passengers with luggage and if a passage is delayed, among others. The authors conclude that the method proposed by Steffen and the model proposed by them in a theoretical way is very effective. These authors also mention that these arrangements for passengers do not seem to be realistic because passengers are not used to them.

The rest of the paper is divided into four sections. Section 10.2 presents the boarding challenges. Next, Section 10.3 describes the used strategies. Section 10.4 shows the results obtained. Finally, Section 10.5 draws the main conclusions.

10.2 The challenges behind airplane boarding

An airline company generally only makes money when it is flying. The time lost on the ground, either boarding or deboarding passengers, is the turnaround. The turnaround time is a traditional metric used to measure the efficiency of airline operations and is the very basis of the company's sustainability in terms of generated costs. Turnaround is defined as the period between the airplane's arrival, and departure from an airport [6]. Steiner and Philipp [10] estimate that the per-minute cost for each airplane is between $30 and $77 during the turnover period. So, it is valuable time and money that airline companies always try to reduce.

Various strategies [2] are used by airline companies to board passengers as fast as possible. The most used strategies are the Back-to-Front strategy, where passengers seated in the last rows enter first. The random strategy, where the passengers enter the plane without any order. The block strategy, where passengers are divided into blocks, and the boarding procedure is made one block at a time. From a theoretical point of view, these strategies are not very efficient, as shown in various papers. Thus, new boarding strategies have emerged, as seen before, but these strategies are very hard to implement in the real context, mainly because the time and cost of organizing the passengers at the result do not compensate for the time that might be saved. In this work, the main strategies (Random, Back-to-Front, and by Block) are analyzed that can be applied and used by airline companies. Moreover, the Wilma method (the more efficient method) was implemented to serve as a reference in the comparison [8].

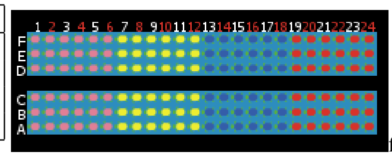
This work focused on implementing a simulation system for the vast majority type of airplanes currently in operation are narrow-body airplanes. These airplanes can take around 250 passengers and mainly do short and medium-haul flights. It was also considered one boarding gate because many airports used by these airplanes have only one boarding gate attached to the aircraft.

10.3 Strategies to solve the problem

The NetLogo tool is used to achieve the objective of this work. This tool is a Multi-Agent programmable model environment and can be used to create and program simulation models that can be very useful for analyzing efficiency problems, par-

Color	Rows	Block
Pink	1 – 6	1
Yellow	7 – 12	2
Blue	13 – 18	3
Red	19 – 24	4

Table 10.1 Block colors



figurePassengers Blocks

ticularly airplane boarding. This tool has a counter named *tick* considered the time unit.

This work simulates passengers boarding an aircraft. In this simulation, an example of a narrow-body aircraft like Boeing B737 or Airbus A320. This aircraft only has one entrance, which is in the front, to all passengers enter. There are 24 rows of seats, and each row of seats has three seats on the left and three seats on the right, with the aisle in the middle.

To make the simulation more accurate, each passenger will take 2 or 3 units of time to sit in their seat if the row that he will seat is not already occupied by another passenger. If the row that the passenger will sit is already with any other passenger seat, it will be added 2 more time units to their sitting time. For example, if a passenger has a window seat on the row number 1 on the left side of the plane and if the aisle and the middle seats are not occupied, the passenger will take 2 or 3 units of time, but if the aisle or middle seat has already someone seated, it will be added 2-time units to their sitting time. If the aisle and middle seats are already occupied, 4-time units will be added to their sitting time. The sitting time variable applied to any passenger has the objective of simulating the time that passengers will be occupying the aisle, fitting their luggage in the overhead bins, and sitting.

To better understand the behavior of the passengers, colors will be assigned colors to their representation, based on their seat row as indicated in Table 10.1 and Fig. 10.3.

The strategies analyzed in this work are the following:

- Random Entrance: All the passengers enter the airplane without any order.
- Sequential Entrance: The passengers enter by blocks from the end of the plane to the beginning (Block 4 → Block 3 → Block 2 → Block1).
- Alternated entrance: The passengers enter by blocks but alternated (Block 4 → Block 2 → Block 3 → Block1).
- Window-Middle-Aisle Entrance: The passengers enter by blocks from the end of the plane to the beginning (Block 4 → Block 3 → Block 2 → Block1).

10.4 Simulation results

Several simulations were executed to assess the models' performance. In figures 10.1 and 10.2, two strategies simulations are illustrated, namely: sequential and random strategies.

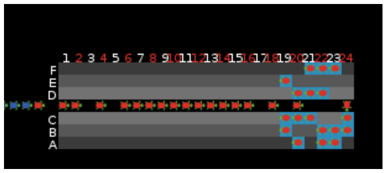


Fig. 10.1 Sequential boarding strategy

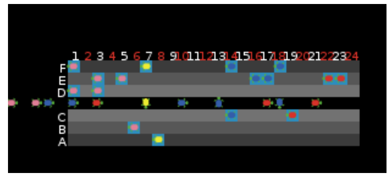


Fig. 10.2 Random boarding strategy

compat=1.15, every axis plot/.append style=line width=0.2pt, every axis plot post/.append style= every mark/.append style=line width=0.2pt,mark size=0.1pt , legend cell align=left, legend style= at=(0.0,-0.35), anchor=south west, legend columns=2,
[scale=0.65] [legend pos=south east, legend columns = 1, ylabel=Seated passagers, xlabel=Time (Ticks), title=Evolution of seated passagers,] table[col sep=comma,x index=0,y index=1] fig/data$_{m}edian.txt; table[colsep = comma, xindex = 0, yindex = 2]fig/data_{m}edian.txt; table[colsep = comma, xindex = 0, yindex = 3]fig/data_{m}edian.txt; table[colsep = comma, xindex = 0, yindex = 4]fig/data_{m}edian.txt;$
Alternate,Sequential,Random,Wilma

Fig. 10.3 Evolution of Seated Passengers

compat=1.15, every axis plot/.append style=line width=0.2pt, every axis plot post/.append style= every mark/.append style=line width=0.2pt,mark size=0.1pt ,
[scale=0.65] [ylabel=Probability density, xlabel=Time (Ticks), title=Seated passagers distributions,] table[col sep=comma,x index=0,y index=1] fig/normal.txt; table[col sep=comma,x index=0,y index=2] fig/normal.txt; table[col sep=comma,x index=0,y index=3] fig/normal.txt; table[col sep=comma,x index=0,y index=4] fig/normal.txt;
Alternate,Sequential,Random,Wilma

Fig. 10.4 Probability of Seated Passengers

To analyze all the models, it is necessary to run each strategy multiple times due to its stochastic nature. Therefore, to achieve this, the "Behavior Space", an integrated software tool of the NetLogo, is used to perform experiments with models, running each multiple times and storing the results of each run. Each strategy was performed 601 times, with their results recorded. This value was a representative number of runs, allowing particularly fast or slow variations of random passenger allocation to be averaged out while also not taking very long to compute all the runs. After recording the runs, the median number of passengers seated at each tick was evaluated for each strategy, depicted in Fig. 10.3, and the density probability and density were calculated, illustrated in Fig. 10.4.

With the results obtained (see Figs. 10.3-10.4), the following analysis can be made:

- **Sequential Entrance**: Although this strategy seems, at first glance, a good strategy by ensuring that the passengers are seated from back to front, we can see by the simulation results that it is, in fact, one of the slowest strategies. This method took, on average, 450-time units (ticks) to seat all the passengers.
- **Alternate Entrance**: This strategy is equivalent with the Sequential Entrance, with an average of 440-time units (ticks) to seat all passengers.

- **Wilma Entrance**: This is by far the fastest strategy, taking only an average of 300-time units (ticks) to sit all passengers properly. Although it can see by the graph in Fig. 10.3 that there are significant time gaps where no one is sitting (for example, from tick ~50 to ~70), it is followed by a period where passengers quickly get to their seats. This strategy is split between a preparation and a sitting phase, where each row of passengers walks to their seats (preparation), followed by a quick sitting phase where the entire row seats almost simultaneously.
- **Random Entrance**: Almost counter-intuitively, this is the second-fastest strategy of those tested, with an average of 409-time units (ticks) to seat every passenger.

The *random entrance* strategy has a better performance than both the *alternate* and *sequential* entrances because, as the passengers and distributed randomly, it allows for more parallelism in sitting – passengers are more likely to reach their seat at the same time, cutting down on time waiting in line – while this is rarer with the *alternate* and *sequential* strategies, as there are a fewer number of rows *ready to sit* (only one Block is *ready to sit* at any given time).

10.5 Conclusions

Airplane boarding strategies are an area that is being studied with some interesting and efficient strategies proposed. However, the most efficient technique can not be applied in the real world since the time and cost necessary to organize the people to enter the airplanes in the right order. Analyzing the experiments performed in this work leads to conclude that the various boarding strategies used by airline companies do not differ a lot from each other. However, in practice, the *random boarding* strategy is more efficient because it introduces some aleatory to the boarding allowing, for example, three or more passengers can sit down at the same time in different rows. Also, the MAS used in this work, recurring to the tool NetLogo, can simulate and analyze efficient strategies that can be applied to many other industries and problems.

References

1. Audenaert, J., Verbeeck, K., Berghe, G.V.: Multi-agent based simulation for boarding. In: The 21st Belgian-Netherlands Conference on Artificial Intelligence. pp. 3–10 (2009)
2. Bazargan, M.: A linear programming approach for aircraft boarding strategy. European Journal of Operational Research **183**(1), 394–411 (2007)
3. Bidanda, R., Winakor, J., Geng, Z., Vidic, N.: A review of optimization models for boarding a commercial airplane. In: Proccedings of 24th International Conference on Production Research. Poznan, Poland. pp. 1–6 (2017)
4. Bureau of Transportation Statistics: Bts updates national transportation statistics with sales price of transportation fuels, condition of u.s. highway bridges, vehicle miles traveled, u.s. air carrier delays & cancellations, and other historical tables (access: 2022), https://www.bts.gov/

5. Cimler, R., Kautzká, E., Olševičová, K., Gavalec, M.: Agent-based model for comparison of aircraft boarding methods. In: Proceedings of 30th International Conference Mathematical Methods in Economics Agent-Based. pp. 73–78 (2012)
6. Delcea, C., Cotfas, L.A., Paun, R.: Agent-based evaluation of the airplane boarding strategies' efficiency and sustainability. Sustainability **10**(6), 1879 (2018)
7. Norvig, P.R., Intelligence, S.A.: A modern approach. Prentice Hall Upper Saddle River, NJ, USA: (2002)
8. Nyquist, D.C., McFadden, K.L.: A study of the airline boarding problem. Journal of Air Transport Management **14**(4), 197–204 (2008)
9. Steffen, J.H.: Optimal boarding method for airline passengers. Journal of Air Transport Management **14**(3), 146–150 (2008)
10. Steiner, A., Philipp, M.: Speeding up the airplane boarding process by using pre-boarding areas. In: 9th Swiss Transport Research Conference, Ascona, 9-11 September (2009)
11. Wong, K.: How different airlines board their passengers (2017), https://lifehacker.com/how-different-airlines-board-their-passengers-1794474413

Chapter 11
Artificial Intelligence for Advanced Manufacturing Quality Control

Fátima A. Saiz[0000-0001-6065-7029], Iñigo Barandiaran [0000-0002-8080-5807] and Manuel Graña

Abstract This paper describes the work plan to be carried out in my PhD thesis, which deals with the development of quality control systems for the manufacturing industry by applying advanced computer vision techniques based on artificial intelligence. The development of these systems involve different challenges that need to be addressed throughout the thesis, such as the difficulty of acquiring images of very specular surfaces, the lack of defective samples or the design of reliable and fast defect detection networks. This paper shows an introduction to these problems and the motivation that leads me to research on these challenges, as well as the planning and progress developed so far.

11.1 Introduction

The inspection stage in the quality control processes of the manufacturing industry is fundamental to avoid a defective final product. This quality control must be not only assured in the final products but also must be present throughout the entire production process.

Techniques such as statistical process control (SPC) are often used in production processes. However, the arrival of automated computer vision and Artificial Intelligence (AI) systems involve a significant change, as they allow the inspection of all products in a reduced timeframe. This generates significant benefits for organisations

FA Saiz and I Barandiaran
Vicomtech Foundation, Basque Research and Technology Alliance (BRTA), Mikeletegi 57, 20009 Donostia-San Sebastián, Spain

M Graña
Computational Intelligence Group, Computer Science Faculty, University of the Basque Country, UPV/EHU, 20018 Donostia-San Sebastian, Spain e-mail: {fsaiz,ibarandiaran}@vicomtech.org
e-mail: manuel.grana@eh.es

due to the reduction of waste, rejects, claims and returns, resulting in lower costs and higher productivity.

Currently, many of these inspection procedures are carried out manually or only a few samples of each batch are inspected. The lack of full automation of the process has several drawbacks:

- No data traceability.
- Human error due to fatigue and other factors.
- Limited analysis.
- Data does not generate process knowledge.

However, by using automated computer vision and AI systems the data is collected properly. Its exploitation allows decision-making at different stages of the industrial process, leading the organisation towards Industry 4.0 [18].

The trend of organisations moving towards the integration of AI into their processes is a reality. The emerging impact in terms of labour productivity, personalisation, time savings and quality is on a growing trend and is expected to continue to increase in the coming years, as shown in the study "Harnessing the economic benefits of Artificial Intelligence" by the European Commission. [1].

11.2 Motivation

The motivation of this PhD thesis is to develop AI-based machine vision systems that are capable of improving current quality control systems throughout all stages of the production process. More precisely, the thesis is focused on the research of the current Deep Learning challenges. To this end, the following key points will be addressed:

11.2.1 Image acquisition methods

The success of any vision system depends largely on the quality of the data, images in this context. The selection of a suitable acquisition system is not an easy task, especially in an industrial environment, which has different particularities and challenges, such as:

- Metallic specular surfaces.
- Small defect size.
- High production rate.

All these characteristics require the use and/or the development of image acquisition techniques that provide optimal "raw material", i.e. images, for training machine learning algorithms. To avoid these problems, the application of the photometric-stereo technique will be studied [8].

11.2.2 Synthetic data generation

In any given components manufacturing process, the number of defective parts is very low. This fact negatively impacts the process of training a defect detection model, due to the lack of defective samples. Given this limitation, synthetic data generation is a powerfull and a promising tool. Thanks to approaches in AI, major advances are being made in the development of controllable generative image models.

One of the goals in this task is the quantification of uncertainty. Current generative models such as GANs do not provide good estimates of uncertainty in image generation. This problem leads to a lack of reliability in the outputs or decisions of these models. To improve reliability progressive uncertainty-driven learning techniques are currently being applied to improve the generalization and performance of these networks [2].

11.2.3 Anomaly detection

The aim of deepening into anomaly detection is to minimise the number of training samples needed. As stated previously, due to a limitation in the number of defective samples is essential to research methods for minimising requirements such as the number of required samples or the dataset size. A growing number of studies propose anomaly detection approaches to get over these limitations, such us [9][10]. The result of this technique is to select/identify and/or segment that which is not normal.

11.2.4 Model visualization

This task involves understanding visually which image features are the target of the model during training and inference. The generation of visualizations allows decision-making on model architecture design and on the image processing operations in order to achieve optimal results. [3]

Neural networks are becoming a fundamental tool in the manufacturing industry to obtain a human-like level of learning in certain tasks, such as classification or detection. Yet, due to their opaque character, it becomes impossible to understand why a neural network makes a decision and whether that decision is trustworthy [4]. Similarly, we expect that the neural network will make decisions that we consider legal, ethical and fair, trying to minimize the biases that may exist.

Ideally, these decisions should be supported by some kind of justification so that the user can successfully understand what process has led the neural network to make that decision. This makes applicability and auditability features currently in demand. For this reason, in recent years a large number of studies and methods have been developed with the aim of providing an explanation of decision making and giving auditability to a neural network [5][6][7], thus giving rise to a new field of study.

11.2.5 Model Optimization

The models need to be adapted to the environment in which they will be deployed and therefore there are several actions that need to be carried out in order to optimize their performance. One of them is to apply transfer learning techniques [13]. This type of approaches are based on developing models that take advantage of the experience of other models to use their knowledge to solve related problems. This type of techniques help to reduce the number of training samples, improve performance, reduce training time and memory requirements among many other advantages. Another applicable technique is the adjustment of training hyperparameters, which have an important significance in the final performance of the model. The search of these hyperparameters can be tedious if done manually, so there is a tendency to automate this study using genetic, bayesian or evolutionary algorithm techniques [11] [12], among others.

11.3 Planned approach

The Gantt chart shown in Figure 11.1 shows the planning of the thesis.

The proposed development approach is as follows: first a design of the thesis is made, where the objectives to be covered as well as the projects that can help to achieve these objectives are defined. Subsequently, a state of the art is made, which will be updated periodically throughout the thesis. In addition, the possible contributions to be made are defined. In the next task, the experiments to be performed are defined. These experiments are designed to solve real problems of component quality control in industrial manufacturing. Once completed, validation will be performed on real industrial prototypes. The last tasks are focused on the dissemination results and the writing of the thesis.

Below, the milestones of the Gantt chart are described:

- T1: Initial version of the Research planning.
- T2: Thesis objectives definition.
- S1: State of the art.
- P1: Paper about the contributions of the acquisition techniques for specular surfaces.
- P2: Paper about the dataset generation (synthetic data).
- P3: Paper about data augmentation techniques using GANs.
- P4: Results of the training of anomaly detection models on industrial components.
- P5: Paper about the developed tools to visualize the activations of each layer.
- P6: Paper about the mathematical optimization of models for their convergence.

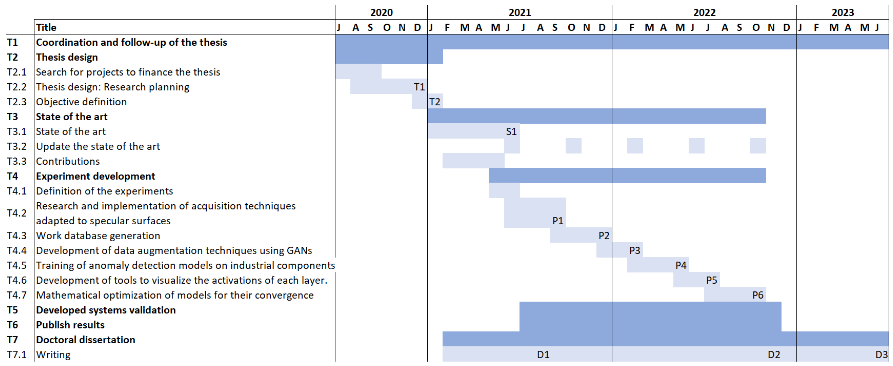


Fig. 11.1 Gantt Chart of the thesis

- D1: State of the art section of the thesis report completed.
- D2: Description of the results written in the thesis report.
- D3: Submit complete thesis.

11.4 Current State and Future Direction

As previously mentioned, advances in experimentation were developed and the results obtained were disseminated. The publications are listed in the following lines, together with a brief description of the work:

- Photometric Stereo-Based Defect Detection System for Steel Components Manufacturing using a deep Segmentation Network [14]: This paper presents an automatic system for the quality control of metallic components using a photometric stereo-based sensor and a customized semantic segmentation network. This system is designed based on interoperable modules, and allows capturing the knowledge of the operators to apply it later in automatic defect detection. A salient contribution is the compact representation of the surface information achieved by combining photometric stereo images into a RGB image that is fed to a convolutional segmentation network trained for surface defect detection. We demonstrate the advantage of this compact surface imaging representation over the use of each photometric imaging source of information in isolation. An empirical analysis of the performance of the segmentation network on imaging samples of materials with diverse surface reflectance properties is carried out, achieving Dice performance index values above 0.83 in all cases. The results support the potential of photometric stereo in conjunction with our semantic segmentation network.

- An Inspection and Classification System for Automotive Component Remanufacturing Industry Based on Ensemble Learning [15]: This paper presents an automated inspection and classification system for automotive component remanufacturing industry, based on ensemble learning. The system is based on different stages allowing to classify the components as good, rectifiable or rejection according to the manufacturer criteria. A study of two deep learning-based models' performance when used individually and when using an ensemble of them is carried out, obtaining an improvement of 7% in accuracy in the ensemble. The results of the test set demonstrate the successful performance of the system in terms of component classification.

- Synthetic Data Set Generation for the Evaluation of Image Acquisition Strategies Applied to Deep Learning Based Industrial Component Inspection Systems [16]: Automated visual inspection is an ongoing machine vision challenge for industry. Faced with increasingly demanding quality standards it is reasonable to address the transition from a manual inspection system to an automatic one using some advanced machine learning approaches such as deep learning models. However, the introduction of neural models in environments such as the manufacturing industry find certain impairments or limitations. Indeed, due to the harsh conditions of manufacturing environments, there is usually the limitation of collecting a high quality database for training neural models. Also, the imbalance between non-defective and defective samples is very common issue in this type of scenarios. To alleviate these problems, this work proposes a pipeline to generate rendered images from CAD models of industrial components, to subsequently feed an anomaly detection model based on Deep Learning. Our approach can simulate the potential geometric and photometric transformations in which the parts could be presented to a real camera to faithfully reproduce the image acquisition behaviour of an automatic inspection system. We evaluated the accuracy of several neural models trained with different synthetically generated data set simulating different transformations such as part temperature or part position and orientation with respect to a given camera. The results shows the feasibility of the proposed approach during the design and evaluation process of the image acquisition setup and to guarantee the success of the real future application.

- A robust and fast deep learning-based method for defect classification in steel surfaces [17]: The final product quality control is critical for any manufacturing process. In the case of steel products, there are different inspection methods that are able to classify the defects, but they usually require human intervention. In this context, a deep learning-based automatic defect classifier method for steel surfaces is proposed. The method combines some traditional Machine Learning techniques with a Convolutional Neural Network (CNN). Different experiments were carried out in order to obtain the best classifier parameter setup. To verily the robustness of the classifier some additional experiments were done, obtaining high classification rate against some sources of noise such as illumination changes or occlusions. The proposed method achieves a classification rate of 99.95% taking 0.019 seconds to classify a single image. The method is compared with seventeen related methods and outperforms them on a publicly available dataset, with six types of defects and 300 samples for each class. The source code of the proposed method is publicly available.

The future direction of this thesis is the publication of an optimised defect segmentation network tested on stereo photometric images. Different experiments in the research direction of training hyperparameters optimization using genetic algorithms are being carried out. We are also developing performance experiments in order to compare with well-known segmentation architectures of the current state of the art.

References

1. European Comission - Advanced Technologies for Industry - Harnessing the economic benefits of Artificial Intelligence, https://ati.ec.europa.eu/reports/technology-watch/harnessing-economic-benefits-artificial-intelligence. Last accessed 4 May 2022
2. Upadhyay, U., Chen, Y., Hepp, T., Gatidis, S., Akata, Z. (2021, September). Uncertainty-Guided Progressive GANs for Medical Image Translation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 614-624). Springer, Cham.
3. Husain, S. S., Ong, E. J., Bober, M. (2021). ACTNET: end-to-end learning of feature activations and multi-stream aggregation for effective instance image retrieval. International Journal of Computer Vision, 129(5), 1432-1450.
4. Xie, N., Ras, G., van Gerven, M., Doran, D. (2020). Explainable deep learning: A field guide for the uninitiated.
5. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., ... Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information fusion, 58, 82-115.
6. Carvalho, D. V., Pereira, E. M., Cardoso, J. S. (2019). Machine learning interpretability: A survey on methods and metrics. Electronics, 8(8), 832.
7. Mueller, S. T., Hoffman, R. R., Clancey, W., Emrey, A., Klein, G. (2019). Explanation in human-AI systems: A literature meta-review, synopsis of key ideas and publications, and bibliography for explainable AI. arXiv preprint arXiv:1902.01876.
8. Klette, R., Kozera, R., Schlüns, K. (1998). Shape from shading and photometric stereo methods. Handbook of computer vision and applications, 2, 531-590.
9. Liu, K., Li, A., Wen, X., Chen, H., Yang, P. (2019, September). Steel surface defect detection using GAN and one-class classifier. In 2019 25th International Conference on Automation and Computing (ICAC) (pp. 1-6). IEEE.
10. Heger, J., Desai, G., El Abdine, M. Z. (2020). Anomaly detection in formed sheet metals using convolutional autoencoders. Procedia CIRP, 93, 1281-1285.
11. Xiao, X., Yan, M., Basodi, S., Ji, C., Pan, Y. (2020). Efficient hyperparameter optimization in deep learning using a variable length genetic algorithm. arXiv preprint arXiv:2006.12703.
12. Victoria, A. H., Maragatham, G. (2021). Automatic tuning of hyperparameters using Bayesian optimization. Evolving Systems, 12(1), 217-223.
13. Zhu, W., Braun, B., Chiang, L. H., Romagnoli, J. A. (2021). Investigation of transfer learning for image classification and impact on training sample size. Chemometrics and Intelligent Laboratory Systems, 211, 104269.
14. Saiz, F. A., Barandiaran, I., Arbelaiz, A., Graña, M. (2022). Photometric Stereo-Based Defect Detection System for Steel Components Manufacturing Using a Deep Segmentation Network. Sensors, 22(3), 882.
15. Saiz, F. A., Alfaro, G., Barandiaran, I. (2021). An Inspection and Classification System for Automotive Component Remanufacturing Industry Based on Ensemble Learning. Information, 12(12), 489.

16. Saiz, F. A., Alfaro, G., Barandiaran, I., Garcia, S., Carretero, M. P., Graña, M. (2021). Synthetic Data Set Generation for the Evaluation of Image Acquisition Strategies Applied to Deep Learning Based Industrial Component Inspection Systems.
17. Saiz, F. A., Serrano, I., Barandiarán, I., Sánchez, J. R. (2018, September). A robust and fast deep learning-based method for defect classification in steel surfaces. In 2018 International Conference on Intelligent Systems (IS) (pp. 455-460). IEEE.
18. Posada, J, Toro, C., Barandiaran, I., Oyarzun, D., Stricker, D., de Amicis, R., Pinto, E.B., Eisert, P., Döllner, J., and Vallarino, I. (2015). Visual Computing as a Key Enabling Technology for Industrie 4.0 and Industrial Internet. IEEE Computer Graphics and Applications 35(26-40).

Chapter 12
Towards the study of high-throughput quantitative imaging in breast imaging

João Pedro Fontes[1 3]0000-0003-4773-5576 and Miguel Angel Guevara Lopez[2 3]0000-0001-7814-1653

Abstract In the last ten years, the computing revolution in artificial intelligence (AI), driven mostly by deep learning and convolutional neural networks, has also pervaded the field of automated breast cancer early detection and diagnosis, which represent a significant contribute to improve traditional computer-aided diagnosis systems (CADx). CADx are utilised in medicine to assist doctors in their diagnosis tasks. Early diagnosis is crucial to identify relevant factors before therapy and might support implementing suitable and personalised conditions for patients and bypass potential side effects. Aligned with the new paradigm in radiology, this work aims to highlight the impact of quantitative imaging for enhancing the prediction algorithms in breast cancer early detection and diagnosis. In particular, the use of quantitative imaging biomarkers (QIBs) has proved helpful to improve the performance of different algorithms in identifying and analysing pathological lesions. The identification of QIBs directly linked to clinical outcomes (i.e., tumour malignancy, cancer phenotypes, metastases) for breast cancer can be a more suitable way to use with different imaging modalities, i.e. without focusing on a specific screening approach, thus improving with the usage of diverse sources of information: digital content (mammography, MRI, ultrasound, CT-scan, or hybrid imaging) and associated metadata.

João Pedro Fontes
Departamento de Sistemas de Informação, Universidade do Minho, 4800 Guimarães
e-mail: mailto://id8968@alunos.uminho.ptid8968@alunos.uminho.pt

Miguel Angel Guevara Lopez
Escola Superior de Tecnologia de Setúbal, Instituto Politécnico de Setúbal, 2910 Setúbal
e-mail: mailto://miguel.lopez@estsetubal.ips.ptmiguel.lopez@estsetubal.ips.pt

João Pedro Fontes and Miguel Angel Guevara Lopez
Centro ALGORITMI, Universidade do Minho, 4800 Guimarães

12.1 Introduction

Medical imaging has evolved drastically since its first use in the 19th century. The use of medical imaging is vital to avoid exploratory surgeries that cause harm to patients, avoiding potentially severe side effects. Beginning with the application of more intrusive methods like X-rays until modern techniques like hybrid imaging, i.e., Positron Emission Tomography–Computed Tomography (PET-CT) and Positron Emission Tomography–Magnetic Resonance Imaging (PET–MRI)). These modern techniques have aided doctors in achieving earlier diagnosis and choosing personalised (more specific) treatment options for their patients. Among the recent developments, we can find the emergent advances in the field of Quantitative Imaging (QI), which seeks to discover numerical or statistical features representing a particular biological process. To those features, we call QI Biomarkers (QIB). A QIB is a characteristic that is measured and evaluated in medical images, which can identified / describe the occurrence of a biological process (e.g. a pathogenic process or the response to a prescribed treatment) [12]. QI and in particular radiomics techniques has taken an essential part in the new "multi-omics paradigm" in medicine. Multi-omics data integration is currently an emergent trend in the field of precision medicine.

This position paper aims to highlight the importance of using QIBs in breast cancer diagnosis (BCD). This work's main proposal lies in the possibility of QIBs from current mainstream imaging modalities in breast cancer diagnosis, benefiting from an expanded availability of data to improve its performance. The use of combined imaging techniques can provide different perspectives and, alongside the extraction of QIBs, give a valuable source of information in the diagnosis task. Relying on an assured diagnosis using imaging techniques can avoid or reduce employing more invasive exams and exploratory surgeries that cause harm to patients. Alongside this, current BCD systems can also benefit from modern improvements in Artificial Intelligence (AI) algorithms to perform better. Deep Learning (DL) approaches represent a logical evolution of classical Machine Learning (ML) methods, giving a theoretical boost in precision imaging and thus, precision medicine tasks performance. ML techniques profited from recent advances in technology (i.e. the introduction of faster CPUs and GPUs) and data availability to overcome their earlier restrictions and achieve results already predicted some years back. The discrepancy between modern DL and older ML algorithms performance is significant, performance-wise, when we increase the amount of data available for training.

In the remaining sections of the manuscript are presented the main ideas behind the hypothesis stated above. Also, it is provided some counter-arguments to the declared position and submits reliable references in its contrary. Finally, we present our view of current and future research trends in image-based breast cancer diagnosis. Some conclusions of the work are also described.

12.2 Background

Medical image analysis is a field mainly supported on computer vision that concerns on extracting critical knowledge and describing medical images/exams like X-ray, ultrasound or MRI [6]. After gathering medical images, computers analyse them to ascertain the presence or absence of lesions, which are difficult to see to naked eye. The influence of medical imaging diagnosis has experienced a continuous increase in clinical decision support systems in recent years, providing a "second opinion" to medical experts, and thus increasing the possibility of physicians to save more lives.

The process of examining these images started to be made only by experts, who would decide the existence or absence of these lesions through the visual analysis of the images collected. With the progress of technology, medical image analysis through computer systems become an aid to specialists, since it helped diagnostic processes using computer vision algorithms. Certain processes became easier, such as organ segmentation [11] or lesion classification [19]. These algorithms and methods proved their usefulness between the 1970s and the 1990s. With the advent of ML algorithms, it was possible to "teach" the machines how to diagnose based on extracted features from the images captured using medical imaging techniques. Medical experts extracted a certain set of handcrafted features from the images and fed those features to ML algorithms to help in the medical expert's diagnosis. By using variety-rich datasets, covering a significant number of possible causes, it ends up being possible to obtain good diagnostic results. The example proposed in [14] shows how to identify cases of breast cancer using machine learning tools.

Traditional machine learning algorithms reached a point where it was not conceivable to obtain better results than the ones obtained at the time. With this, Deep Learning algorithms became the first line of research in medical image analysis. As shown below, Deep Learning algorithms can learn features directly from raw data (e.g. image pixels), so the intervention of an expert for hand-crafted feature extraction is no longer necessary. The medical expert is nevertheless required to increase the algorithm's performance by adding human knowledge to the one extracted by the algorithm. These algorithms can achieve a better performance than traditional machine learning techniques, which will lead to better results in diagnostic and an increase in successful cancer prevention.

12.2.1 Existing approaches

Currently, to improve the performance of Breast cancer CADx systems is one of the biggest research interests in medical image analysis. Since the introduction of DL methods, these have been used both for early detection tasks in screening images and assisting medical experts on classification / diagnosis tasks with a second opinion.

The first uses of DL methods in CADx systems to support breast cancer diagnosis were made by Sahiner et al. in 1996 [17], where a CNN was used to classify between mass and normal breast tissue. This approach had a prepreprocessing stage, where a region of interest was selected, followed by a series of operations (e.g. filters) to facilitate the learning process by the CNN. This two-stage approach was used by other solutions, varying the preprocessing techniques and the input image type (by using mammographies, MRI or tomosynthesis (3D mammography)). Mammographies are the most used exams for detection / classification tasks [17, 8, 10, 7], although MRI [4] and more lately tomosynthesis [9] are starting to get some approaches, despite not being so expressive for lack of samples to train the DL models.

Although recent strategies aim to use feature learning (features extracted automatically) approaches to train classification models, there are some approaches [7, 13, 1] that combine DL feature extraction with handcrafted features designed by experts that demonstrated an increase in the performance / accuracy of the models. The objective is use all the knowledge extracted to train ML and DL models, which in general achieves better results than using more discrete groups of features for training.

12.3 Benefits of Quantitative Imaging in Radiology

The potential of using Quantitative Imaging and discovering QIBs comes with increased benefits: reduced diagnosis time, a standardised way of analysing images, increased automation in getting an initial response and a broader range of possibilities for discovering patterns previously impossible to see for the human eye [18]. The measurement of QIB values is useful to determine the direction that a particular treatment is taking and, consequently, to define future steps based on the observed evolution. QIBs are very useful in cancer treatment, where a robust assessment of the tumour's progression is necessary to adapt the treatment to it [20]. Figure 12.1 shows an example of a QIBs measurement workflow in breast cancer MRI [15]. Moreover, quantitative imaging holds unique potential for radiotherapy purposes but is still not used routinely. Thus, and with the potential described, the research community seeks to develop QIBs that can assist radiologists in their tasks and achieve better diagnostic results.

With already several published research works on the development of QIBs that allow evaluating several diseases of the human body [3, 2, 5, 16], breast cancer is still one of the type of cancer that does not represent a high percent of these works. Therefore, the research community seeks solutions to enhance the results in one of the deadly diseases with high incidence in the world. Since the use of Quantitative Imaging is a core part of the new "quantitative paradigm" in radiology aims to improve personalised medicine, this approach will be useful for enhancing the performance of current CADx methods. This is why, we propose a systematic and combined effort in the development of QIBs, but linked to clinical outcomes to address breast tumours for improving the performance of current Machine Learning methods. Since current methods (usually convolutional neural networks) already detect features using filters, the resulting images can be a starting point to boost the feature discovery and development of new QIBs.

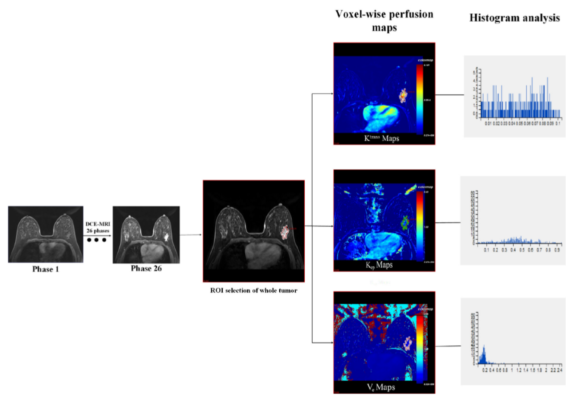


Fig. 12.1 Example of a QIB measurement workflow in breast cancer MRI [15]. The radiologist starts by manually segmenting the tumour, which is then analysed as a single 3D volume. Hence, 3 quantitative features are extracted.

With the embodiment of new discovered QIBs in biomedical data analysis pipelines / clinical workflows, healthcare professionals will have at their disposal a series of new tools that will easy their tasks. QIBs are expected to provide new precise knowledge and potentially accelerating the diagnosis of suspected breast cancer pathological lesions .

Taking into account the arguments presented, it is expected that the development of QIBs can provide a useful tool for radiologists / physicians in their diary tasks. Moreover, the potential shown by quantitative imaging techniques leads us to believe that the CADx systems already used today can benefit from a boost in their performance, leading to better decisions and better results in the treatment of different types of cancer, but also in others chronic non-transmissible diseases like neurodegenerative diseases.

12.4 Discussion

Breast cancer diagnosis systems have already shown that they can be an asset to healthcare professionals. Thus, the construction and validation of a greater number of these systems will be essential to further increase the treatment given to patients and increase their expectation of correct diagnosis and recovery. Since these systems are already highly accepted by the medical community, they are still not 100% utilised on a day-to-day basis because they are catalogued as "substitutes". Thus, the scientific community must seek to improve these systems to show themselves as real aids and facilitators of diagnosis, and not as the so-called substitutes in which solutions based on artificial intelligence are now being called. In addition, the correct validation and testing of these new systems are fundamental in overcoming the reluctance of health professionals to use these systems, showing that they can facilitate their work and provide a better quality of service to their patients.

12.5 Conclusions

This work emphasises the importance of developing new systems to help radiologists/physicians through the formalisation of advanced structures called Quantitative Imaging Biomarkers. QIBs through the calculation of properly tested and validated quantitative features allow the establishment of threshold values indicative of a certain variables / clinical outcomes. Validated QIBs have demonstrated to be very precise and capable of providing health professionals with a variety of potentially useful data (indicators) in the discovery or validation of a diagnosis.

The biggest challenge presented by the QIBs approach is to ensure a direct link between the discovered QIBs and clinical variables, i.e. clinical endpoints and/or surrogate endpoints. Another major challenge concerns the acceptance of QIBs, which involves accurate and extensive validation prior to their deployment in clinical workflows.

Summing-up, we believe in the theoretical and practical power of QIBs-based solutions to improve medical imaging based diagnosis tasks. In particular, breast cancer is a case study that is at the forefront as one of the most promising applications where this kind of imaging techniques can have a high impact. Mainly, when it is need to examine some imaging modalities like breast MRI and breast CT.

References

1. Arevalo, J., González, F.A., Ramos-Pollán, R., Oliveira, J.L., Guevara Lopez, M.A.: Representation learning for mammography mass lesion classification with convolutional neural networks. Computer methods and programs in biomedicine **127**, 248–257 (2016)
2. Barendregt, A.M., Bray, T.J.P., Hall-Craggs, M.A., Maas, M.: Emerging quantitative MR imaging biomarkers in inflammatory arthritides. European journal of radiology **121**, 108707 (12 2019). https://doi.org/10.1016/j.ejrad.2019.108707
3. Beaumont, H., Maas, M., Wormanns, D., Zaim, S., Klifa, C., Faye, N., Iannessi, A.: Enhanced vertebra to disk ratio as a new semi-quantitative imaging biomarker for Gaucher disease patients. European Journal of Radiology **129**, 109091 (8 2020). https://doi.org/10.1016/j.ejrad.2020.109091
4. Dalmış, M.U., Litjens, G., Holland, K., Setio, A., Mann, R., Karssemeijer, N., Gubern-Mérida, A.: Using deep learning to segment breast and fibroglandular tissue in mri volumes. Medical Physics **44**(2), 533–546 (2017). https://doi.org/10.1002/mp.12079, http://dx.doi.org/10.1002/mp.12079
5. Deng, J., Wang, Y.: Quantitative magnetic resonance imaging biomarkers in oncological clinical trials: Current techniques and standardization challenges. Chronic Diseases and Translational Medicine **3**(1), 8–20 (3 2017). https://doi.org/10.1016/j.cdtm.2017.02.002
6. Dhawan, A.P.: Medical Image Analysis. IEEE Press series in biomedical engineering., John Wiley & Sons - IEEE Press, 2ª edn. (2011)

7. Dhungel, N., Carneiro, G., Bradley, A.P.: The Automated Learning of Deep Features for Breast Mass Classification from Mammograms, pp. 106–114. Springer International Publishing, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8$_1$3, $https$: $//doi.org$/10.1007/978 − 3 − 319 − 46723 − 8$_1$3
8. Fonseca, P., Mendoza, J., Wainer, J., Ferrer, J., Pinto, J., Guerrero, J., Castaneda, B.: Automatic breast density classification using a convolutional neural network architecture search procedure. In: Proc. of SPIE Vol. vol. 9414, pp. 941428–1 (2015)
9. Fotin, S.V., Yin, Y., Haldankar, H., Hoffmeister, J.W., Periaswamy, S.: Detection of soft tissue densities from digital breast tomosynthesis: comparison of conventional and deep learning approaches. In: Proc. SPIE. vol. 9785 (2016)
10. Huynh, B.Q., Li, H., Giger, M.L.: Digital mammographic tumor classification using transfer learning from deep convolutional neural networks. Journal of Medical Imaging **3**(3), 034501–034501 (2016). https://doi.org/10.1117/1.JMI.3.3.034501, http://dx.doi.org/10.1117/1.JMI.3.3.034501
11. Karssemeijer, N., Leon, J.T.O., Eijkman, E.G.: Recognition of organs in CT-image sequences: a model guided approach. Computers and biomedical research **21**(5), 434–448 (1988)
12. Kessler, L.G., Barnhart, H.X., Buckler, A.J., Choudhury, K.R., Kondratovich, M.V., Toledano, A., Guimaraes, A.R., Filice, R., Zhang, Z., Sullivan, D.C.: The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. Statistical Methods in Medical Research **24**(1), 9–26 (2 2015). https://doi.org/10.1177/0962280214537333
13. Kooi, T., Litjens, G., van Ginneken, B., Gubern-Mérida, A., Sánchez, C.I., Mann, R., den Heeten, A., Karssemeijer, N.: Large scale deep learning for computer aided detection of mammographic lesions. Medical Image Analysis **35**, 303 – 312 (2017). https://doi.org/https://doi.org/10.1016/j.media.2016.07.007, http://www.sciencedirect.com/science/article/pii/S1361841516301244
14. López, Y., Novoa, A., Guevara, M.A., Silva, A.: Breast cancer diagnosis based on a suitable combination of deformable models and artificial neural networks techniques. In: Iberoamerican Congress on Pattern Recognition. pp. 803–811. Springer (2007)
15. Luo, H.B., Du, M.Y., Liu, Y.Y., Wang, M., Qing, H.M., Wen, Z.p., Xu, G.H., Zhou, P., Ren, J.: Differentiation between Luminal A and B Molecular Subtypes of Breast Cancer Using Pharmacokinetic Quantitative Parameters with Histogram and Texture Features on Preoperative Dynamic Contrast-Enhanced Magnetic Resonance Imaging. Academic Radiology **27**(3), e35–e44 (3 2020). https://doi.org/10.1016/j.acra.2019.05.002
16. Piraud, M., Wennmann, M., Kintzelé, L., Hillengass, J., Keller, U., Langs, G., Weber, M.A., Menze, B.H.: Towards quantitative imaging biomarkers of tumor dissemination: A multi-scale parametric modeling of multiple myeloma. Medical Image Analysis **57**, 214–225 (10 2019). https://doi.org/10.1016/j.media.2019.07.001
17. Sahiner, B., Chan, H.P., Petrick, N., Wei, D., Helvie, M.A., Adler, D.D., Goodsitt, M.M.: Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. IEEE Transactions on Medical Imaging **15**(5), 598–610 (Oct 1996). https://doi.org/10.1109/42.538937
18. Sullivan, D.C.: Imaging as a quantitative science (8 2008). https://doi.org/10.1148/radiol.2482080242
19. Wolberg, W.H., Street, W.N., Mangasarian, O.L.: Machine learning techniques to diagnose breast cancer from image-processed nuclear features of fine needle aspirates. Cancer letters **77**(2-3), 163–171 (1994)
20. Yankeelov, T.E., Mankoff, D.A., Schwartz, L.H., Lieberman, F.S., Buatti, J.M., Mountz, J.M., Erickson, B.J., Fennessy, F.M.M., Huang, W., Kalpathy-Cramer, J., Wahl, R.L., Linden, H.M., Kinahan, P.E., Zhao, B., Hylton, N.M., Gillies, R.J., Clarke, L., Nordstrom, R., Rubin, D.L.: Quantitative Imaging in Cancer Clinical Trials. Clinical Cancer Research **22**(2), 284–290 (1 2016). https://doi.org/10.1158/1078-0432.CCR-14-3336

Chapter 13
Trends and regulation in drone delivery logistics

Julian Estevez

Abstract Unmanned aerial vehicles (UAV) or drone last mile deliveries are receiving a lot of attention from startups and big companies like Google, Amazon, and UPS. Around the world there are already examples of successful small scale delivery operations utilizing drones. Despite all the hype and excitement around this topic, there is surprisingly little research done on understanding the capabilities and technical tradeoffs among drones that are currently available in the market

13.1 Introduction

From filming movies or researching a pod of whales to delivering medication or delivering an explosive payload, Unmanned Aerial Vehicles (UAVs) are increasingly being utilized for a wide range of tasks. Since 2002, when the Predator drone was first used by the U.S. military in Afghanistan [8], drones have become smaller and cheaper, making it feasible for people to imagine alternate uses for UAVs, like delivering freight. Since 2011, big names such as UPS, Amazon, and Google [3] have thrown their hat into the UAV delivery ring, while other lesser known companies, such as Matternet and Zipline, have actually started delivery service in Rwanda, Australia, Switzerland, and Bhutan. UAVs have become a popular topic of conversation and an exciting source of speculation around how they might change the status quo of many businesses. UAVs that deliver cargo are already in operation in several different countries. Mostly, these UAVs were specifically tailored to meet the specific demands of the job or service. For example, in Rwanda, there is a great need for life-saving blood medicines in the rural parts of the country, but the road infrastructure is very poor. A company called Zipline has started using fixed-wing autonomous drones to deliver these medicines via parachute faster than any other kind of transportation available [1].

Julian Estevez
University of the Basque Country, UPV-EHU

In Switzerland, the Swiss Post and Matternet are in the final stages of testing the use of drones to transport medications and lab samples between two hospitals in Lugano, a town in southeastern Switzerland near the Italian border [7]. The two hospitals are only about 1.5 km apart from each other in the middle of the town, but the roads take delivery vans along a circuitous 12 minute route. Drones would make hospital deliveries much faster and much cheaper. Even in the United States, the FAA (Federal Aviation Administration) has allowed a company in Virginia to test the effectiveness of drones delivering medical supplies to low-income rural areas in the Appalachian foothills from mobile medical vans. Some recent reports forecasts that consumer preferences are likely to push new business models and delivery systems where drones are likely to have a major role in future last-mile logistics systems [2].

The academic literature has already documented the advantages that UAVs can provide to deliver medicines in remote locations [9]. Other researchers have analyzed UAVs potential applications and challenges for smart city applications such as data collection and some authors have focused on the regulatory barriers that can preclude large UAV deployments. The logistics services company DHL has identified higher last-mile efficiency, reduction of accidents, and faster deliveries as key potential UAV benefits; key potential challenges associated to UAVs are security, privacy, congestion, and regulatory concerns [6]. UAVs may also present clear benefits in terms of sustainability and CO_2 emissions for rural deliveries but the sustainability of drones for urban deliveries are affected by several factors such as customer density, ground vehicle efficiency, and delivery route characteristics [4]. For small payloads and speedy deliveries, UAVs may be competitive with ground vehicles especially in remote areas. Some of the existing literature compares the capabilities of an UAV against the current method of performing a certain task, while others compare UAV's particular components, such as propulsion systems and orientation sensors. For example, a recent 2017 paper [5] surveys UAV technology and their subsystems (frame, propellers, motors and batteries, payloads, and data processing) with a brief discussion of potential applications in spraying of liquids, and logistics.

Although there is a lot of excitement about how drones could change how freight transportation networks are designed and operated, there is very little research on the actual capabilities of the products that are currently on the shelf. The main goal of this research is to analyze, based on a survey state of the art UAVs, main capabilities and limitations of UAVs in the freight industry. As discussed throughout the paper, the potential is high but currently the limitations are serious in terms of capabilities and unsolved problems for urban deliveries such as the tradeoffs among UAV logistical capabilities, sustainability (in terms of energy consumption and CO_2 emissions), and safety concerns. The paper ends with conclusions and a discussion of future research avenues.

13.2 Regulation

In 2016, the Federal Aviation Administration (FAA) issued restrictions on the non-recreational use of unmanned aerial vehicles which effectively prohibited freight delivery using drones in the USA. Some restrictions do not affect the drones surveyed (400' maximum altitude, 45 m/s (100 mph) maximum land speed). However, other restrictions prevent any business from currently utilizing drones in a freight delivery service. For example, drones must be flown using VLOS (visual line of sight) at all times which would greatly reduce the size of the service area, especially in forested hilly terrains or dense areas with skyscrapers, and reduce the economic benefit of not having a human pilot in the UAV. Furthermore, the communication range of most of the surveyed drones is shorter than the theoretical flying range. Hence, a VLOS mandate further restricts the range for most drones currently available in the market. According to 2016 FAA rules, drones must not be flown over populated areas, less than 400' from any structure, when visibility is a less than three miles, and reduced daytime visibility. These restrictions allow freight to be delivered in rural environments over short distances and on very clear days. The 25 kg (55 lb) weight limit, which includes payload, does affects only one of the drones we surveyed, the Vader HL, most are well under that limit. In summary, most of the available multicopter drones' basic capabilities, e.g. speed, altitude, and payload; do not violate FAA's restrictions. However, restrictions governing where/what the drone can fly over, how it can be piloted (beyond line of sight or autonomously), and how far it can fly from its origin severely limit UAVs business and geographical scope.

The FAA is partnering with NASA to study when drones can be used in U.S. National Airspace and in what capacities. NASA is working on an air traffic management system for drones similar to what exists for today's air traffic, except that the UAV air space resides mainly within altitudes from 200' to 500'. This is critical to ensure that the digital aviation infrastructure that would be designed to organize the many different paths of the UAVs prevents drones from crashing into one another or flying into a restricted zone. Other countries where regulation is more "relaxed" are likely to see more progress. In Australia, a commercial operator can procure a Remotely Piloted Aircraft Operator Certificate and operate within 15 meters of a non-company person but requesting their consent first. Also, the Australian certificate enables the operator to apply for permission for other flight procedures such as night flying and Beyond Visual Line of Sight flying. In September 2016, the Australian government reduced the kinds of commercial operations that would require a remote pilot's license or operator's certificate, making it easier to utilize drones in certain circumstances. In the UK, drones weighing 7 kg or less including maximum payload are allowed to fly in all airspaces including those around the busiest airports over "congested areas".

In the UK, drones are regulated into three groups according to their mass and the potential harm that can result from a midair malfunction. For example, the heaviest category, 150 kg or higher, is regulated at the same level as manned aircraft. Privacy, safety, and noise are issues that may problematic in urban areas though not a major (or so important) problem in rural areas. A privacy backlash may restrict the flying space in urban areas if some neighbours do not allow unannounced or unauthorized aircrafts over their properties. Safety is also a major concern, different kinds of problems could result from thousands of drones zipping through the sky. Not only is there potential for collisions to power lines, buildings, and monuments resulting in property damage, but also the more important risk to people. Falling from the sky or crashing at maximum speed could be fatal. For a drone to be allowed to fly within 15 meters from a non-company person, the Australian government required that it be fitted with a kind of motor redundancy that would prevent it from falling from the sky.

In the UK, any flight that will go over a populated area must submit a "safety case" that includes the potential Kinetic Energy Limits that a free-fall from 400 feet high or that a collision at maximum speed would create. In addition to danger from a system malfunction, there is also the threat of people using it as a terrorism device. In response, some American manufacturers have created multicopter drones that can seek and capture enemy drones with netting.

Finally, noise is another aspect that can be subject to regulation. There is a concern in populated areas of added noise pollution coming from drones flying nearby. A few drone manufacturers state the noise level of the aircrafts as a function of distance or flying speed. In general, noise levels increase with the size of the drone and with the proximity to the aircraft. According to one study the movements from hexcopter and quadcopter style UAVs can emit decibel levels in the low to mid 70s, which is about as loud as a vacuum cleaner or "living room music.". There is ongoing research by NASA to lessen the noise pollution emitted by drones by improving propeller technology. One way being explored is by having each propeller rotate at varying speeds, which prevents the collective whines of the propellers from amplifying themselves. That particular kind of technology, though, is not available on the market yet.

13.3 Conclusions

This research presented novel data and linear relationships among current multicopters payload, take-off weight, and energy efficiency. A novel contribution of this research is the discussion of tradeoffs among UAV logistical capabilities, sustainability (in terms of energy consumption), safety, and last yard delivery constraints. The data shows that there are clear economies of scale in terms of energy consumed per unit distance travelled and per kilogram of payload delivered. Based on our comparative analysis range, payload, size, and cost are positively correlated and tend to increase together. Unfortunately, potential safety, noise, and last yard constraints also increase as drone capabilities and size increase. The survey seems to indicate that currently available UAVs can fill a delivery service niche in sparsely populated areas with a low number of customers and density. In rural areas, the regulatory landscape and last yard delivery constraints are also more relaxed. The majority of existing applications for delivery drones have been in rural areas, e.g. rural Africa, the Appalachian mountain villages or islands near a mainland.

In rural areas the economic benefit brought about by reducing the cost of a driver to visit remote customers are obvious but in this environment UAV range is a key consideration. UAVs for package delivery have a lot of potential to improve logistics productivity and reduce environmental externalities such as trucking diesel engine pollution. However, safety concerns and last yard constraints are likely to limit the benefits that can be achieved through economies of scale. Future research efforts can compare UAVs and ground vehicles in terms of safety and last mile efficiency. It is expected that multicopter UAV technology, capabilities, and costs will improve substantially in the near future. Hence, there are still many areas to research and model in terms of UAVs costs, markets, potential benefit, and supply chain impacts.

13.4 Acknowledgement

This work has been partially supported by MICIN funds through project PID2020-116346 GB-I00.

References

1. Paul Adepoju. Africa's struggle with inadequate covid-19 testing. *The Lancet Microbe*, 1(1):e12, 2020.
2. Parcel Delivery. The future of last mile. *McKinsey & Company*, 2016.
3. Malik Doole, Joost Ellerbroek, and Jacco Hoekstra. Estimation of traffic density from drone-based delivery in very low level urban airspace. *Journal of Air Transport Management*, 88:101862, 2020.
4. Miguel A Figliozzi. Lifecycle modeling and assessment of unmanned aerial vehicles (drones) co2e emissions. *Transportation Research Part D: Transport and Environment*, 57:251–261, 2017.
5. Higinio González-Jorge, Joaquin Martínez-Sánchez, Martín Bueno, and Pedor Arias. Unmanned aerial systems for civil applications: A review. *Drones*, 1(1):2, 2017.
6. Matthias Heutger and Markus Kückelhaus. Unmanned aerial vehicles in logistics a dhl perspective on implications and use cases for the logistics industry. *DHL Customer Solutions & Innovation, Troisdorf, Germany*, 2014.
7. Mireia Roca-Riu and Monica Menendez. Logistic deliveries with drones: State of the art of practice and research. In *19th Swiss Transport Research Conference (STRC 2019)*. STRC, 2019.
8. IWM Staff. A brief history of drones. *Imperial War Museum*, 2018.
9. Cornelius A Thiels, Johnathon M Aho, Scott P Zietlow, and Donald H Jenkins. Use of unmanned aerial vehicles for medical product transport. *Air medical journal*, 34(2):104–108, 2015.

Chapter 14
A Path Planning Approach for an Autonomous Wheeled Mobile Robot

Imane Nizar and Mohammed Mestari

Abstract The approach we present in this short paper, aims to solve the problem of path planning and optimization for a four-wheeled robot. Our goal is to enable the robot to avoid collisions in a dynamic environment, under the assumption that the environment is known, and the upcoming obstacles can be identified a priori by a control unit. The navigation starts from an initial position, where the first safe and optimal path is previously planned, and due to the changing environment, the path could become dangerous. At that moment, the control unit will compute other intermediate paths, allowing the mobile robot enough flexibility to avoid collisions in almost real-time, and navigate safely toward the final position. To resolve this nonlinear problem, we make use of the Decomposition-Coordination Method (DCM). The principal of this method consists in breaking down the nonlinear system into several interconnected subsystems, which allows the non-linearity to be treated at a local level, then we use the Lagrange multipliers to achieve coordination. Consequently, parallel processing minimizes massively the computation time. A demonstration is given to verify the convergence and the stability of the method, and we present as well the simulation results to confirm the potential of our approach.

14.1 Introduction

Autonomous robots have an important place in our lives, nowadays robots have become increasingly used in the most vital areas like medicine , military, industry, and more [1]. After the emergence of the industry 4.0 and its solutions that lie in the complete automation of the production processes, we can clearly see that man's life depends highly on robotics. Thus, if you are not using one yourself, you are most

Imane Nizar and Mohammed Mestari
Laboratory: Informatique, Intelligence Artificielle et Cyber Sécurité (IIACS), Ecole Normale Supérieure de l'Enseignement Technique (ENSET), University Hassan II, Casablanca. Mohammedia 20800, Morocco.

likely using products and devices that have been manufactured by means of robots [2]. In this paper, we focus on the four-wheeled mobile robot, which is the most popular robot used laterally for all the previously mentioned applications and more . A mobile robot can be defined as a mechanical artificial agent, usually a machine that is guided by a computer program or circuit board. An important problem in mobile robotics is the path planning that allows the robot to reach a particular position from its current position [3]. The path represents the set of points through which the center of the robot passes during its movement towards final position. Yet, there may be different paths to the target position, depending on the constraints imposed by the environment or by the robot, however, the objective is always to reach the destination in the most safe and optimal way [4].

Autonomous navigation in dynamic environments is still a technological challenge, the study described in this paper focuses on the development of a global control strategy to address this problem. Using a decomposition-coordination algorithm [5], we succeeded in computing the control that ensures optimal and safe navigation. First, we define our multi-objective optimization problem using the kinematic model of the four-wheeled mobile robot [6], taking into account the constraints and objective functions.

The multi-objective optimization problem is transformed to a single-objective optimization problem, and then broken down into smaller separable sub-problems that may be addressed simultaneously and in a reasonable time. The DC supports a decomposition technique that handles non-linearity at the local level and accomplishes coordination by utilizing Lagrange multipliers [7].

A second algorithm addresses the collision avoidance task, based on the information obtained from the robot's exteroceptive sensors. Its principle is simple: whenever an obstacle appears, the robot veers from the unsafe path by a small value $deltaq$, and proceeds to the re-computation of the new optimal path using the DC algorithm, starting from its current state. We achieve the optimal and obstacle-free path once the robot reaches the desired state [8].

This paper is structured as follows: Section II depicts the modeling of the mobile robot. Section III concentrates on problem formulation and transformation. Section VI, is dedicated to the analysis and resolution of the scalar optimization problem, followed by the discussion of appropriate convergence and stability criteria for the proposed technique. The obstacle avoidance approach is introduced in section V. Section VI shows how resolving such a complicated system allows us to compute the control input and compute the optimal path. The last section provides a brief conclusion.

14.2 Model of the robot

To model the four-wheeled mobile robot we used the following assumptions:

- The robot navigates on a plan surface,
- The rolling is slip-free,

- The velocity is null when the wheel is in direct contact with the surface at a geometrical point,
- The tires deformation is neglected.

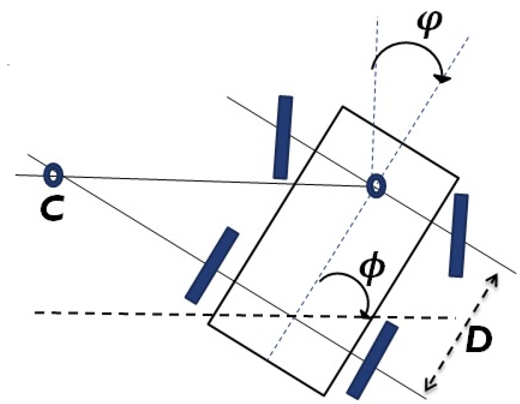


Fig. 14.1 Four wheel robot

The central point of the rear axis is defined by O of coordinates (x, y), the robot heading orientation ϕ, the steering angle φ, and D the distance between the front wheels and rear axes. Then we put the state vector $q = (x, y, \phi, \varphi)^T$, and the control vector $u = (v, \eta)^T$ (See Fig.14.1).

For a small time step δt, the four-wheeled robot would move in a direction enforced by the rear wheels direction. As the time interval approaches 0, it implies that: $dy/ds = \tan(\phi)$. Since the the rolling is slip free, the wheel can only move perpendicularly to the direction to its rotation axis.

The non-holonomic constraints may therefore be written in parametric form as follows:

$$\dot{q} = \begin{pmatrix} \cos(\phi) & 0 \\ \sin(\phi) & 0 \\ \frac{1}{D}\tan(\varphi) & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} v \\ \eta \end{pmatrix} \tag{14.1}$$

We put: $B(q) = \begin{pmatrix} \cos(\phi) & 0 \\ \sin(\phi) & 0 \\ \frac{1}{D}\tan(\varphi) & 0 \\ 0 & 1 \end{pmatrix}$

So that the kinematic model can be simply represented by:

$$\dot{q} = B(q).u \tag{14.2}$$

To implement our method of DC correctly, we need to express the robot's kinematic model in discrete time.

$$\begin{cases} \dot{x} = \dfrac{x_{k+1} - x_k}{\delta t} \\ \dot{y} = \dfrac{y_{k+1} - y_k}{\delta t} \\ \dot{\phi} = \dfrac{\phi_{k+1} - \phi_k}{\delta t} \\ \dot{\varphi} = \dfrac{\varphi_{k+1} - \varphi_k}{\delta t} \end{cases} \tag{14.3}$$

Which implies:

$$\frac{q_{k+1} - q_k}{\delta t} = \begin{pmatrix} \cos \phi_k & 0 \\ \sin \phi_k & 0 \\ \frac{\tan \varphi_k}{D} & 0 \\ 0 & 1 \end{pmatrix} \times \begin{pmatrix} \nu_k \\ \eta_k \end{pmatrix} \tag{14.4}$$

We obtain the following kinematic model in discrete time:

$$q_{k+1} = B(q_k)u_k + q_k = f(q_k, u_k) \tag{14.5}$$

With $B(q_k) = \delta t \begin{pmatrix} \cos \phi_k & 0 \\ \sin \phi_k & 0 \\ \frac{\tan \varphi_k}{D} & 0 \\ 0 & 1 \end{pmatrix}$

The control matrix at time t_k, $q_k = (x_k, y_k, \Phi_k, \varphi_k)^T$ and $u_k = (\nu_k, \eta_k)^T$, are respectively the state and the control vector at time t_k.

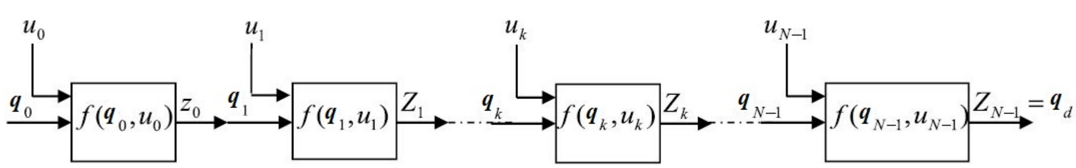


Fig. 14.2 Overall system made up of N interconnected subsystems.

14.3 Problem formulation

The discrete-time nonlinear system described below is assumed to be an appropriate formulation of the problem:

$$\begin{cases} q_{k+1} = f(q_k, u_k) \\ q_0 \quad given \end{cases} \tag{14.6}$$

With $q_k \in IR^n$ and $u_k \in IR^m$ are respectively the system's state and control at time t_k. We want to find the best control input that will get us to the target state in time t_k. Every problem involves a set of objective functions J_i that integrate the constraints to be satisfied. Since the objective functions are conflicting, the final design must achieve a balance between them, given that in general, there is no absolute optimal solution that meets all the objectives at once;

$$J_1(y^*) \geq J_1(y), J_2(y^*) \geq J_2(y), ... , J_p(y^*) \geq J_p(y).$$

The solution to a collection of competing objective functions is dependent on locating a set of acceptable solutions in the decision space. The Minimax technique [5], which offers the lowest value of the maximum values of all the objective functions J_i, is one important method. ω_k is defined as the weight of the objective functions J_k, where $\sum_{k=0}^{p} \omega_i = 1$.
With:

$$E(q,u) = max_(1 \leq i \leq p)\{w_i J_i(q,u)\} \tag{14.7}$$

The associated optimization problem is then defined as:

$$\begin{cases} min_{\{u_k^*/0 \leq k \leq N-1\}} E(q,u) \\ s.t \quad q_{k+1} = f(q_k, u_k) \\ \quad q_0 = q(0) \quad given, \end{cases} \tag{14.8}$$

Where $q_k = (x_k, y_k, \Phi_k, \varphi_k)^T$ and $u_k = (v_k, \eta_k)^T$ are respectively the state and the control at time t_k. The significant amount of processing required, which might grow exponentially, makes addressing such problem a tough task.

14.4 Problem analysis

The first step in addressing the problem (14.8) is to decompose the system into a set of N connected subsystems arranged in a basic serial structure (See Fig.14.2). We describe z_k as the output of the subsystem k.

$$z_k = f(q_k, u_k), \quad k = 0, \ldots, N-2 \tag{14.9}$$

$$q_k = z_{k-1}, \quad k = 0, \ldots, N-1 \tag{14.10}$$

So that, the problem (14.8) may be expressed as follows::

$$\begin{cases} min_{\{u_k^* | 0 \leq k \leq N-1\}} E(q,u) \\ s.t \quad z_k = f(q_k, u_k) \\ q_0 = q(0) \quad given, \end{cases} \tag{14.11}$$

It should be noted that even if all functions $J_i(q,u))$ are differentiable, the cost function $E(q,u)$ will have corners where differentiability fails, which means that the cost function has discontinuous partial first-order derivatives at points where two or more functions $J_i(q,u)$ are equal to $max_(1 \leq i \leq p)\{w_i J_i(q,u)\}$. To solve this derivation problem for the cost function,we construct the ordinary Lagrange function:

$$L_0 = \frac{1}{N} E(q,u) + \mu_0^T (f(q_0, u_0) - z_0) \tag{14.12}$$

$$L_k = \frac{1}{N} E(q,u) + \mu_k^T (f(q_k, u_k) - z_k) + \beta_k^t (q_k - z_{k-1}) \qquad for \quad 1 \leq k \leq N-2 \tag{14.13}$$

$$L_{N-1} = \frac{1}{N} E(q,u) + \mu_{N-1}^T (f(q_{N-1}, u_{N-1}) - q_d) + \beta_{N-1}^T (q_{N-1} - z_{N-2}) \tag{14.14}$$

The Lagrange multiplier $\mu_k \in IR^n$ and $\beta_k \in IR^n$ take into account the equality constraints (14.8) and (14.9). The equality-constrained minimization problem (14.11) can be transformed into a set of differential equations through the derivation of the ordinary Lagrange function (14.13). The equilibrium point (q_k^*,u_k^*,μ_k^*,β_k^*,t_k^*) meets the KKT conditions and satisfies:

$$\nabla_{q_k} L = \frac{1}{N} \frac{\partial E}{\partial q_k} + {\mu_k^*}^T + {\beta_k^*}^T = 0 \tag{14.15}$$

$$\nabla_{u_k} L = \frac{1}{N} \frac{\partial E}{\partial u_k} + \mu_k^* B = 0 \tag{14.16}$$

$$\nabla_{\mu_k} L = B u_k^* + q_k^* - z_k^* = 0 \tag{14.17}$$

$$\nabla_{z_k} L = -\mu_k^* - \beta_{k+1}^* = 0 \tag{14.18}$$

$$\nabla_{\beta_k} L = q_k^* - z_{k-1}^* = 0 \tag{14.19}$$

The equality constrained minimization problem (14.11) is addressed by solving the related system of differential equations (14.15) - (14.19).

14.4.1 Decomposition-Coordination (DC) method

Based on the serial architecture of the coordinating subsystems, and using our DC approach [5], the problem is split into two layers of difference equations that coordinate to find the best solution. The procedure begins by dividing the treatment of the linked system of differential equations (14.15) - (14.19) into two levels (see Fig.14.3). An upper level that addresses the equations (14.18) and (14.19) and fixes z_k and β_k, then transfers these values to the lower level, to solve equations (14.15)-(14.17). The global problem (14.11) is solved locally after the equations are satisfied (14.15)-(14.17).

Using forward Euler rule, the system of differential equations (14.15) - (14.17) is turned into a system of difference equations to process the system in discrete-time.

$$q_k^{(j+1)} = q_k^{(j)} + \lambda_q \left(\frac{1}{N} \frac{\partial E}{\partial q_k} + \mu_k^{(j)T} + \beta_k^{(j)T} \right) \tag{14.20}$$

$$u_k^{(j+1)} = u_k^{(j)} + \lambda_u \left(\frac{1}{N} \frac{\partial E}{\partial u_k} + \mu_k^{(j)} B \right) \tag{14.21}$$

$$\mu_k^{(j+1)} = \mu_k^{(j)} + \lambda_\mu \left(B u_k^{(j)} + q_k^{(j)} - z_k^{(j)} \right) \tag{14.22}$$

In the resolution process, the coordination between the upper and lower level is critical. It allows the transfer of the required information to the lower level, in order to achieve the overall optimization (i.e., satisfaction of (14.15)-(14.19)). To achieve such coordination, the first level works on the variables $\beta_k^{(j)}(k = 1, \ldots, N-1)$ and $z_k^{(j)}(k = 0, \ldots, N-2)$ at the same time. $z_k^{(j)}$ and $\beta_k^{(j)}$ are the coordination variables, assumed to be known within the second level, and used to allow for a local resolution of the system of difference equations (14.20) - (14.22) and to find values of the variables $q_k^*(z_k^{(j)}, \beta_k^{(j)})$, $u_k^*(z_k^{(j)}, \beta_k^{(j)})$ and $\mu_k^*(z_k^{(j)}, \beta_k^{(j)})$ that satisfy equations (14.20), (14.21) and (14.22) respectively. The outputs $q_k^*(z_k^{(j)}, \beta_k^{(j)})$ and $\mu_k^*(z_k^{(j)}, \beta_k^{(j)})$ of the lower level are sent back for the upper level, to validate or correct the previously proposed values if needed. $z_k^{(j)}$ and $\beta_k^{(j)}$ are given by the following equations:

$$z_k^{(j+1)} = z_k^{(j)} + \lambda_z \left(\mu_k^*(z_k^{(j)}, \beta_k^{(j)}) + \beta_{k+1}^{(j)} \right) \tag{14.23}$$

$$\beta_k^{(j+1)} = \beta_k^{(j)} + \lambda_\beta \left(q_k^*(z_k^{(j)}, \beta_k^{(j)}) + z_{k+1}^{(j)} \right) \tag{14.24}$$

The system of difference equations (14.20)-(14.22) is therefore solved repeatedly until acceptable coordination is found, i.e. fulfillment of coordination equations (14.23) and (14.24). The defined approach is presented in Fig.14.3.

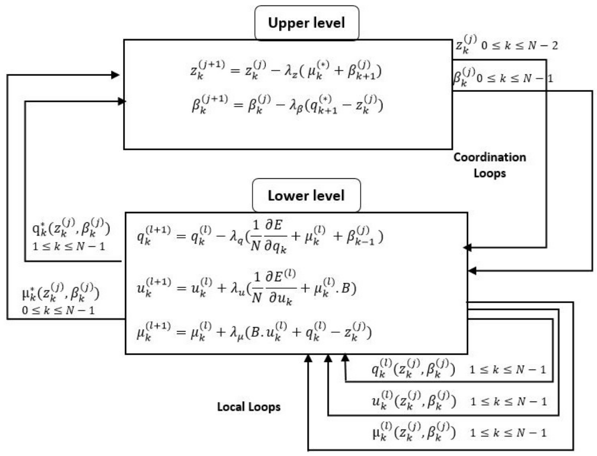


Fig. 14.3 Information transfer between the upper and the lower level

14.4.2 Stability analysis

This section demonstrates that the convergence of the algorithm described in Fig.14.3 may be shortened to that of the coordinating level. To make things easier, we will use the notations shown below:

$v_k = \begin{pmatrix} q_k \\ u_k \end{pmatrix}$, with $v_k^*(z_k^*, \beta_k^*)$,$\mu_k^*(z_k^*, \beta_k^*)$,z_k^* and β_k^* the ideal solution.

The values $v_k^*(z_k^{(j)}, \beta_k^{(j)})$ and $\mu_k^*(z_k^{(j)}, \beta_k^{(j)})$ have to be generated by the lower level (14.20) - (14.22), in order to solve the equations (14.15)-(14.17) locally, and adjust the coordination parameters z_k^* and β_k^* at the upper level through the coordination loop.

We put as well:

$$\begin{cases} G_k = \begin{pmatrix} \nabla_{q_k} L \\ \nabla_{u_k} L \end{pmatrix} \\ P_k = \nabla_{\mu_k} L \\ R_k = \nabla_{\beta_k} L \\ H_k = \nabla_{z_k} L \end{cases} \tag{14.25}$$

Using 14.25, the system of differential equations (14.20)-(14.24) can be represented as follows:

$$\begin{cases} G_k^{(j)} = \begin{pmatrix} \nabla_{q_k} L \\ \nabla_{u_k} L \end{pmatrix} = \begin{pmatrix} \frac{1}{N}\frac{\partial E}{\partial q_k} + \mu_k^{(j)^T} + \beta_k^{(j)^T} \\ \frac{1}{N}\frac{\partial E}{\partial u_k} + \mu_k^{(j)} B \end{pmatrix} = 0 \\ P_k^{(j)} = \nabla_{\mu_k} L = Bu_k^{(j)} + q_k^{(j)} - z_k^{(j)} = 0 \end{cases} \tag{14.26}$$

$$\begin{cases} R_k^{(j)} = \nabla_{\beta_k} L = q_k^*(z_k^{(j)}, \beta_k^{(j)}) - z_{k-1}^{(j)} = 0 \\ \\ H_k^{(j)} = \nabla_{z_k} L = -\mu_k^*(z_k^{(j)}, \beta_k^{(j)}) - \beta_{k+1}^{(j)} = 0 \end{cases} \tag{14.27}$$

The optimal solution is:

$$\begin{cases} G_k^* = \begin{pmatrix} \frac{1}{N}\frac{\partial E}{\partial q_k} + \mu_k^{*T} + \beta_k^{*T} \\ \frac{1}{N}\frac{\partial E}{\partial u_k} + \mu_k^{(j)} B \end{pmatrix} = 0 \\ \\ P_k^* = Bu_k^* + q_k^* - z_k^* = 0 \\ R_k^* = q_k^* - z_{k-1}^* = 0 \\ \\ H_k^* = -\mu_k^* - \beta_{k+1}^* = 0 \end{cases} \tag{14.28}$$

We define the errors calculated at iteration j of the coordination loop as follows:

$$\begin{cases} e_{v_k}^{(j)} = v_k^*(z_k^{(j)}, \beta_k^{(j)}) - v_k^*(z_k^*, \beta_k^*) \\ e_{\mu_k}^{(j)} = \mu_k^*(z_k^{(j)}, \beta_k^{(j)}) - \mu_k^*(z_k^*, \beta_k^*) \\ e_{z_k}^{(j)} = z_k^{(j)} - z^* \\ e_{\beta_k}^{(j)} = \beta_k^{(j)} - \beta^* \end{cases} \tag{14.29}$$

Let us consider the Lyapunov function below:

$$\Theta(j) = \frac{1}{2}\sum_{k=0}^{N-1} e_{z_k}^{(j)^T} e_{z_k}^{(j)} + e_{\beta_k}^{(j)^T} e_{\beta_k}^{(j)} \tag{14.30}$$

With:

$$\Delta e_{z_k}^{(j)} = e_{z_k}^{(j+1)} - e_{z_k}^{(j)} = -\lambda H_k^{(j)} \tag{14.31}$$

$$\Delta e_{\beta_k}^{(j)} = e_{\beta_k}^{(j+1)} - e_{\beta_k}^{(j)} = -\lambda R_k^{(j)} \tag{14.32}$$

Where $\lambda_z = \lambda_\beta = \lambda$.

The change of the Lyapunov function is:

$$\Delta\Theta = \Theta(j+1) - \Theta(j) = A(j)\lambda^2 + B(j)\lambda \tag{14.33}$$

With $A(j) = \sum_{k=0}^{N-1} \Delta e_{z_k}^{(j)T} \Delta e_{z_k}^{(j)} + \Delta e_{\beta_k}^{(j)T} \Delta e_{\beta_k}^{(j)}$

and $B(j) = \sum_{k=0}^{N-1} e_{z_k}^{(j)T} \Delta e_{z_k}^{(j)} + e_{\beta_k}^{(j)T} \Delta e_{\beta_k}^{(j)}$

Using the two theorems below [5], will allow us to prove the convergence of the algorithm:

Theorem 1 Let $e_{v_k}^{(j)}$, $e_{\mu_k}^{(j)}$, $e_{t_k}^{(j)}$ and $e_{\beta_k}^{(j)}$ be the errors computed at the iteration j of the coordination loop. Then: $e_{v_k}^{(j)} \to 0$ and $e_{\mu_k}^{(j)} \to 0$ if $e_{t_k}^{(j)} \to 0$ and $e_{\beta_k}^{(j)} \to 0$.

Theorem 2 The convergence is guaranteed if one of the matrices $\frac{\partial G_k^*}{\partial v_k}$ ($k = 0, 1, \ldots, N-1$) is positive definite and the others are only positive semi-definite and if $A(j) \neq 0$, λ should be chosen as: $0 \leq \lambda \leq \left|\frac{B(j)}{A(j)}\right|$.

The convergence of equations 14.28 is thus dependent on the satisfaction of the criteria stated in the theorems above the detailed mathematical demonstration may be found in [5].

14.5 Obstacle avoidance

The obstacle avoidance module is used to find the best collision-free path. Its primary aim is to perform the optimum control sequence, which will allow the robot to drive autonomously and effectively while avoiding collisions. We suppose that perception of the environment is done through the usage of exteroceptive sensors that collect several types of environmental information and measurements.

The algorithm that creates the route is represented as follows:

- Compute the optimal path T_0 without any obstacle from the initial state $q_0^{(0)}$ to the final state q_f, with $m = 0$.
- As q_f is not reached yet :
  - if an obstacle appears:
    - Then $m = m + 1$
    - Correction of the path, we put:
      - $x_0^{(m)} = x_{obs}^{(m)} + \delta x$
      - $y_0^{(m)} = y_{obs}^{(m)} + \delta y$
      - $\theta_0^{(m)} = \theta_{obs}^{(m)} + \delta\theta$

· $\psi_0^{(m)} = \psi_{obs}^{(m)} + \delta\psi$

· and then compute the optimal trajectory T_m from $q_0^{(m)}$ to q_f.

– Otherwise continue to execute the previous path.

- Optimal safe path reached.

With:

- m: Number of obstacles.
- T_m: The optimal path computed from the m^{th} obstacle.
- $q_0^{(m)}(x_0^{(m)}, y_0^{(m)}, \phi_0^{(m)}, \varphi_0^{(m)})$: Coordinates of the initial starting from the m^{th} obstacle.
- $q_{obs}(x_{obs}^{(m)}, y_{obs}^{(m)}, \phi_{obs}^{(m)}, \varphi_{obs}^{(m)})$: Coordinates of the m^{th} obstacle.

14.6 Simulation

This section describes the simulation findings for the four-wheeled mobile robot, which meet the convergence conditions. The initial position is $q_0 = (0,0,0,0)$, and the robot's target is $q_f = (5,5,0,0)$.

The Fig. 14.4 shows the ideal path formed by the set of optimal states determined at each $t_k = k\delta t$, of the coordination loop iteration.

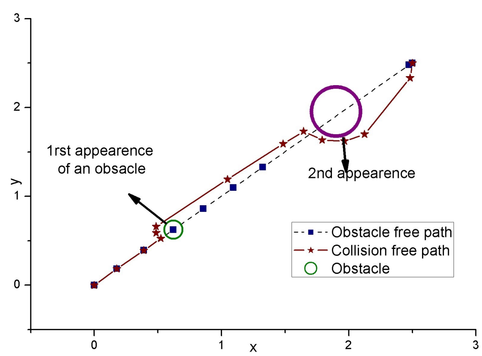


Fig. 14.4 Optimal obstacle free path from initial q_0 to final q_f.

Fig. 14.5 illustrates the adjustment of the control uk for each state qk. u_1 and u_2 are the variation of v_k and η_k respectively.

The curve in Fig. 14.6 confirms the fast convergence of the algorithm.

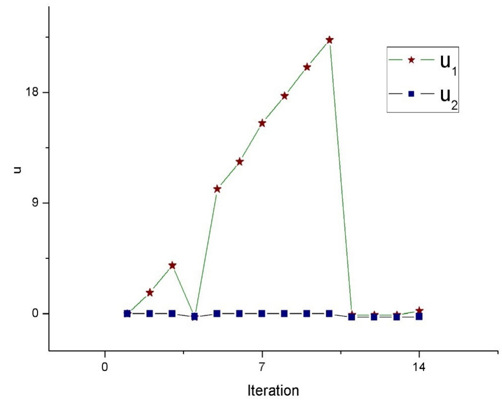


Fig. 14.5 The variation of the optimal control $u_1 = \nu$ and $u_2 = \eta$ for the obstacle free path.

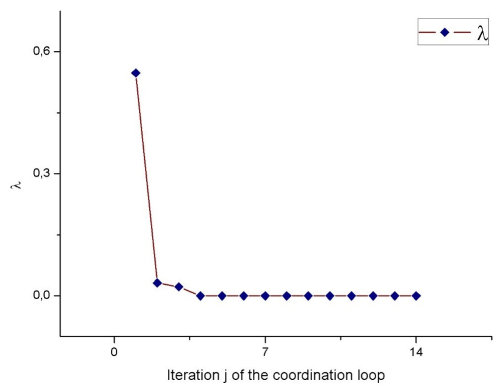


Fig. 14.6 Evolution of the adaptive coefficient λ adjusted according to the condition $0 \leq \lambda \leq \left|\frac{B(j)}{A(j)}\right|$.

14.7 Conclusion

In this paper, we proposed a new solution for a four-wheeled robot to address the problem of optimal navigation within a changing environment by integrating the decomposition coordination method with an obstacle avoidance algorithm. The potential of this technique comes from the problem's decomposition, which lies on the principle of the assignment of computing operations to several interconnected sub-systems, allowing for considerable reductions in processing time and, as a consequence, allowing the mobile robot to be more reactive in a tough environment. Our simulation findings validate the method's efficacy. The fast convergence of the adaptive coefficient demonstrates the stability and convergence of the DC method.

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 777720.

References

1. Y. Zhang and M. Lu, "A review of recent advancements in soft and flexible robots for medical applications," Int J Med Robot, vol. 16, no. 3, Jun. 2020, doi: 10.1002/rcs.2096.
2. S. Bragança, E. Costa, I. Castellucci, and P. M. Arezes, "A Brief Overview of the Use of Collaborative Robots in Industry 4.0: Human Role and Safety," in Occupational and Environmental Safety and Health, P. M. Arezes, J. S. Baptista, M. P. Barroso, P. Carneiro, P. Cordeiro, N. Costa, R. B. Melo, A. S. Miguel, and G. Perestrelo, Eds. Cham: Springer International Publishing, 2019, pp. 641–650. doi: 10.1007/978-3-030-14730-3_68.
3. M. Dirik, A. F. Kocamaz, and O. Castillo, "Global Path Planning and Path-Following for Wheeled Mobile Robot Using a Novel Control Structure Based on a Vision Sensor," Int. J. Fuzzy Syst., vol. 22, no. 6, pp. 1880–1891, Sep. 2020, doi: 10.1007/s40815-020-00888-9.
4. F. H. Ajeil, I. K. Ibraheem, A. T. Azar, and A. J. Humaidi, "Autonomous navigation and obstacle avoidance of an omnidirectional mobile robot using swarm optimization and sensors deployment," International Journal of Advanced Robotic Systems, vol. 17, no. 3, p. 1729881420929498, May 2020, doi: 10.1177/1729881420929498.
5. M. Mestari, M. Benzirar, N. Saber, and M. Khouil, "Solving Nonlinear Equality Constrained Multiobjective Optimization Problems Using Neural Networks," IEEE Transactions on Neural Networks and Learning Systems, vol. 26, no. 10, pp. 2500–2520, Oct. 2015, doi: 10.1109/TNNLS.2015.2388511.
6. Y. Chang, S. Ma, H. Wang, and D. Tan, "A kinematic modeling method for a wheeled mobile robot," in 2009 International Conference on Mechatronics and Automation, 2009, pp. 1100–1105.
7. I. Nizar, Y. Illoussamen, H. El Ouarrak, E. Hossein Illoussamen, M. Grana (Graña), and M. Mestari, "Safe and optimal navigation for autonomous multi-rotor aerial vehicle in a dynamic known environment by a decomposition-coordination method," Cognitive Systems Research, vol. 63, pp. 42–54, Oct. 2020, doi: 10.1016/j.cogsys.2020.05.003.
8. I. Nizar, A. Jaafar, Z. Hidila, M. Barki, E. H. Illoussamen, and M. Mestari, "Effective and Safe Trajectory Planning for an Autonomous UAV Using a Decomposition-Coordination Method," J Intell Robot Syst, vol. 103, no. 3, p. 50, Nov. 2021, doi: 10.1007/s10846-021-01467-2.

Chapter 15
SIMARGL - Secure Intelligent Methods for Advanced RecoGnition of malware and stegomalware

Aleksandra Pawlicka[1,2] and Marek Pawlicki[1,3] and Rafał Kozik[1,3] and Michał Choraś[3,4]

Abstract Ensuring cybersecurity equals protecting the availability, confidentiality and integrity of data. The paper discusses the EU-funded SIMARGL project - one of the initiatives aimed at enhancing the cybersecurity level of Europe. The project has particularly focused on stegomalware, i.e., the novel use of the techniques of hiding information. The SIMARGL project has been successful in achieving both its strictly technical and other, "softer" objectives. The paper briefly presents how SIMARGL contributed to a safer Europe, by providing effective attack detection and mitigation solutions, participating in the scientific research and dissemination, and addressing the cyber skills gaps and raising the awareness of the employees of the law enforcement agencies.

15.1 Introduction

In the recent years, the humankind has completely changed the way they work, shop and socialise, owing to the development of cyberspace. However, the rapid development of the Internet did not bring advantages only. Along with people's increasing dependency on the digital and online services, there emerged a plethora of potential threats to data and assets, no matter if they belong to individual citizens, companies, organisations or critical infrastructures alike. The cyberspace became the new, fertile

A. Pawlicka,Marek Pawlick and Rafał Kozik
ITTI Sp. z o.o., Poznań, Poland

A. Pawlicka
University of Warsaw, Warsaw, Poland

Marek Pawlick, Rafał Kozik and Michał Choraś
Bydgoszcz University of Science and Technology, Bydgoszcz, Poland

Michał Choraś
FernUniversität in Hagen, Hagen, Germany

ground where cybercrime can bloom. The catalogue of cyberthreats is very vast and besides computer viruses, it includes financial frauds, malware, ransomware, and many more. Ensuring cybersecurity equals protecting the availability, confidentiality and integrity of data. The EU-funded SIMARGL project has been one of the initiatives aimed at enhancing the cybersecurity level of Europe.

SIMARGL, standing for Secure Intelligent Methods for Advanced RecoGnition of malware and stegomalware, begun in 2019 and has recently finished. The project has particularly focused on stegomalware, i.e., the novel use of the techniques of hiding information in messages, images or other files (steganography), employed maliciously to cloak the malicious traffic.

The SIMARGL's strategic objectives have been as follows:

1. To propose, implement and validate innovative machine and deep learning methods do detect malware (including stegomalware), ransomware and network anomalies.
2. To significantly improve current solutions (already at advanced TRL) owned by SIMARGL consortium partners such as Thales, Airbus, Pluribus-One and ITTI.
3. To ensure privacy by design and security by design within our solutions, as well as to analyse and meet legal and ethical requirements.
4. To provide training, especially for LEAs and other end-users focusing on threats using information hiding techniques, such as stegomalware.
5. To integrate, deploy, demonstrate and validate our results at realistic use-case sites.
6. To communicate and disseminate the results.
7. To provide relevant training of our innovative solutions, including the raise of awareness about risks of information-hiding-capable threats such as stegomalware.
8. To transfer the results to market and to generate wide impact to European societies and economy

In order to help mitigate cyberattacks, SIMARGL has provided an integrated toolkit, incorporating state-of-the art algorithms and advanced methods of network analysis, such as deep learning techniques and advanced signal processing and transformations, able to adapt and detect novel, emerging malware.

15.2 The SIMARGL Toolkit

One of the innovative features of the SIMARGL project has been that it was delivered int the form of a toolkit. The SIMARGL Toolkit was designed in order to bring together a vast range of high-end components designed and supplied by the partners, as well as leverage the approaches adopted by them. As a result of their collaboration, an advanced malware detection platform was developed. Owing to its design, the platform is able to cover various attack surfaces and detect a great range of attacks, including the ones which employ steganography - which are otherwise particularly difficult to uncover.

The SIMARGL toolkit offers their users effective protection against malware and stegomalware, as well as means of in-depth analysis. The toolkit's high-level overview has been shown in Fig. 15.1.

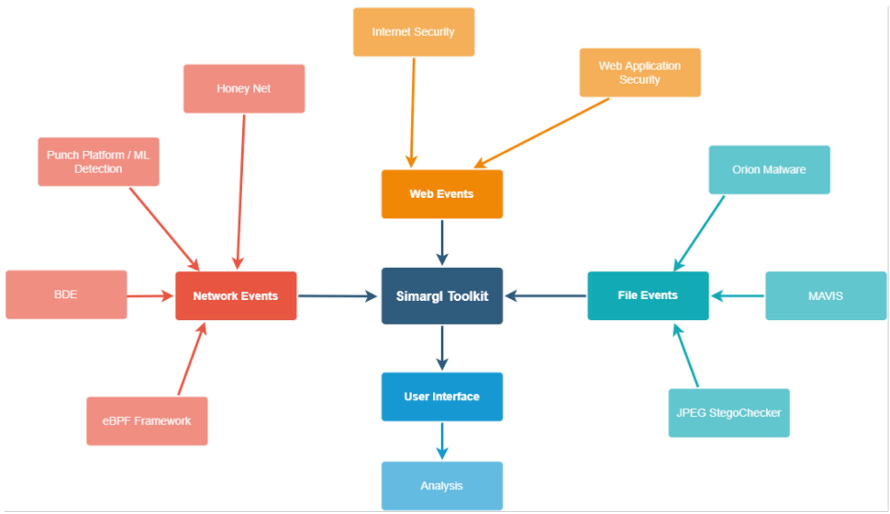


Fig. 15.1 A high-level overview of the SIMARGL toolkit

As shown in the Figure, the Toolkit has been formulated around three broad categories of events: network events, web events and file events.

The solutions protecting the various aspects of network encompass the BDE Platform (a machine-learning-based network intrusion detection component, detecting cyberattacks by using flow-based features), Honey Net (a system simulating a subnetwork of an existing information system, deploying Honey Pots for attackers), the extended Berkeley Packet Filter Framework (eBPF; collecting information on the hosts' behaviour and the network traffic), and the Punch Platform / ML Detection (a component employing various algorithms for threat identification, based on the CYBELS Sensor data).

In order to monitor and protect the web services, the Toolkit employs the Web Application Security component. Safe Internet use is provided by the Internet Security tool, by detecting such threats as malware, phishing or scams.

Lastly, the File Events aspect includes the tools such as the Orion Malware (a solution for file analysis, featuring five antiviruses), MAVIS, (for the detection of the Least Significant Byte Steganography), and JPEG StegoChecker (detecting image manipulation). The Toolkit's User Interface serves as the command and control center for the user.

Before the SIMARGL, there were no available, effective general countermeasures to stegolmaware, so the solutions and concepts that SIMARGL has proposed have already enhanced the European cybersecurity by advancing the cybersecurity tools, contributed to better protection of sensitive data, and as a result, helped reduce the cost of the potential damage done by cyberattacks.

However, the project has also made other significant, non-strictly-technical contributions. The further part of the paper will discuss some of them.

15.3 Real-life Use Cases

The effectiveness of the SIMARGL Toolkit was validated and demonstrated in seven Pilot Use Case Scenarios, among them Orange Poland, part one of the largest European mobile service providers, and RoEduNet, a Romanian educational and research network.

The following scenarios which covered the use cases identified in the project were implemented at the end users' premises:

1. Phishing Attack with cryptomalware
2. IPv6 Network Covert Channels
3. Reconnaissance and Denial-of-Service attacks
4. Internal Network Escalation
5. Production Ingress Traffic Intrusion Detection
6. Banking & Financial malware Banking & Financial malware
7. Least Significant Bit based Steganography in PNG files using Invoke-PSImage

Giving priority to testing the SIMARGL Toolkit within the scenarios covering the identified use cases, and its successful implementation made it possible to confirm the Toolkit's usefulness outside the lab, in real life.

15.4 Project branding

To make the project more recognisable, and disseminate its results in a more effective manner, the SIMARGL project, similarly to market products, has received its unique branding. Its specific identity was enhanced by defining the name, visual identity, color palette, typographies, etc. An important element of SIMARGL's visual identity has been the logo, depicting a winged wolf - the Slavic deity the project got its name from. The logo was used in the common presentation template for use in conferences and meetings with end-users/industry, as well as in a variety of branding products, such as T-shirts, hoodies, face masks and various gadgets.

The SIMARGL logo has been presented in Figure 15.2.

As this approach of treating a project as product has proved successful with SIMARGL, it is a valuable lesson for the future project managers.

Fig. 15.2 The SIMARGL logo

15.5 The research on cyberspace threat actors

It is worth noting that as part of the extensive, multi-faceted research within the SIMARGL activities, the nexus of cyberspace actors was carefully scrutinised and systematised. Thanks to the study, it was possible to identify the broad range of actors who employ cyberattacks as their tool of trade:

1. Nation states
2. Cyberterrorists
3. Hacktivists
4. Cybercriminals
5. Cyber militias
6. Trolls
7. Geeks, script kiddies.

The outcomes of the study and the resulting reflections on the cyberthreat actors have both contributed to understanding the threat landscape better, and inspired further research.

15.6 The scientific impact of the Project

Besides achieving tangible, concrete technical and scientific results, SIMARGL has also been engaged in disseminating the results outside the consortium, both to the scientific community and the wider audience. So far, the results of the various fields of research conducted as part of SIMARGL have been published in a number of high ranking publications. During the project's runtime, close to 50 peer-reviewed publications, including journal articles, contributions to conferences and book chapters were either published or accepted for publication. Even after the project has ended, its results will further be disseminated in future publications.

Some of the publications included: [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], and [11]. Not only have the several dozens publications helped make the results known, but they have also contributed to making them last, as the SIMARGL legacy.

SIMARGL has also attended and took active part in workshop organisation at esteemed, yearly conferences such as: ARES (International Conference on Availability, Reliability and Security) - the CUING workshops, European Interdisciplinary Cybersecurity Conference (EICC) - the DETONATOR workshops, and IEEE International Conference on Network Softwarization (NetSoft) - the SecSoft workshops.

15.7 Law Enforcement Agencies Training

Another of the SIMARGL's objectives was to develop a syllabus and pilot training for a series of modular training courses in cyberattacks management, advanced assurance and protection and provide training for the Law Enforcement Agencies (LEAs), to improve their capabilities. Although the preliminary plan included providing 2 training modules, eventually, eight full training modules were put at the users' disposal. The contents of the modules were tailored to the stakeholders' needs, after having analysed the cyber skills gaps of the LEAs.

As a result, SIMARGL provided the following training modules:

1. S1 Introduction to Cyber Threats
2. S2 Introduction to Cyber Attack Management
3. S3 Introduction to Malware Analysis
4. S4 Cyber Range (live event)
5. S5 Information Sharing Analysis and CTI
6. S6 Stegomalware
7. S7 Hidden Networks
8. S8 SIMARGL toolkit

The provision of training to LEAs has been very successful. The number of participants also greatly exceeded the preliminary anticipated numbers. On top of that, the training was well-received, with the feedback from participants being very favourable.

The training materials and the lessons learnt are though to become another part of the lasting legacy of SIMARGL.

15.8 Conclusions

The SIMARGL project has been successful in achieving both its strictly technical and other, "softer" objectives. By providing effective attack detection and mitigation solutions, participating in the scientific research and dissemination, and addressing the cyber skills gaps and raising the awareness of the LEAs employees, not only has it contributed to a more cybersecure Europe, but also paved the path for the future cybersecurity-related projects and initiatives.

Acknowledgments

This work is funded under the SIMARGL Project – Secure Intelligent Methods for Advanced RecoGnition of malware and stegomalware, with the support of the European Commission and the Horizon 2020 Program, under Grant Agreement No. 833042.

References

1. Caviglione, L., Choras, M., Corona, I., Janicki, A., Mazurczyk, W., Pawlicki, M., Wasielewska, K.: Tight Arms Race: Overview of Current Malware Threats and Trends in Their Detection. IEEE Access **9**, 5371–5396 (2021). https://doi.org/10.1109/ACCESS.2020.3048319, https://ieeexplore.ieee.org/document/9311151/
2. Choraś, M., Pawlicki, M.: Intrusion detection approach based on optimised artificial neural network. Neurocomputing (2020). https://doi.org/https://doi.org/10.1016/j.neucom.2020.07.138, https://www.sciencedirect.com/science/article/pii/S0925231220319032
3. Choraś, M., Pawlicki, M., Puchalski, D., Kozik, R.: Machine Learning – The Results Are Not the only Thing that Matters! What About Security, Explainability and Fairness? In: ICCS, vol. 4, pp. 615–628 (2020). https://doi.org/10.1007/978-3-030-50423-6_46, https://link.springer.com/10.1007/978-3-030-50423-6_46
4. Komisarek, M., Choraś, M., Kozik, R., Pawlicki, M.: Real-Time Stream Processing Tool for Detecting Suspicious Network Patterns Using Machine Learning. In: Proceedings of the 15th International Conference on Availability, Reliability and Security. ARES '20, Association for Computing Machinery, New York, NY, USA (2020). https://doi.org/10.1145/3407023.3409189, https://doi.org/10.1145/3407023.3409189
5. Komisarek, M., Pawlicki, M., Kowalski, M., Marzecki, A., Kozik, R., Choraś, M.: Network Intrusion Detection in the Wild - the Orange use case in the SIMARGL project. In: The 16th International Conference on Availability, Reliability and Security. pp. 1–7. ACM, New York, NY, USA (aug 2021). https://doi.org/10.1145/3465481.3470091, https://dl.acm.org/doi/10.1145/3465481.3470091

6. Komisarek, M., Pawlicki, M., Kozik, R., Choras, M.: Machine Learning Based Approach to Anomaly and Cyberattack Detection in Streamed Network Traffic Data. J. Wirel. Mob. Networks Ubiquitous Comput. Dependable Appl. **12**(1), 3–19 (2021). https://doi.org/10.22667/JOWUA.2021.03.31.003, https://doi.org/10.22667/JOWUA.2021.03.31.003
7. Komisarek, M., Pawlicki, M., Kozik, R., Hołubowicz, W., Choraś, M.: How to Effectively Collect and Process Network Data for Intrusion Detection? Entropy **23**(11), 1532 (nov 2021). https://doi.org/10.3390/e23111532, https://www.mdpi.com/1099-4300/23/11/1532
8. Mihailescu, M.E., Mihai, D., Carabas, M., Komisarek, M., Pawlicki, M., Hołubowicz, W., Kozik, R.: The Proposition and Evaluation of the RoEduNet-SIMARGL2021 Network Intrusion Detection Dataset. Sensors **21**(13), 4319 (jun 2021). https://doi.org/10.3390/s21134319, https://www.mdpi.com/1424-8220/21/13/4319
9. Pawlicka, A., Choraś, M., Pawlicki, M.: Cyberspace threats: Cyberspace threats: not only hackers and criminals. Raising the awareness of selected unusual cyberspace actors - cybersecurity researchers' perspective. In: Proceedings of the 15th International Conference on Availability, Reliability and Security. pp. 1–11. ACM, New York, NY, USA (aug 2020). https://doi.org/10.1145/3407023.3409181, https://dl.acm.org/doi/10.1145/3407023.3409181
10. Pawlicka, A., Choraś, M., Pawlicki, M.: The stray sheep of cyberspace a.k.a. the actors who claim they break the law for the greater good. Personal and Ubiquitous Computing **25**(5), 843–852 (oct 2021). https://doi.org/10.1007/s00779-021-01568-7, https://link.springer.com/10.1007/s00779-021-01568-7
11. Pawlicka, A., Jaroszewska-Choras, D., Choras, M., Pawlicki, M.: Guidelines for Stego/Malware Detection Tools: Achieving GDPR Compliance. IEEE Technology and Society Magazine **39**(4), 60–70 (dec 2020). https://doi.org/10.1109/MTS.2020.3031848, https://ieeexplore.ieee.org/document/9290450/

Chapter 16
SWAROG – fake news classification for the local context

Pawel Ksieniewicz, Rafał Kozik, Pawel Zyblewski, Andrzej Kasprzak, Michał Woźniak

Abstract Disinformation has recently become one of the most important propaganda tools to steer public opinion, and thus a contemporary source of political power in a way that pursues the interests of specific groups, which are not necessarily consistent with the well-being of the general public. Therefore, the fake news classification has become a critical issue, going beyond the standard framework of natural language processing methods. Proper solutions to this problem require a specific approach to content annotation, taking into account the phenomenon of concept drift and – above all – the local specificity of content in national languages. The following article presents the concept of the SWAROG project, which, based on the latest achievements of NLP, proposes processing methods that will allow – in the example of the Polish language – to build a reliable fake news classification model resistant to temporal degeneration.

16.1 Introduction

In the light of the brutal attack by the Russian Federation on independent Ukraine that has been going on for over two months now, it is difficult not to notice the critical role of the media, especially social media, in building a political mood of society. Both in Ukraine – where reports from the front become one of the main elements building the morale, in Russia and Belarus – where government propaganda tries to hide the reality of its actions from citizens. It affects also the Western world, where come both uplifting news about the heroic defending the homeland, as well as unconfirmed information about aggressive refugees, depleting supplies at gas

Pawel Ksieniewicz, Pawel Zyblewski, Andrzej Kasprzak and Michał Woźniak
Wroclaw University of Science and Technology, Wrocław, Poland e-mail: [pawel.ksieniewicz,michal.wozniak,pawel.zyblewski]@pwr.edu.pl

Rafał Kozik
Bydgoszcz University of Science And Technology, Bydgoszcz, Poland, e-mail: rkozik@pbs.edu.pl

stations, or the utterly unjustified necessity of taking Lugol's iodine, to mention examples from the first weeks of the war.

Even though we can intuitively associate this specific information war front with modern times, disinformation, even in the armed conflicts, has been with us basically from the beginning of written history. Around 1280 BC, one of the best-documented battles of the ancient world took place in what is now Syria at Kadesh. The Hittite armies, led by Muwatallis II, clashed with the Egyptian army led by Ramesses II. On the eve of the battle, the Egyptians arrested two Hittite Bedouins who were spies sent there deliberately to mislead the Egyptians about the location of hostile troops. This crafted information led to a partial dispersion of the Egyptian forces, thanks to which the battle itself ended with a draw. The information influenced the course of military operations.

Of course, the potential for disinformation grows with each new invention that improves the replication and distribution of messages. Since the printing press, through the telegraph, radio, television, and finally, the Internet, we are gaining faster and more accessible tools that allow us to reach the other end of the world with our message in fractions of a second, regardless of the author's attitude to the truth. In modern times, propaganda and disinformation are spreading on the Internet, where it does not have to wait for the printing press to end its work, nor is it limited by newspaper circulation. Of course, this phenomenon has been noticed by both public opinion and scientists, who have given it the keyword *fake news* [4]. Since 2017, we can see a significant increase in publications that try to tackle this topic [1]. After five years of intensified efforts by the scientific community, we can already conclude that it has been initially recognized.

In line with this preliminary diagnosis, a SWAROG (*Artificial Intelligence Disinformation Detection System*) project proposal has been developed, which is motivated by three main processes and challenges:

- The development and characteristics of social networks – including the biggest problem, i.e. the fact that the opinion of an expert is currently as important and significant as the opinion of an amateur (non-expert), and even of a person who intentionally misleads the recipients of the content.
- Changes in traditional media - both in connection with recent events and due to dynamically changing financial requirements, one can observe changes and processes in traditional media that reduce the role of the *editorial office*, making a gradual emphasis on remote work of journalists, forcing more and more frequent use of online sources than reporting.
- The impact of fake news on the security of the state and citizens – social networks, as well as traditional media, have become an element of state security because disinformation and fake news produced by internal and external actors may manipulate readers, creating an unnecessary debate in topics organically irrelevant to society, having a degenerating effect on the quality of discourse, causing fear of citizens, and threatening the security of the state.

16.2 SWAROG project overview

The project involves the development of artificial intelligence algorithms for the automatic classification and detection of fake news. As a result of the project, a service based on artificial intelligence algorithms will be created, determining the credibility of information published in public space – particularly on the Internet and social media. The solution will be prepared to take into account the current model of information distribution – in which the recipient more and more often functions in his closed *information bubble*, and synchronized activities of many small sources produce false information.

The solution will be based on the latest achievements in the fields of *Natural Language Processing*, *social network analysis* and *machine learning*, mainly in the field of data classification. Advanced supervised learning models, including classifier ensembles, will be used. The target solution will also consider the streaming nature of fake news and thus the non-stationarity of the model [2], i.e., the occurrence of the model drift phenomenon that requires the predictive model to adapt to the changing parameters of the stream of analyzed messages [3].

In order to counter the phenomenon of fake news, the SWAROG project proposes the following solutions:

- Open architecture for computer fake news prevention systems.
- Advanced machine learning methods for detecting fake news based on text analysis in English – the most frequently used language by journalists, even in a traditionally non-English-speaking world (including Al Jazeera or RussiaToday).
- Development of a representative database of annotated texts in Polish (with false and true information).
- Advanced machine learning methods for detecting fake news based on text analysis in Polish.

Commercial recipients of the SWAROG project may include traditional polish editorial offices and media, incl. PAP, TVP, RMF FM, TVN or Polsat, modern media like bloggers, podcast authors, social network operators, and independent journalists.

16.3 Project aims and goals

The target result of the project will be the development of a website that allows verifying the belonging of texts available at a given URL to the fake news category. The website will provide a standard user interface in the form of a web application and a programming interface (Web API) consistent with the REST architectural design pattern. Implementing the programming interface will extend the availability of the website functionality, enabling the development of native and mobile apps and plugins for popular web browsers using the resources and methods offered by the SWAROG service.

After feeding SWAROG with the URL of the article, the user will receive information about the belief that it belongs to the fake news category. In the case of an extended form of service subscription, information about the probability of belonging to the thematic fields and support for the classification of fake news in each of them will be provided.

The critical element of the service will be the recognition model using the supervised learning paradigm. As part of the planned architecture of the predictive system, it will be:

- HIERARCHICAL – using the ensemble learning paradigm, allowing both the processing of heterogeneous data and the prediction of many objective functions for the implementation of domain models,
- STREAMABLE – following the incremental learning paradigm, which will allow both effective work on large volumes of data and the introduction of adaptive mechanisms, increasing the potential survival rate of models.

It should be emphasized that the SWAROG service will not directly use the models developed as part of the design work, and they will only constitute the initial state of the production system. A significant disadvantage of the vast majority of publicly available predictive systems is a stationary approach to the problems analyzed by them. While they can guarantee correct solutions, for example, in classical medical systems, where the definition of a disease entity is not dependent on time, in the case of problems defined in the social or cultural context, which include the detection of disinformation, they lead to the construction of models whose the maximum efficiency is achieved for the moment of acquisition of the training set. Such a system – already at the moment of publication – is outdated. During its use, it further degenerates, in parallel and proportionally, to the changes in the properties of the dynamic recognition problem classes, called the *concept drift*.

A partial solution to this problem is the use of the incremental learning paradigm proposed in the SWAROG project, in which we do not look at the problem stationary, and in solving it, we can observe, analyze and take into account the dynamics of the concept changes in the prediction provided to the user. The predictive model in such an approach is often not a single classifier. Its decisions are agreed upon by (*a*) a diverse and (*b*) mutually independent set of historical classifiers using an adequately constructed combination rule. Such fuser is taking into account the measurable competencies of each member classifier at the level of general predictive ability, at the level of individual classes, and individual regions of the attribute space.

A necessary complement to this partial solution is to provide the system with access to real labels for new cases of the evolving disinformation problem. A recognition system that accumulates and constantly revises its knowledge as part of incremental learning must be able to constantly, uninterruptedly verify the current predictive ability. We must be aware that it is impossible to manually develop labels that constitute the so-called *human bias* necessary to update the supervised learning model for all documents resulting from frequent acquisitions to ensure the system's viability (lifelong learning). It is, therefore, necessary, separated in the dedicated project task, to develop an effective acquisition and annotation module, which, using the active learning strategy, will allow for the identification of key documents for updating the model, allowing for rationalization of the work of annotators, making maximum use of the human factor necessary to ensure the continuity of service provision [5]. It is not possible to treat this task as a routine work, because in order to ensure the maximum possible and – above all – lasting quality provided by the prediction service, it will be necessary to actively select data for annotation using the knowledge previously accumulated in the predictive model, which will be a critical research challenge.

In order to build an additional processing context and increase the predictive potential of the system, the description of each document (both for model training and for documents verified in the system by users) will be extended with the content spread analysis, included in a separate subtask of the project. The SWAROG system, in every element of its functionality, will be able to handle documents in Polish and English, having the ability to identify the document's language and process mixed bilingual content.

16.4 Conclusions

Funded by *The National Centre for Research and Development's* program INFOSTRATEG, SWAROG aims to develop and utilize advanced *natural language processing* and *machine learning* algorithms in order to perform automatic detection and classification of fake news in both English and Polish.

The prepared solutions will be made available to users in the form of a website and a programming interface, enabling easy analysis of the credibility of the texts at the given URL address. It will also allow the development of mobile applications and browser extensions. More information about the project is available at its webpage[1].

Acknowledgement

This publication is cofinanced by the National Center for Research and Development within INFOSTRATEG program, number of application for funding: INFOSTRATEG-I/0019/2021-00.

References

1. Michał Choraś, Konstantinos Demestichas, Agata Giełczyk, Álvaro Herrero, Paweł Ksieniewicz, Konstantina Remoundou, Daniel Urda, and Michał Woźniak. Advanced machine learning

[1] https://www.kssk.pwr.edu.pl/projects/swarog

techniques for fake news (online disinformation) detection: A systematic mapping study. *Appl. Soft Comput.*, 101(C), mar 2021.

2. Joanna Komorniczak, Pawel Zyblewski, and Pawel Ksieniewicz. Prior probability estimation in dynamically imbalanced data streams. In *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE, July 2021.
3. Pawel Ksieniewicz, Pawel Zyblewski, Michal Choras, Rafal Kozik, Agata Gielczyk, and Michal Wozniak. Fake news detection from data streams. In *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, July 2020.
4. Gordon Pennycook and David G Rand. The psychology of fake news. *Trends in cognitive sciences*, 25(5):388–402, 2021.
5. Paweł Zyblewski, Paweł Ksieniewicz, and Michał Woźniak. Combination of active and random labeling strategy in the non-stationary data stream classification. In *Artificial Intelligence and Soft Computing*, pages 576–585. Springer International Publishing, 2020.

Chapter 17
SocialTruth - content verification for the digital society

Michał Choraś, Andrzej Kasprzak, Rafał Kozik, Paweł Ksieniewicz, Michał Woźniak

Abstract The fight against the fake news phenomenon has become one of the most critical challenges facing democratic governments, societies, and scholars in various fields. As seen in the conflict caused by Russia's attack on Ukraine, disinformation is one of the primary weapons. Therefore, the issue of fake news has become one of the most critical problems affecting society. Fake news spreads rapidly through social media. That is why they should be identified as soon as possible to avoid their negative impact on decisions made by citizens (e.g., during elections). The fight against disinformation is no longer local but requires a holistic approach to multi-source data analysis. This paper will present the selected results of the SocialTruth project, which, based on an innovative data processing architecture, offers a user-friendly environment to support fact-checking. The system architecture uses state-of-the-art technologies emphasizing performance, scalability, and openness, allowing for additional components and analyzers developed by the fact-checking community.

17.1 Introduction

Fake news has been with us for hundreds of years. Most historians indicate that one of the first instances of fake news used in political combat occurred over 2,000 years ago when the Roman Republic was threatened with civil war by Octavian and Mark Antony. To gain public support, Octavian began spreading fake news against Mark Antony, claiming that Mark Antony disrespected traditional Roman values,

MichaÅ ChoraÅ and RafaÅ Kozik
Bydgoszcz University of Science And Technology, Bydgoszcz, Poland, e-mail: [chorasm,rkozik]@pbs.edu.pl

Andrzej Kasprzak, PaweÅ Ksieniewicz and MichaÅ WoÅºniak
Wroclaw University of Science and Technology, WrocÅaw, Poland
e-mail: [pawel.ksieniewicz,michal.wozniak]@pwr.edu.pl

was addicted to Egypt's Queen Cleopatra, and abused alcohol. Octavian used coins on which he placed the information mentioned above as a way to reach the public [1].

With the advent of digital media, including social networks, fake news became a real scourge. However, the real boom of fake news is connected with the emergence of mass media and its popularity began to grow. On the one hand, they are often used to increase the interest of Internet users to encourage the increase of traffic to a given website, and thus revenues related to the sale of their products or advertisements. On the other hand, fake news has been recognized as an essential tool of political struggle, which was brutally and blatantly used by Donald Trump's team during the 2016 election campaign. Today, disinformation is used widely including during the current conflict triggered by Russia's brutal assault on Ukraine. Nowadays, in addition to fake articles, falsified images or videos are increasingly used, often modified using DeepFake models, which aim to replace one person's face with another person. Such a technique was also used during the current conflict in Ukraine [2].

Although most intuitively understand fake news, there is no universal agreement on its definition [1]. However, it can be assumed that fake news is false information disseminated in the media, in print, or on online social media sites. There are several methods to combat fake news. However, the most popular method is manual fact-checking. In many cases, this method is difficult, time-consuming and expensive. For example, in Lithuania, a network of elves has been set up to gather volunteers who manually detect fake news and report it to the authorities [3]. Furthermore, there is a risk that manual fact-checking may be biased, as it is subject to the subjective judgment of the annotator and thus may be unreliable. One can, of course, strive to have each message annotated by a group of annotators or choose some sort of dispute resolution rule of the 2+1 type of annotation. However, this significantly increases the cost and time of such an annotation. Therefore, the community is looking for various automatic detection solutions to streamline and speed up this process. Since manual annotation is not possible by a human, so machine learning methods have begun to be used to detect fake news.

Therefore, the goal of the SocialTruth [2] project is to create a system for classifying digital data and certifying its trustworthiness. It is aimed at both professionals and ordinary users of media and social networks. The technical solution proposed and developed by the project consortium is based on emerging technologies and trends such as NLP (Natural Language Processing), sentiment analysis, image content modification detection algorithms and ML-based text classification.

---

[1] https://storymaps.arcgis.com/stories/03206d2d092f4aa6bbc5a8827bf6d6a9

[2] https://www.youtube.com/watch?v=CPqGI3Hg83s

[3] Michael Peel, Fake news: How Lithuania'sLithuania's "elves" take on Russian trolls, Financial Times, Feb. 4, 2019, https://www.ft.com/content/b3701b12-2544-11e9-b329-c7e6ceb5ffdf

17.2 SocialTruth project overview

The SocialTruth system is designed to combat fake news using the latest technology. The project's originality is based on creating (in a fully decentralized environment) an open, democratic, pluralistic, and distributed ecosystem for easy access to various verification services, ensuring scalability, and building trust. High reliability is the system's distinctive feature based on blockchain technology for distributed reputation and trust and enhanced security.

Moreover, SocialTruth is a multicultural, multidisciplinary European project. It is developed by an international consortium of 11 partners and led by ICCS from Athens. The consortium covers 6 European countries, namely Greece, Poland, France, Italy, Romania and Great Britain, which pose the required variety of challenges and great importance for the EU. Onboard of the consortium, there are:

- Three research institutions: ICCS - Institute of Communication and Computer Systems (Athens - GR); UTP (Bydgoszcz - PL); LSBU-London South Bank University (London - United Kingdom).
- Three technical partners, developers, and software vendors: Thales (Paris - France); Expert system France (Paris - FR), and Qwant (Paris - FR); together with two experienced management consulting companies: Tecoms (Rome - IT) and Zanasi & Partners (Modena - IT).
- Two main news and media groups: Adnkronos (Rome - IT) and DeAgostini Scuola (Rome - IT) and InfocoCons (Bucharest - RO), a very active consumer organization.

17.3 Platform architecture

While designing the SocialTruth solution, we have adapted the microservice architecture style that promotes the single responsibility paradigm and recommends a loosely coupled module system.

In general, there are different approaches to the problem of monolithic systems decomposition. This process aims to divide the original application into many smaller autonomous components (minimizing dependencies, decoupling possibilities, business relevance and maintenance, etc.).

When the application is broken down into a set of separate services, they need to communicate to provide complex business capabilities [4]. That capability usually needs to assemble the results obtained from multiple services. Depending on the specific application usage, this could impose several difficulties. There are two forms of organizing how the different microservices will communicate and work with one another to produce the final requested result.

If an application is split into separate services, they must communicate to provide complex business functions. This feature typically requires collating the results from multiple services. Depending on the specific design of the application, this may pose several difficulties.

To overcome these challenges, we adapted a modern technology stack (Fig. 17.1). In general, it can be broken down into the following logical elements:

- physical elements (nodes) and their orchestration,
- messaging and event processing.
- verification services,

The physical layer is located on top of an infrastructure composed of virtual and hardware machines. This layer is intended to implement automated resource management, and thus it facilitates the entire platform with such capabilities as flexibility, scalability, and fault tolerance. It is the responsibility of the orchestration layer to effectively deploy the services on the available computational nodes (both physical and virtual). It is achieved thanks to sandbox containers containing the implemented service and all the software dependencies (libraries and execution environment). The services can be easily migrated between the computational nodes and deployed in this form.

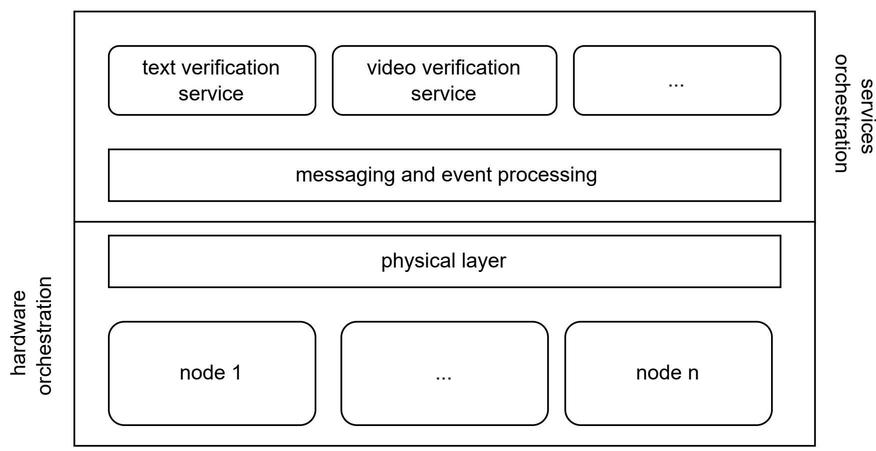


Fig. 17.1 The Proposed Platform - a general overview of the architecture and technology stack

For messaging and event processing we have adapted *Apache Kafka* technology. Our goal was to establish bilateral and persistent communication channels between different parts of the SocialTruth systems. Apache Kafka is a distributed streaming platform implementing the publish-subscribe model. Once the ingested data is published to one of *Kafka* topics, it can be simultaneously consumed by various verification services and/or stream processing applications. Once the services finalize their computations, they make the results available on another *Kafka* topic, which can be consumed by other services again. Moreover, on top of the *Apache Kafka* system, the *Complex Event Processing* engine was deployed. It is used to pre-process the data before storing and presenting it to the end-user or the verification service. For example, this technology was used to join the data streams produced by verification services of the same type. In that regard, the user can be presented with analysis results obtained from different classifiers.

As for the verification services are the key building blocks of the system and are deployed as dedicated micro-services. For instance, the verification services enable the SocialTruth system to detect fake news in text and images. We mainly focused on developing innovative solutions for unstructured text analysis and fake news detection (e.g., Transformer-based solutions [3]).

17.4 Conclusions

Funded by the EU research and innovation program Horizon 2020, SocialTruth aims to benefit individual users by testing content credibility on social media and stopping disinformation in their footsteps. In addition, the project aims to help media organizations, content authors, and journalists improve their research resources with a better option to match different media sources and create a more sustainable, quality, and safety-oriented network and social media ecosystem. The advantage of the architecture lies in its openness so that other services (methods, e.g., for image analysis) can be added and run on the deployed platform. More information are available at the project's webpage [4] and Twitter [5].

Acknowledgement

This work is funded under SocialTruth project, which has received funding from the European Unionâs Horizon 2020 research and innovation programme under grant agreement No. 825477.

References

1. Michał Choraś, Konstantinos Demestichas, Agata Giełczyk, Álvaro Herrero, Paweł Ksieniewicz, Konstantina Remoundou, Daniel Urda, and Michał Woźniak. Advanced machine learning techniques for fake news (online disinformation) detection: A systematic mapping study. *Appl. Soft Comput.*, 101(C), mar 2021.
2. Michał Choraś, Marek Pawlicki, Rafał Kozik, Konstantinos Demestichas, Pavlos Kosmides, and Manik Gupta. Socialtruth project approach to online disinformation (fake news) detection and mitigation. In *Proceedings of the 14th International Conference on Availability, Reliability and Security*, ARES â19, New York, NY, USA, 2019. Association for Computing Machinery.

[4] http://www.socialtruth.eu/

[5] https://twitter.com/SocialTruth

3. Sebastian Kula, Rafał Kozik, Michał Choraś, and Michał Woźniak. Transformer based models in fake news detection. In Maciej Paszynski, Dieter Kranzlmüller, Valeria V. Krzhizhanovskaya, Jack J. Dongarra, and Peter M.A. Sloot, editors, *Computational Science – ICCS 2021*, pages 28–38, Cham, 2021. Springer International Publishing.
4. James P Lawler and Hortense Howell-Barber. *Service-oriented architecture: SOA strategy, methodology, and technology*. Auerbach Publications, 2007.

Chapter 18
Machine Learning Methods for Predicting Adverse Drug Reactions in Hospitalized Patients

Patricio Wolff, Sebastian Rios,and Claudio Gonzales

Abstract Adverse Drug Reaction (ADR) in hospitalized patients have significant impacts on costs as well as on morbidity and mortality. A subset of ADR is preventable at the moment of prescribing. The use of supervised learning tools has been shown to be a good alternative to predict ADR, but more research is needed to make it a standard. A study was conducted to demonstrate how the use of machine learning methods can strengthen the ability to predict ADR in hospitalized patients. Binary classification models were constructed using three class balancing, feature selection and supervised learning approaches known as deep learning, random forest, and gradient boosting trees. Five-fold cross-validation and various evaluation metrics such as AUC, recall and F-measure were used to evaluate the models' performance. Gradient boosting trees models was the approach that performed best, exhibiting a recall level of 78.3% and an ROC curve AUC of 0.81. Findings on attribute importance were also generated, which should thus prove useful in developing tools for clinical applications. Our predictive models' results outperform those reported in the recent literature using traditional statistical methods. In this work we have shown the enormous potential of ML tools in this field and their great contribution to the development of information systems that use them.

18.1 Introduction

An adverse drug reaction (ADR) is defined by the World Health Organization (WHO) as any clinical manifestation or response to a drug that is noxious and unintended, and which occurs at doses normally used in patients (World Health Organization, 1975).

Patricio Wolff, Sebasti
Business Intelligence Research Center, Industrial Engineering Department, University of Chile, Chile

Claudio Gonzs
Clinical Pharmacology Unit (CPU), Barros Luco Hospital, Chile

ADRs in hospitalized patients have significant impacts on costs (due to increased hospital stays and treatment complexity) as well as on morbidity and mortality. Their occurrence puts additional pressure on health systems given that hospital staff must treat them as well as the affected patients' primary pathologies. Although not all ADRs originate in a medication error, a subset of them is preventable at the moment of prescribing [7,8]. ADRs are often confused with the symptoms of a patient's underlying clinical condition [39,40], thus complicating timely intervention; and due to the multifactorial nature of ADRs, multiple interventions are required if they are to be detected in time [9]. For these reasons, there is considerable interest in developing and improving models for predicting ADRs in hospitalized patients. The existing academic literature on building ADR or ADE (adverse drug event) risk prediction models focuses mainly on identifying risk factors for use in developing a risk scoring method [15]. The use of such models has been reported for subsets of patients such as the elderly [15, 16, 17,18]. According to a recent study published in the Joint Commission Journal [12], alerting systems based on advanced data science methods can identify medication prescribing errors that might be missed by traditional clinical decision support (CDS) systems using statistical or expert-opinion models. This is so because existing CDS systems are typically rule-based and thus can only detect errors that have been previously identified and programmed into their alerting logic. Various automatic learning approaches for the problem of predicting ADRs or ADEs already exist, with variations depending on their intended function, the available data, the target user, etc. Another alternative is supervised learning models [13], which have demonstrated excellent predictive abilities for other patient risk prediction problems [14]. However, this approach involves the construction of data labels, a complex procedure based on expert opinion or clinical results that are often unreliable or simply unavailable. Most publications on ADR or ADE calculate risk using strategies such as logistic regression [15, 16,17,18,19,20]. Segal et al. [21] present a model based on outlier detection to predict prescription errors and ADE at the moment of prescribing. Another study uses deep learning to predict ADEs with data labeled by the International Classification of Diseases (ICD-10) code group [22]. Jamal et al. (2019) [23] used machine learning-based (Random Forest) approach for cardiovascular (CV) ADRs prediction. The performance of models reported in the recent literature (see Table 1) has generally been evaluated in terms of the Area Under the ROC Curve (AUC) indicator. The score achieved by these models is in every case less than 0.81. Other studies attempting to predict ADE or ADR attained a high AUC, but its focus was confined exclusively to surgical patients [24], pediatric patients [25], or a specific type of ADR [26].

The present article reports on a study in which a number of ADR binary prediction models were constructed using machine learning methods and implemented with real-world ADR data on hospitalized patients. The results were evaluated to determine which type of model performed best and how they compared with the results reported in the recent literature. Our findings suggest how the use of such techniques could improve ADR prediction. The outcomes are also discussed in terms of usability in real-world clinical environments.

18.2 Materials and Methods

18.2.1 Data

Our data were takenfrom a recent study conducted at the Onofre Lopes University Hospital in Natal, Brazil [19], which are available on the website of the journal where the study was published. The dataset related to adult patients hospitalized for more than 24 hours who had received at least one medication during their hospital stay. Only first hospitalizations of patients registered between June 2016 and December 2017 were considered. The average patient age was 52.3 years and 58.8% were women. Excluded were patients admitted for ADRs, transplants or chemotherapy, as well as those hospitalized in the intensive care unit and pregnant women.

The data gathered for the original Brazilian study were obtained using observational methods that identified a set of 343 patients with ADRs and 686 controls. Each ADR case was detected through an active search of all patients carried out by three clinical pharmacists and four pharmacy students. Various sources of information were checked such as medical and nursing records, laboratory tests and prescription changes. Cases were considered only if they were definite, probable or possible according to the Liverpool Adverse Drug Reaction Causality Assessment Tool [27] and moderate, severe or lethal on the Hartwig's Severity Assessment Scale [28]. For each suspected case, two other patients were chosen at random to serve as controls. Patients in whom an ADR was confirmed were assigned to the set of positive cases; all others were put in the control group. For each patient identified as an ADR case, the original study collected data on 84 variables, including age, sex, race, self-reported daily intake of alcohol and annual cigarette packs smoked, intravenous substance use, number of prior hospitalizations, admission type (medical, urgent or elective surgery), previous ADR history, body mass index, ICD-10 classification chapter of main diagnosis, comorbidities included in the Charlson Comorbidity Index [29], Charlson Comorbidity Index score, Glasgow Coma Score, assisted ventilation, urine output, laboratory data, the total number of drugs prescribed, number of drugs prescribed for each of the first levels of the Anatomical Therapeutic and Chemical (ATC) Classification system, number of drugs prescribed by route of administration, number of drug interactions at different risk levels, and number of pharmaceutical incompatibilities.

18.2.2 Study design

One of the first difficulties to arise in this type of problem is that the dataset is intrinsically imbalanced. This is due mainly to the fact that ADRs are not as frequent as the number of prescriptions written by a hospital. The result is that there are datasets for low-risk classes with many cases and others for high-risk classes with few or very few cases. However, there area number of techniques for addressing the class-imbalance problem [30] that have been used in a variety of applications. In particular, the synthetic minority over-sampling technique (SMOTE) [31] has been employed to increase the minority classes. The SMOTE algorithm generates new minority-class samples through random interpolation between k randomly chosen minority-class neighbors (we choose k=5 and equalize the number of cases in both classes). The interpolation is performed within the convex hull of the reference samples. In the present study, we used a genetic algorithm method to choose a subset of the most important characteristics for our prediction problem [32]. This class of algorithms has proved to perform well in many areas including pharmacovigilance [33].We test the genetic algorithm operators with our three different approaches and select the best result of AUC. The performance indicator employed was AUC, evaluated after five-fold cross-validation of the best model. A minimum of 10 attributes and a maximum of 300 generations were considered. The selection scheme used was Tournament with 0.25 size. The initial probability to be switched on and crossover were defined as 0.5. The probability of mutation was defined 1/n where n was the number of attributes. uniform crossover-type was selected. We used Rapidminer Studio 9.6 package for the feature selection stage (Evolutionary optimize selection operator).

Two of the approaches we implemented are based on decision trees, which were first introduced in scientific work in 1984 [34] and recognized in the machine learning literature two years later [35].Random Forest (RF) [36] is an approach based on an ensemble of decision tree sets trained on random subsets of the original data, each subset trained independently. Each node of a tree is a division rule for a specific attribute where the number of variables on each tree is smaller than the total number in the whole model and are chosen randomly. The trees are then combined in order to reduce variance by combining the results of various weak classifiers (bagging). In this study, the maximum tree depth used with RF was 8 and forests of 100 (RF100), 500 (RF500), 1,000 (RF1000), 3,500 (RF3500) and 5,000 (RF5000) trees were tested. For each experiment and each case we tested 15 RF models, applying three different splitting criteria: information gain (i), Gini index (gi) and gain ratio (gr). We used Rapidminer Studio 9.6 package for RF models implementation.

The second tree-based approach, known as gradient-boosted trees (GBT) [37] is widely used due to its excellent performance on supervised learning problems. As with RF, it is based on constructing weak decision trees that are added to the model at each iteration in order to build a more exact formulation. The tree parameters to be added are determined by calculating the gradient descent that minimizes a loss function (which is pre-defined). GBT differs from bagging models mainly in that, in the former case, tree construction adapts to the predictor at different segments of the training set sequentially. With this method the maximum tree depth used was 5 and forests of 100 (GBT100), 500 (GBT500), 1,000 (GBT1000), 3,500 (GBT3500) and 5,000 (GBT5000) trees were tested in 5 different models. We used the H2O version 3.8.2.6 package [38] for GBT models implementation.

The third approach we implemented was deep learning (DL). This family of methods is considered to be state of the art based on its excellent performance in various pattern-recognition tasks [39]. To obtain good results, DL requires high volumes of data for training. We used the H2O version 3.8.2.6 package [38], a library that uses a multi-layer artificial neural network trained with back-propagation (using the stochastic gradient descent). The two activation functions of the neurons in the hidden layers were the rectifier (rect) and the hyperbolic tangent (tanh). We explored different network topologies, choosing the ones that delivered the best results. These were: (DL2x50), with two hidden layers made up of 50 neurons each; (DL10x50), with ten hidden layers made up of 50 neurons each; and (DL2x500), with two hidden layers made up of 500 neurons each. The 6 resulting models were trained over 50 epochs. The adaptive learning rate method (ADADELTA) [40] was utilized to avoid slow convergence. Other parameter values used were epsilon=10 e-8 and rho=0.99.

The experimental setup shown in Figure 1 is the framework for the evaluation of machine learning-based ADR prediction. We used the H2O version 3.8.2.6 package [38].

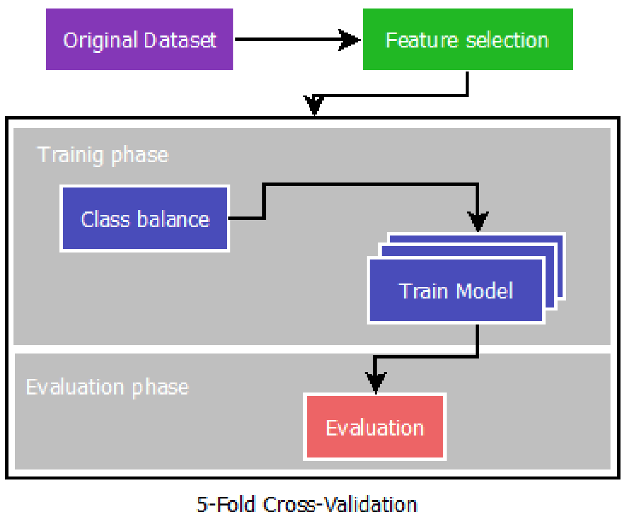


Fig. 18.1 Study design schema.

18.2.3 Evaluation metrics

To evaluate our results, we employed five-fold cross-validation with random stratified partitions of the initial set and a number of prediction error evaluation functions that are widely used in the literature. For each model developed we evaluated its performance in predicting an ADR, using the following metrics for dichotomous classes: recall, F-Score (or F-measure), precision by class and accuracy. Also used were measures such as the ROC curve and the area under the ROC curve (or the c-statistic) [41].

18.3 Results

The preliminary application of an attribute selection algorithm found just 23 of the 84 variables in the original study. Thus, for each patient only the following variables were used: age, sex, race, body mass index, height, weight, heart rate, systolic pressure, temperature, ICD-10 chapter classification of main diagnosis, previous ADR history, days since admission, mechanical ventilation, and the presence of the following: acquired immune deficiency syndrome, cardiovascular disease, diabetes, hemiplegia, leukaemia, liver disease, lymphoma, metastases, peripheral vascular disease and renal disease. With these variables thus selected, we ran 26 models based on the above-described techniques plus logistic regression. The best result as measured by AUC was 0.81, generated by the GBT model with 5,000 trees . The best result as measured by AUC was 0.81, generated by the GBT model with 5,000 trees. The ROC curve for that model is shown in red in Figure 2.

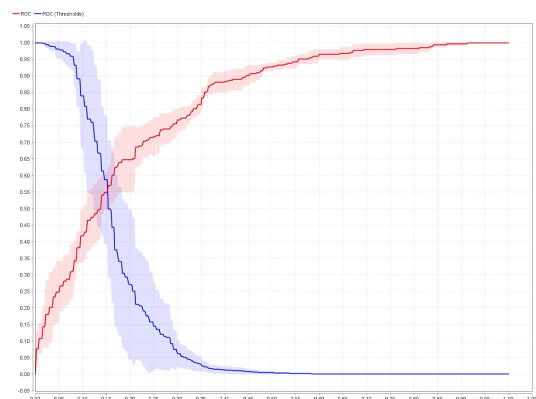


Fig. 18.2 ROC Curve for the Best Model (GBT5000)

The models were run on the same computer, powered by an Intel Core i7-8750H up to 4.1 GHz processor with 8 GB DDR3 of RAM. For comparison purposes, the ROC curves for the best model obtained with each method employed in the study are plotted in Figure 3. As can be seen, on the AUC metric GBT5000 outperformed not only logistic regression but also RF and DL (pairwise t-test $p<0.005$). The statistical significance of the difference in results between DL and GBT is minor, as is suggested by their respective curves in the figure. Also evident is that the two decision tree-based models (GBT and RF) perform better (pairwise t-test $p<0.005$) than either DL or logistic regression. GBT5000 exhibited a high recall level (true positive rate) of 78.3% with a relatively low level of false positives of 35%. By contrast, at that same recall level the corresponding false positives for DL and logistic

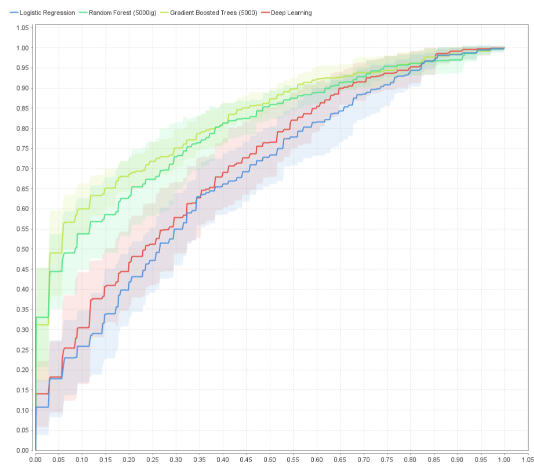


Fig. 18.3 ROC Curves for the Best Result by Approach

18.4 Conclusions

Models for predicting adverse drug reactions in hospitalized patients were constructed under several different machine learning methods and implemented with real-world data. The results of the models showed that they significantly outperform traditional statistical methods, which typically attain an AUC level of about 72%. The approaches that performed best in our experiments were gradient boosting trees (AUC=81.1% +/- 1.4%) and random forest (AUC=80.1% +/- 3.0%). These outcomes surpass the best-performing model we found in the recent literature [22], which reported an AUC for deep learning of 77%. Unfortunately, few earlier studies report their results for standard deviation so no comparisons on that indicator could be presented. Hopefully, future published work in this area will include standard deviation results given the importance of determining model reliability for application in real clinical settings. A noteworthy finding in the present study was that for the best GBT model, 40% of total attribute importance was imputable to only three variables: ICD-10, age and heart rate. A simple model using only these attributes attained AUC levels of about 70%. This is of particular interest given that the other attributes used in our models are, for the most part, biometric variables subject to capture errors and often costly to obtain. By contrast, ICD-10 and age should be relatively simple to determine accurately while heart rate is a biometric that is routinely measured, though it may depend on factors not related to medication. The implementation of electronic methods for clinic record-keeping and drug prescription has demonstrated that they can be powerful tools for reducing incidents relating to the use of medication incidence of medication errors that may pose a serious danger to patients. The potential of these methods can be significantly increased if they are combined with clinical decision support systems. The present study has shown that machine learning methods are capable of greater predictive ability than traditional statistical methods and their incorporation in such systems should therefore be seriously considered in efforts to improve the quality and safety of medical attention.

Acknowledgements

This research was funded by National Agency for Research & Development of Chilean Government. Under FONDEF project title: “Advanced Data Science Methods for medication error prevention”, project CODE: ID20I10001.

References

1. N.C.C for Medication Error Reporting,.; Prevention About Medication Errors. What is a Medication Error? 2016.
2. Ghaleb, M.A.; Barber, N.; Franklin, B.D.; Yeung, V.W.S.; Khaki, Z.F.; Wong, I.C.K. Systematic Review of Medication Errors in Pediatric Patients. Ann. Pharmacother. 2006, 40, 1766–1776, doi:10.1345/aph.1G717.
3. Rishoej, R.M.; Almarsdr, A.B.; Christesen, H.T.; Hallas, J.; Kjeldsen, L.J. Medication errors in pediatric inpatients: a study based on a national mandatory reporting system. Eur. J. Pediatr. 2017, 176, 1697–1705, doi:10.1007/s00431-017-3023-8.
4. Donaldson, M.S.; Corrigan, J.M.; Kohn, L.T.; others To err is human: building a safer health system; National Academies Press, 2000; Vol. 6;.
5. Organization, W.H.; others Reporting and learning systems for medication errors: the role of pharmacovigilance centres. 2014.
6. Hoonhout, L.H.F.; de Bruijne, M.C.; Wagner, C.; Asscheman, H.; van der Wal, G.; van Tulder, M.W. Nature, occurrence and consequences of medication-related adverse events during hospitalization. Drug Saf. 2010, 33, 853–864.
7. Bates, D.W.; Leape, L.L.; Petrycki, S. Incidence and preventability of adverse drug events in hospitalized adults. J. Gen. Intern. Med. 1993, 8, 289–294.
8. Kaushal, R.; Bates, D.W.; Landrigan, C.; McKenna, K.J.; Clapp, M.D.; Federico, F.; Goldmann, D.A. Medication errors and adverse drug events in pediatric inpatients. Jama 2001, 285, 2114–2120.
9. Barber, N.; Rawlins, M.; Franklin, B.D. Reducing prescribing error: competence, control, and culture. BMJ Qual. Saf. 2003, 12, i29--i32.
10. Sethuraman, U.; Kannikeswaran, N.; Murray, K.P.; Zidan, M.A.; Chamberlain, J.M. Prescription errors before and after introduction of electronic medication alert system in a pediatric emergency department. Acad. Emerg. Med. 2015, 22, 714–719.
11. Gates, P.J.; Hardie, R.-A.; Raban, M.Z.; Li, L.; Westbrook, J.I. How effective are electronic medication systems in reducing medication error rates and associated

harm among hospital inpatients? A systematic review and meta-analysis. J. Am. Med. Informatics Assoc. 2021, 28, 167–176, doi:10.1093/jamia/ocaa230.
12. Rozenblum, R.; Rodriguez-Monguio, R.; Volk, L.A.; Forsythe, K.J.; Myers, S.; McGurrin, M.; Williams, D.H.; Bates, D.W.; Schiff, G.; Seoane-Vazquez, E. Using a Machine Learning System to Identify and Prevent Medication Prescribing Errors: A Clinical and Cost Analysis Evaluation. Jt. Comm. J. Qual. Patient Saf. 2020, 46, 3–10, doi:10.1016/j.jcjq.2019.09.008.
13. Ouchi, K.; Lindvall, C.; Chai, P.R.; Boyer, E.W. Machine Learning to Predict, Detect, and Intervene Older Adults Vulnerable for Adverse Drug Events in the Emergency Department. J. Med. Toxicol. 2018, 14, 248–252, doi:10.1007/s13181-018-0667-3.
14. Wolff, P.; Gra.; R, S.A.; Yarza, M.B. Machine Learning Readmission Risk Modeling: A Pediatric Case Study. Biomed Res. Int. 2019, 2019, 9.
15. Trivalle, C.; Burlaud, A.; Ducimeti, P. Risk factors for adverse drug events in hospitalized elderly patients: A geriatric score. Eur. Geriatr. Med. 2011, 2, 284–289, doi:10.1016/j.eurger.2011.07.002.
16. Onder, G.; Petrovic, M.; Tangiisuran, B.; Meinardi, M.C.; Markito-Notenboom, W.P.; Somers, A.; Rajkumar, C.; Bernabei, R.; van der Cammen, T.J.M. Development and Validation of a Score to Assess Risk of Adverse Drug Reactions Among In-Hospital Patients 65 Years or Older. Arch. Intern. Med. 2010, 170, doi:10.1001/archinternmed.2010.153.
17. McElnay, J.C.; McCallion, C.R.; Al-Deagi, F.; Scott, M.G. Development of a Risk Model for Adverse Drug Events in the Elderly. Clin. Drug Investig. 1997, 13, 47–55, doi:10.2165/00044011-199713010-00006.
18. Tangiisuran, B.; Scutt, G.; Stevenson, J.; Wright, J.; Onder, G.; Petrovic, M.; van der Cammen, T.J.; Rajkumar, C.; Davies, G. Development and validation of a risk model for predicting adverse drug reactions in older people during hospital stay: Brighton Adverse Drug Reactions Risk (BADRI) model. PLoS One 2014, 9, e111254.
19. Lima, S.I.V.C.; Martins, R.R.; Saldanha, V.; Silbiger, V.N.; dos Santos, I.C.C.; Arajo, I.B. de; Oliveira, A.G. Development and validation of a clinical instrument to predict risk of an adverse drug reactions in hospitalized patients. PLoS One 2020, 15, e0243714, doi:10.1371/journal.pone.0243714.
20. O'Mahony, D.; O'Connor, M.N.; Eustace, J.; Byrne, S.; Petrovic, M.; Gallagher, P. The adverse drug reaction risk in older persons (ADRROP) prediction scale: derivation and prospective validation of an ADR risk assessment tool in older multi-morbid patients. Eur. Geriatr. Med. 2018, 9, 191–199.
21. Segal, G.; Segev, A.; Brom, A.; Lifshitz, Y.; Wasserstrum, Y.; Zimlichman, E. Reducing drug prescription errors and adverse drug events by application of a probabilistic, machine-learning based clinical decision support system in an inpatient setting. J. Am. Med. Informatics Assoc. 2019, 26, 1560–1565, doi:10.1093/jamia/ocz135.
22. Rebane, J.; Samsten, I.; Papapetrou, P. Exploiting complex medical data with interpretable deep learning for adverse drug event prediction. Artif. Intell. Med. 2020, 109, 101942, doi:10.1016/j.artmed.2020.101942.

23. Jamal, S.; Ali, W.; Nagpal, P.; Grover, S.; Grover, A. Computational models for the prediction of adverse cardiovascular drug reactions. J. Transl. Med. 2019, 17, 171.
24. Bos, J.M.; Kalkman, G.A.; Groenewoud, H.; van den Bemt, P.M.L.A.; De Smet, P.A.G.M.; Nagtegaal, J.E.; Wieringa, A.; van der Wilt, G.J.; Kramers, C. Prediction of clinically relevant adverse drug events in surgical patients. PLoS One 2018, 13, e0201645, doi:10.1371/journal.pone.0201645.
25. Yu, Z.; Ji, H.; Xiao, J.; Wei, P.; Song, L.; Tang, T.; Hao, X.; Zhang, J.; Qi, Q.; Zhou, Y.; et al. Predicting Adverse Drug Events in Chinese Pediatric Inpatients With the Associated Risk Factors: A Machine Learning Study. Front. Pharmacol. 2021, 12.
26. On, J.; Park, H.-A.; Yoo, S. Development of a prediction models for chemotherapy-induced adverse drug reactions: A retrospective observational study using electronic health records. Eur. J. Oncol. Nurs. 2022, 56, 102066.
27. Gallagher, R.M.; Kirkham, J.J.; Mason, J.R.; Bird, K.A.; Williamson, P.R.; Nunn, A.J.; Turner, M.A.; Smyth, R.L.; Pirmohamed, M. Development and Inter-Rater Reliability of the Liverpool Adverse Drug Reaction Causality Assessment Tool. PLoS One 2011, 6, e28096, doi:10.1371/journal.pone.0028096.
28. 28. Haile, D.B.; Ayen, W.Y.; Tiwari, P. Prevalence and assessment of factors contributing to adverse drug reactions in wards of a tertiary care hospital, India. Ethiop. J. Health Sci. 2013, 23, 39–48.
29. Charlson, M.E.; Pompei, P.; Ales, K.L.; MacKenzie, C.R. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J. Chronic Dis. 1987, 40, 373–83, doi:10.1016/0021-9681(87)90171-8.
30. L, V.; Fernez, A.; Garc S.; Palade, V.; Herrera, F. An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. Inf. Sci. (Ny). 2013, 250, 113–141, doi:10.1016/j.ins.2013.07.007.
31. Chawla, N. V; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: synthetic minority over-sampling technique. J. Artif. Intell. Res. 2002, 16, 321–357.
32. Leardi, R.; Boggia, R.; Terrile, M. Genetic algorithms as a strategy for feature selection. J. Chemom. 1992, 6, 267–281, doi:10.1002/cem.1180060506.
33. Ghaheri, A.; Shoar, S.; Naderan, M.; Hoseini, S.S. The Applications of Genetic Algorithms in Medicine. Oman Med. J. 2015, 30, 406–416, doi:10.5001/omj.2015.82.
34. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. Classification and Regression Trees. Cole Publ. Monterey 1984, 535, 358, doi:10.1002/cyto.990080516.
35. Quinlan, J.R. Induction of Decision Trees. Mach. Learn. 1986, 1, 81–106, doi:10.1023/A:1022643204877.
36. Breiman, L. Random Forests. Mach. Learn. 2001, 45, 5–32.
37. Friedman, J.H. Greedy function approximation: a gradient boosting machine. Ann. Stat. 2001, 1189–1232. 38. Cook, D. Practical Machine Learning with H2O: Powerful, Scalable Techniques for Deep Learning and AI; O'Reilly Media, 2016; ISBN 9781491964606.
38. Goodfellow, I.; Bengio, Y.; Courville, A. Deep Learning; MIT Press, 2016.
39. Zeiler, M.D. ADADELTA: An Adaptive Learning Rate Method. CoRR 2012, abs/1212.5.

40. McCoy, A.B.; Waitman, L.R.; Lewis, J.B.; Wright, J.A.; Choma, D.P.; Miller, R.A.; Peterson, J.F. A framework for evaluating the appropriateness of clinical decision support alerts and responses. J. Am. Med. Informatics Assoc. 2012, 19, 346–352..

Chapter 19
Time-based reasoning in medical data: from visualization and data mining to artificial intelligence and deep learning methods

Jon Kerexeta[0000−0002−6516−8619]

Abstract This article describes both the work done to date and the next steps to be taken in this thesis. The thesis deals with different methods of exploiting clinical time-based data. First, the two main works in which visual analysis and data mining methods have been used to exploit data are presented. Then, the next steps to be taken in order to perform data mining with artificial intelligence methods applied to monitored data are presented. In conclusion, the article shows the introduction of innovative data mining methods for different clinical use cases with a temporal component in different medical scenarios.

19.1 Introduction

On the one hand, data mining and visual analytics methods are very demanding for the correct interpretation and exploitation of data, both for static and time-series data. If we analyze literature reviews [1], we see that there are many types of visual analytics, and data mining techniques for data analysis in medicine, and if we go into more detail, we can find articles such as Fadel et al's [2], that shows how a visualization dashboard improves guideline implementations in hypertension. In general, this is a well-studied field, nevertheless, contributions are still required, especially in time-based data. This is the focus of the first part of the thesis, wherein two projects we have focused on solving the exploitation of temporal data by means of visual analytics and data mining. On the other hand, artificial intelligence (AI) has proven to be of great benefit in medicine in the last few decades. In the state of the art, we can find many classification models that diagnose different diseases or predict readmissions based on patient's static data [3, 4], diagnoses or segmentations based on images such as CT scans [5], and even risk estimations or diagnoses based on monitored data [6]. This last case is the focus of the second part of the thesis: the development of artificial intelligence models of classification that best fit monitored medical data, which are mathematically considered time series.There are limited examples of scientific research works on AI models in medicine that are suitable for time series data. There are many cases where time series are de-temporalized (extract mean, standard deviation etc.) and analyzed as if they were cross-sectional data. Although this approach can deliver good results in certain scenarios, much of temporality information is lost in the data that could improve the model's suitability. There are several examples of time-related variables. An example could be the weight increment that can be destemporised by performing moving averages and similar methods, i.e. with a doc technique. However, here there are many parameters on how to perform this destemporisation such as, original weight (the first weight date) and final weight (for example, three-day's weight average). This results in work that if a model is able to explore automatically, it may even be able to perform these techniques with a better fit for the dataset. Also, weight gain is a known problem, but there may be many unexplored significant cases that a time-series specialized IA model could find. Due to this problem, it is seen that there is the possibility to improve those models making use of the temporal reasoning such as the used in natural language processing. Moreover, in the last few years we have seen a significant increase in the use of such models in medicine.

This thesis aims to investigate and attempt to improve the reasoning in time series in medicine, the first part is focused on visualization and data mining and the second part, which is ongoing, in the improvements of the AI methods in medical time series data.

19.2 Visual analytics and data mining

In this section, we show two innovative methods that we have proposed in two different projects, both of which are based on the exploitation of time through visual analytics and data mining methods.

19.2.1 COLAEVA

This is a study carried out in the European CAPTAIN project, which makes use of data mining methods, unsupervised classification, and visual analytics to offer the best possible advice (nutritionally, physically, socially, or cognitively) to older adults. The study was published in sensors [7]. In the project, older people have been monitored for, social, physical and nutritional data, as well as being tested in these areas as well as in the cognitive field at the beginning and end of the study. COLAEVA [7] is a tool (accessible online it [8]), which exploits this information in an interactive way so that the coach can recommend different recommendations to future users. Firstly, it groups patients with semi-supervised techniques based on the results of the first assessment. Secondly, it visualizes the activity they have had on the basis defined fields, and with the color if there has been improvement or not.

An example that can be found in the tool presented in the same article [7] is shown in Figure 1. After the first steps in which the patients are grouped according to their initial state, in Figure 1 the evolution of the participants who a priori have a bad diet is shown. In the vertical axis, the nutritional activity is shown, and in color if they have improved/improved physically. With this tool, we draw conclusions such as that in this group, those who have maintained a good diet, have improved physically.

Fig. 19.1 Example extracted from [7] showing an example of COLAEVA's final graph. This graph shows how patients in group 0 of a previous grouping (low nutritional level) obtain improvements in the final physical assessments if they maintain a good nutritional level throughout the follow-up.

19.2.2 Disease Trajectories

Disease trajectories attempt to estimate the probability that a patient will suffer from one disease after being diagnosed with another. In this study, using data mining techniques, the most representative trajectories have been selected and visualized using networks to find previously unknown trajectories. The study has been carried out in a cohort of 50k patients. It is an ongoing work with preliminary results, and it is intended to be published in a journal, so no further details will be given in order not to cause conflicts at the time of publication. Similar studies have been carried out in the following articles [9, 10].

19.3 Time-series-based AI models

To create classification AI models with static data, we have a very wide range of models: Random Forest, Naive Bayes, Support Vector Machines, etc. However, for monitored data the possibilities are considerably reduced. The field of Natural Language Processing (NLP) is generally where the most powerful time series models are being used because of their type, need and large sample size. Deep-learning models have long been used to learn from temporality, more specifically those based on recurrences. In the same way that convolutions are used in imaging, these recurrences have been used in time series. This allows the model to have a type of memory that learns from the latest steps. In recent years, Transformers [11, 12] are the models that are obtaining the best results both in the field of NLP [13] and in some fields of medicine. In this Ph.D. thesis, the intention is to investigate these methodologies in depth (as well as others in the State of the Art), understand them, attempt to improve them and later apply them in the most optimal way to the different scenarios we will study in the course of the Ph.D. (see Section 3.1).

19.3.1 Different Scenarios

There is a large area of diseases that could benefit greatly from the use of such technologies, but in this doctoral thesis the focus will be on the following three Use Cases:

19.3.1.1 Intensive Care Unit (ICU)

Intensive care unit (ICU): In the case of ICUs, patients enter intensive care at high risk of deterioration and/or death. These patients are continuously monitored to detect any deterioration and so that clinical are able to intervene quickly and effectively. In this case, AI models based on time series could both anticipate the patients' decline effectively and even diagnose different types of secondary diseases that the patient may suffer from in the future (e.g., type of shock suffered by a patient in ICU) [14]. Moreover, with the current global pandemic, these methods could be of great help in the case of COVID-19 in the ICU. For this particular use case, we have identified this set of large public datasets:

- AmsterdamUMCdb [15]
- HiRID [16]
- MIMIC III [17]

These datasets will enable us to do different modelling studies in the thesis and will help us to compare the results with other studies. Upon achieving adequate results, we will attempt to apply of adjust the models to the patients in the Basque Country, Spain.

19.3.1.2 Telemonitoring

Telemonitoring aims to have daily information on patients without needing for them to go to the hospital, and simultaneously allowing for detection of patient decline based on the variables being monitored. In this case, we have experience in prior projects in which we have monitored patients with heart failure in order to detect cardiac decompensation [6]. In these models, we de-temporize the information and train AI models. We believe that in this Ph.D., based on what we learn from the technologies that can be applied to these cases, we can improve the models we currently use, even those in the state-of-the-art.

19.3.1.3 Disease trajectories

As commented in Section 2.2, disease trajectories attempt to estimate the probability that a patient will suffer from one disease after being diagnosed with another. Generally, these problems are treated with statistical methods [9, 10], but we have seen that there are also some cases that use AI methods based on time series and transformers [18]. In this case, it imitates methods used in natural language processing and manages to take advantage of temporal information such as prior diagnoses, age, and other variables. We estimate that we will have some projects related to disease trajectory and involving even more data so the methods studied in the PhD may be well suited to this scenario as well.

19.4 Summary

This paper shows the steps followed so to date, and the steps to be followed in the thesis. In the first phase, we have worked on data mining and visualization with clinical data. Two cases have been shown in which these techniques have been used. In the second phase, we will deal with AI models that are based on temporal data and have a clinical component. For this last step, we have shown some identified use cases where these AI methods can be very beneficial and where there is room to propose innovation methods such as the ones addressed in the thesis.

References

1. West, V.L., Borland, D., Hammond, W.E.: Innovative information visualization of electronic health record data: a systematic review. J Am Med Inform Assoc. 22, 330–339 (2015). https://doi.org/10.1136/amiajnl-2014-002955.
2. Fadel, R.A., Ross, J., Asmar, T., Sridasyam, K., Demertzis, Z., Ahluwalia, G., Roumayah, T., Scott, M., Ibrahim, H., Hammoudeh, R., Gandhi, N., Flynn, M.,

Haftka-George, A., Heidemann, D., Sims, S., Levy, P., Miller, J.: Visual Analytics Dashboard Promises to Improve Hypertension Guideline Implementation. American Journal of Hypertension. 34, 1078–1082 (2021). https://doi.org/10.1093/ajh/hpab081.
3. Kerexeta, J., Artetxe, A., Escolar, V., Lozano, A., Larburu, N.: Predicting 30-day Readmission in Heart Failure Using Machine Learning Techniques. In: HEALTHINF (2018).
4. Lei, Y., Yang, B., Jiang, X., Jia, F., Li, N., Nandi, A.K.: Applications of machine learning to machine fault diagnosis: A review and roadmap. Mechanical Systems and Signal Processing. 138, 106587 (2020). https://doi.org/10.1016/j.ymssp.2019.106587.
5. López-Linares Román, K., García Ocaña, M.I., Lete Urzelai, N., González Ballester, M.Á., Macía Oliver, I.: Medical Image Segmentation Using Deep Learning. In: Chen, Y.-W. and Jain, L.C. (eds.) Deep Learning in Healthcare: Paradigms and Applications. pp. 17–31. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-32606-7_2.
6. Larburu, N., Artetxe, A., Escolar, V., Lozano, A., Kerexeta, J.: Artificial Intelligence to Prevent Mobile Heart Failure Patients Decompensation in Real Time: Monitoring-Based Predictive Model. Mobile Information Systems. 2018, (2018).
7. Sensors | Free Full-Text | COLAEVA: Visual Analytics and Data Mining Web-Based Tool for Virtual Coaching of Older Adult Populations, https://www.mdpi.com/1424-8220/21/23/7991, last accessed 2022/05/02.
8. Kerexeta, J., Beristain, A., Alvarez, R.: (May, 2021) ICEWEA: Visual analytics and data mining web-based tool for virtual coaching of older adult populations: https://captain-elderly-evolution.herokuapp.com/.
9. Paik, H., Kan, M.J., Rappoport, N., Hadley, D., Sirota, M., Chen, B., Manber, U., Cho, S.B., Butte, A.J.: Tracing diagnosis trajectories over millions of patients reveal an unexpected risk in schizophrenia. Scientific Data. 6, 201 (2019). https://doi.org/10.1038/s41597-019-0220-5.
10. Jensen, A.B., Moseley, P.L., Oprea, T.I., Ellesøe, S.G., Eriksson, R., Schmock, H., Jensen, P.B., Jensen, L.J., Brunak, S.: Temporal disease trajectories condensed from population-wide registry data covering 6.2 million patients. Nature Communications. 5, 4022 (2014). https://doi.org/10.1038/ncomms5022.
11. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need. arXiv preprint arXiv:1706.03762. (2017).
12. Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., Zhang, W.: Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. arXiv:2012.07436 [cs]. (2021).
13. Braşoveanu, A.M.P., Andonie, R.: Visualizing Transformers for NLP: A Brief Survey. In: 2020 24th International Conference Information Visualisation (IV). pp. 270–279 (2020). https://doi.org/10.1109/IV51561.2020.00051.
14. Harutyunyan, H., Khachatrian, H., Kale, D.C., Ver Steeg, G., Galstyan, A.: Multitask learning and benchmarking with clinical time series data. Scientific Data. 6, 96 (2019). https://doi.org/10.1038/s41597-019-0103-9.
15. Thoral, P.J., Peppink, J.M., Driessen, R.H., Sijbrands, E.J.G., Kompanje, E.J.O., Kaplan, L., Bailey, H., Kesecioglu, J., Cecconi, M., Churpek, M., Clermont, G.,

van der Schaar, M., Ercole, A., Girbes, A.R.J., Elbers, P.W.G.: Sharing ICU Patient Data Responsibly Under the Society of Critical Care Medicine/European Society of Intensive Care Medicine Joint Data Science Collaboration: The Amsterdam University Medical Centers Database (AmsterdamUMCdb) Example*. Crit Care Med. 49, e563–e577 (2021). https://doi.org/10.1097/CCM.0000000000004916.
16. 16. Faltys, Martin, Zimmermann, Marc, Lyu, Xinrui, Hüser, Matthias, Hyland, Stephanie, Rätsch, Gunnar, Merz, Tobias: HiRID, a high time-resolution ICU dataset, https://physionet.org/content/hirid/1.1.1/. https://doi.org/10.13026/NKWC-JS72.
17. MIMIC-III, a freely accessible critical care database | Scientific Data, https://www.nature.com/articles/sdata201635?sou
--------------------------, last accessed 2021/03/31.
18. Li, Y., Rao, S., Solares, J.R.A., Hassaine, A., Ramakrishnan, R., Canoy, D., Zhu, Y., Rahimi, K., Salimi-Khorshidi, G.: BEHRT: transformer for electronic health records. Scientific reports. 10, 1–12 (2020).

Chapter 20
A digital twin strategy for sustainable development of smart protected natural areas through ecological indicators

Jose Garcia Rodriguez, Jose Norberto Mazon, Esther Salvador, Jose Maria Cecilia-Canales, Pietro Manzonni

Abstract CHAN-TWIN will develop a digital twin strategy to promote innovative techniques based on low-cost monitoring for creating and using improved models of the interplay between ecological and societal systems that can be adapted to develop smart natural areas across Europe, with special emphasis in tourism sustainability. The CHAN-TWIN digital twin will address research challenges combining different sensing and modeling approaches to feed a data consumption pipeline, i.e., from raw input data through data integration, to data analysis and visualization, thus providing information to different users, including policy-makers, tourism managers or re-searchers

20.1 Introduction

Protected Natural Areas (PNAs) are places to be managed sustainably for the long-term conservation of their biodiversity [1] and their promotion of both direct and indirect social benefits[2]. These benefits are grouped into three broad categories: (1) environmental, such as the ability to regulate both local and global climate; (2) social, such as the opportunity for recreation; and (3) economic, related to the development of business activities [3]. Thus, PNAs contribute to all spheres of human well-being [4], either through their influence on health, freedom, security, and social relations or, simply, by providing resources [5]. However, the level of welfare provision of a PNA is highly conditioned by the effects of anthropic pressure [6], since tourism activity, pollution, overexploitation of resources, and unsustainable uses are the main causes of biodiversity loss as well as other forms of degradation [7] .

Jose Garcia Rodriguez, Jose Norberto Mazon, Esther Salvador
University of Alicante, Alicante

Jose Maria Cecilia-Canales, Pietro Manzonni
Polytechnic University of Valencia, Valencia, Spain

Among PNAs, wetlands are one of the most globally threatened ecosystems of the world [8], having reduced their surface by 60-70 % worldwide in the last century [9]. Intriguingly, wetlands also provide many functions to the ecosystems and services to humans. For example, they are responsible for maintaining and regulating water quality and atmospheric gases, providing cultural and recreational resources and sequestering carbon [10]. These ecological functions make wetlands key to increase ecosystem resilience to climate change [11]. Many aquatic animal and plant species also rely completely on wetlands for their survival, making some well-preserved wetlands, hotspots of biodiversity where large concentrations of animals, such as waterbids, can be found [12]. The presence of these aquatic species increases the ecological value of wetlands by adding new functions to the ecosystem. For example, waterbirds are known to control pests, be indicators of ecosystem health, maintain ecosystem diversity by propagule dispersal and provide recreational and cultural services to humans [13]. One of the main reasons for deterioration of wetlands status is the lack of awareness by the general public about the economic value of the benefits they provide, but also the anthropogenic impacts such as tourism. Agriculture or mining they receive. Moreover, resources for the conservation and management of PNAs are limited, and often there is a lack of long-term data, which may be essential for the monitoring of the ecological functions of these ecosystems. Therefore, continuous assessment of the interplay between ecological and societal systems is required to ensure the sustainability of PNAs but also the maintenance of economic activities on them. Key information technologies (IT) for managing tourism in PNAs are reviewed in [14] where it is stated that ITs are highly required to support monitoring ecotourism developments and associated impacts, thus preserving biodiversity and embracing principles of sustainable tourism; largely concerned for the economic, social and, most importantly, environmental impacts. Consequently, the concept of Smart PNAs emerged as the application of smart city concepts to protected natural areas. As stated by INVATTUR in 2019, Smart PNA makes use of certain enabling technologies to facilitate sustainable socio-economic development through the interaction and integration of visitors with natural resources and biodiversity, adapting to the guidelines defined for the conservation of natural heritage to improve the quality and experience of tourism in an innovative way.

20.2 Proposal

In view of the above, the starting hypothesis of CHAN-TWIN is as follows: it is feasible to develop intelligent systems based on digital twins to build a systemic understanding of the socio-ecological processes affecting PNAs systems and enable assessment of the outcomes of future management options. The CHAN-TWIN digital twin will address all research challenges combining different sensing and modelling approaches that will be distributed in different modules (see Figure 1), which will produce data that feed a data consumption pipeline. From a technical standpoint, data consumption, either for policy makers, researchers, watchdog organisations or citizenships, is about the pipeline from raw input data through data integration, to data analysis and visualisation, to eventually get insights. This could be achieved, for instance, by means of generating novel mechanisms to interact with a large number of potential users of CHAN-TWIN digital twin for consuming the right data at the right moment when browsing the Web, as a universal and ubiquitous means of accessing information, and providing a seamless visualisation.

Particularly, CHAN-TWIN will be application-driven based on the "Lagunas de la Mata y Torrevieja" Natural Park (Alicante, Spain); a protected area whose environmental value has been globally recognized since 1972 in the RAMSAR convention. Composed of two salt lagoons; the pink lagoon (i.e., Torrevieja lagoon), where the Torrevieja salt mine is located, and the green lagoon (i.e., La Mata Lagoon) where the park's Interpretation Centre is located. It was chosen as an example for rolling out the CHAN-TWIN's approach as it has intensified pressures (tourism, urbanisation, agriculture, eutrophication, contamination, flash flooding and climate change) that have led to ecosystem deterioration, which through negative feedback impacts some of the socioeconomic drivers, mainly tourism, which accounts for the largest share of GDP of surrounding municipalities. This natural park is also important ecologically, as it is used by many waterbirds, including large concentrations of some species such as the flamingo (Phoenicopterus ruber) or the black-headed grebe (Podiceps nigricollis). In addition, the "La Mata y Torrevieja" Natural park is a challenging scenario from a technological point of view because of its large geographical spread (3.743 ha), which limits connectivity. It is also worth pointing out that the existing legal framework of this PNA must be considered; this PNA is owned by the Spanish State, but due to the importance of the existing salt production, it is leased to a private company for its exploitation and commercialisation, and at the same time, it is a natural area protected by the Government of the Valencian Community as a Natural Park. There are regulations to be analysed (such as "Plan Ordenaci los Recursos Naturales, PORN" or "Plan Rector de Uso y GestiRUG"), Habitats Directive and Birds Directive, Law 42/07 on the conservation of nature and natural heritage, Law on Natural Spaces of the Regional Ministry of the Valencian Community, Data protection law, Animal Welfare Act, as well as Law on the Sustainable Development of the Rural Environment. Consequently, the application of CHAN-TWIN project to a real case study intends to adapt results to the regulations of the corresponding environmental impact assessment law, thus identifying the existing environmental regulations for the design of a smart PNA.

Finally, stakeholders will play a fundamental role in CHAN-TWIN as they will provide qualitative information based on their main concerns, opinions and subjective perspectives of their reality. They will be involved at several stages and for different purposes, including participatory model and scenario development, data collection, interpretation of results, etc.

In contrast to a large number of scientific papers and reports which highlight the use of AI for environmental or socio-economic modelling, CHAN-TWIN, through its conceptual approach, will go one step further by developing an integrative tool to produce data for supporting the decision-making process in such complex scenarios. Synergies with international and national research and innovation activities will be sought, where CHAN-TWIN will build further on relevant past and ongoing (inter)national research activities and programs deployed across Europe, especially those relating to water management, water quality monitoring, best farm management practises, and use of IoT and sensing technologies, such as those previously demonstrated in the projects SMARTLAGOON, CollectionCare EOSC, WaterJPI, PROGNOS, ClimateJPI WATExR, FAIRWAYS, WaterOT or WATERSensing.

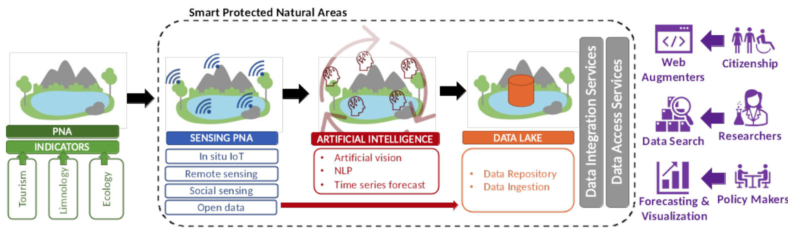


Fig. 20.1 Overview of CHAN-TWIN Digital Twin Strategy for Smart Protected Natural Areas.

CHAN-TWIN focuses on changing the current tourism management model towards sustainability by developing a digital twin to build a systemic understanding of the socio-ecological inter-relationships affecting PNAs (see Figure 1). The CHAN-TWIN digital twin will address research challenges combining different sensing and modeling approaches to feed a data consumption pipeline, i.e., from raw input data through data integration, to data analysis and visualization, thus providing information to different users, including policy-makers, tourism managers or researchers.

20.3 Objetives

Protected Natural Areas are ecosystems with great ecological and socioeconomic value, but they are also especially vulnerable to climatic and anthropogenic pressures, such as intensive agriculture and extensive urbanization because of tourist development. The vulnerability and complexity of these ecosystems requires a systemic understanding of the socio-ecological inter-relationships affecting PNAs and their ecosystem. Real-time monitoring, analysis and management of these critical resources can provide a mechanism for finding novel solutions to assess the trade-offs between socioeconomic and ecological aspects. Indeed, these solutions must be based on the data analysis of different sources, including social (e.g., surveys, social-media, legislation), biological (e.g., animal population monitoring) and physical (e.g., meteorological, historical) data. However, the analysis of such heterogeneous data is not straight-forward as it requires the design and development of advanced physical, artificial intelligence and system dynamics models for accurate forecasting. These computational expensive workloads have a strong impact on the carbon footprint as they usually require to be executed in large clusters of computers. he overall objective of CHAN-TWIN is to develop cross-cutting and green technology for monitoring and modelling socio-environmental processes across different temporal and spatial scales. This will be achieved through a digital twin strategy (see Figure 1) that allows researchers, stakeholders and policy-makers to collect data in a more cost-effective way, and to create more precise models and predictions to support better decision making, paving the way for the development of the smart PNAs model. As a case study, this project uses the PNA of La Mata and Torrevieja lagoons (Alicante, Spain), an ecosystem that supports a great variety of human activities encompassing tourism, agriculture and salt exploitation that need to be monitored to avoid deterioration. The proposed project is divided into five specific objectives:

Objective 1 Enable real-time tracking of socio-environmental data, obtained from different sources, including low-power sensors, social sensing and remote sensing using green information and communication technology (ICT).

Objective 2 Develop new ML modelling approaches, which combine sensing data and artificial intelligence techniques for the digitalization of ecosystem variables and functions in order to monitor ecosystem functioning and ultimately preserve smart PNAs

Objective 3 Develop a sustainable tourism approach, which provides insights into the coupled socio-ecological dynamics for the example of the Torrevieja destination to enable and support better management of PNAs.

Objective 4 Enable real-time visualization, data integration, management and forecasting of the interplay between environmental and societal systems for sustainable tourism of smart PNAs.

Objective 5 Disseminate project outcomes through tailored activities and formats targeting key audiences (stakeholders, policy-makers, the general public), including empowering existing networks to exchange knowledge, improving skills related to the use of CHAN-TWIN methodology. Moreover, the project aims to increase local and citizen awareness of environmental impacts through citizen sciences activities.

An important novelty of the CHAN-TWIN approach relies on the holistic view of the protected natural areas. Developing a data-centric technology solution will provide real-time monitoring and forecasting of socio-environmental tradeoffs in these ecosystems, aimed at helping policy-makers in their decision-making. It will also enable to increase people's awareness of these tradeoffs, which will benefit the understanding and enforcement of applicable legislation (such as the aforementioned. Birds Directive: Directive 2009/147/EC or Habitats Directive: Council Directive 92/43/EEC, RD 2/2019, DOGV nm. 2892, de 18.12.96).

20.4 Methodology

CHAN-TWIN is an interdisciplinary project whose main objective is to bridge the gap between the distant disciplines of ecology, tourism, advanced sensor research (ASR), social sciences (SS) and Artificial Intelligence (AI) to achieve the development of a digital twin platform to build the concept of smart PNAs. Methodology to be followed will be participatory, allowing collaboration among researchers and participation of stakeholders (i.e., in the sense of Action Research). To do so, specifically, our methodology focuses on involving stakeholders by considering a real situation to solve real problems. Particularly, the following approaches will be combined:

1. Efficient ICT techniques for data collection: the project will develop efficient tools and algorithms for crawling socio-ecological information in real time. From the hardware perspective, efficient IoT infrastructures with novel AI-based sensing technologies will be developed and eventually deployed in La Mata lagoon and its watershed.
2. Machine learning (ML) driven models;new approaches to creating ecological and environmental models will be developed based on the capacity of ML techniques to estimate nonlinear relationships between variables. Moreover, ML models will be used to develop novel sensors.
3. Participatory system dynamics modelling for strategic policy and management analysis: using the project's strengthened modelling capabilities in a multi-model set-up, CHAN-TWIN will analyze, through systems dynamic modelling and scenario studies, multiple dimensions of future pathways, environmental risks and management practices particularly focused on changing tourism towards sustainability.
4. Citizen engagement, stakeholder dialogue and interactions: multiple participatory activities (workshops, expert seminars, science coffees) will be organized with stakeholders and citizens to capture perspectives and knowledge and actively involve them in an iterative process aimed at both understanding needs, as well as presenting results and co-developing results and products.

20.5 Summary of socio-economic and scientific impact

CHAN-TWIN intents to advance the state of the art in computer science, ecology and tourism, both scientifically and technologically, while promoting the internationalization thanks to the European context in which the proposal is being developed. In the field of digitalization, CHAN-TWIN aims to develop a digital twin strategy that promotes sustainable tourism, paving the way to managing smart PNAs and the optimal ecological state of their ecosystems. To this end, an energy-efficient computing strategy based on continuum computing will be developed, enabling the computation of data at different levels of the IoT infrastructure. Specifically, CHAN-TWIN will design (1) artificial vision -based sensors for monitoring, classification and census of different species of birds, (2) new IoT architectures with adaptive mesh-based and fault-tolerant topologies, using low-power and long-range connectivity protocols, covering the geographical and technological conditions of the PNAs, (3) data management procedures of socio-environmental data from different sources, including remote sensing strategies, social sensors, in-situ infrastructures and open data to generate a comprehensive data lake on the PNA's socio-environmental status, and (4) a service layer for the management of ecological and tourism data to ease the creation of data-driven products and services to foster the PNA's sustainability, considering the data lifecycle (acquisition, preparation and consumption).

These advances in the domain of digitalization will, in turn, enable breakthroughs in the ecological transition. CHAN-TWIN will develop tourism sustainability indicators for smart PNAs, focusing on preserving the value of ecosystems, while optimizing the economic activity of the sector. Finally, CHAN-TWIN will establish a co-creation process with different institutions, organisations and citizens to raise public awareness on the relevance of preserving our natural resources through participatory events, and open innovation initiatives, which will enable the steady transition towards sustainable manage of socio-economic activities affecting natural resources.

Acknowledgement

This research has been partially funded by project "Desarrollo de un ecosistema de datos abiertos para transformar el sector turico" (GVA-COVID19/2021/103) funded by Conselleria de Innovaciniversidades, Ciencia y Sociedad Digital de la Generalitat Valenciana and "A way of making Europe" European Regional Development Fund (ERDF) and MCIN/AEI/10.13039/501100011033 for supporting this work under the "CHAN-TWIN" project (grant TED2021-130890B-C21.

References

1. Kettunen, M., & ten Brink, P. (Eds.). (2013). Social and economic benefits of protected areas: an assessment guide. Routledge.
2. Rode, J., Wittmer, H., Emerton, L., & Schrter-Schlaack, C. (2016). 'Ecosystem service opportunities': A practice-oriented framework for identifying economic instruments to enhance biodiversity and human livelihoods. Journal for nature conservation, 33, 35-47.
3. Chen, X., Lupi, F., & Liu, J. (2017). Accounting for ecosystem services in compensating for the costs of effective conservation in protected areas. Biological Conservation, 215, 233-240.
4. Iniesta-Arandia, I., GarcLlorente, M., Aguilera, P. A., Montes, C., & MartL, B. (2014). Socio-cultural valuation of ecosystem services: uncovering the links

between values, drivers of change, and human well-being. Ecological economics, 108, 36-48.

5. Liu, J., Mooney, H., Hull, V., Davis, S. J., Gaskell, J., Hertel, T., ... & Li, S. (2015). Systems integration for global sustainability. Science, 347(6225).
6. Sarmento, W. M., & Berger, J. (2017). Human visitation limits the utility of protected areas as ecological baselines. Biological Conservation, 212, 316-326.
7. Bonet-Garc F. J., Pz-Luque, A. J., Moreno-Llorca, R. A., Pz-Pz, R., Puerta-Pi, C., & Zamora, R. (2015). Protected areas as elicitors of human well-being in a developed region: A new synthetic (socioeconomic) approach. Biological Conservation, 187, 221-229.
8. Davidson et al. (2018) Marine and Freshwater Research, 69(4):620.
9. Davidson NC (2014) Marine and Freshwater Research, 65(10):934.
10. Dise NB 2009. Peatland response to global change. Science 326: 810–811.
11. Moomaw WR, Chmura GL, Davies GT, Finlayson CM, Middleton BA, Natali SM, Perry JE, Roulet N, Sutton-Grier AE. 2018. Wetlands in a changing climate: science, policy and management. Wetlands, 38:183-205.
12. Clarkson, B. R., Ausseil, A. G. E., & Gerbeaux, P. (2013). Wetland ecosystem services. Ecosystem services in New Zealand: conditions and trends. Manaaki Whenua Press, Lincoln, 192-202.
13. Green, A. J., & Elmberg, J. (2014). Ecosystem services provided by waterbirds. Biological reviews, 89(1), 105-122.
14. Katsoni, V., & Dologlou, N. (2017). ICT applications in smart ecotourism environments. In Smart cities in the mediterranean (pp. 225-244). Springer.

Chapter 21
Dynamic Window Approach for Autonomous Vehicles with Mecanum Wheels

Mireya Cabezas-Olivenza1[0000–0001–5025–1711], Iker Azurmendi-Marquinez1[0000–0002–8959–0742], Ekaitz Zulueta1[0000–0001–6062–9343] and Ander Sanchez-Chica1[0000–0002–0271–6825]

Abstract The Dynamic Window Approach algorithm is one of the most widely used methods for collision-free navigation. This algorithm searches for the optimal control parameters for the velocity and rotation of autonomous vehicles. Such AGVs often have varying dynamics, as in the case of this study, where the autonomous robot has Mecanum wheels that allow for omnidirectional navigation. This implies that the AGV can move in any direction while rotating in the plane. The study analyses the DWA algorithm applied to autonomous vehicles with this equipment. The energy consumption is closely associated with these wheels and must also be taken into consideration. A separation of the dynamic window into velocities and angle is proposed in order to allow the problem to be solved more efficiently. Simplifications are carried out to reduce the computational time, and therefore, a fast response to objects in navigation is produced. Moreover, the proposed design has only been executed in simulation, obtaining the appropriate results.

21.1 Introduction

Automated guided vehicles (AGVs) are recognised as requiring intelligent navigation systems. In the field of industrial applications, these systems must have the ability to adapt to various environments. In addition, it is necessary to provide such vehicles with decision-making capabilities, thus validating actions to avoid collisions. This is achieved by immediate readings from the sensors implemented in the AGV. These measurements are provided by means of a navigation algorithm, which also has to consider dynamic characteristics of the vehicle.

Mireya Cabezas-Olivenza et al.
System Engineering and Automation Control Department, University of the Basque Country (UPV/EHU), Nieves Cano, 12, 01006 Vitoria-Gasteiz, Spain e-mail: [mireya.cabezas004, iker.azurmendi, ekaitz.zulueta, ander.sanchez]@ehu.eus

There are different techniques for handling the navigation of an AGV. As demonstrated by Wang et al. [1], it is possible to make route planning decisions using Virtual Force Field (VFF). Once the trajectory is defined, MPC predictive control is applied to track the trajectory. By combining the VFF with the Potential Force Field (PFF) Burgos et al. [2] manage to build a viable environment for navigation. The union of the two methods provides angle commands for collision-free navigation.

Continuing with vector fields, there also exists the Vector Field Histogram (VFH) where the grid of the two-dimensional cartesian histogram is used as a map of the environment, continuously updating data from the sampled range. With data minimisation, a viable means of navigation can be constructed, but the algorithm needs to stop in front of obstacles (see Borenstein et al. [3]). With the use of a two-dimensional laser range finder (LRF) Ye [4] studies the traversability field histogram (TFH), correcting shortcomings of the VFH. In conjunction with a traversability index, it is able to decide the speed and direction of the AGV.

Lateral Control is another technique for navigation. It employs a discrete prediction model that anticipates future states of the vehicle by identifying the closest point between the vehicle and the trajectory. The algorithm that performs this type of navigation is Stanley (see Hoffmann et al. [5]). As Cabezas-Olivenza et al. [6] have demonstrated, together with computer vision, the Stanley algorithm is highly robust and produces a minimum trajectory deviation error. As a substitute for computer vision, AbdElmoniem et al. [7] use a local Time Elastic Band (TEB) planner to provide the controller with a collision-free path.

One of the most common methods for AGV navigation is the Dynamic Window Approach (DWA). As Fox et al. [8] argue, it is necessary to deal with constraints imposed by the synchronised drive mobile robots themselves, involving velocities and accelerations. These speeds must therefore be permissible within a dynamic window. In addition, inertia may need to be taken into account for the control of high velocities. Thus, by combining motion planning methods with real-time obstacle avoidance methods, navigation at high speeds is achievable (see Brock [9]).

To make the DWA avoid obstacles, Saranrittichai et al. [10] propose to modify the objective function to consider obstacles close to the trajectory by means of the histogram grid representation (F-DWA). A need is consequently recognised to also know the movement of obstacles to predict future collisions, as demonstrated by Missura et al. [11]. Furthermore, a treatment of the convergence properties of the algorithm is possible, merging both the concept of DWA and the concept of convergence. This idea is inspired by an MPC/CLF control (see Ögren et al. [12]).

Based on the Dijkstra algorithm it is also possible to implement navigation with the DWA. The former method is used for global route planning, the latter for local path planning. It requires a map constructed through SLAM (Simultaneous localization and mapping), being able to avoid obstacles and reach the target position, as studied by Liu et al. [13]. Another option for hybridisation of algorithms is presented by Kashyap et al. [14] where the DWA works together with a teaching-learning-based optimisation (TLBO) technique.

Continuing with learning techniques, employing a convolutional neural network to address the choice of parameters of the DWA cost function is another option (see Dobrevski et al. [15]). Parameters are predicted dynamically, considering instantaneous sensor readings. The network is trained with reinforcement learning algorithms, using the proximal policy optimisation (PPO) method. Reinforcement learning is also applied by Chang et al. [16] in their study, where they improve the DWA algorithm based on Q-learning. In this way, they add two new evaluation functions to improve the performance of global navigation and propose that the DWA parameters are adaptively learned.

There are AGVs with less conventional wheels such as mecanum wheels. In order to achieve an improvement in energy consumption in these cases, it is possible to extend the DWA algorithm with an added term in performance, as Xie et al. [17] studied. A combination of objectives is therefore realised, being low power consumption and high speeds. It has to be taken into account that these wheels suffer from slippage and vibrations at high speeds, which involve positioning errors and energy consumption as demonstrated by Xie et al. [18].

An inertial navigation system (INS) designed for vehicles with this type of wheels has been studied, which integrates an encoder, an accelerometer and a gyro sensor together with Kalman filters in purpose of correcting the mentioned errors (see Kim et al. [19]). Due to the A* algorithm, which has good expansibility and adaptability, Li et al. [20] improved navigation in unknown environments. Furthermore, the development of a platform including the mechanical design, the design of the system itself and the construction of the robot is necessary, as Piemngam et al. [21] point out.

In the present work, the use of the DWA algorithm for autonomous vehicles that are equipped with mecanum wheels is proposed. The energy consumption of the vehicle is simplified, thus simplifying the application of the collision-free navigation algorithm.

21.2 Materials and Method

Automatic guided vehicles need a control algorithm; in this study it concerns the DWA. In order to realize its application, it is necessary to take into account that the AGV is equipped with mecanum wheels, which requires considering some aspects such as velocity or energy consumption. Furthermore, this study has been carried out in simulation format only.

21.2.1 Vehicle dynamics

The mecanum wheels permit the AGV to execute translational movements in either direction while being concurrently capable of rotating in the plane. This is depicted in Fig. 1., with the blue square representing the autonomous robot.

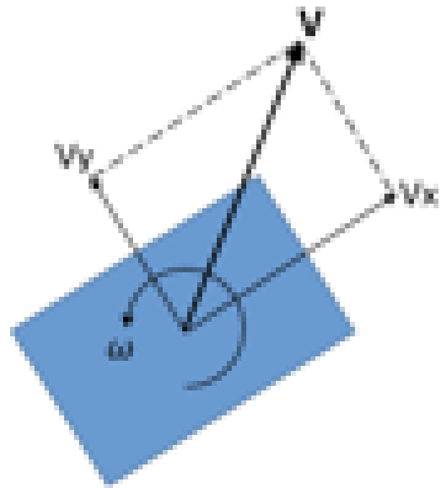


Fig. 21.1 Different speeds involved in the movement of the AGV.

Once the dynamic equations of the autonomous vehicle have been defined, a division of the dynamic window of the DWA algorithm is proposed. This is done in order to decouple problems and simplify the navigation algorithm. The window division is shown in Fig. 2. On the left side is the window which represents the velocities of both axes of the AGV. These will have an admissible value in a range. On the right side, the window involving the rotation angle. As the vehicle does not necessarily have to start the trajectory with the same desired final orientation, it is considered that in the first instance, the angle of the vehicle should be in accordance with the angle of the trajectory, i.e. [rad]. As the AGV approaches the target position, this angle will be adapted to achieve the [rad].

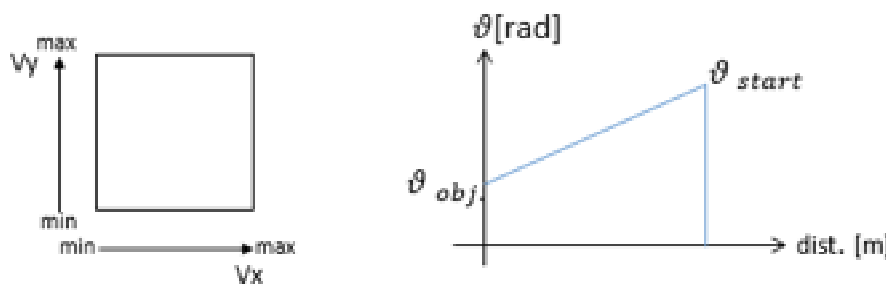


Fig. 21.2 Individual dynamic window approaches

Hence, there is equally a simplification in the energy consumption of the vehicle by setting the energy equation at any position [J] as follows, where [kg] corresponds to the mass of the vehicle and [kg/m] to the inertia.

As discussed, the second component of Equation 7 does not depend on adjustable parameters, because inertia is inherent to the vehicle and the angular velocity will depend on the distance between the vehicle and the target position, as depicted in Fig. 2. As for the first term, the mass of the vehicle is an attribute of the vehicle and only the velocities of both axles can be manipulated. Therefore, with this approach, the energy consumption of the AGV will depend on the limits set in the dynamic velocity window.

21.2.2 Proposed algorithm

A collision-free trajectory is needed. To this purpose, and continuing with the concept of the DWA algorithm, a cost function appropriate to the characteristics of the problem is presented. On the one hand, there is a term that evaluates the minimum distance to obstacles, defined as , in units of [m]. This distance to the nearest object is accumulated over the iterations. In the graph below it is shown how the minimum distance only reflects the distance to the closest obstacle.

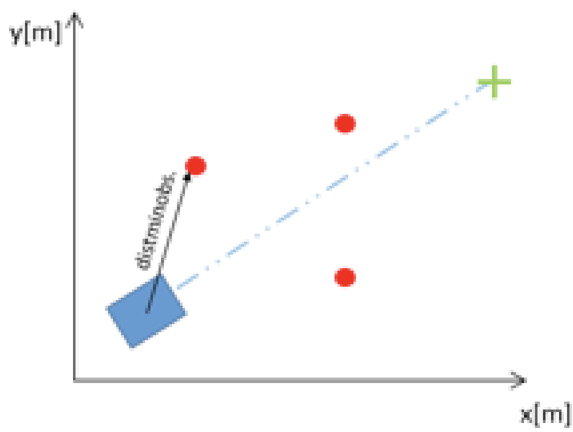


Fig. 21.3 Nearest obstacle representation.

Therefore, the equation to calculate $distminobs.$ relies exclusively on the position of the nearest obstacle and the current position of the AGV, being independent values of the angle.

$distminobs.{=}norm(obs{-}actual)$

The second term of Equation 8 evaluates the [m] of the vehicle to the target position. It is performed by normalising the distance between the point where the vehicle is located () and the target (), having regard also to the starting point (). Once again, just the and positions are considered, the equations remaining independent of the angle. Therefore, the following equation is defined.

It is also necessary to implement an equation that captures the dynamic window related to the vehicle angle described in Fig. 2. For this purpose, an equation is proposed where if the distance is far away the angle tends to 1 and the closer the AGV gets to the destination the closer it tends to 0. The following relationship is formulated which satisfies these conditions

As seen in Equation 5 and Equation 6, the subsequent positions of the AGV will depend on the value of . Equation 12 computes the relative position, i.e. where the AGV is currently located with respect to its target. Equation 13 implements what is described in Equation 11, with the introduction of a new term that evaluates the relative position defined as [m].

Finally, the calculation of the value of the vehicle angle is considered as follows.

This allows both proposed dynamic windows to be taken into account, allowing for collision-free navigation without the need to calculate a route in advance. The only information required is that associated with the target and the position of obstacles.

21.3 Results

A simulation is performed in which the AGV is located at a position = 0[m] and = 0[m] with an angle of = 0[rad]. It is desired to reach the position = 10m and = 10[m] with a rotation of = pi/2[rad]. A maximum linear velocity is fixed at = 1[m/s], as is the angular velocity, where = 1[rad/s]. For the DWA parameters, = 5, = 0.6 and = 0.1 are set. These parameters are constant and unitless. With these values the navigation shown in Fig. 4. is achieved. The black cross represents the starting point and the black circle the target position. The blue crosses indicate the obstacles, while the red circles represent the position of the AGV at each instant.

It can be observed that the vehicle is able to reach its objective while avoiding obstacles. Furthermore, it can be seen that as it approaches the target, its movement gradually decelerates in order not to overshoot the destination. In the following figures, different moments of the navigation are shown. They illustrate the vehicle's direction of motion at each step.

In the left image is captured the initial moment of the simulation where it can be appreciated that the angle of the AGV is similar to the direction it is steering as a trajectory. The central image shows how in curves the angle of the vehicle adapts, but it is progressively more uncoupled from the steering. Finally, in the right image, it can be seen how the vehicle already adopts the value of the angle with which it wants to reach the target.

21.4 Conclusions

The suggested method works well, as it is able to navigate in a collision-free trajectory, this being the main goal. It has the advantage that it is merely necessary to

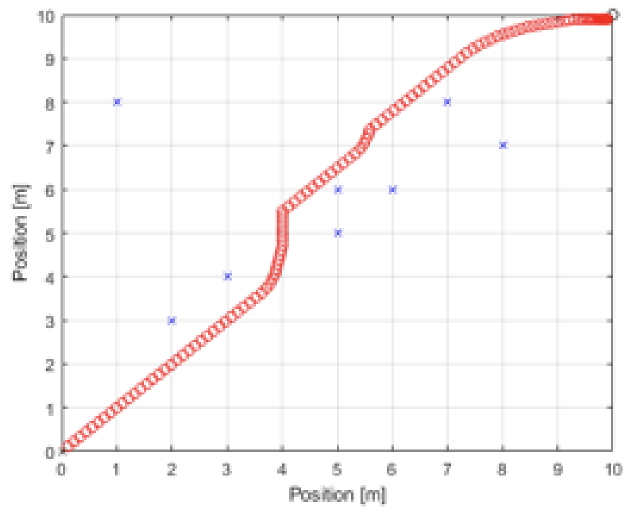


Fig. 21.4 Simulation of AGV navigation.

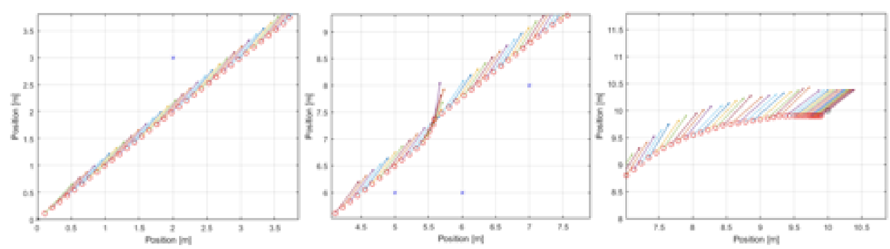


Fig. 21.5 Simulation moments with the directional control of the vehicle.

provide the algorithm with the target position to be reached together with the final angle and the position of obstacles in the environment. As the execution is done in simulation form only, it is uncertain how wheel slip affects the process, as it has not been measured or included in the development of the algorithm. Implementing the actual method in a real AGV would be a future work, where precision measurements can be performed. Furthermore, as the algorithm has not been implemented in the vehicle, it has also not been possible to measure the energy consumption generated and how much this factor would limit the navigation speed. This is also a future work to be carried out. Overall, for the simulation, a simplification of the algorithm is made for vehicles with mecanum wheels that fulfils the desired collision-free navigation.

References

1. Wang, T., Yan, X., Wang, Y., Wu, Q.: A distributed model predictive control using virtual field force for multi-ship collision avoidance under COLREGs. 2017 4th Int. Conf. Transp. Inf. Safety, ICTIS 2017 - Proc. 296–305 (2017).
2. Burgos, E., Bhandari, S.: Potential flow field navigation with virtual force field for UAS collision avoidance. 2016 Int. Conf. Unmanned Aircr. Syst. ICUAS 2016. 505–513 (2016).
3. Borenstein, J., Koren, Y.: The Vector Field Histogram - Fast obstacle avoidance for mobile robots. IEEE J. Robot. Autom. 7, 278–288 (1991).
4. 4. Ye, C.: Navigating a mobile robot by a traversability field histogram. IEEE Trans. Syst. Man, Cybern. Part B Cybern. 37, 361–372 (2007).
5. Hoffmann, G.M., Tomlin, C.J., Montemerlo, M., Thrun, S.: Autonomous automobile trajectory tracking for off-road driving: Controller design, experimental validation and racing. Proc. Am. Control Conf. 2296–2301 (2007).
6. Cabezas-Olivenza, M., Zulueta, E., Sanchez-Chica, A., Teso-fz-Betoño, A., Fernandez-Gamiz, U.: Dynamical Analysis of a Navigation Algorithm. 1–20 (2021).
7. AbdElmoniem, A., Osama, A., Abdelaziz, M., Maged, S.A.: A path-tracking algorithm using predictive Stanley lateral controller. Int. J. Adv. Robot. Syst. 17, 1–11 (2020).
8. Fox, D., Burgard, W., Thrun, S.: The dynamic window approach to collision avoidance. IEEE Robot. Autom. Mag. 4, 23–33 (1997).
9. Brock, O.: High-speed Navigation Using the Global Dynamic Window Approach. (1999).
10. Saranrittichai, P., Niparnan, N., Sudsang, A.: Robust local obstacle avoidance for mobile robot based on Dynamic Window approach. 2013 10th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol. ECTI-CON 2013. (2013).
11. Missura, M., Bennewitz, M.: Predictive collision avoidance for the dynamic window approach. Proc. - IEEE Int. Conf. Robot. Autom. 2019-May, 8620–8626 (2019).
12. Ögren, P., Leonard, N.E.: A convergent dynamic window approach to obstacle avoidance. IEEE Trans. Robot. 21, 188–195 (2005).
13. Liu, L.S., Lin, J.F., Yao, J.X., He, D.W., Zheng, J.S., Huang, J., Shi, P.: Path Planning for Smart Car Based on Dijkstra Algorithm and Dynamic Window Approach. Wirel. Commun. Mob. Comput. 2021, (2021).
14. Kashyap, A.K., Parhi, D.R., Muni, M.K., Pandey, K.K.: A hybrid technique for path planning of humanoid robot NAO in static and dynamic terrains. Appl. Soft Comput. J. 96, 106581 (2020).
15. Dobrevski, M., Skocaj, D.: Adaptive dynamic window approach for local navigation. IEEE Int. Conf. Intell. Robot. Syst. 6930–6936 (2020).
16. Chang, L., Shan, L., Jiang, C., Dai, Y.: Reinforcement based mobile robot path planning with improved dynamic window approach in unknown environment. Auton. Robots. 45, 51–76 (2021).

17. Xie, L., Henkel, C., Stol, K., Xu, W.: Power-minimization and energy-reduction autonomous navigation of an omnidirectional Mecanum robot via the dynamic window approach local trajectory planning: 15, (2018).
18. Xie, L., Scheifele, C., Xu, W., Stol, K.A.: Heavy-duty omni-directional Mecanum-wheeled robot for autonomous navigation: System development and simulation realization. Proc. - 2015 IEEE Int. Conf. Mechatronics, ICM 2015. 256–261 (2015).
19. Kim, J., Woo, S., Kim, J., Do, J., Kim, S., Bae, S.: Inertial navigation system for an automatic guided vehicle with Mecanum wheels. Int. J. Precis. Eng. Manuf. 2012 133. 13, 379–386 (2012).
20. Li, Y., Dai, S., Shi, Y., Zhao, L., Ding, M.: Navigation Simulation of a Mecanum Wheel Mobile Robot Based on an Improved A* Algorithm in Unity3D. Sensors 2019, Vol. 19, Page 2976. 19, 2976 (2019).
21. Piemngam, K., Nilkhamhang, I., Bunnun, P.: Development of Autonomous Mobile Robot Platform with Mecanum Wheels. 2019 1st Int. Symp. Instrumentation, Control. Artif. Intell. Robot. ICA-SYMP 2019. 90–93 (2019).

Chapter 22
Dynamic Window Approach for Autonomous Vehicles with Mecanum Wheels

Mireya Cabezas-Olivenza1[0000−0001−5025−1711], Iker Azurmendi-Marquinez1[0000−0002−8959−0742], Ekaitz Zulueta1[0000−0001−6062−9343] and Ander Sanchez-Chica1[0000−0002−0271−6825]

Abstract The Dynamic Window Approach algorithm is one of the most widely used methods for collision-free navigation. This algorithm searches for the optimal control parameters for the velocity and rotation of autonomous vehicles. Such AGVs often have varying dynamics, as in the case of this study, where the autonomous robot has Mecanum wheels that allow for omnidirectional navigation. This implies that the AGV can move in any direction while rotating in the plane. The study analyses the DWA algorithm applied to autonomous vehicles with this equipment. The energy consumption is closely associated with these wheels and must also be taken into consideration. A separation of the dynamic window into velocities and angle is proposed in order to allow the problem to be solved more efficiently. Simplifications are carried out to reduce the computational time, and therefore, a fast response to objects in navigation is produced. Moreover, the proposed design has only been executed in simulation, obtaining the appropriate results.

22.1 Introduction

Automated guided vehicles (AGVs) are recognised as requiring intelligent navigation systems. In the field of industrial applications, these systems must have the ability to adapt to various environments. In addition, it is necessary to provide such vehicles with decision-making capabilities, thus validating actions to avoid collisions. This is achieved by immediate readings from the sensors implemented in the AGV. These measurements are provided by means of a navigation algorithm, which also has to consider dynamic characteristics of the vehicle.

Mireya Cabezas-Olivenza et al.
System Engineering and Automation Control Department, University of the Basque Country (UPV/EHU), Nieves Cano, 12, 01006 Vitoria-Gasteiz, Spain e-mail: [mireya.cabezas004, iker.azurmendi, ekaitz.zulueta, ander.sanchez]@ehu.eus

There are different techniques for handling the navigation of an AGV. As demonstrated by Wang et al. [1], it is possible to make route planning decisions using Virtual Force Field (VFF). Once the trajectory is defined, MPC predictive control is applied to track the trajectory. By combining the VFF with the Potential Force Field (PFF) Burgos et al. [2] manage to build a viable environment for navigation. The union of the two methods provides angle commands for collision-free navigation.

Continuing with vector fields, there also exists the Vector Field Histogram (VFH) where the grid of the two-dimensional cartesian histogram is used as a map of the environment, continuously updating data from the sampled range. With data minimisation, a viable means of navigation can be constructed, but the algorithm needs to stop in front of obstacles (see Borenstein et al. [3]). With the use of a two-dimensional laser range finder (LRF) Ye [4] studies the traversability field histogram (TFH), correcting shortcomings of the VFH. In conjunction with a traversability index, it is able to decide the speed and direction of the AGV.

Lateral Control is another technique for navigation. It employs a discrete prediction model that anticipates future states of the vehicle by identifying the closest point between the vehicle and the trajectory. The algorithm that performs this type of navigation is Stanley (see Hoffmann et al. [5]). As Cabezas-Olivenza et al. [6] have demonstrated, together with computer vision, the Stanley algorithm is highly robust and produces a minimum trajectory deviation error. As a substitute for computer vision, AbdElmoniem et al. [7] use a local Time Elastic Band (TEB) planner to provide the controller with a collision-free path.

One of the most common methods for AGV navigation is the Dynamic Window Approach (DWA). As Fox et al. [8] argue, it is necessary to deal with constraints imposed by the synchronised drive mobile robots themselves, involving velocities and accelerations. These speeds must therefore be permissible within a dynamic window. In addition, inertia may need to be taken into account for the control of high velocities. Thus, by combining motion planning methods with real-time obstacle avoidance methods, navigation at high speeds is achievable (see Brock [9]).

To make the DWA avoid obstacles, Saranrittichai et al. [10] propose to modify the objective function to consider obstacles close to the trajectory by means of the histogram grid representation (F-DWA). A need is consequently recognised to also know the movement of obstacles to predict future collisions, as demonstrated by Missura et al. [11]. Furthermore, a treatment of the convergence properties of the algorithm is possible, merging both the concept of DWA and the concept of convergence. This idea is inspired by an MPC/CLF control (see Ögren et al. [12]).

Based on the Dijkstra algorithm it is also possible to implement navigation with the DWA. The former method is used for global route planning, the latter for local path planning. It requires a map constructed through SLAM (Simultaneous localization and mapping), being able to avoid obstacles and reach the target position, as studied by Liu et al. [13]. Another option for hybridisation of algorithms is presented by Kashyap et al. [14] where the DWA works together with a teaching-learning-based optimisation (TLBO) technique.

Continuing with learning techniques, employing a convolutional neural network to address the choice of parameters of the DWA cost function is another option (see Dobrevski et al. [15]). Parameters are predicted dynamically, considering instantaneous sensor readings. The network is trained with reinforcement learning algorithms, using the proximal policy optimisation (PPO) method. Reinforcement learning is also applied by Chang et al. [16] in their study, where they improve the DWA algorithm based on Q-learning. In this way, they add two new evaluation functions to improve the performance of global navigation and propose that the DWA parameters are adaptively learned.

There are AGVs with less conventional wheels such as mecanum wheels. In order to achieve an improvement in energy consumption in these cases, it is possible to extend the DWA algorithm with an added term in performance, as Xie et al. [17] studied. A combination of objectives is therefore realised, being low power consumption and high speeds. It has to be taken into account that these wheels suffer from slippage and vibrations at high speeds, which involve positioning errors and energy consumption as demonstrated by Xie et al. [18].

An inertial navigation system (INS) designed for vehicles with this type of wheels has been studied, which integrates an encoder, an accelerometer and a gyro sensor together with Kalman filters in purpose of correcting the mentioned errors (see Kim et al. [19]). Due to the A* algorithm, which has good expansibility and adaptability, Li et al. [20] improved navigation in unknown environments. Furthermore, the development of a platform including the mechanical design, the design of the system itself and the construction of the robot is necessary, as Piemngam et al. [21] point out.

In the present work, the use of the DWA algorithm for autonomous vehicles that are equipped with mecanum wheels is proposed. The energy consumption of the vehicle is simplified, thus simplifying the application of the collision-free navigation algorithm.

22.2 Materials and Method

Automatic guided vehicles need a control algorithm; in this study it concerns the DWA. In order to realize its application, it is necessary to take into account that the AGV is equipped with mecanum wheels, which requires considering some aspects such as velocity or energy consumption. Furthermore, this study has been carried out in simulation format only.

22.2.1 Vehicle dynamics

The mecanum wheels permit the AGV to execute translational movements in either direction while being concurrently capable of rotating in the plane. This is depicted in Fig. 1., with the blue square representing the autonomous robot.

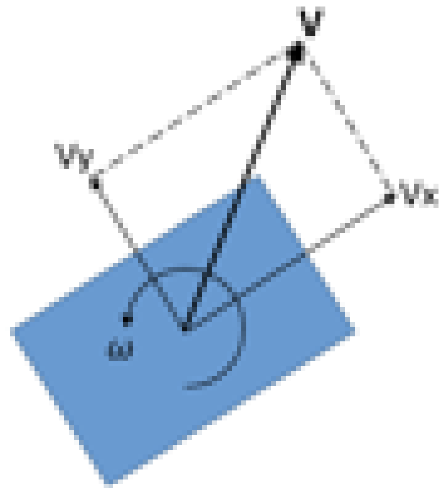


Fig. 22.1 Different speeds involved in the movement of the AGV.

Once the dynamic equations of the autonomous vehicle have been defined, a division of the dynamic window of the DWA algorithm is proposed. This is done in order to decouple problems and simplify the navigation algorithm. The window division is shown in Fig. 2. On the left side is the window which represents the velocities of both axes of the AGV. These will have an admissible value in a range. On the right side, the window involving the rotation angle. As the vehicle does not necessarily have to start the trajectory with the same desired final orientation, it is considered that in the first instance, the angle of the vehicle should be in accordance with the angle of the trajectory, i.e. [rad]. As the AGV approaches the target position, this angle will be adapted to achieve the [rad].

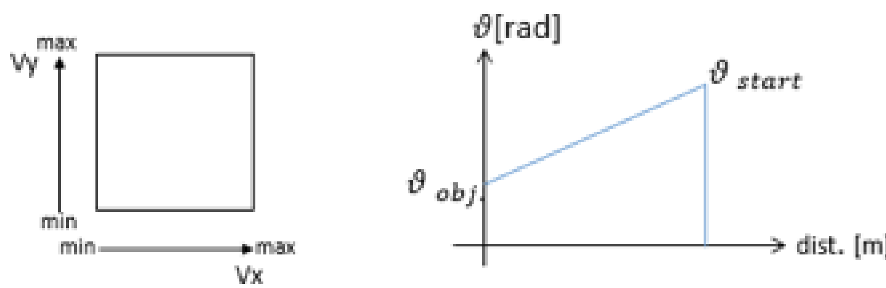


Fig. 22.2 Individual dynamic window approaches

Hence, there is equally a simplification in the energy consumption of the vehicle by setting the energy equation at any position [J] as follows, where [kg] corresponds to the mass of the vehicle and [kg/m] to the inertia.

As discussed, the second component of Equation 7 does not depend on adjustable parameters, because inertia is inherent to the vehicle and the angular velocity will depend on the distance between the vehicle and the target position, as depicted in Fig. 2. As for the first term, the mass of the vehicle is an attribute of the vehicle and only the velocities of both axles can be manipulated. Therefore, with this approach, the energy consumption of the AGV will depend on the limits set in the dynamic velocity window.

22.2.2 Proposed algorithm

A collision-free trajectory is needed. To this purpose, and continuing with the concept of the DWA algorithm, a cost function appropriate to the characteristics of the problem is presented. On the one hand, there is a term that evaluates the minimum distance to obstacles, defined as , in units of [m]. This distance to the nearest object is accumulated over the iterations. In the graph below it is shown how the minimum distance only reflects the distance to the closest obstacle.

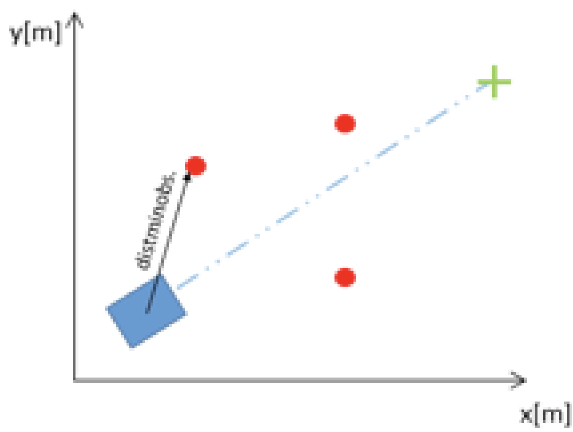


Fig. 22.3 Nearest obstacle representation.

Therefore, the equation to calculate $distminobs.$ relies exclusively on the position of the nearest obstacle and the current position of the AGV, being independent values of the angle.

$distminobs.{=}norm(obs{-}actual)$

The second term of Equation 8 evaluates the [m] of the vehicle to the target position. It is performed by normalising the distance between the point where the vehicle is located () and the target (), having regard also to the starting point (). Once again, just the and positions are considered, the equations remaining independent of the angle. Therefore, the following equation is defined.

It is also necessary to implement an equation that captures the dynamic window related to the vehicle angle described in Fig. 2. For this purpose, an equation is proposed where if the distance is far away the angle tends to 1 and the closer the AGV gets to the destination the closer it tends to 0. The following relationship is formulated which satisfies these conditions

As seen in Equation 5 and Equation 6, the subsequent positions of the AGV will depend on the value of . Equation 12 computes the relative position, i.e. where the AGV is currently located with respect to its target. Equation 13 implements what is described in Equation 11, with the introduction of a new term that evaluates the relative position defined as [m].

Finally, the calculation of the value of the vehicle angle is considered as follows.

This allows both proposed dynamic windows to be taken into account, allowing for collision-free navigation without the need to calculate a route in advance. The only information required is that associated with the target and the position of obstacles.

22.3 Results

A simulation is performed in which the AGV is located at a position = 0[m] and = 0[m] with an angle of = 0[rad]. It is desired to reach the position = 10m and = 10[m] with a rotation of = pi/2[rad]. A maximum linear velocity is fixed at = 1[m/s], as is the angular velocity, where = 1[rad/s]. For the DWA parameters, = 5, = 0.6 and = 0.1 are set. These parameters are constant and unitless. With these values the navigation shown in Fig. 4. is achieved. The black cross represents the starting point and the black circle the target position. The blue crosses indicate the obstacles, while the red circles represent the position of the AGV at each instant.

It can be observed that the vehicle is able to reach its objective while avoiding obstacles. Furthermore, it can be seen that as it approaches the target, its movement gradually decelerates in order not to overshoot the destination. In the following figures, different moments of the navigation are shown. They illustrate the vehicle's direction of motion at each step.

In the left image is captured the initial moment of the simulation where it can be appreciated that the angle of the AGV is similar to the direction it is steering as a trajectory. The central image shows how in curves the angle of the vehicle adapts, but it is progressively more uncoupled from the steering. Finally, in the right image, it can be seen how the vehicle already adopts the value of the angle with which it wants to reach the target.

22.4 Conclusions

The suggested method works well, as it is able to navigate in a collision-free trajectory, this being the main goal. It has the advantage that it is merely necessary to

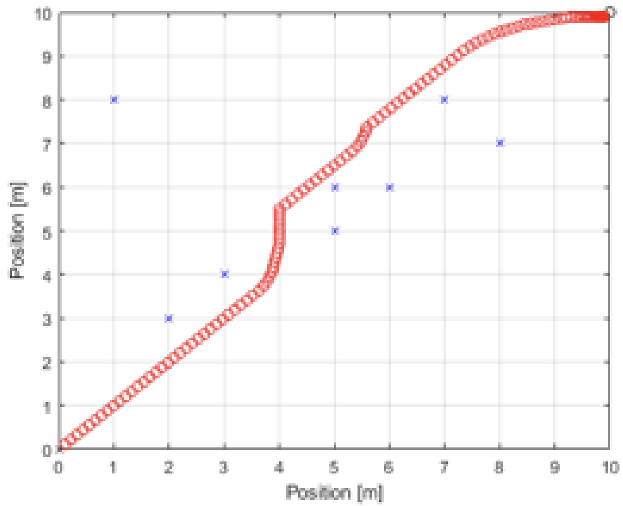


Fig. 22.4 Simulation of AGV navigation.

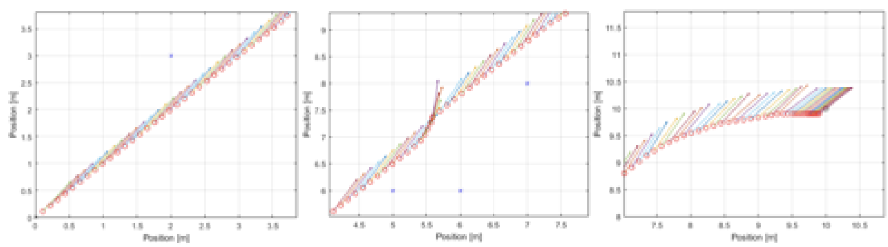


Fig. 22.5 Simulation moments with the directional control of the vehicle.

provide the algorithm with the target position to be reached together with the final angle and the position of obstacles in the environment. As the execution is done in simulation form only, it is uncertain how wheel slip affects the process, as it has not been measured or included in the development of the algorithm. Implementing the actual method in a real AGV would be a future work, where precision measurements can be performed. Furthermore, as the algorithm has not been implemented in the vehicle, it has also not been possible to measure the energy consumption generated and how much this factor would limit the navigation speed. This is also a future work to be carried out. Overall, for the simulation, a simplification of the algorithm is made for vehicles with mecanum wheels that fulfils the desired collision-free navigation.

References

1. Wang, T., Yan, X., Wang, Y., Wu, Q.: A distributed model predictive control using virtual field force for multi-ship collision avoidance under COLREGs. 2017 4th Int. Conf. Transp. Inf. Safety, ICTIS 2017 - Proc. 296–305 (2017).
2. Burgos, E., Bhandari, S.: Potential flow field navigation with virtual force field for UAS collision avoidance. 2016 Int. Conf. Unmanned Aircr. Syst. ICUAS 2016. 505–513 (2016).
3. Borenstein, J., Koren, Y.: The Vector Field Histogram - Fast obstacle avoidance for mobile robots. IEEE J. Robot. Autom. 7, 278–288 (1991).
4. 4. Ye, C.: Navigating a mobile robot by a traversability field histogram. IEEE Trans. Syst. Man, Cybern. Part B Cybern. 37, 361–372 (2007).
5. Hoffmann, G.M., Tomlin, C.J., Montemerlo, M., Thrun, S.: Autonomous automobile trajectory tracking for off-road driving: Controller design, experimental validation and racing. Proc. Am. Control Conf. 2296–2301 (2007).
6. Cabezas-Olivenza, M., Zulueta, E., Sanchez-Chica, A., Teso-fz-Betoño, A., Fernandez-Gamiz, U.: Dynamical Analysis of a Navigation Algorithm. 1–20 (2021).
7. AbdElmoniem, A., Osama, A., Abdelaziz, M., Maged, S.A.: A path-tracking algorithm using predictive Stanley lateral controller. Int. J. Adv. Robot. Syst. 17, 1–11 (2020).
8. Fox, D., Burgard, W., Thrun, S.: The dynamic window approach to collision avoidance. IEEE Robot. Autom. Mag. 4, 23–33 (1997).
9. Brock, O.: High-speed Navigation Using the Global Dynamic Window Approach. (1999).
10. Saranrittichai, P., Niparnan, N., Sudsang, A.: Robust local obstacle avoidance for mobile robot based on Dynamic Window approach. 2013 10th Int. Conf. Electr. Eng. Comput. Telecommun. Inf. Technol. ECTI-CON 2013. (2013).
11. Missura, M., Bennewitz, M.: Predictive collision avoidance for the dynamic window approach. Proc. - IEEE Int. Conf. Robot. Autom. 2019-May, 8620–8626 (2019).
12. Ögren, P., Leonard, N.E.: A convergent dynamic window approach to obstacle avoidance. IEEE Trans. Robot. 21, 188–195 (2005).
13. Liu, L.S., Lin, J.F., Yao, J.X., He, D.W., Zheng, J.S., Huang, J., Shi, P.: Path Planning for Smart Car Based on Dijkstra Algorithm and Dynamic Window Approach. Wirel. Commun. Mob. Comput. 2021, (2021).
14. Kashyap, A.K., Parhi, D.R., Muni, M.K., Pandey, K.K.: A hybrid technique for path planning of humanoid robot NAO in static and dynamic terrains. Appl. Soft Comput. J. 96, 106581 (2020).
15. Dobrevski, M., Skocaj, D.: Adaptive dynamic window approach for local navigation. IEEE Int. Conf. Intell. Robot. Syst. 6930–6936 (2020).
16. Chang, L., Shan, L., Jiang, C., Dai, Y.: Reinforcement based mobile robot path planning with improved dynamic window approach in unknown environment. Auton. Robots. 45, 51–76 (2021).

17. Xie, L., Henkel, C., Stol, K., Xu, W.: Power-minimization and energy-reduction autonomous navigation of an omnidirectional Mecanum robot via the dynamic window approach local trajectory planning: 15, (2018).
18. Xie, L., Scheifele, C., Xu, W., Stol, K.A.: Heavy-duty omni-directional Mecanum-wheeled robot for autonomous navigation: System development and simulation realization. Proc. - 2015 IEEE Int. Conf. Mechatronics, ICM 2015. 256–261 (2015).
19. Kim, J., Woo, S., Kim, J., Do, J., Kim, S., Bae, S.: Inertial navigation system for an automatic guided vehicle with Mecanum wheels. Int. J. Precis. Eng. Manuf. 2012 133. 13, 379–386 (2012).
20. Li, Y., Dai, S., Shi, Y., Zhao, L., Ding, M.: Navigation Simulation of a Mecanum Wheel Mobile Robot Based on an Improved A* Algorithm in Unity3D. Sensors 2019, Vol. 19, Page 2976. 19, 2976 (2019).
21. Piemngam, K., Nilkhamhang, I., Bunnun, P.: Development of Autonomous Mobile Robot Platform with Mecanum Wheels. 2019 1st Int. Symp. Instrumentation, Control. Artif. Intell. Robot. ICA-SYMP 2019. 90–93 (2019).

Chapter 23
Integrity of electronic transmitted images for forensic investigations

Robert Cristian Beloiu[0000−0002−5815−4008]

Abstract In this article it is analyzed the impact of sending image for forensic purposes considering various non-professional means available to laypersons. In day to day life, there are many situations when arises the need of investigations and the only available probe is an image taken at the scene and sent to the investigator. Some means are more reliable than others and some should be avoided when sensitive information is required, especially when this information is the base for serious consequences: trials, insurance policy coverage, fire investigation, fraud investigations, or even academic related issues. Images contain signature codes and GPS location information alongside with other parameters, besides the colored pixels. Some of the identification parameters are used in forensic authentication investigations. The receiver should be aware of the modifications that the transmission means do to the image. Besides the forensic specialist, other professionals should be aware of these influences. In my opinion, in any investigation of any kind that uses images as probes, a forensic specialist should be called to be sure the probe image could be used or not, especially in trials, insurance claims, and other similar situations.

23.1 Introduction

23.1.1 Image transmission necessity

A forensic investigator needs probes to be able to detect what was asked by the person or institution that asked for the investigation. At the same time, his/hers probes have to be solid and stand the questions raised by professionals or other people. There are many times when the probes are digital images provided professionally or not [1]. An accident passed and there was a witness that grabbed it in his/her mobile

RC Beloiu
University of Pitesti, Pitesti, 110040, Romania, e-mail: robertbeloiu@gmail.com

phone. The images can be provided via electronic transmission: email or social media. There is no other way to get access to those images. A fire started affecting the property or even worse, the health of some person. A witness provides images long after the fire was extinguished and any sign of it was removed. The witness stored some images of the event in his phone and can send them to the investigator. One patient was hospitalized. In the medical institution she noticed the use of some products that raised suspicion. She takes a picture with her phone and shares it with a friend. A while after, results that those products were not good and the medical institution denies its use in order to escape accusations and financial consequences. The image the patient took proves the facts. All of the above mentioned cases of image transmission require analysis [2]. There are many other situations when an investigator needs to analyze digital images taken by witnesses with the means at their reach at the time of the event. Many times, during these days, the mobile phone camera is the most used device to acquire images. Most of the times, the witnesses do not have forensic preparation and they do the best they can help clarify certain situation

23.2 Probes

A mobile phone is a device that many people own or have access to [3]. According to [4] in 2020 61,51% of population own a smart or feature phone, which makes the total number of devices to be 4,78 billion. In 201-2020 the number of smartphone grew by 40%. The estimations are that by 2023 there will be 7,33 billion mobile devices in use. According to [5] over half of the world's population own smartphones. The general public has become extremely dependent on smartphones. This means that a great number of people have access to mobile camera that can record events. Some of these recordings will be used by professional forensics to solve investigations. Most of these recordings will be transmitted using available means which are not in most cases professional.

23.2.1 Transmission ways

23.2.1.1 Forensic

When a professional forensic either individual or laboratory is faced with the situation of images or video transmission from one device to another, they are considering digital forensically tools. There is a variety of programs built for this purpose

23.2.1.2 Lay person

When a regular user transfer an image or movie acquired by his/hers phone they think at the means at their reach. Many times this involves free or open source tools: email, social media, file transfer, etc. Most of the time regular transference of images do not involve specialized nor professional hardware or software tools [6].

23.3 Probes

In most image enhancement and authentication courses, the discussion is about analyzing what you have assuming it was properly (I mean professionally) transmitted. The transmission process is covered in less situations. An image analysis deals many times with authenticity question: is this material genuine? The answer to this question is sometimes easy to give using dedicated software such as [7], [8] or other. The analysis, weather using professional or free software [9], [10], etc., starts with metadata analysis of EXIF format [11].

23.3.1 EXIF information

"The EXIF standard organizes metadata into five file directories (IFDs): (1) Primary; (2) Exif; (3) Interoperability; (4) Thumbnail and (5) GPS." [11]. EXIF contains information about date and time of recording and modification, camera parameters, camera brand, etc. [12]. The EXIF can be altered very easy and should be analyzed in combination with other information about the image. In order to check the integrity of an image, there are some parameter of interest in the EXIF list: MD5, SHA [13]. These codes are like signatures for images and in this article are used to check the integrity of transmitted images using different non-professional procedures.

23.3.1.1 Probes

For the experiments in this article were used three samples taken in close locations, as indicated in Table 1. The EXIF were collected using free tools such as Irfanview [9], Exif Pilot [14] as well as professional tools such as Amped Authenticate [15]. In Table 1 are indicated the test images with some of their EXIF information. In total, Amped Authenticate [15] identifies 89 EXIF parameters in an image. Besides EXIF a regular jpg image contains a lot more information used in image authentication [7].

Table 23.1 Probes images and EXIF information

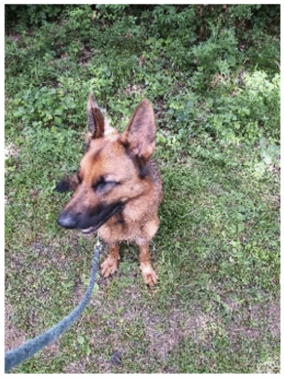


Figure 1. Probe 20200617_165425

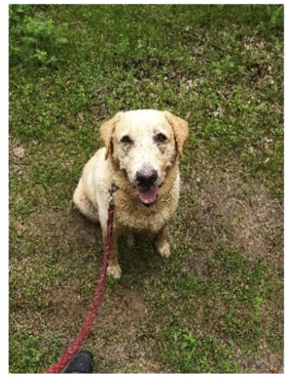


Figure 2. Probe 20200617_165432

23.3.2 Transmission means

In practice, there are various ways of communications between a sender and a receiver analyzed in this paper:

- Online tools [16]
  - Instagram. Instagram is a social media tools for image and video sharing [17]. More than 70% of people between 12 and 24 years old are Instagram users [18].
  - Whatsapp. Whatsapp is an instant communication application that was downloaded and installed on almost 120 billion Android OS based mobile phones and the number is increasing [19]. this makes it one of the most popular applications on mobile phones.
  - Skype. Skype [20] is one of the first applications that facilitates communication between people over the Internet.
  - Moodle. Moodle is an application that helps online teaching through many facilities it offers [21]. Students can share images among themselves and/or with the academic staff for several reasons. One of them is to check their location while assisting to a lecture, especially during COVID-19 restrictions.
  - Email
    - Gmail [22]/Yahoo!Mail [23]. These are one of the most common free email providers used by people around the world.
    - Proprietary email. It was tested with a public educational and private company emails.

– WiFi File transfer [24]. WiFi File Transfer is an application that allows the users to transfer files from phone or tablet to PC over a wireless connection. It has about 60.000 users.
– Google drive is a free solution offered by Google to share and store digital information online. It is used by about 1 billion users. o Dropbox is Microsoft solution to online storage. It has 700.000 registered users [25].

- Offline tools
  – DVD. Nowadays this media storage is less used and available on the market. Nevertheless, it is still used when transmitting sensitive information that it is not supposed to be lost.
  – USB. This is the wider share offline transmitting media. It is easy to use and cheap. Nevertheless, accidents can happen and information can be lost very easy.
  – Image integration
    · MS word. This is one of the widest used document processor [26]. There are very few computers running Windows OS that does not have installed MS Office package. This word processor it is used to incorporate images for various reasons. The documents are either saved as docx or are transformed in a different type.
    · Pdf. The pdf files can be exchanged between different computers with confidence of maintaining document formatting. By far, the pdf format is the most common used to exchange documents between users of all kind (private or public). Most of software computer programs, regardless of their application domain, have the ability of saving or exporting their results in pdf format. Therefore, documents that include images transformed in pdf are (sometimes) considered proofs of authenticity and a safe way to transmit images.

23.4 Discussion

For this work the reference image was acquired from the camera (both Samsung SM-A920F and Sony DSC-HX350) with Amped software. Then these images were sent using the indicated procedures. Table 2 shows the probes' signatures MD5, SHA1, SHA256, SHA384 and SHA512. It is not the purpose of this work to detail how these signatures are generated. Therefore, this topic is not presented. The results show that the online tools: Instagram, Whatsapp, Skype modify the structure of the transmitted images. At the same time emailing and cloud services do not report changes in the image structure. The transmission with offline methods does not report changes in the image structure. Also image integration both in editable documents (ex. docx) as well as universal format (ex. pdf) reports image structure modifications. The results indicate that in spite of their popularity, the communication tools are not reliable as far as maintaining the image structure. They introduce modifications in the signature information as well as GPS location, among other parameters.

Table 23.2 EXIF image signatures

Table 23.3 Integrity analysis

Sending procedure	File format different parameters	File format deleted parameters	EXIF introduced parameters	EXIF different parameters	EXIF deleted parameters	Signature related changes
Instagram	15	3	4	2	79	yes
Whatsapp	13	4	4	2	79	Yes
Skype	8	3	1	1	17	yes
Moodle	0	0	0	0	0	no
Google gmail / Yahoo mail	0	0	0	0	0	no
Google gmail 7zip	0	0	0	0	0	no
Educational proprietary email	0	0	0	0	0	no
Private proprietary email	0	0	0	0	0	no
WiFi file transfer	0	0	0	0	0	no
Cloud GoogleDrive / MS Dropbox	0	0	0	0	0	no
DVD / USB	0	0	0	0	0	no
MS Word	16	4	4	2	79	yes
Acrobat pdf	14	4	4	2	79	yes

23.5 Conclusions

In spite of the fact that emailing services does not report image structure modifications, the recommendation is not to use them to send sensitive images. Image integration in editable documents as well as pdf introduce image structure modifications. Nevertheless, there are tools and means that an image included in a pdf can be further analyzed for authentication. The most reliable way to transmit sensitive images to be processed is DVD. The DVD has the advantage that once recorded, the information can be stored for a later investigation. Many ordinary computers nowadays do not come with DVD unit.

Some second reliable ways are USB and password protection archives. When password protection archives are used, they can be sent via email for future analysis. As people that need to process and make difficult decisions based on image analysis many times are not aware of image authentication procedures, they can make erroneous appreciations and sentences. These can have a very big impact in many lives and budgets. Therefore, not all means of image transmissions are reliable and advisable to be used for analysis of sensitive of events. It is advisable that when such events are analyzed a knowledgeable forensic specialist should be consulted. In this paper the original probe was extracted from the camera using dedicated forensic software. Image extraction from the camera is a process that has to be done following certified procedures.

Acknowledgments

I would like to thank to the Amped Software company for granting educational licenses for Amped FIVE and Amped Authenticate tools. Also, I would like to thank Intech Integration Technologies for granting access to both software as well as hardware resources.

References

1. S. Madhisetty, M.-A. Williams, J. Massy-Greene, L. Franco, and M. El Khoury, 'How to Manage Privacy in Photos after Publication':, in Proceedings of the 21st International Conference on Enterprise Information Systems, Heraklion, Crete, Greece, 2019, pp. 162–168. doi: 10.5220/0007614001620168.
2. Y. Xiong, Y. Zhang, D. Wang, and S. Feng, 'Picture or it didn't happen: catch the truth for events', Multimed. Tools Appl., vol. 76, no. 14, pp. 15681–15706, Jul. 2017, doi: 10.1007/s11042-016-3864-6.
3. C. Montag et al., 'Smartphone usage in the 21st century: who is active on WhatsApp?', BMC Res. Notes, vol. 8, no. 1, p. 331, Dec. 2015, doi: 10.1186/s13104-015-1280-z.
4. A. Turner, 'HOW MANY SMARTPHONES ARE IN THE WORLD? AUGUST 2020 MOBILE USER STATISTICS: DISCOVER THE NUMBER OF PHONES IN THE WORLD & SMARTPHONE PENETRATION BY COUNTRY OR REGION.' https://www.bankmycell.com/blog/how-many-phones-are-in-the-world (accessed Aug. 22, 2020).
5. '2017 Global Internet Use Survey Report: Internet, community, mobile phones are our daily life', Feb. 23, 2017. https://www.smartm.com.tw/article/33313939cea3 (accessed May 11, 2022).
6. T. L. Wagner and A. Blewer, '"The Word Real Is No Longer Real": Deepfakes, Gender, and the Challenges of AI-Altered Video', Open Inf. Sci., vol. 3, no. 1, pp. 32–46, Jan. 2019, doi: 10.1515/opis-2019-0003.
7. 'Amped Blog, Forensic Image and Video Processing'. https://blog.ampedsoftware.com/
8. 'Foclar'. https://www.foclar.com/articles.html
9. I. Skiljan, 'IrfanView'. https://www.irfanview.com/
10. P. Harvey, 'ExifTool by Phil Harvey, Read, Write and Edit Meta Information!', Aug. 10, 2020. https://exiftool.org/
11. H. Farid, Photo Forensics. England: Massachutes Institute of Technology, 2016.
12. X. Sun, Y. Li, S. Niu, and Y. Huang, 'The detecting system of image forgeries with noise features and EXIF information', J. Syst. Sci. Complex., vol. 28, no. 5, pp. 1164–1176, Oct. 2015, doi: 10.1007/s11424-015-4023-2.
13. D. P. Gangwar and A. Pathania, 'Authentication of Digital Image using Exif Metadata and Decoding Properties', Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol., pp. 335–341, Dec. 2018, doi: 10.32628/CSEIT183815.
14. 'Exif Pilot'. https://www.colorpilot.com/exif.html (accessed Apr. 30, 2022).

15. ‘Amped Authenticate’. https://ampedsoftware.com/authenticate (accessed Apr. 30, 2022).
16. E. Raad, B. Al Bouna, and R. Chbeir, ‘Preventing sensitive relationships disclosure for better social media preservation’, Int. J. Inf. Secur., vol. 15, no. 2, pp. 173–194, Apr. 2016, doi: 10.1007/s10207-015-0278-9.
17. J. P. Carpenter, S. A. Morrison, M. Craft, and M. Lee, ‘How and why are educators using Instagram?’, Teach. Teach. Educ., vol. 96, p. 103149, Nov. 2020, doi: 10.1016/j.tate.2020.103149.
18. Y.-T. Huang and S.-F. Su, ‘Motives for Instagram Use and Topics of Interest among Young Adults’, Future Internet, vol. 10, no. 8, p. 77, Aug. 2018, doi: 10.3390/fi10080077.
19. ‘WhatsApp’, https://play.google.com/store/apps/details?id=com.whatsapp.
20. M. Matulin, Š. Mrvelj, B. Abramović, T. Šoštarić, and M. Čejvan, ‘User Quality of Experience Comparison Between Skype, Microsoft Teams and Zoom Videoconferencing Tools’, in Future Access Enablers for Ubiquitous and Intelligent Infrastructures, vol. 382, D. Perakovic and L. Knapcikova, Eds. Cham: Springer International Publishing, 2021, pp. 299–307. doi: 10.1007/978-3-030-78459-1_22.
21. J. Pereira, ‘Motivating users to online participation. A practice- based comparison between Moodle Forums and Telegram groups’, p. 9.
22. C. Petrov, ‘52 Gmail Statistics To Show How Big It Is In 2022’, Apr. 26, 2022. https://techjury.net/blog/gmail-statistics/ (accessed Apr. 30, 2022).
23. ‘Yahoo! Mail’. https://en.wikipedia.org/wiki/Yahoo!_Mail (accessed Apr. 30, 2022).
24. ‘WiFi File Transfer’. https://play.google.com/store/apps/details?id=com.smarterdroid.wififiletransfer&gl=ES (accessed Apr. 30, 2022).
25. B. Dean, ‘Dropbox Usage and Revenue Stats’, Apr. 23, 2021. https://backlinko.com/dropbox-users (accessed Apr. 30, 2022).
26. V. Koutsokostas et al., ‘Invoice #31415 attached: Automated analysis of malicious Microsoft Office documents’, Comput. Secur., vol. 114, p. 102582, Mar. 2022, doi: 10.1016/j.cose.2021.102582.

Chapter 24
Estimation of dynamics of a piezoelectric wind energy harvester via neural network

Erol Kurt, Jose Manuel Lopez-Guede, Mehmet Tekerek, Mehmet Gök and Yunus Uzun

Abstract A new classification tool has been designed and implemented for the identification of dynamic responses of a recently built wind energy harvester, namely piezoelectric wind energy harvester (PWEH). The classification tool indeed uses an artificial neural network algorithm in order to determine movement and acceleration of the tip of the piezoelectric layer. For this purpose, it trains time series data of the displacement and velocity of the piezoelectric layer tips and processes them into a two-layer neural network for a given set of excitation frequency and amplitude. According to one of our previous work in Ref. [1], it was proven that such a system creates different types of dynamic responses from periodic to hyper chaotic ones. All those dynamic responses yield to different power generation for the harvester naturally; thereby an estimation of the dynamic character (i.e. vibration motion) of the piezoelectric layer is vital task to estimate the output power of the harvester. Following the network training procedure, it has been proven that the classification tool performs with a high accuracy compared to the experimental findings, thereby we recommend this neural network scheme to estimate such piezoelectric-based

Erol Kurt
Department of Electrical and Electronics Engineering, Technology Faculty, Gazi University, Ankara, Turkey

Jose Manuel Lopez-Guede
Systems Engineering and Automatic Control Department, Faculty of Engineering Vitoria-Gasteiz, University of the Basque Country (UPV/EHU), Vitoria, Spain

Mehmet Tekerek
Kahramanmaraş Sütçüimam University, Computer Education and Instructional Technology Department,

Mehmet Tekerek
Kahramanmaraş, Turkey d Kahramanmaraş İstiklal University, Kahramanmaraş, Turkey,

Yunus Uzun
Department of Electrical and Electronics Engineering, Engineering Faculty, Aksaray University, Aksaray, Turkey

harvesters for wide engineering applications including vibration fed transducers and sensors.

24.1 Introduction

The limitation over the energy resources and day-by-day increase in energy demand motivate the engineers and basic scientists to design and create new and more effective energy harvesting systems [1] [2]. These systems mainly focus to meet the power requirements of the small-scale electro-mechanical devices such as sensor networks, actuators and systems operating at milliwatt or nanowatt levels. Because the harvesters make it possible to extend the lifecycle of these kind of systems by contributing to power of the batteries [3]. For some systems, harvesters even enable the systems to operate without any battery, therein it presents regenerative and sustainability operations [4] [5]. On the one hand, the batteries for low powered portable electronic devices can have longer operation life and have less repetitive charging intervals by the help of the harvesting technologies. By looking at the literature, there have been many types of energy harvesters getting its operation power from mechanical vibrations [6], solar lights [7] and wind-based vibrations [8] . In the case of vibration-based ones, the mechanical energy produced by the ambient vibrations is converted into a useable electric energy via a piezoelectric layer [4] [5] [9] [10]. However, the burden for such mechanical excitation is that the layers may damage in a short time due to mechanical contact. However, in one of our previous study, we have fixed that problem by applying magnetic forces to the tip of the layers by eliminating the mechanical contact [11] [12]. Thus, the vibration-based piezoelectric harvesters have the highest power density and they exhibit much longer life span due to contactless work by the help of magnetic field [11] [12]. A piezoelectric layer has a high efficiency at its natural frequency [13] [14]. For this reason, it is the most important issue to convert the frequencies of natural vibrations of ambient to the resonance frequency of the layer in order to obtain the maximum power. Indeed, the random vibrations having different frequencies cause loss of power and decrease the efficiency. The wide band spectrum of the ambient vibrations gives us chance for the explorations of distinctive feature of the pendulum-like layers by changing the magnetic force. By doing so, one can clarify the dynamics of energy harvesters under different frequency and amplitude for the aim of having higher energy values from the magnetic media.

In this study, it is focused on the classification of chaotic data obtained from the previously performed wind energy harvester with the Artificial neural networks (ANN). The novel classification tool for the evaluation of the dynamic responses of wind energy harvester is implemented with the artificial neural network. As far as we know, the first study in the literature for classification of the chaotic statues of the dynamic system.

24.2 Theoretical background

The PWEH has 4 units: A freely rotating propeller for flowing wind. The shaft unit transmitting the rotational motion to the magnet located on the shaft end, a piezoelectric unit generating voltage. Lastly, an output port regulating the harvested electrical power as shown in Fig. 1. The other physical properties and important parameters of the harvester can also be learned from our previous work [1].

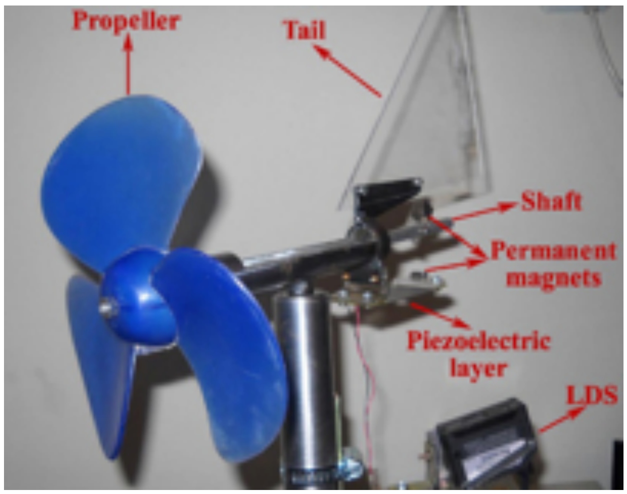


Fig. 24.1 The piezoelectric wind energy harvester (PWEH).

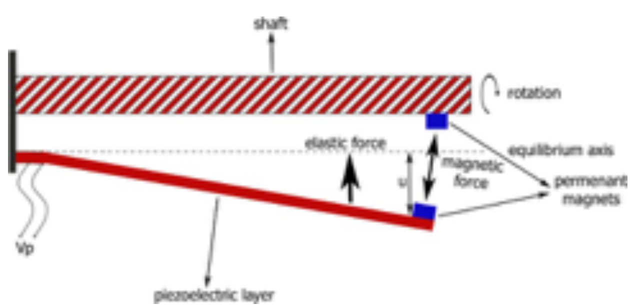


Fig. 24.2 The operation mechanism of the wind harvester under elastic and magnetic forces.

The piezoelectric layer in Figure 2 can be modelled as in Fig. 3. In this model, a ferromagnetic knob attached to the PZT is activated with the effect of the saw-teeth-like magnetic force near the tip, which is generated by the electromagnet. The pendulum-like beam made up of the elastic force and magnetic force as in Ref. [15].

Note that if the repulsive magnetic effect is much larger than the gravitational force, one can eliminate gravitational effect as in [15] [16]. Indeed, if elastic forces are strong enough, such an effect do not make obvious change in the motion in many systems [15]- [17].

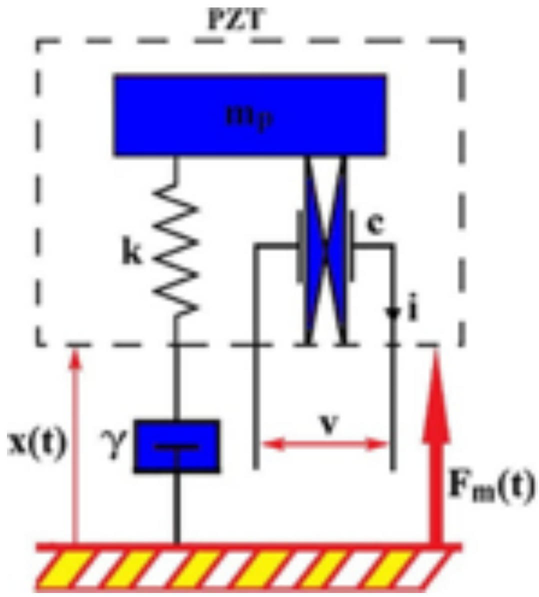


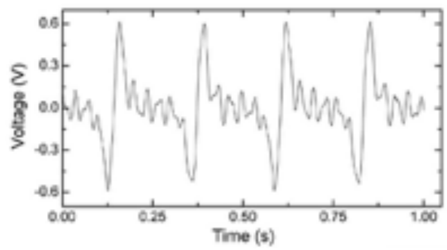
Fig. 24.3 The electromechanical modelling of the PWEH.

24.2.1 The vibration analyses of the piezoelectric layer

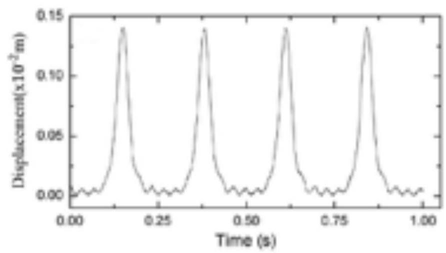
In order to describe the experimental vibration characteristics of a single piezoelectric layer under a periodic magnetic field, a specific test setup is organized. This will assist us to learn characteristic responses of the layer and classify them correctly for the ANN algorithm. In this respect, the test setup consists of a ferromagnetic beam with a piezoelectric layer, the magnetic excitation unit, AC/ DC converter, a power circuit, a signal generator, a laser displacement sensor (LDS) and a data acquisition/monitoring unit as seen in Fig. 4.

Fig. 24.4 The experimental test setup for the piezoelectric layer dynamics as a teaching tool for ANN.

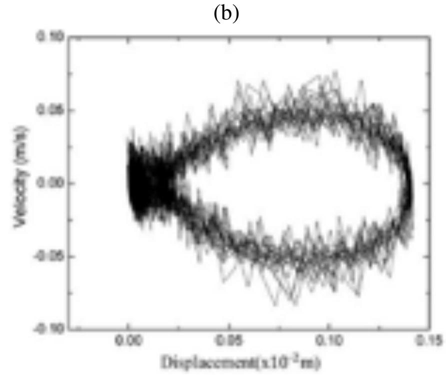
The experimental data obtained from the test setup in Fig. 4 is presented in Figs. 5 and 6. The figures illustrates voltage, displacement and velocity waveforms of the layer. Note that such responses have been received for 3.5 m/s wind speed from the PWEH, too. In order to classify the dynamics, the main criteria are the type of attractors as shown in Figs. 5(c) and 6(c). Note that we have a rich attractor response for different motion types. Thus, the test setup has led to get an overall dynamic perspective on the piezoelectric motion under periodic magnetic field.



(a)



(b)



(c)

Fig. 24.5 Experimental data which is generated for 3.5 m/s wind speed.

The generated voltage waveforms have been explored during the present work. It is obvious that the present system can be used to analyze the dynamical vibration characteristics of the pendulum like piezoelectric layers under different magnetic excitation frequencies and the magnetic force strength parallel with the results of

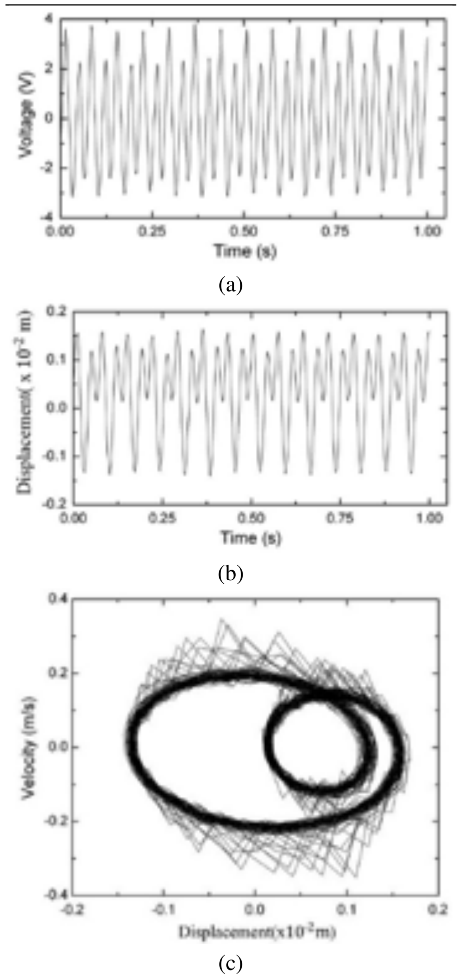


Fig. 24.6 Experimental data which is generated for 8 m/s wind speed.

PWEH. The displacement occurs in the system with the effect of the magnetic force generated by the applied periodic signal, which creates voltage and current on the load.

24.3 Artificial neural networks

The ordinary estimation algorithms use the mathematical programs systematically for obtaining the desired results with the iteration method. However, it is a reality that sometimes this kind of systematic approach fail for solving the problem and estimation could not work. This reality gives the idea that pure mathematical model of the system cannot exhibit the real reply of the system. In that point, artificial Neural Networks (ANNs) give the opportunity for solving the nonlinear problems with sufficiently high accuracy (Fig. 7). With the help of the training process, any system can be modelled and trained network represents the desired system with high precision. Generally, an ANN consists of three types of layers, the input layer where external information is applied to the network, hidden layers that are capable of processing information and an output layer that exports the neural network's decisions. The number of the layers of the ANN and the number of the neurons in the layers, the type of the activation function are usually application dependent [18] [19]. Artificial neural networks are an auxiliary tool for classification problems and have been successfully applied in many classification problems in industrial and scientific fields. In this study, "Back Propagation Network" is used as artificial neural network [20] [21] [22].

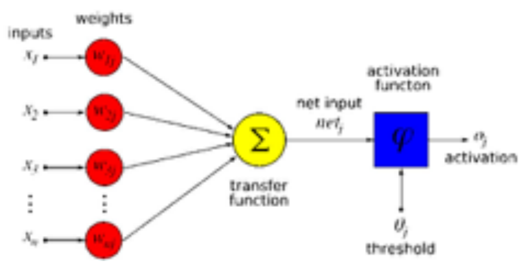


Fig. 24.7 The general structure of an artificial nerve cell.

24.4 Training procedure and the results of ANN Experiments

In this part of the paper, the process of the classification of the attractors via the back propagation neural network are presented rather than visually classifying. The excitation voltages and frequencies are used for the training of the ANN. The data set used for the ANN training is reapplied for the verification process of the system at the experimental study (Table 1). RMS-D and THD-V values are used as a third input variable respectively. When RMS-D is used as the third feature, the achieved success rate of the ANN system is 81.05 % for the classification of the chaotic attractors. It is seen that the correct classification could not be made in any way for case 6 (triple scroll chaotic) since the number of the data in the training set is low. However, it is clearly seen that the classification rate of the 3rd third case (double scroll chaotic) is higher than the 6th case since due to the large amount of data.

Table 24.1 Sample assessments of classification

Frequency value of excitation voltage (F)	Excitation voltage (U)	RMS value of displacement values (RMS-D)	THD values of velocity values (THD-V)	Class
190	30	0.1542	0.7313	multiple scroll chaotic
190	90	0.0971	0.2528	multiple scroll chaotic
190	2	0.0731	0.1011	periodic regular
190	4	0.0736	0.1041	periodic regular
190	8	0.0747	1.6042	periodic regular
190	12	0.0761	0.1093	periodic regular
190	20	0.0804	0.1079	periodic regular
200	25	0.1493	0.0734	multiple scroll chaotic
200	90	0.1681	0.2025	multiple scroll chaotic
200	16	0.0891	0.1172	quasi periodic
200	50	0.0877	0.2151	double scroll chaotic
200	60	0.0928	0.1008	double scroll chaotic

24.5 Results and discussion

It is obvious that ANN based classification method is successful. It can predict the dynamical motion types with a high accuracy (Figs. 8 & 9). This ANN tool provides a great advantage in terms of prediction without time consuming and experimental setup construction. Especially the chaotic regimes occur for similar wind speed (i.e. angular speed of rotor) and field strengths. Quasi-periodic region appears at lower speeds as in the experiments.

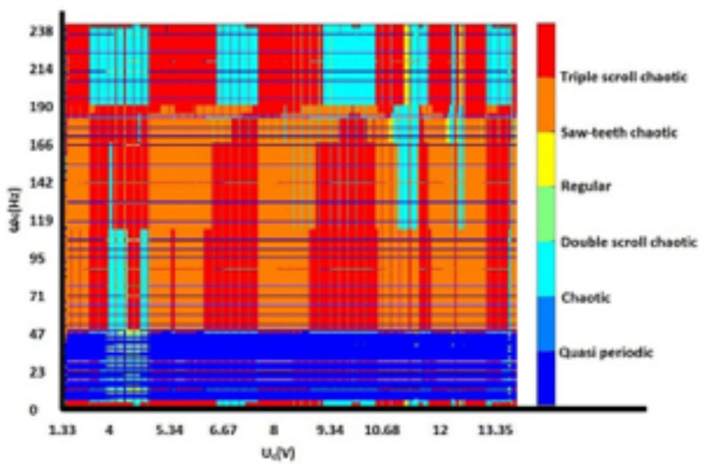


Fig. 24.8 The ANN dynamic classification graph on the plane of magnetic field strength equivalent to the applied voltage to the electromagnet and angular speed of rotor.

24.6 Conclusions

In this paper, a novel dynamic classification tool has been designed and implemented for a new built PWEH. For this, the dynamic parameters have been scanned in a

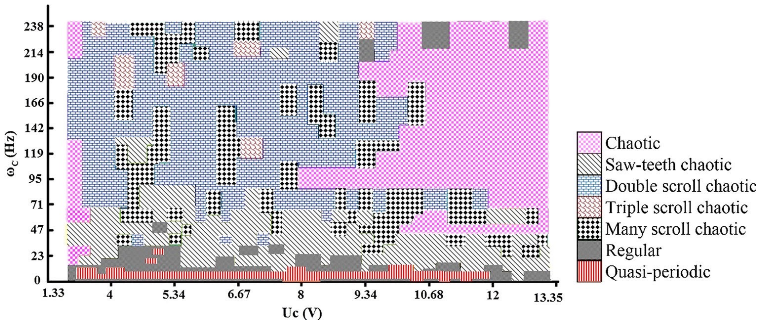


Fig. 24.9 The experimental dynamic classification graph on the plane of magnetic field strength equivalent to the applied voltage to the electromagnet and angular speed of rotor.

broad area. The displacement and velocity values of the piezoelectric tip have been recorded and analyzed. The attractors have been formed for many different motion types. The information of the system is processed in the designed 2-layer Back Propagation Network. The obtained results prove that the designed system can be used for the dynamic classification of the harvester with a good accuracy.

Acknowledgment

The authors are grateful to the European Union Ministry of Turkey, National Agency of Turkey, for the support of this project under the Project Code: 2015-1-TR01-KA203-021342 entitled Innovative European Studies on Renewable Energy Systems.

References

1. Çelik, K., Kurt, E., & Uzun, Y. (2017). Experimental and theoretical explorations on the buckling piezoelectric layer under magnetic excitation. Journal of Electronic Materials, 46(7), 4003-4016.
2. Bizon, N., Tabatabaei, N. M., Blaabjerg, F., & Kurt, E. (2017). Energy harvesting and energy efficiency. In Technology, Methods, and Applications, vol. 37 of Lecture Notes in Energy. Springer International Publishing.
3. Dondi, Denis, et al. "Modeling and optimization of a solar energy harvester system for self-powered wireless sensor networks." IEEE Transactions on industrial electronics 55.7 (2008): 2759-2766.
4. Ferrari, M., Ferrari, V., Guizzetti, M., Ando, B., Baglio, S., & Trigona, C. (2010). Improved energy harvesting from wideband vibrations by nonlinear piezoelectric converters. Sensors and Actuators A: Physical, 162(2), 425-431.

5. Cook-Chennault, K. A., Thambi, N., & Sastry, A. M. (2008). Powering MEMS portable devices—a review of non-regenerative and regenerative power supply systems with special emphasis on piezoelectric energy harvesting systems. Smart materials and structures.
6. Sun, Shilong, and W. Tse Peter. "Modeling of a horizontal asymmetric U-shaped vibration-based piezoelectric energy harvester (U-VPEH)." Mechanical Systems and Signal Processing 114 (2019): 467-485.
7. Saha, C. R., et al. "Photovoltaic (PV) and thermo-electric energy harvesters for charging applications." Microelectronics Journal 96 (2020): 104685.
8. Hu, Gang, et al. "Experimental investigation on the efficiency of circular cylinder-based wind energy harvester with different rod-shaped attachments." Applied energy 226 (2018): 682-689.
9. Mitcheson, P. D., Yeatman, E. M., Rao, G. K., Holmes, A. S., & Green, T. C. (2008). Energy harvesting from human and machine motion for wireless electronic devices. Proceedings of the IEEE, 96(9), 1457-1486.
10. James, E. P., Tudor, M. J., Beeby, S. P., Harris, N. R., Glynne-Jones, P., Ross, J. N., & White, N. M. (2004). An investigation of self-powered systems for condition monitoring applications. Sensors and Actuators A: Physical, 110(1-3), 171-176.
11. Uzun, Yunus, S. Demirbas, and E. Kurt. "Implementation of a new contactless piezoelectric wind energy harvester to a wireless weather station." Elektronika ir Elektrotechnika 20.10 (2014): 35-39.
12. Karami, M. Amin, Justin R. Farmer, and Daniel J. Inman. "Parametrically excited nonlinear piezoelectric compact wind turbine." Renewable energy 50 (2013): 977-987.
13. UZUN, YUNUS; KURT, Erol. Implementation and modeling of a piezoelastic pendulum under a harmonic magnetic excitation. In: 11. Int. Conf. Applications of Electrical Engineering. 2012. p. 184-189.
14. Mateu, Loreto, and Francesc Moll. "Review of energy harvesting techniques and applications for microelectronics." VLSI Circuits and Systems II. Vol. 5837. International Society for Optics and Photonics, 2005.
15. Kurt, E. (2006). Nonlinear responses of a magnetoelastic beam in a step-pulsed magnetic field. Nonlinear Dynamics, 45(1), 171-182.
16. E. Kurt, R. Kasap, S. Acar, Effects of periodic magnetic field to the dynamicsof vibrating beam, Mathematical and Computational Applications Journal.
17. CELIK, Emre, et al. A neural network design for the estimation of nonlinear behavior of a magnetically-excited piezoelectric harvester. Journal of Electronic Materials, 2018, 47.8: 4412-4420.
18. Huang, Guang-Bin, and Haroon A. Babri. "Upper bounds on the number of hidden neurons in feedforward networks with arbitrary bounded nonlinear activation functions." IEEE transactions on neural networks 9.1 (1998): 224-229.
19. Karlik, Bekir, and A. Vehbi Olgac. "Performance analysis of various activation functions in generalized MLP architectures of neural networks." International Journal of Artificial Intelligence and Expert Systems 1.4 (2011): 111-122.
20. LeCun, Yann, et al. "Handwritten digit recognition with a back-propagation network." Advances in neural information processing systems. 1990. Hecht-Nielsen,

Robert. "Theory of the backpropagation neural network." Neural networks for perception. Academic Press, 1992. 65-93.
21. Nekovei, Reza, and Ying Sun. "Back-propagation network and its configuration for blood vessel detection in angiograms." IEEE Transactions on Neural Networks 6.1 (1995): 64-72.

Chapter 25
Haptic Technology: A Valuable Paradigm to Reinforce Learning Processes in Special Education*

Eloy Irigoyen 0000−0002−9119−1652, Mikel Larrea0000−0002−9223−9355 and Manuel Graña0000−0001−7373−4097

Abstract At present, it can be seen that all kinds of technology are gradually being incorporated to help human beings in the performance of their frequent and daily tasks. Technological tools provide means to facilitate, improve, help and support any activity that, in itself, can be complicated, either because of its difficulty in performing it, or because the person who performs it has limited capabilities due to not being prepared the environment to their abilities and skills. In this sense, Special Education is currently receiving special attention from researchers, especially that aimed at minors. Although the proposals that have appeared recently provide perfect solutions to improve educational processes with ingenious didactic applications, there remains an interesting line of study on how to interact with these people in said educational processes. In this sense, haptic technologies can play a very important role, complementing all these solutions by incorporating tools to facilitate direct contact with people, especially in its initial phases, which are the most complicated. This paper aims to offer a panoramic view of haptic technologies and how they present solutions that can be integrated into the aforementioned educational processes.

Eloy Irigoyen, Mikel Larrea
University of the Basque Country (UPV/EHU), Systems Engineering and Automatic Control Department, 48013, Bilbao, Spain
e-mail: {eloy.irigoyen,m.larrea}@ehu.eus

Manuel Graña
University of the Basque Country (UPV/EHU), Computer Science and Artificial Intelligence Department, 20018, Donostia-San Sebastián, Spain
e-mail: manuel.grana@ehu,eus

* Supported by EUROPEAN COMMISSION, H2020-MSCA-RISE-2017: Under the project *Cyber-Physical Systems for PEdagogivcal Rehabilitation in Special EDucation*.

25.1 Introduction

The efforts that are being made today to improve the conditions of educational processes in multiple scientific areas, from Pedagogy to Engineering, through Psychology, Art, Medicine or Electronics, make this field very productive in very different solutions [1, 2]. Proof of all this are projects like CybSPEED "Cyber-Physical Systems for PEdagogivcal Rehabilitation in Special EDucation" [3], European project funded by the European Commission in the framework of the call for projects Marie Skłodowska-Curie Research and Innovation Staff Exchange, where the main objective is to advance a novel framework for analysis, modelling, synthesis and implementation of Cyber-Physical Systems for pedagogical rehabilitation in special education, based on a combination of brain-aware robotics, cognitive biometrics, computational intelligence and reasoning in humanoid and non-humanoid robots for education.

In this type of project, a good group of researchers from various universities, research centers and directly interested entities meet to study and develop solutions that facilitate, improve, help and support educational activities, especially those aimed at people with difficulties in integrating into traditional educational processes, as the educational environment is usually not prepared for their abilities and skills. Although the solutions provided in this type of project help to improve technologically and procedurally the means and methods of teaching, it is not easy to find proposals that focus on the process of interaction between people and the media, from a basically physical level. In other words, the nature of the human being when contacting the elements of his environment (devices, human beings, signals, etc.) is very special and in people with certain reduced capacities, it is even more so.

On the other hand, the progress that haptic technologies have experienced in recent times, with more innovative solutions, of reduced dimensions and working in fields of application unimaginable until now, opens up a new and promising line of study that aims to complement all the educational solutions discussed above, so that the interaction between the person and the help or support system can be facilitated, preventing it from being such an insurmountable barrier in many cases [4, 5].

For this reason, this paper aims to present a brief introduction to some of the new solutions that have appeared in the field of haptic technologies, offering a panoramic view where possible future lines of study for new projects and developments can be found. In the next section, the technological solutions that have been proposed to date will be presented. In this section, it will be found various areas in which haptic solutions have been developed. Subsequently, different fields of application where these technologies are currently used will be classified. Finally, as conclusion, the most relevant aspects presented in this paper will be gathered and listed.

25.2 Haptic Technologies

Haptic technology is located within the areas of sensorization and perception, but conceptually it has a much broader meaning, considering contact and interaction with the human being as its main basis. This discipline not only provides measurements of certain magnitudes related to tactile development, but also gives meaning to the type of contact established by intensity, duration, location, displacement and composition, among others. Next, some of the disciplines that have been studied in the field of haptic technology will be briefly presented.

25.2.1 Artificial Skin

The human being, through the skin, perceives much more information than what is commonly called "touch". That is, a person can discern between temperature, vibration, pressure (force), surface (area), to interact with the environment. Devices capable of transmitting or capturing different temperatures on the surface of the skin can provide information that could easily be used to perceive the degree of activity and confidence of the person when performing some type of educational task [6] . The combination of the pressure applyed on an object, together with the type of grip that can be measured with the vibration generated by the force exerted by the person, could be used to determine the degree of interaction and intention that a person has when dealing with some type of educational communication device [7]. Aspects such as the quality of the contact in the interaction of the person with the environment achieved through measurements of the contact surfaces used could improve the definition of the degree of decision making and the conviction with which a certain item is selected [8, 9, 10, 11].

25.2.2 Artificial Vision

The occurrence of new haptic devices based on vision can be transferred to educational processes, offering a new dimension of the sensation achieved in the person by educational tools designed for human-machine interaction in aforementioned processes [12, 13].

25.2.3 Artificial Hearing

New devices that reproduce the sense of hearing have recently been introduced into the world of haptic technologies. These systems can provide relevant information on aspects directly related to the attention and intention of interaction of the person with the educational tool. These disciplines, once again, come to complement the already existing developments, providing them with new variables that, correctly analyzed, can make more objective the choice of the level of communication with the person [14, 15].

25.2.4 Physical models

Every day there are more haptic devices that incorporate measurements of magnitudes that the human being perceives, on many occasions unconsciously. One of the most common examples would be the perception of certain chemical elements in the environment [6]. Smell is another of the human senses that provides information and translates into sensations of comfort, intention and location [16, 17]. On the other hand, the way of perceiving the flow of air currents is also being studied.This is another dimension that could be added to the analysis processes in human-machine interaction. In special education situations, everything that could be useful to achieve an environment of trust and comfort for students will result in better designs and, above all, in efficient, applicable and effective solutions [18].

25.3 Haptic Applications

Studies and developments with haptic solutions have been in vogue for a long time. But new technologies, together with the ability to miniaturize and make the designed elements more flexible, mean that this technology continues to be of great interest among researchers from very diverse areas. Precisely, thanks to these new advances, the proposals that have appeared are especially interesting in the field of Special Education. Taking into account all the proposals that have appeared in recent times, a good collection of ideas can be taken to transfer to learning systems, where interaction with students is crucial, especially in people with reduced abilities or for whom the educational environment is not adapted. Next, some of the fields in which new and original haptic solutions are being incorporated to various problems that require human interaction with the environment will be presented.

25.3.1 Medicine

The field of Medicine is one of the most prolific in haptic devices, both at an educational and training level, as well as at a basically operational level. Regarding educational development, there are multiple areas where this type of haptic solutions can be found: Surgery, Cardiology, Oncology, Diagnosis, etc. [19, 11, 20]. With respect to functional solutions in any of these areas, there is also a continuous emergence of proposals based on haptic devices [21, 7, 22].

25.3.2 Games and Gamification

The world of games on digital platforms, supported by virtual reality and augmented reality, has grown enormously in recent years. New computing and visualization technologies have opened the door to the generation of new and spectacular interaction devices with the gaming environment [8, 19, 23]. On the other hand, this advance in the gaming world made many researchers in the educational field look at devices and procedures to integrate them into new educational plans. Due to this, many gamification proposals are known that incorporate these haptic technologies into their solutions [24, 25]. This allows to propose a new line of study where all these technologies come to reinforce the learning processes in Special Education.

25.3.3 Services

Nowadays, the environment, in which the human being develops, is full of tools that help him to decide and interact from haptic devices. It is common to find many solutions supported by these technologies and that are transparent to users [26]. Based on this concept, efforts should be made to develop non-invasive or intrusive solutions that allow improving educational processes that are sensitive to people with reduced capabilities.

25.3.4 Rehabilitation

The innumerable proposals that have been appearing recently in the field of rehabilitation of people with mobility and communication problems offer haptic tools for the performance and improvement of the quality of life of these people. These devices manage to improve the functionality of the motor and sensory apparatus of these people, either by assisting them in the execution of the movements they must make, as well as in the perception of the elements of the environment [7, 9]. These developments present interesting solutions that can be transferred to the field of special education, by allowing learning processes to be complemented with devices and tools that bring people closer to a world of education, in a more friendly way.

25.3.5 Personal Assistance

Currently, there are many developments to give support and support to people with disabilities. Many of these solutions incorporate haptic technology, although it is mostly aimed at offering help in performing daily tasks [27, 15, 22]. In this sense, it would be of great interest that the haptic tools used in this field serve, in a complementary way, to identify the emotional state of the person, as well as the stress levels that can be reached when performing these tasks. This would be especially relevant if it could be transferred to learning processes in Special Education. Currently, it is not easy to find proposals and references that, based on haptic technology, present solutions for the improvement of said educational and formatives processes. Particularly, for children who have identified Autism Spectrum Disorder (ASD), to the knowledge of these researchers, there are hardly any reference works where haptic technologies have been incorporated into educational processes, which leaves open a very promising field of study.

25.4 Conclusions

In this work, certain ideas and concepts have been presented to encourage researchers in the study of how to incorporate new haptic technologies into teaching processes in the field of Special Education. A set of possible new technologies to be used and examples that, from various fields, can be analyzed to integrate them into the aforementioned educational processes have been briefly exposed. Haptic technology can facilitate, improve, help and support any educational activity. Especially, in children with special needs in the educational context, this type of technology can complement and improve the procedures, methods and tools that all these processes currently use.

25.4.0.1 Acknowledgements

Project 777720 supported by EUROPEAN COMMISSION, H2020-MSCA-RISE-2017. Project title: Cyber-Physical Systems for PEdagogical Rehabilitation in Special EDucation.

References

1. Ariza, Jonathan Á. and Baez, Heyson: Understanding the role of single-board computers in engineering and computer science education: A systematic literature review. Computer Applications in Engineering Education **1**(30), 304–329 (2022)

2. Almusawi, Hashem A. and Durugbo, Christopher M. and Bugawa, Afaf M.: Wearable Technology in Education: A Systematic Review. IEEE Transactions on Learning Technologies **4**(14), 540–554 (2021)
3. CybSPEED Homepage, http://www.ehu.eus/ccwintco/cybSPEED/. Last accessed 16 May 2022
4. Zhu, M. et al.: Soft, Wearable Robotics and Haptics: Technologies, Trends, and Emerging Applications. Proceedings of the IEEE **2**(110), 246–272 (2022)
5. Sanfilippo, F. et al.: A Perspective Review on Integrating VR/AR with Haptics into STEM Education for Multi-Sensory Learning. Robotics **2**(11), 2218–6581 (2022)
6. Lu, J., et al.: Chemical Haptics: Rendering Haptic Sensations via Topical Stimulants. In: Association for Computing Machinery (eds.) UIST '21, vol. 1, pp. 239–257. New York, NY, USA (2021). 10.1145/3472749.3474747
7. EPFL Artificial Skin Homepage, https://actu.epfl.ch/news/artificial-skin-could-help-rehabilitation-and-enha/. Last accessed 27 Sep 2019
8. "Inside Reality Labs Research: Bringing Touch to the Virtual World" homepage, https://about.fb.com/news/2021/11/reality-labs-haptic-gloves-research/. Last accessed 16 Nov 2021
9. "Electronic "Skin" Brings Sense of Touch and Pain to Prosthetic Hands" homepage, https://www.techbriefs.com/component/content/article/tb/pub/briefs/bio-medical/37624. Last accessed 1 Sep 2020
10. Piacenza, P., et al.: A Sensorized Multicurved Robot Finger With Data-Driven Touch Sensing via Overlapping Light Signals. IEEE/ASME Transactions on Mechatronics **5**(25), 2416–2427 (2020)
11. Schiatti, L., et al.: A Novel Wearable and Wireless Device to Investigate Perception in Interactive Scenario. Annu Int Conf IEEE Eng Med Biol Soc **1**(1), 3252–3255 (2020)
12. Romeo, K., et al.: The MAPS: Toward a Novel Mobility Assistance System for Visually Impaired People. Sensors **9**(22), 1424–8220 (2022)
13. Wan, C., Cai, P., Guo, X. et al.: An artificial sensory neuron with visual-haptic fusion. Nat Commun **11**, 4602 (2020). https://doi.org/10.1038/s41467-020-18375-y
14. "Teenage Engineering Rumble: a bolt-on haptic subwoofer for the OP-Z" homepage, https://www.gearnews.com/teenage-engineering-rumble-a-bolt-on-haptic-subwoofer-for-the-op-z/. Last accessed 12 Sep 2019
15. Sensory Substitution Homepage, https://eagleman.com/science/sensory-substitution/. Last accessed 17 May 2022
16. Spencer, B.S.: Incorporating the sense of smell into patient and haptic surgical simulators. IEEE Transactions on Information Technology in Biomedicine **1**(10), 168–173 (2006)
17. Chi Thanh Vi: Not just seeing, but also feeling art: Mid-air haptic experiences integrated in a multisensory art exhibition. International Journal of Human-Computer Studies **108**, 1071–5819 (2017)
18. Takei, k., et al.: Highly sensitive electronic whiskers based on patterned carbon nanotube and silver nanoparticle composite films. Proceedings of the National Academy of Sciences **5**(111), 1703–1707 (2014)
19. TESLASUIT Homepage, https://teslasuit.io/. Last accessed 17 May 2022
20. Felix G. HAMZA-LUP, Dorin M. POPOVICI, Crenguta M. BOGDAN: Haptic Feedback Systems in Medical Education. Journal of Advanced Distributed Learning Technology **1**(2), 7–16 (2013)
21. HeroSurg Homepage, https://www.deakin.edu.au/about-deakin/news-and-media-releases/articles/deakin-builds-robotic-surgical-system-with-sense-of-touch. Last accessed 13 Oct 2016
22. Oh, S., Jung, Y., Kim, S. et al.: Remote tactile sensing system integrated with magnetic synapse. Sci Rep **7**, 16963 (2017). https://doi.org/10.1038/s41598-017-17277-2
23. Hapticast Homepage, https://rss.com/community/categories/leisure/games/. Last accessed 17 May 2022
24. Conecta Homepage, https://conecta.tec.mx/es/noticias/ciudad-de-mexico/educacion/traves-del-sentido-del-tacto-buscan-mejorar-aprendizaje-de-fisica. Last accessed 13 Abr 2021

25. Gearnews Homepage, https://www.gearnews.com/teenage-engineering-rumble-a-bolt-on-haptic-subwoofer-for-the-op-z/. Last accessed 12 Sep 2019
26. Van Baelen, D., et al.: Design of a Haptic Feedback System for Flight Envelope Protection. Journal of Guidance, Control, and Dynamics **4**(43), 700–714 (20220)
27. WeWALK Homepage, https://wewalk.io/en/about/. Last accessed 20 Oct 2021

Chapter 26
The Lattice Computing Paradigm for Modeling Intelligence in Cyber-Physical Systems

Vassilis Kaburlasos, Christodoulos Lytridis, Christos Bazinas, Fotios Panagiotopoulos, Eleni Vrochidou, and George Papakostas

Abstract This article reviews the research contribution of beneficiary IHU (former EMaTTech) to the CybSPEED Action with emphasis on lattice computing (LC) models of intelligence for social robot applications toward improving cyber-physical system (CPS) interaction with humans. Work extensions are also described.

26.1 1. Introduction

The CybSPEED Action [1] aimed at creating an international and inter-sectorial network of organizations in order to perform research advancing a novel framework for analysis, modelling, synthesis and implementation of Cyber-Physical Systems (CPSs) for pedagogical rehabilitation in special education. The pursued research [2] was a combination of the best of experience and achievements of the participant institutions in the domains of robotic technology design, human-robot interaction with humanoid and non-humanoid robots in education, brain-inspired robotic and agent design, systems science, computational intelligence (CI) and Lattice Computing (LC) algorithms.

The participating individuals have been trained in the management of complex projects in multidisciplinary teams of quite diverse backgrounds. In particular, the career of the participating IHU researchers (secondees) was boosted towards further professional achievements also by participating as well as co-organizing several events in the context of the CybSPEED Action, including: network-wide synergy trainings organized locally by different beneficiaries, a workshop in Japan, a training course on LC in Morocco, a summer school in Kavala, and the concluding conference in San Sebastian.

all authors
HUMAIN-Lab, Department of Computer Science, International Hellenic University (IHU), 65404 Agios Loukas, Kavala, Greece Correspondence: vgkabs@teiemt.gr; Tel.: +30-2510-462-320

The contribution of beneficiary IHU (former EMaTTech) has been in the areas of human-robot interaction with humanoid robots in education, brain-inspired robotic and agent design, systems science, computational intelligence (CI) and, mostly, in modeling Cyber-Physical Systems (CPSs) based on the LC paradigm.

Note that a Cyber-Physical System (CPS) is defined as a device with both sensing and reasoning capacities. CPSs need to be driven by suitable models preferably with a capacity to deal effectively with both numerical and nonnumerical data, where numerical data are inherently associated with the physical" system component(s), whereas nonnumerical data are inherently associated with the cyber" system component(s). A typical method for dealing with nonnumerical data is by transforming them to numerical ones using ad-hoc" techniques thus risking a non-reversible performance deterioration. An innovation of beneficiary IHU in the context of the CybSPEED Action was the engagement of the Lattice Computing (LC) paradigm [3-8] for modelling based on numerical and/or nonnumerical data per se without transforming the latter data to the former data – Recall that LC has been defined as an evolving collection of tools and methodologies that process lattice ordered data per se including logic values, numbers, sets, symbols, graphs and other". An ensuing advantage is that LC enables computing with semantics [3], [9] instead of number crunching"; computing with semantics is sensible for human-robot interaction applications.

During the CybSPEED Action beneficiary IHU (former EMaTTech) has organized two special sessions on LC in established conferences [10], [11]. Furthermore, two established journal special issues have been compiled [12], [13].

Appendix A displays all publications with a contribution from IHU, with an explicit Acknowledgement to CybSPEED; note that an item in Appendix A is referred to below prefixed by an A, e.g. [A01]. Part of the publications in Appendix has been gold open access, the remaining had been either green open access or presentations. Furthermore, Appendix B displays the distribution of IHU publications in Appendix A over beneficiaries /partners. More specifically, a dot mark is shown in a Table cell when a co-author from a beneficiary /partner is included in the corresponding publication. The Table in Appendix B confirms that beneficiary 6 (i.e., IHU) has, obviously, participated in all 32 publications. Furthermore, approximately 1/3 of IHU's 32 publications, in particular 11 publications, were compiled in collaboration with partner 11 (i.e., UH2C, Morocco); approximately 1/3 of IHU's 32 publications, in particular 1+2+7+3= 13 publications, were compiled in collaboration with another four beneficiaries /partners, namely UPV/EHU, Theater Tsvete, IR-BAS and KyuTech, respectively; the remaining of IHU's 32 publications involved exclusively IHU researchers.

IHU's collaborating beneficiaries and partners are committed to sustain the partnership after the end of the project not only for the benefit of Europe but also for the benefit of other parts of the world including North/West Africa and Far-East Asia. For instance, two Erasmus+" agreements have already been contracted between IHU and UH2C (Morocco) [14], [15].

The layout of this work is as follows. Section 2 summarizes the mathematical background with emphasis on logic. Section 3 presents certain data /knowledge representations. Section 4 reviews computational experiments by logic-based models as well as other models. Section 5 concludes by summarizing IHU's contribution as well as future work extensions. Appendix A displays all the publications compiled by beneficiary IHU with an Acknowledgement to the CybSPEED Action. Appendix B displays the distribution of IHU publications over beneficiaries /partners.

26.2 Mathematical

Background This section summarizes lattice theory background, followed by a fuzzy lattice reasoning reminder. Useful mathematical instruments are presented with emphasis on logic [3].

26.2.1 Elementary Lattice Theory Definitions

Recall that a lattice is a partially ordered set (L,⊑) with the additional property that any two elements x,yL have both a greatest lower bound, namely infimum (meet), denoted by x⊓y and a least upper bound, namely supremum (join), denoted by x⊔y. A lattice is called complete if and only if all subsets have both a supremum and an infimum. Given a lattice (L,⊑), a valuation is a real function given by v(x⊓y) + v(x⊔y) = v(x) + v(y), x,yL. A valuation is called monotone if x⊑y v(x)v(y), and positive if x⊏y < v(x)<v(y). A positive valuation in (L,⊑) defines a metric distance given by d(x,y) = v(x⊔y) - v(x⊓y). A parametric valuation function v(.) introduces tunable nonlinearities.

26.2.2 Fuzzy Lattice Reasoning (FLR)

Let (L,⊑) be a lattice. An inclusion measure function : LL[0,1] is defined by the following two conditions:

$$\begin{aligned} &1. \quad u \sqsubseteq w \Leftrightarrow \sigma(u,w) = 1 \\ &2. \quad u \sqsubseteq w \Leftrightarrow \sigma(x,u) \leq \sigma(x,w) \end{aligned} \tag{1}$$

An inclusion measure : LL[0,1] can be interpreted as a fuzzy order relation; hence, the notations (u,w) and (u⊑w) are used interchangeably. Any use of an inclusion measure (.,.) is called Fuzzy Lattice Reasoning, or FLR for short, and it supports two different modes of reasoning, namely generalized modus ponens and reasoning by analogy. There are at least two different functions for defining an inclusion measure : LL[0,1], both are based on a positive valuation function in (L,⊑), as follows.

One more inclusion measure can be defined as explained in the following. It is well known that a probability space is a triplet (W,S,P), where W is an abstract space, S is a sv-algebra of subsets of W, and P is a probability measure on S. It turns out that S is a Boolean lattice. Moreover, P is a monotone valuation in lattice S, hence function d(A,B)= P(AB)-P(AB) defines a pseudometric. In addition, it is known that the set S of equivalence classes in S, defined by the equivalence relation d(A,B)=0, is a lattice; in conclusion, the probability measure P is a positive valuation in lattice (S,). The latter conclusion holds, in particular, when W is a finite set W= {c1,. . . ,cN}, S= 2W is the power set of W, and P: 2W[0,1] is a probability measure. Hence, the positive valuation function P in lattice (2W,) results in two inclusion measure functions (.,.) and (.,.) given by Eq. (2) and Eq. (3), respectively.

In a (non-void) complete lattice with least element o and greatest element i, the following two reasonable constraints" are assumed: first, ⊔(x,o) = 0 = ⊓(x,o) for all x⊐o and, second, d(x,i) < + for all x. The former constraint implies v(o) = 0, whereas the latter one implies v(i) < +. Inclusion measure can be extended to Cartesian products of lattices.

26.3 Data and Knowledge Representation

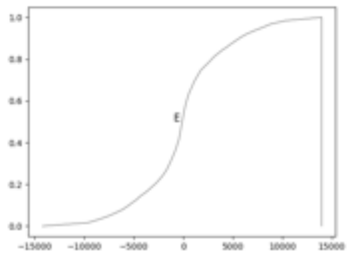
This section reviews selected (lattice-ordered) data and knowledge representations developed (or, simply, used) in the context of the LC paradigm during the CybSPEED Action. References are cited for the interested reader for further study.

26.3.1 Intervals' Numbers (INs)

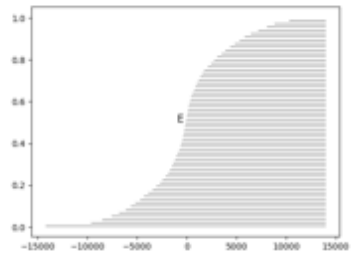
Fig.1 displays the two equivalent representations of an Intervals' Number (IN) E, namely the membership-function representation, and the interval representation. A number of applications involving INs have been presented in [A16], [A25], [16], [17], [18]. For instance, novel LC models, including lattice-ordered Interval Numbers (IN) have been engaged to fit cognitive biometrics signals recorded by brain-computer interfaces. A known practical advantage of an IN is its capacity to represent all order data statistics. A novelty of the CybSPEED Action is the engagement of INs for modelling cognitive biometrics activities including EEG signals.

26.3.2 Tree Data Structures

A tree data structure has been considered for representing a human face as illustrated in Fig.2. Fig.3 displays the corresponding (lattice-ordered) tree data structure. Details regarding the aforementioned tree data structures are presented in [A28].



(a)



(b)

Fig. 26.1 The two equivalent representations of an Intervals' Number (IN) E: (a) the membership-function representation, and (b) the interval representation.

26.3.3 Ontologies Autonomous

CPSs in agriculture have been considered for grape harvesting. Note that a grape is considered 100% mature when indices TSS and TA values are within certain ripeness (interval) ranges. If, in addition, index pH value is in a (interval) range hen a grape is of high quality; otherwise, a grape is of lower quality. The following four inequality constraints: a) TSSmin < TSS < TSSMAX, b) TAmin < TA < TAMAX, and c1) pH < pHMAX, and c2) pHmin < pH have been considered after the work in [19] by the following rule R0: If (TSS[TSSmin= 21, TSSMAX= 27]).and.(TA[TAmin= 4, TAMAX= 7]). and.(pH < pHMAX).and.(pHmin < pH) then (grape maturity is 1.00). The powerset 2{a,b,c1,c2} is a crisp Boolean lattice ontology, as shown in Fig.4, that includes all the combinations of the constraints a, b, c1 and c2. A subset in 2{a,b,c1,c2} is interpreted as a conjunction of inequalities, e.g. the set {a,b,c1,c2} is interpreted as a^b^c1^c2, etc.

26.4 Computational Experiments and Results

Special attention has been given to novel CPS models based on logic/reasoning. From a computational point of view the LC models are promising for CPS applications because they (a) can deal with disparate types of data (including both numerical data for the physical" system component(s) and non-numerical data for the cyber" system component(s)), (b) can compute with semantics, (c) can rigorously deal with ambiguity, (d) can naturally engage logic and reasoning and (e) can process data fast; furthermore, the employment of LC models in pedagogical context is innovative.

Logic-based LC models as well as other LC models, such as IN regressor models, implementable in software on CPSs, have been both developed and applied; furthermore, the corresponding results have been detailed in a number of publications as described in sections 1 and 3.

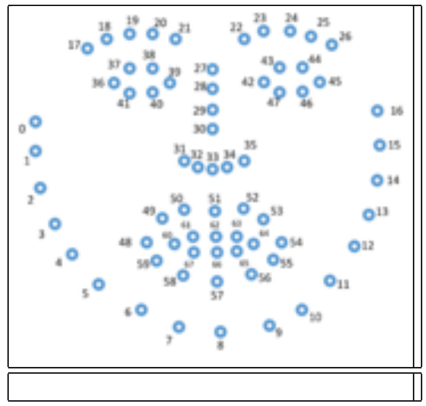


Fig. 26.2 (a) 68 facial landmark points on a human face. (b) The unit vector defined along the nose. The first three primary vectors are also shown to the centers of the eyes and mouth as well as secondary vectors from the left eye center.

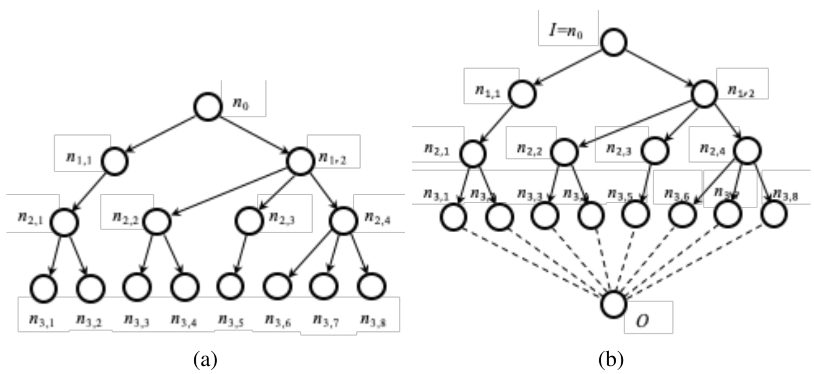


Fig. 26.3 (a) A tree data structure example. (b) A lattice results in by inserting an additional level including the least lattice element O. The tree root corresponds to the greatest lattice element I.

The following Algorithm 1 presents a generic FLR classifier model for training involving a single decision-making rule R0; whereas, Algorithm 2 presents a generic FLR classifier model for testing involving rule R0; the latter (rule) is the product of intervals in a Cartesian produce lattice L= L1LN, where L1,,LN are constituent lattices involving data of interest.

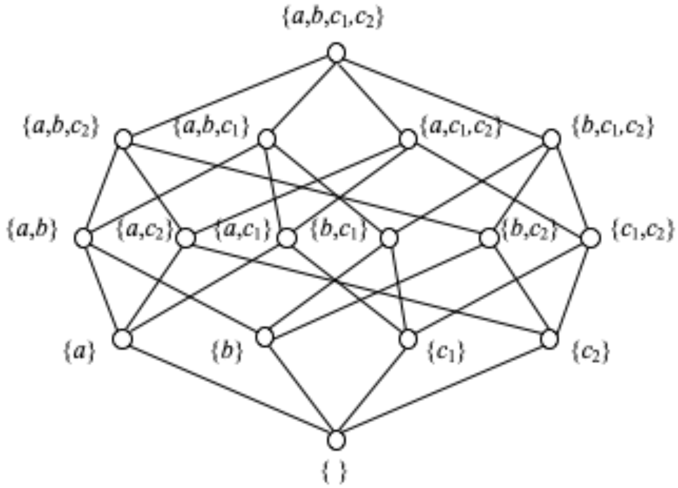


Fig. 26.4 A Boolean lattice ontology of the four inequality constraints a: TSSmin < TSS < TSS-MAX, b: TAmin < TA < TAMAX, c1: pH < pHMAX and c2: pHmin < pH, regarding grape maturity indices TSS, TA and pH.

Algorithm 1 A generic single rule FLR classifier model for training

---

0. Let n_{trn} be the number of the training data $(\bar{x}_i, r_i)$, $i\in\{1,\ldots,n_{trn}\}$, where $\bar{x}_i\in L_1\times\ldots\times L_N$, $r_i\in[0,1]$; let Q_{thres} be a user-defined threshold for the fitness function Q; let n_G be the maximum number of generations of the GENETIC algorithm; let R_0 be a single decision-making rule;
1. m = 1; Q ← Q_{thres} + 1;
2. **while** ($m < n_G$).and.(Q > Q_{thres}) **do**
3. Q ← 0;
   for i = 1 to n_{trn} **do**
4. Q ← Q + $|\sigma(\bar{x}_i \sqsubseteq R_0) - r_i|$;
5. **end for**
6. GENETIC optimization of positive valuation function parameters;
7. m ← m + 1;
8. **end while**

Algorithm 2 A generic single rule FLR classifier model for testing

---

0. Let n_{tst} be the number of the testing data $\bar{x}_i$, $i\in\{1,\ldots,n_{tst}\}$, where $\bar{x}_i\in L_1\times\ldots\times L_N$, $r_i\in[0,1]$; let R_0 be a decision-making rule;
1. Q ← 0;
2. **for** i = 1 to n_{tst} **do**
3. Maturity r_i equals $r_i = \sigma(\bar{x}_i \sqsubseteq R_0)$;
5. **end for**

26.5 Discussion and Conclusion

The CybSPEED Action [1] emphasized a similar to the intrinsic-motivation approach to learning by designing human-robot situations (games, pedagogical cases, and artistic performances) and advanced interfaces (brain-computer, eye-gaze tracking and virtual reality) where children and students interact with the novel technology to enhance the underlying self-compensation and complementarity of brain encoding during learning. Special attention was given to modelling Assistive CPSs in order to account for both levels of the human counterpart – as physical presence of the human and as social presence of the human, thus triggering different decision making algorithms.

This article has reviewed the scientific contribution of beneficiary IHU (former EMaTTech) of the CybSPEED Action [1] consortium especially in the context of the collaborative Work Package WP#2 entitled Research on Analysis, Modelling and Synthesis of CPSs for Pedagogical Rehabilitation in Education" [2]. More specifically, the Lattice Computing (LC) information processing paradigm has been proposed for Cyber-Physical System (CPS) modeling.

Mathematical lattice-ordered data representations, including INs, tree data structures and ontologies, have been considered in applications of interest to the CybSPEED project as explained in section 3. New partnerships have been developed between our institution, i.e. IHU, and other institutions especially from Morocco (UH2C), Bulgaria (IR-BAS and Theater Tsvete) and Japan (KyuTech). Special attention was given in applications of a humanoid robot, e.g. NAO, in (special) education [A09], [A10], [A11], [A13], [A23], [A24] as well as in theater [A14].

Of particular interest has been our (i.e. IHU's) collaboration with Theatre Tsvete introducing humanoid robots as actors in theatrical plays in order to implement innovative special education approaches. A large scale testing, including NAO robots in the theater, scheduled 8-9 April 2020, was cancelled due to the COVID-19 pandemic lockdown. Alternative tests with NAOs in Sofia, Bulgaria has revealed deficiencies of NAOs in the theater mainly due to the limited capacity on NAO in interacting with humans as well as in recognizing emotions. More future work needs to be carried out in the aforementioned directions. An employment of psychology's concept Gestalt" in innovative human-robot interaction applications was also investigated [20] but without resulting in any handy tools.

Our proposed LC models drew the attention of cutting-edge research, in particular toward composing music inspired by sculpture [21]. Furthermore, new projects have been accessed [22-25] with a potential for regional applications with robots along the Greek-Bulgarian border [26]. The proposed techniques have also been extended to CPSs in Agriculture 4.0 applications [27], [28]. Note that CPSs (robots) in Agriculture 4.0 applications are called Agrobots [29]. Teams of Agrobots have been proposed based on LC models [30] toward increasing effectiveness.

Acknowledgement

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 777720.

Appendix A

(List of IHU publications with an acknowledgement to CybSPEED)

[A01] V. Kaburlasos, C. Bazinas, G. Siavalas, G. Papakostas, Linguistic social robot control by crowd-computing feedback, No. 18-2, Proceedings of the 2018 JSME Conference on Robotics and Mechatronics (ROBOMECH 2018), Kitakyushu, Japan, 2-5 June 2018, poster 1A1-B13.

[A02] C. Lytridis, E. Vrochidou, V. Kaburlasos, Emotional speech recognition toward modulating the behavior of a social robot, No. 18-2, Proceedings of the 2018 JSME Conference on Robotics and Mechatronics (ROBOMECH 2018), Kitakyushu, Japan, 2-5 June 2018, poster 1A1-B14.

[A03] M. Graña, M. Dimitrova, V. Kaburlasos, CybSPEED project description: aims and means, No. 18-2, Proceedings of the 2018 JSME Conference on Robotics and Mechatronics (ROBOMECH 2018), Kitakyushu, Japan, 2-5 June 2018, poster 1P1-A13.

[A04] G. Papakostas, G. Sidiropoulos, M. Bella, V. Kaburlasos, Social robots in special education: current status and future challenges, No. 18-2, Proceedings of the 2018 JSME Conference on Robotics and Mechatronics (ROBOMECH 2018), Kitakyushu, Japan, 2-5 June 2018, poster 1P1-A15.

[A05] C. Lytridis, E. Vrochidou, S. Chatzistamatis, V. Kaburlasos, Social engagement interaction games between children with autism and humanoid robot NAO, Proceedings of the 9th International Conference on EUropean Transnational Educational (ICEUTE'18), San Sebastian, Spain, 6-8 June 2018. In: Graña M. et al. (eds) International Joint Conference SOCO'18-CISIS'18-ICEUTE'18. SOCO'18-CISIS'18-ICEUTE'18 2018. Advances in Intelligent Systems and Computing (AISC), vol 771, pp. 562-570, 2019. Springer, Cham.

[A06] G.A. Papakostas, V.G. Kaburlasos, Modeling in cyber-physical systems by lattice computing techniques: the case of image watermarking based on intervals' numbers, Proceedings of the World Congress on Computational Intelligence (WCCI) 2018, FUZZ-IEEE Program, Rio de Janeiro, Brazil, 8-13 July 2018, pp. 491-496.

[A07] E. Vrochidou, M. Manios, V. G. Kaburlasos, F. Panagiotopoulos, C. Aitsidis, V. Ferelis, Design of social robots using open-source robotic platforms, Proceedings of the International Conference on Robotics & Mechatronics and Social Implementations, Varna, Bulgaria, 28 August - 01 September 2018, pp. 21-26.

[A08] M. Dimitrova, H. Wagatsuma, V. Kaburlasos, A. Krastev, I. Kolev, Towards social cognitive neuropsychology account of human-robot interaction Proceedings of the International Conference on Robotics, Mechatronics and Social Implementation, Varna, Bulgaria, 28 August - 01 September 2018, pp. 12-16.

[A09] E. Vrochidou, A. Najoua, C. Lytridis, M. Salonidis, V. Ferelis, G. Papakostas, Social robot NAO as a self-regulating didactic mediator: a case study of teaching/learning numeracy, Proceedings of the 26th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2018), Sym-

posium on: Robotic and ICT assisted wellbeing, Split, Croatia, 13-15 September 2018.

[A10] E. Vrochidou, M. Manios, G. A. Papakostas, C. N. Aitsidis, F. Panagiotopoulos, Open-source robotics: investigation on existing platforms and their application in education", Proceedings of the 26th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2018), Symposium on: Robotic and ICT assisted wellbeing, Split, Croatia, 13-15 September 2018.

[A11] G. A. Papakostas, A. K. Strolis, F. Panagiotopoulos, C. N. Aitsidis, Social robot selection: a case study in education", Proceedings of the 26th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2018), Symposium on: Robotic and ICT assisted wellbeing, Split, Croatia, 13-15 September 2018.

[A12] L. Moussiades, C. Kiourt, S. Papadimitriou, Towards a robot assistant in vocabulary learning, Proceedings of the International Conference on Robotics, Mechatronics and Social Implementation, Varna, Bulgaria, 28 August - 01 September 2018, pp. 32-45.

[A13] V. Kaburlasos, E. Vrochidou, Social robots for pedagogical rehabilitation: trends and novel modeling principles. In: Cyber-Physical Systems for Social Applications, M. Dimitrova & H. Wagatsuma (Eds.), pp. 1-21, 2019. IGI Global: Pennsylvania, USA. ISBN13: 9781522578796, DOI: 10.4018/978-1-5225-7879-6.

[A14] C. Lytridis, V. Vassileva-Aleksandrova, M. Youssfi, C. Bazinas, V. Ferelis, A. Jaki, M. Mestari, V. G. Kaburlasos, Social robots as cyber-physical actors in entertainment and education, Proceedings of the 27th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2019), Symposium on: Robotic and ICT Assisted Wellbeing, Split, Croatia, 19-21 September 2019.

[A15] A. Jaafar, Y. Illoussamen, E. Vrochidou, T. Pachidis. V. G. Kaburlasos, M. Mestari, Multi-agent parallel implementation to solve nonlinear equality constrained multiobjective optimization problem – Case of unmanned aerial vehicle (UAV)", Proceedings of the 3rd International Conference on Intelligent Computing in Data Sciences (ICDS 2019), Marrakech, Morocco, 28-30 October 2019.

[A16] C. Lytridis, A. Lekova, C. Bazinas, M. Manios, V.G. Kaburlasos, WINkNN: Windowed Intervals' Number kNN classifier for efficient time-series applications, Mathematics, vol. 8, no. 3, 413, 2020. https://www.mdpi.com/2227-7390/8/3/413

[A17] M. Youssfi, O. Bouattane, V. Kaburlasos, G. Papakostas, Generic distributed polymorphic learning model for a community of heterogeneous cyber physical social robots in MAS environment and GPU architecture, 4th International Conference on Intelligent Systems and Computer Vision (ISCV 2020), Fez, Morocco, 9-11 June 2020.

[A18] F. Ezzahra Ezzrhari, H. Bensag, M. Youssfi, O. Bouattane, V. Kaburlasos, Scalable multi agent system middleware for HPC of big data applications", 4th International Conference on Intelligent Systems and Computer Vision (ISCV 2020), Fez, Morocco, 9-11 June 2020.

[A19] V.G. Kaburlasos, The Lattice Computing (LC) paradigm". In: Francisco J. Valverde-Albacete, Martin Trnecka (Eds.), Proceedings of the 15th International Conference on Concept Lattices and their Applications (CLA 2020), Tallinn, Estonia, 29 June - 1 July 2020, pp. 1-8. Tallinn University of Technology, Estonia, ISBN: 978-9949-83-557-7. http://ceur-ws.org/Vol-2668/

[A20] M. Qbadou, I. Salhi, H. El Fazazi, K. Mansouri, M. Manios, V. Kaburlasos, Human-robot multilingual verbal communication – the ontological knowledge and learning-based models, Advances in Science, Technology and Engineering Systems (ASTES) Journal, vol. 5, no. 4, pp. 540-547, 2020. https://astesj.com/v05/i04/p64/

[A21] V. G. Kaburlasos, C. Lytridis, C. Bazinas, S. Chatzistamatis, K. Sotiropoulou, A. Najoua, M. Youssfi, O. Bouattane, Head pose estimation using lattice computing techniques, 28th International Conference on Software, Telecommunications and Computer Networks (SoftCOM 2020), Hvar, Croatia, 17-19 September 2020.

[A22] V. G. Kaburlasos, C. Lytridis, C. Bazinas, G. A. Papakostas, A. Naji, M. Hicham Zaggaf, K. Mansouri, M. Qbadou, M. Mestari, Structured human-head pose representation for estimation using fuzzy lattice reasoning (FLR), The Fourth International Conference on Intelligent Computing in Data Sciences (ICDS 2020), Fez, Morocco, 21-23 October 2020.

[A23] M. Dimitrova, A. Krastev, R. Zahariev, E. Vrochidou, C. Bazinas, T. Yaneva, E. Blagoeva-Hazarbassanova. Robotic technology for inclusive education: a cyber-physical system approach to pedagogical rehabilitation. In Proceedings of the 21st International Conference on Computer Systems and Technologies (CompSysTech 20). 19 - 20 June 2020, pp. 293-299. University of Ruse, Bulgaria, Association for Computing Machinery (ACM) Digital Library, https://doi.org/10.1145/3407982.3408019

[A24] M. Dimitrova, S. Kostova, A. Lekova, E. Vrochidou, I. Chavdarov, A. Krastev, R. Bostova, A. Andreeva, V. Stancheva-Popokostadinova, L. Ozeata. Cyber-physical systems for pedagogical rehabilitation from an inclusive education perspective, BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 11(2Sup1), pp. 186-207, 2020. https://doi.org/10.18662/brain/11.2Sup1/104

[A25] E. Vrochidou, C. Lytridis, C. Bazinas, G.A. Papakostas, H. Wagatsuma, V.G. Kaburlasos, Brain signals classification based on fuzzy lattice reasoning, Mathematics, vol. 9, no. 9, 1063, 2021. https://doi.org/10.3390/math9091063 (Special Issue on Numerical Analysis and Scientific Computing – Section Computational and Applied Mathematics. Guest Editors: Theodore E. Simos, Charampos Tsitouras)

[A26] N. S. T. Hirata, G. A. Papakostas, On Machine-Learning Morphological Image Operators, Mathematics, vol. 9, no. 16, p. 1854, Aug. 2021. https://doi.org/10.3390/math9161854 (Special Issue on Lattice Computing: A Mathematical Modelling Paradigm for Cyber–Physical System Applications – Section Computational and Applied Mathematics. Guest Editor: Vassilis G. Kaburlasos)

[A27] M. Dimitrova, H. Wagatsuma, A. Krastev, E. Vrochidou, J. D. Nunez-Gonzalez, A Review of Possible EEG Markers of Abstraction, Attentiveness and Memorisation in Cyber-Physical Systems for Special Education, Frontiers in Robotics and AI, 31 August 2021. https://www.frontiersin.org/articles/10.3389/frobt.2021.715962/full

(Section Biomedical Robotics – Research Topic Emerging Technologies for Assistive Robotics: Current Challenges and Perspectives. Guest Editors: Lyuba Alboul, Jacques Penders, Peter Mitrouchev, Maya Dimitrova, Anna Lekova, Vassilis Kaburlasos)

[A28] V. G. Kaburlasos, C. Lytridis, E. Vrochidou, C. Bazinas, G. A. Papakostas, A. Lekova, O. Bouattane, M. Youssfi, T. Hashimoto, Granule-based-classifier (GbC): a lattice computing scheme applied on tree data structures, Mathematics, vol. 9, no. 22, 2889, 2021. https://www.mdpi.com/2227-7390/9/22/2889 (Special Issue on Lattice Computing: A Mathematical Modelling Paradigm for Cyber–Physical System Applications – Section Computational and Applied Mathematics. Guest Editor: Vassilis G. Kaburlasos)

[A29] V. G. Kaburlasos (Guest Editor), Special Issue on Lattice Computing: A Mathematical Modelling Paradigm for Cyber-Physical System Applications, Mathematics, vol. 10, no. 2, 271, 2022. https://www.mdpi.com/2227-7390/10/2/271 (Section Computational and Applied Mathematics) https://www.mdpi.com/journal/mathematics/special (8 papers).

[A30] I. Salhi, M. Qbadou, S. Gouraguine, K. Mansouri, C. Lytridis, V. Kaburlasos, Towards robot-assisted therapy for children with autism - the ontological knowledge models and reinforcement learning-based algorithms, Frontiers in Robotics and AI – section Biomedical Robotics, 06 April 2022, vol. 9, Article 713964. doi: 10.3389/frobt.2022.713964 (Special Issue on Emerging Technologies for Assistive Robotics: Current Challenges and Perspectives. Guest Editors: Lyuba Alboul, Jacques Penders, Peter Mitrouchev, Maya Dimitrova, Anna Lekova, Vassilis Kaburlasos) https://www.frontiersin.org/research-topics/14419/emerging-technologies-for-assistive-robotics-current-challenges-and-perspectives

[A31] M. Youssfi, F. Ezzahra Ezzrhari, Y. Hajoui, O. Bouattane, V. Kaburlasos, Multi micro-agent system middleware model based on event sourcing and CQRS patterns. In: Intelligent Distributed Computing for Trajectories: Metamodeling, Reactive Architecture for Analytics, and Smart Applications, A. Boulmakoul, L. Karim, B. Bhushan (Eds.), 2022. CRC Press, Taylor & Francis Group: Florida, USA.

[A32] P. Delias, L. Moussiades, V. G. Kaburlasos, Potentials for decision support in business processes through a multi-layer network embeddings approach, 32nd EURO Conference, Espoo, Finland, 3-6 July 2022.

Appendix B

(Distribution of IHU publications over Beneficiaries/Partners)

pub	Beneficiaries								Partners		
no.	1	2	3	4	5	6	7	8	9	10	11
A01						•					
A02						•					
A03	•				•	•					
A04						•					
A05						•					
A06						•					
A07						•					
A08					•	•			•		
A09						•					•
A10						•					
A11						•					
A12						•					
A13						•					
A14				•		•					•
A15						•					•
A16					•	•					
A17						•					•
A18						•					•
A19						•					
A20						•					•
A21						•					•
A22						•					•
A23				•	•	•					
A24					•	•					
A25						•			•		
A26						•					
A27					•	•			•		
A28					•	•					•
A29						•					
A30						•					•
A31						•					•
A32						•					
Total	1	-	-	2	7	32	-	-	3	-	11

LEGEND

Beneficiaries

1. UPV/EHU (UNIVERSIDAD DEL PAIS VASCO/ EUSKAL HERRIKO UNIBERTSITATEA)
2. UGA (UNIVERSITE GRENOBLE ALPES)
3. PRAXIS (AGGELIDIS I.-PALLIKARIDIS X. OE)
4. Theater Tsvete (TEATAR TSVETE)
5. IR-BAS (INSTITUT PO ROBOTIKA)
6. IHU (INTERNATIONAL HELLENIC UNIVERSITY) [former EMaTTech (TECHNOLOGIKO EKPEDEFTIKO IDRIMA ANATOLIKIS MAKEDONIAS & THRAKIS)]
7. CHU (CENTRE HOSPITALIER UNIVERSITAIRE DE GRENOBLE)
8. CVC (CENTRO DE VISION POR COMPUTADOR)

Partners

9. KyuTech (KOKURITSUDAIGAKUHOJINKYUSHUKOGYODAIGAKU) Japan
10.UNIVERSIDAD DE CHILE (UNIVERSIDAD DE CHILE) Chile
11. UH2C (UNIVERSITE HASSAN II DE CASABLANCA) Morocco

References

[1] H2020-MSCA-RISE-2017 project no. 777720 entitled “Cyber Physical Systems for PEdagogical Rehabilitation in Special Education” (CybSPEED), 1 December 2017 - 31 July 2022. [2] CybSPEED Action’s Work Package WP#2 entitled “Research on Analysis, Modelling and Synthesis of CPSs for Pedagogical Rehabilitation in Education”.

[3] V. G. Kaburlasos, Towards a Unified Modeling and Knowledge-Representation based on Lattice Theory, vol. 27. Berlin, Heidelberg: Springer Berlin Heidelberg, 2006.

[4] P. Sussner, “Lattice fuzzy transforms from the perspective of mathematical morphology,” Fuzzy Sets Syst., vol. 288, pp. 115–128, Apr. 2016.

[5] P. Sussner and T. Schuster, “Interval-valued fuzzy morphological associative memories: Some theoretical aspects and applications,” Inf. Sci. (Ny)., vol. 438, pp. 127–144, Apr. 2018.

[6] P. Sussner, E. Esmi, and J. Florindo, “A New Algorithm for Designing J-Fuzzy Associative Memories Based on Subsethood Measures,” in 2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), 2019, pp. 1-6. [7] Sussner, P.; Campiotti, I. Extreme learning machine for a new hybrid morphological/linear perceptron. Neural Networks 2020, 123, 288–298, doi:10.1016/j.neunet.2019.12.003.

[8] Sussner, P.; Caro Contreras, D.E. Generalized morphological components based on interval descriptors and n-ary aggregation functions. Inf. Sci. (Ny). 2022, 583, 14-32, doi:10.1016/j.ins.2021.10.012.

[9] V. G. Kaburlasos and G. A. Papakostas, “Learning Distributions of Image Features by Interactive Fuzzy Lattice Reasoning in Pattern Recognition Applications,” IEEE Comput. Intell. Mag., vol. 10, no. 3, pp. 42–51, Aug. 2015.

[10] Peter Sussner, George Papakostas (chairmen) Special Session on “Recent Advances in Lattice Computing”, World Congress on Computational Intelligence (WCCI) 2018, FUZZ-IEEE Program. 8-13 July 2018, Rio de Janeiro, Brazil.

[11] Peter Sussner, Vassilis Kaburlasos (chairmen) Special Session on “Recent Approaches Toward Lattice Computing”, 2022 North American Fuzzy Information Processing Society (NAFIPS 2022) Conference. 31 May -3 June 2022, Halifax, Nova Scotia, Canada.

[12] V. G. Kaburlasos (Guest Editor), Special Issue on “Lattice Computing: A Mathematical Modelling Paradigm for Cyber-Physical System Applications”, Mathematics, vol. 10, no. 2, 271, 2022. https://www.mdpi.com/2227-7390/10/2/271 (Section “Computational and Applied Mathematics”) https://www.mdpi.com/journal/mathematics/special_issues/Latti (8 papers).

[13] Lyuba Alboul, Jacques Penders, Peter Mitrouchev, Maya Dimitrova, Anna Lekova, Vassilis Kaburlasos (Guest Editors), Special Issue on “Emerging Technologies for Assistive Robotics: Current Challenges and Perspectives”, Frontiers in Robotics and AI (Section “Biomedical Robotics”) https://www.frontiersin.org/research-topics/14419/emerging-technologies-for-assistive-robotics-current-challenges-and-perspectives (4 papers).

[14] Erasmus+, Reference No: 2019-1-EL01-KA107-061958 (Greece - Morocco).

[15] Erasmus+, Reference No: 2020-1-EL01-KA107-077911 (Greece - Morocco).
[16] V. G. Kaburlasos, E. Vrochidou, F. Panagiotopoulos, Ch. Aitsidis, A. Jaki, "Time series classification in cyber-physical system applications by intervals' numbers techniques", Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2019), New Orleans, Louisiana, USA, 23-26 June 2019.
[17] C. Bazinas, E. Vrochidou, C. Lytridis, V. G. Kaburlasos, "Time-series of distributions forecasting in agricultural applications: an intervals' numbers approach", 7th International Conference on Time Series and Forecasting (ITISE 2021), Gran Canaria, Spain, 19-21 July 2021. In: MDPI Engineering Proceedings 2021, 5 (1), 12; https://www.mdpi.com/2673-4591/5/1/12
[18] V. G. Kaburlasos, C. Bazinas, E. Vrochidou, E. Karapatzak, "Agricultural yield prediction by difference equations on data-induced cumulative possibility distributions", 2022 North American Fuzzy Information Processing Society (NAFIPS 2022) Conference, Halifax, Nova Scotia, Canada, 31 May - 3 June 2022.
[19] V. G. Kaburlasos, C. Lytridis, G. Siavalas, T. Pachidis, S. Theocharis, "Fuzzy lattice reasoning (FLR) for decision-making on an ontology of constraints toward agricultural robot harvest", 15th International FLINS (Fuzzy Logic and Intelligent Technologies in Nuclear Science) Conference (FLINS 2022) on Machine learning, Multi agent and Cyber physical systems, Tianjin, China, 26-28 August 2022. (submitted)
[20] A. Amanatiadis, V. Kaburlasos, E. Kosmatopoulos, "Understanding deep convolutional networks through gestalt theory", Proceedings of the 2018 IEEE International Workshop on Imaging Systems & Techniques (IST 2018), Kraków, Poland, 16-18 October 2018, pp. 312-317.
[21] F. Braga and H. S. Pinto, "Composing Music Inspired by Sculpture: A Cross-Domain Mapping and Genetic Algorithm Approach", Entropy 24, no. 4: 468, 2022. https://doi.org/10.3390/e24040468
[22] EU Danube Strategic Project Fund (DSPF), Interreg Danube Transnational Programme, project no. 07_ECVII_PA07_RONNI entitled "Increasing the well being of the population by RObotic and ICT based iNNovative educatIon" (RONNI), 1 January 2018 - 31 January 2019.
[23] Action "RESEARCH – DEVELOP - INNOVATE", cycle A, Intervention II, NSRF of Greece 2014-2020, project no. T1EDK-00929 entitled "Social Robots as Tools in Special Education" (SRTSE), 28 June 2018 - 27 June 2022.
[24] Action "RESEARCH – DEVELOP - INNOVATE", cycle A, Intervention II, NSRF of Greece 2014-2020, project no. T1EDK-00300 entitled "Personalized Optimal Grape Harvest by Autonomous Robot" (POGHAR), 28 June 2018 - 27 June 2022.
[25] Action "Reinforcement of the Research and Innovation Infrastructure", NSRF of Greece 2014-2020, project no. MIS 5046047 entitled "Technology for Skillful Viniculture" (SVtech), 20 November 2020 - 19 May 2023.
[26] INTERREG V-A GREECE-BULGARIA 2014-2020, 6th Call, project entitled "Greek Bulgarian Business Partnership by Assistance, Services, Solutions to Promote Open Regions Team" (GR BG BUSINESS PASSPORT), 1 January 2021 - 31 December 2022.

[27] V. R. Sampath Kumar et al., "Ontologies for Industry 4.0," Knowl. Eng. Rev., vol. 34, p. e17, Nov. 2019.

[28] M. De Clercq, A. Vats, and A. Biel, "Agriculture 4.0: The future of farming technology.," in The World Government Summit, Dubai, UAE, 2018, pp. 11-13.

[29] E. Mavridou, E. Vrochidou, G. A. Papakostas, T. Pachidis, and V. G. Kaburlasos, "Machine Vision Systems in Precision Agriculture for Crop Farming," J. Imaging, vol. 5, no. 12, p. 89, Dec. 2019.

[30] C. Lytridis, V. G. Kaburlasos, T. Pachidis, M. Manios, E. Vrochidou, T. Kalampokas, S. Chatzistamatis, "An overview of cooperative robotics in agriculture" https://www.mdpi.com/2073-4395/11/9/1818, Agronomy 2021, vol. 11, iss. 9, 1818; https://www.mdpi.com/2073-4395/11/9.

Chapter 27
Cyber- Physical System for language therapy for children with communication disorders

Snezhana Kostova[0000−0002−9274−9611], Anna Lekova[0000−0002−9012−4714], Tanio Tanev[0000−0002−8076−0928], Anna Andreeva[0000−0002−9012−4714], Miglena Simonska[0000−0003−0655−6125], Peter Mitrouchev[0000−0001−8495−9240]

Abstract This paper presents robot-based Cyber-Physical System (CPS) for language therapy for children with communication disorders that was created and experimented in frame of the CybSPEED H2020-MSCA-RISE project. The CPS includes humanoid and semi-humanoid robots and is based on a combination of the best of experience, achievements and practices of the researchers from the domains of robotics, AI, system science and speech therapists. All experiments were conducted in accordance with the experimental protocol for measuring of listening, understanding and speaking skills for the verbal language in children with communication disorders, approved by the Ethics Committee for Scientific Research (ECSR) of the IR-BAS. In the paper, detailed description of the CPS, the organisation and conduction of the experiments are given. The first published results are commented which confirm the effectiveness of the robots - based CPS for the children with communication disorders.

27.1 Introduction

The term cyber-physical system (CPS) refers to a new generation of systems which integrated computational and physical capabilities that can interact with humans through many new modalities. In general, CPS is a heterogeneous system, which integrates physical, computer and communication components [12]. These components are interconnected by feedback loops and act autonomously to produce a

S Kostova, A Lekova, T. Tanev
Institute of Robotics, Bulgarian Academy of Science, Acad. G. Bonchev str. 1113 Sofia, Bulgaria
e-mail: kostovasp@yahoo.com

Anna Andreeva, Miglena Simonska
South-West University "Neofit Rilski", 66 Ivan Mihaylov str. 2700, Blagoevgrad, Bulgaria

Peter Mitrouchev
Univ. Grenoble Alpes, CNRS, G-SCOP, 38000 Grenoble, France

consistent response [8]. In other words, computer systems recognize characteristics from the physical world, analyses them and produce a pre-programmed appropriate reaction. Computers recognize and react to physical activity by embedded systems, such as actuators and sensors [12]. Sensors detect physical activity and convert it into information. Actuators receive signals and convert their energy into motion. Physical processes affect cyber mechanisms and vice versa.

The application of CPS both in mass and special education is a big challenge today. The main goal of the CybSPEED H2020-MSCA-RISE project is to contribute to analysis, modelling, synthesis and implementation of Cyber-Physical Systems for pedagogical rehabilitation in special education. Learning by designing human-robot situations (games, pedagogical cases, artistic performances) and advanced interfaces (brain-computer, eye-gaze tracking and virtual reality) where children and students interact with the novel technologies to improved pedagogical rehabilitation in education is the approach used in the project. The robots are a central component in a CPS for pedagogical rehabilitation in special education. Addressing Special Educational Needs (SEN) with CPS and in particular with CPS involving humanoid, semi-humanoid and non-humanoid robots is a challenging approach with proven good results. Robots have an ability to repeat the same set of words and actions over and over again, which can help children with communication disorders to remember and use the learned vocabulary in everyday life [1]. Robots could support the learning of children [13], could be personalized to the child needs, could reduce the teachers' workload, could complement, improve and even replace the work of therapists, especially in situations when there is a lack of therapists or access to kindergartens or schools like the one we had during the Covid-19 pandemic [9, 1]. The lack of therapists could be critical for children with communication disorders.

In the proposed robot-based CPS various appropriate scenarios of the games were developed before the sessions, experimented during the sessions, analysed and evaluated after the sessions. The important issue of the CPS is how to measure its effectiveness. This can be done: (1) subjectively by observation, analysis of behaviour, questionnaires (for parents and/or therapist's) and (2) objectively by detection, processing and analysis of gaze, speech and brain signals.

27.2 Experiments – idea, planning and conducting

After the idea of conducting the experiments with robot-based CPS for language therapy for children with communication disorders during the first CybSPEED training in Sofia, April 2018, the research protocol was prepared from the research team and approved by the IR-BAS Ethics Committee for Scientific Research (ECSR) in May 2018. The experiments were postponed by 2 years due to the COVID-19 pandemic. In January 2020 the pilot experiment was conducted in Day-care Center for Children with Disabilities "Zdravetc" in Bansko. The several scenarios with CPS involving NAO and EmoSan robots were developed, experimented and analysed with the help of special pedagogues, psychologists and parents. After the detailed discussion and analysis, the robot-based CPS was improved and adapted for the goal of this research, namely to confirm the hypotheses that communication skills like speaking, listening and understanding are enhanced in the robotic vs human special educator conditions because the robot reduce the stress in the children and emotionally, quickly and automatically capture attention and enhance the perceptual processes.

Fig. 27.1 Experiments in the Centre of Logopedics at the South-West University "Neofit Rilski"

The main group of experiments was conducted at the Centre of Logopedics, part of the South-West University "Neofit Rilski", Bulgaria in the period November 2021- April 2022. The research team conduct the experiments in which child-robot interaction in play-like structured speech and language therapy for children with communication disorders is realized within six consecutive months, 3 days each month. They are organized like a group session. Participants are a speech and language therapist, a child and two robots - NAO and its "friend" the emotion-expressive robot EmoSan. Two engineers who developed and deployed the technical scenarios observe the process and control the robots. Parents of involved children with communication disorders stayed in the room and they observed the robot-assisted session.

Fifteen children with communication disorders took part in the study, age 3 to 10. The group of children consisted of seven boys and eight girls. Their speech and language therapists referred them to participate in the study. All of them were diagnosed with language disorders – six children were with Developmental Language Disorder and nine were with Language Disorder as a secondary condition (1 with Down Syndrome, 4 with Hearing Impairment and 4 with Autism Spectrum Disorders). All children studied at mainstream kindergartens and schools and used speech and language therapy services.

The results after the first session was published in [1] and they represent parents' opinion about the interest, engagement and communication of their child when interacting with robots. Researchers developed an online questionnaire for evaluation of interaction between the child and the robots. Results show that children are interested in playing and learning with robots. There is a correlation between the type of children's communication disorder and the use of verbal communication in interaction. According to parents' answers of the questionnaire, most of the children liked to interact with robots during the play-like structured activities in the speech and language therapy session. The processing of the results from the other 5 sessions is to be done.

27.3 Description of the robot-based CPS for language therapy for children with communication disorders

As it was mentioned above CPS are composed of physical systems that affect computations, and vice versa, in a closed loop. By tightly integrating and interacting computing with physical systems one can design CPS that are smarter, cheaper, more reliable, efficient and environmentally friendly than systems based on physical design alone. The presence of feedback loops supported by a sensing infrastructure is the common characteristic to all proposals on CPS [12]. The CPS enables the physical world to be monitored, controlled and influenced both adaptively and intelligently. Recently, the humans are added somewhere in the CPS in order to be more productive, more efficiently and more reliable. The systems that consider humans as part of the physical world are known as Human-in-the-loop systems. Three types of Human-in-the-loop systems are considered [4] - (i) systems where humans may control the operation; (ii) systems where humans are only passively monitored and (iii) hybrid systems of the previous ones.

The main parts of proposed robot-base CPS are:

27.3.1 DEVICES

- Humanoid robot NAO, equipped with many sensors and actuators that allow different modalities for interactions. It can perform gestures, play sounds, recognize objects, words or landmarks and barcodes. NAO has a preinstalled Operating System NAOqi. The programming of the robot is performed in two ways: by using graphical interface of the Choregraphe environment [11] and/or in Python programming language inside Choregraphe or from external IDE (integrated development environment). For active interactions NAO uses pre-programed gestures and movements, however via the animation mode in the robot software [10] custom movements can be programed. The Timeline box that contains Motion and Behavior layers enables for synchronization of boxes with movements, movements with each other and/or Boxes with each other. NAOqi Application Programming Interface (API) provides preprogrammed modules and methods by which the interactive behavior of NAO is programed. The useful NAOqi module that we used is ALLandMarkDetection [10]. The robot recognizes special landmarks with specific patterns on them by this vision module. Each picture or 3D object used in the experiments had unique landmark stick on it and it used the ALLandMarkDetection methods, NAO recognizes whether the child's response is correct or not. The ALMemory module provides information about the current state of the actuators and sensors in a format of key-value variables. By the ALMemory APIs this information can be read or written, custom events as key-value pairs can be defined and publish-subscribe model for exchanging messages between publishers and subscribers can be applied. Thus, we manage the intervention session and synchronize the communication of both robots. For

each play, we set a custom event as a key-value pair by writing it in the internal memory of the robot. Then NAO updates the stimulus of this event according to the goals of the play and the child's progress.
- Semi-humanoid Emotion-Expressive Robot EmoSan [6] – EmoSan is a new, semi-humanoid and emotion-expressive robot, designed in frame of the CybSPEED project. The main advantage of this robot is that it can represent different emotional reactions by movement of the head. EmoSan looks like an emoticon, but it moves, it has eyes, eyebrows and the mouth. The goal of EmoSan is to be "partner" of NAO and to complement it in the expression of emotions (which the NAO does not have). Expressing emotions is extremely important when working with children, and even more so for children with SEN. Emotional intelligence helps to enhance the learned abilities of children with SEN. The head movements in communication is an inevitable candidate to analyze the unconscious emotional expression. The movement of the head is realized by a parallel kind of robot based on the Gough-Stewart platform. EmoSan is designed to have an affective interaction between the human and robot, including a coupling through brain signals such as electroencephalography, known as EEG. EmoSan enhances the teaching of children in basic emotions through playing them in an exciting away.
- Synchronization between NAO and EmoSan - Each one second both robots read the memory in order to trigger and synchronize their interactions. NAO reads and writes into its memory via python scripts inside in Choregraphe, while EmoSan reads the NAO memory by python scripts from external IDE.
- Laptop, Smartphone, Emotive, Kinect, Eye tracking and other devises can be added depending from the scenario.

27.3.2 HUMANS

Two humans are involved in the proposed CPS:

- Speech and language therapist and
- Child with communication disorders.

27.3.3 COMMUNICATION

For the purpose of these experiments, we have developed a control framework which consists of Nod-RED, EMOTIV BCI, humanoid robot NAO and semi-humanoid robot EmoSan. Node-RED [14] allows to wire up IoT (Internet of Things) as nodes in flows and it is an open-source development tool built by IBM. Node-RED can be run on laptops, PCs, small single-board computers (e.g. Raspberry Pi), or cloud environments (e.g. IBM Cloud). The Node-RED connectivity allows nodes to collect and exchange data and its flow-based programming allows to wire up the robots, EMOTIV BCI, laptops and people. Both robots read the NAO memory (NAO using

Choregraphe, EmoSan using Node-RED) in one-second interval in order to trigger and synchronize their interactions.

Games

Before the game we checked whether the child could hear well the whole range of speech sounds using the Ling Six-Sound Test. NAO played a sound and the child gave feedback if he/she had heard it. Every game began with an introduction sentence "EmoSan does not know X (e.g. farm animals) and it makes him unhappy." The emotion-expressive robot EmoSan moved its eyebrows and its mouth in order to imitate a sad face. Then NAO continued: "But I know X (e.g. farm animal). Let's teach him together. Now we are going to play a game." During the game, when the child's choice of a task was correct, NAO said "Well done!". However, when the respond was wrong the robot motivated the child to "Try again!" and played the task instruction again. At the end of the game NAO concluded positively "Well done! I like to play the game X with you! I am happy that EmoSan has learned X with us!". Then the emotion-expressive robot EmoSan imitated a happy face with movement of eyebrows and mouth and said "I'm happy!".

27.3.4 Feedback

The feedback in this CPS has three dimensions:

- Feedback from NAO sensors and from other devises – this reaction is key to the smooth running of the game. For example, if the child give correct/incorrect answer during the game, the sensors detect it and react with corresponding action – the game continues, stopped, repeat, give more information etc. This reaction also provokes a subsequent emotional reaction of the EmoSan robot;
- Feedback from Therapist – The therapist constantly monitors the child behaviour and reaction and can take adequate action if necessary;
- Feedback from the child – The child gives feedback if he/she can't heard properly, if he/she don't understand the rules of the game etc.

The figure 2 illustrates the proposed robot-base CPS.

27.4 Conclusion

After receiving encouraging results from the experiments, we can conclude that the developed robot-based CPS can be successfully applied to children with communication disorders. The proposed system has potential to be applied for therapy of people in remote areas or in case of pandemic when the concept of IoT can be used.

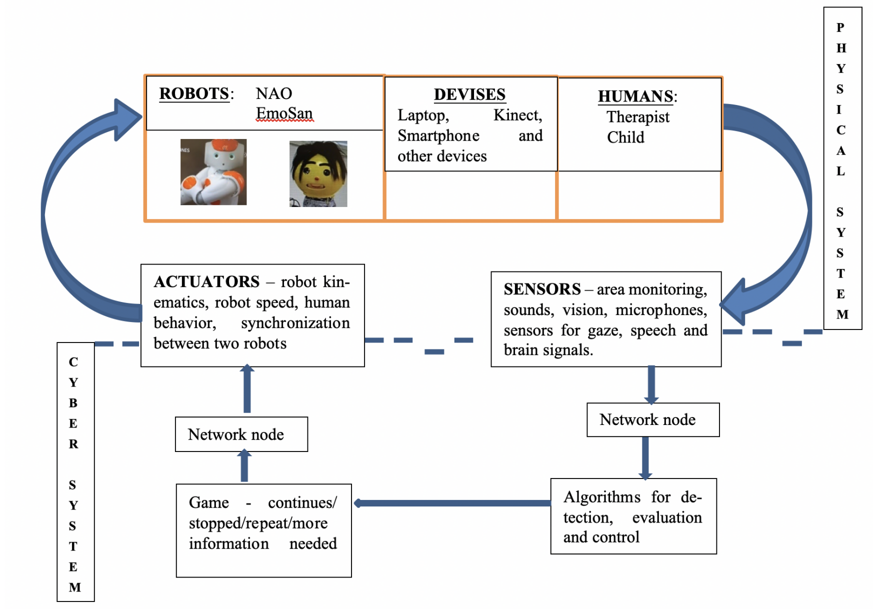


Fig. 27.2 Robot-based CPS for children with communication disorders with human in the loop

We believe that this is a step ahead to enhance communication skills of children with communication disorders with a positive effect on their future quality of life.

Acknowledgements

This work was supported by the CybSPEED H2020-MSCA-RISE-2017 grant 777720 and by European Regional Development Fund within the OP "Science and Education for Smart Growth 2014 - 2020", Project CoC "Smart Mechatronic, Eco- And Energy Saving Systems And Technologies", № BG05M2OP001-1.002-0023.

References

1. Andreeva A., Lekova A., Simonska M. and Tanev T., Parents' Evaluation of Interaction between Robots and Children with Neurodevelopmental Disorders. The Smart Digital Futures 2022 multi-theme hybrid conference, Rhodes, Greece, 20-22 June, 2022 (accepted).

2. Atanasova, Aneta, and Aleksandra Yosifova. "Adressing Special Educational Needs in Classroom with Cyber Physical Systems." International Journal of Cyber-Physical Systems (IJCPS) 1.1 (2019): 71-90.
3. Baheti, Radhakisan, and Helen Gill. "Cyber-physical systems." The impact of control technology 12.1 (2011): 161-166.
4. B. Bordel, R. Alcarria, T. Robles and D. Martín, Cyber–physical systems: Extending pervasive sensing from control theory to the Internet of Things, Pervasive and Mobile Computing, Elsevier Vol. 40, 2017, 156–184.
5. Cifuentes, C.A., Pinto, M.J., Céspedes, N.& Munera, M.: Social Robots in Therapy and Care. Current Robotics Report https://doi.org/10.1007/s43154-020-00009-2, (2020).
6. Dachkinov P., Tanev T. K., Lekova A., Batbaatar D., Wagatsuma H.: Design and Motion Capabilities of an Emotion-Expressive Robot EmoSan. Joint 10th International Conference on Soft Computing and Intelligent Systems and 19th International Symposium on Advanced Intelligent Systems SCIS&ISIS2018, Toyama, Japan, (2018)
7. Dimitrova, Maya, et al. "Cyber-Physical Systems for Pedagogical Rehabilitation from an Inclusive Education Perspective." BRAIN. Broad Research in Artificial Intelligence and Neuroscience 11.2Sup1 (2021): 187-207.
8. Lee, Edward A. "Cyber physical systems: Design challenges." 2008 11th IEEE international symposium on object and component-oriented real-time distributed computing (ISORC). IEEE, 2008.
9. Leyzberg, Daniel, Samuel Spaulding, and Brian Scassellati. "Personalizing robot tutors to individuals' learning differences." 2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, 2014.
10. NAOqi Vision - Landmark Detection. Available Online: https://developer.softbankrobotics.com/nao-naoqi-2-1/naoqi-developer-guide/naoqi-framework/naoqi-apis/naoqi-vision/allandmarkdetection#allandmarkdetectio
11. Softbank Robotics Choregraphe. Available Online: https://developer.softbankrobotics.com/nao-naoqi-2-1/naoqi-developer-guide/programming/choregraphe-user-guide/main-panels/flow-diagram (accessed February 2022).
12. Tan, Ying, Mehmet C. Vuran, and Steve Goddard. "Spatio-temporal event model for cyber-physical systems." 2009 29th IEEE International Conference on Distributed Computing Systems Workshops. IEEE, 2009.
13. Westlund, Jacqueline M. Kory, et al. "Children use non-verbal cues to learn new words from robots as well as people." International Journal of Child-Computer Interaction 13 (2017): 1-9. 14. IBM NodeRED Flow-based programming for the Internet of Things. https://nodered.org/

, Jun Wang Univ. Grenoble Alpes, GIPSA-Lab, CNRS, 38000 Grenoble, France

, Julia Dvorak, Amélie Svensson Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany

, Snezhana Kostova, Tanio Tanev, Anna Lekova, Maya Dimitrova, Roman Zahariev, Nina Vlachkova Bulgarian Academy of Sciences, 1000 Sofia, Bulgaria

, Tsvete Yaneva, Violina Markova Tsvete Theater, 1000 Sofia, Bulgaria
Peter.Mitrouchev@grenoble-inp.fr

Chapter 28
Sensory input and its treatment in the context of different illusions. Application for size-weight perception

Jun Wang, Peter Mitrouchev, Julia Dvorak, Amélie Svensson, Snezhana Kostova, Tanio Tanev, Anna Lekova, Tsvete Yaneva, Violina Markova, Maya Dimitrova, Roman Zahariev, Nina Vlachkova and Franck Quaine

Abstract Today, the interaction of illusion and truth, which is often based on psychophysical phenomena, represents important research subject. An important role is played by the human in different situations in his/her daily life such as: artistic or sports performances, or assembly/disassembly (A/D) operations in production, ergonomics, training, health, service or recycling stages. Thus, the aim of this review/opinion paper about our current research is to determine how sensory input is captured and treated by the "human" and how both aspects are affected by different illusions as for example size-weight perception. This is of interest in a larger context in general and in the context of assembly/disassembly operations simulations in particular where the question is how weight perception impacts the physiology of the operators which is related with their muscle fatigue. In this literature review, a share of weight illusions is presented. Some concepts in the context of different illusions and their treatment are presented as well. These are classified into four categories: i). concepts considering the integration of various inputs, ii). concepts treating the importance of different information inflows, iii). concepts treating the modification of input and iv). measurable effects in the human body. In addition, biological aspects, being the foundation of the approach, are considered, including which brain regions are active for different illusions.

28.1 Introduction

When manipulating an object, the brain uses, for example, colour visual stimuli to estimate the object's weight and scales muscle force by the muscular activity control system accordingly. The idea is to show that even when we visually perceive an illusion - for example, whether a black object (dumbbell for instance) is heavier than a white one, which are otherwise of the same weight, the motor system controlled by the motor area of the brain makes an adequate assessment of the stimulus - the weight of the object.

The purpose of these psychophysical experiments being generally to substantiate the ability of the cognitive system to perceive and communicate truth - for example, in the perception of physical exercises, theatrical sketches or assembly/disassembly (A/D) operations simulation, in this paper, we estimated the Sensory input in a particular illusion related with size-weight perception. For this purpose, a series of experiments was performed in an Augmented reality environment. The results of the performed experiments show that the visual stimuli based on the colour of an object for weight perception, can modify the muscle activity.

Compared to 1890, the life expectancy of a child born in Germany, for instance, has doubled by 2020 from 40.32 years to 81.1 years [1]. In line with an aging population and an over aging society, the age until which people still have to work by law before being eligible for retirement support increases as well. Depending on the nature of the task, the impact of work on personal health is not to be underestimated. While handling an object, for example, during A/D operations some movements may be harmful to the operator. Those damages are due to the high load of the manipulated objects or unergonomic conditions of work and consequently to incorrect movements. In order to detect such movements, various models have been established. For example Puthenveetil et al. [2] suggested a model based on body segments allowing to analyse the ergonomics of operations by using specific software.

In the context of assembly and disassembly operations, lifting objects is a frequently executed task. The weight of an object to be handled, and especially the perception of its weight may greatly influence the fatigue of the person handling it. The fatigue of a person's muscles can, on the one hand, be felt by the subject, on the other hand it can be estimated physiologically. A common method to conclude on muscle fatigue is the analysis based on electromyography (EMG) [3, 4].

The question of whether muscle fatigue can be manipulated by the modification of the visual appearance of an object (e.g. changing its color, brightness, size, shape or material) is subject of many studies [5]. The idea is to modify objects' visual appearance using new technologies such as augmented reality, and examine how muscle fatigue is impacted. As early as the end of the 19th century, so-called weight illusions were studied in which a correlation between the perceived weight and modifications of a property of the object were observed [6]. The size-weight illusion (SWI) implies that small objects are perceived as heavier than large objects of equal weight. Other illusions to mention within this context are the material-weight illusion (MWI), visual-haptic illusion (VHI) and brightness-weight illusion (BWI). It has been shown by [3, 4] that weight illusions can be used to influence muscle fatigue. Objects, which are perceived as lighter, resulted in less muscle activity during lifting than dark objects of the same weight and shape [3].

Considering illusions, it is of interest to understand how does multi-model perception work and which of sensorial modalities are stimulated in the context of the four illusions mentioned above. Studies examining weight illusions frequently ask test subjects on their perception, but physiological impacts on the human's body are less considered [7-13]. In a first approach the aim of these studies is to investigate the reflection of different illusions in the nervous system consisting of the peripheral nervous system (PNS) the muscles and the central nervous system (CNS) the brain. The CNS is the complementary part of the nervous system to the PNS consisting of the brain and the spinal cord. It processes information from the PNS and sensory organs to send output signals to the muscles and the autonomic nervous system regulating organs [14]. Thus, the literature review focused on the various sensory inputs in the context of weight perception. Consequently, this study focuses on various approaches which have already shown the correlation between human perception and the impact of actions on the peripheral nervous system. If available, studies considering weight perception and their impact on the peripheral nervous system are preferred. Thus, the first aim of this study is to provide a representative overview of the current state of research.

The question is: how visual appearance of an object interacts with other stimuli to form a percept and how the information input and percept is represented in the human body.

The next of this study is as follows: in Section 2 the methodology and the search strategy are presented. Section 3 shows the different weight-illusions and provides some knowledge on human perception. In addition, the different concepts found in the context of weight illusions human's sensory analysis are presented and clustered. Section 4 concludes on potential avenues for future research.

28.2 Methodology

The pipeline of the proposed approach is presented in Figure 1. I1 to In represent that different inputs which are weighted according to coefficients w1 to wn. The proposed process is assumed to be impacted by the illusion that is created (e.g. size-weight illusion).

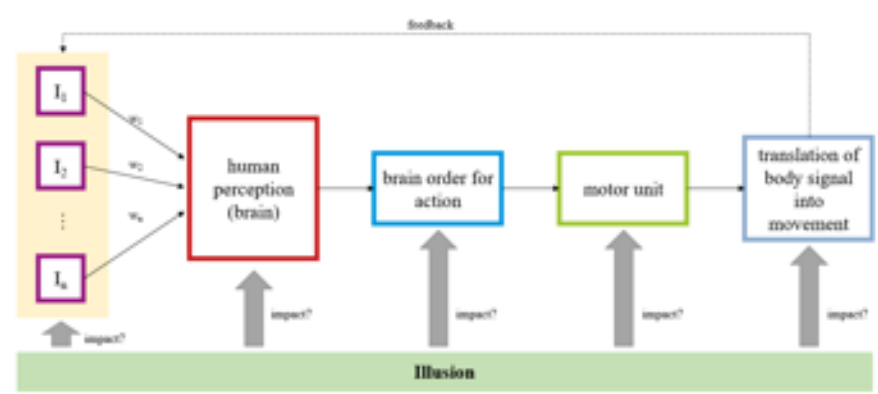


Fig. 28.1 Interaction of sensory inputs and process to translate into a perception and an action.

28.2.1 Search strategy

In a first step, information on different weight illusions is collected. This information is then used to assess its impact on different sensory inflows of the human body. As previously stated, the aim is to only consider the peripheral nervous system, but it has then been notified that literature was scarce in this way. Nonetheless, since the term "peripheral nervous system" is of great importance in biology as well as in medicine, the search has been limited to results that explicitly consider one of the weight illusions found beforehand.

28.2.2 Research question

Based on the approach presented previously it is aimed to answer the following research question: Which sensory input is of importance as it comes to weight illusions and how is this input treated? This question has been frequently answered, but the aim of this study is also to provide avenues for future research. These might be indicated based on the framework that is derived at the end of the manuscript.

28.3 Results

First, different weight illusions, some of which are presented in the literature, are summarized here below. Second, the perception of objects' weight is approached from two sides: the sensorimotor and cognitive perception. Third, a framework is suggested that shows concepts which are related to the body-internal perception system in the context of weight illusions. These results are presented in the following section.

28.3.1 Weight illusions

The relation between physical and dynamical properties of an object and its apparent weight has been studied since the late 19th century by Charpentier. Table 1 summarizes some studies. Thereby, the size-weight illusion is very likely the illusion that has been studied widely [15]. Even if its use within an A/D context is questionable, it is expected to find more literature on this illusion in what concerns the connection between perceived weight and the nervous system. In addition, it is reported as a very strong illusion [15-17].

Table 28.1 Overview of a share of weight illusion studies

Illusion	Description and effects	Studies
Size-weight illusion (SWI) or Charpentier illusion	Large objects are expected to be heavier and the discrepancy between its real and expected weight leads to a perceived heaviness or lightness.	[3, 12]
Visual-haptic illusion (VHI)	The visual perception outweighs the haptic one. As, for example, the visual angle of an inclination is changed by Augmented Reality (AR), the angle is perceived as the seen one and not the felt one.	[11, 12]
Material weight illusion (MWI)	Based on previous experiences on materials and the corresponding density, a certain weight of an object is assumed. These expectations lead to different weight perceptions for objects of the same mass and size but made from different materials.	[8, 16]
Brightness weight illusion (BWI)	Brighter objects are assumed to be lighter than dark ones, which has been shown in an experiment modifying the brightness of objects with AR.	[3, 15, 17]

28.3.2 Cognitive perception

In the following, the question shall be followed of how heaviness is perceived. Therefore, some weight matching experiences are presented. It is important to mention that the visual perception biases the proprioception [18]. The proprioception may be compared to the transmittance of a picture to the brain that reflects the status of the self-concerning position and movements [19]. Guerraz et al. [18], for instance, studied the influence of the visual impression through a mirroring experiment where test subjects looked at a reflection of their healthy extremities instead of the amputation. When moving the healthy extremity an illusion was created and they perceived that the amputated part was moving. Van Polanen and Davare [13] assumed that the weight perception is for example biased by previous experience that is called sensorimotor memory. This is understood as the knowledge of the system (in this case: the human). Before the grasp movement is executed, a weight estimation of an object is made, and the grasp is planned. Based on the estimated weight, a certain force to lift the object is determined which is then applied. An object whose heaviness is overestimated, may be perceived as lighter. In the following trial, more information was available and the subject learnt to adjust the applied force [13].

However, it is of great importance to distinguish between the perceived heaviness and the measured forces really applied. Buckingham et al. [16] claim that for objects with the same mass, but made from different material, the applied forces are equal after several trials. By contrast, Ban et al. [3] state that as lighter objects are lifted, the less the muscles do fatigue. It is to be kept in mind that the former treats force, whilst the latter treats fatigue – this contradiction fosters the impression that fatigue may be related to more than only the force applied.

28.3.3 Sensorimotor perception

The aspects described above arise the question of how heaviness of objects is perceived on a physiological level. Flanagan et al. [20] state that the "control of action and making of perceptual judgements rely on neural processes that use sensory information in different ways". It is assumed that the weight perception is related to the PNS as well as the CNS [21]. In the context of sensorimotor perception, the golgi tendon organs are of importance. The golgi tendon are organs to measure the tension in muscles and are found on the connection between muscles and tendons [22]. Later, Luu et al. [19] pointed out that the measurements of the tensions were forwarded to the brain. Other than the golgi tendon organs, muscle spindles and cutaneous receptors which may be found in the skin are part of the sensorimotor system [19]. The friction between object and skin is reported by the skin receptors and has an impact on the force that must be applied to lift an object [23]. The muscle spindles are responsible for proprioception [19]. These three types of sensors detect sensations and send information as afferent inflow to the brain. The brain processes the information and generates an efferent flow which will trigger an action. Throughout this process the force needed for lifting an object is determined [19]. In combination with the visual impression a perception of weight is formulated. However, it is not clear to which share each of the sensorial inputs are of importance.

28.3.4 Concepts in the context of the different illusions

The aim of this section is to provide a framework that gives a first orientation of concepts related to the body-internal perception system and that are of interest in the context of weight illusions. The concepts found are clustered, assigned to one of the illusions and set in relation to measurable biological aspects. However, this framework is neither exhaustive nor exclusive, not at least since it is "not fully understood how perception and sensorimotor action are integrated" [10]. Note, that the field of research is very dynamic. Consequently, the aim of this study is not to provide an answer, but rather to suggest a structure and to organise the results found concerning weight illusions, sensory input and its treatment and effects. First, the groups of concepts are presented. Second, the concepts found for each group are presented and in a third step the results are synthesized to provide a framework as basis for future analysis.

28.3.4.1 Suggested Clustering

The concepts found for the different illusions are assigned to the following groups:

1. Concepts considering the integration of various inputs.
2. Concepts treating the importance of different information inflows.

3. Concepts that treat the modification of input.
4. Measurable effects in the human body.

These concepts are all underlined by the biological basics that were presented previously in section 3.1.

28.3.4.2 Concepts considering the integration of various input streams.

Concerning human perception, various sources of input are considered. The main sources that have been found in literature are: i). previous experiences made (previously learned) [24]; ii). information that is provided by the sensory system of the body (including proprioception) [8] and iii). visual and haptic information (obtained in the situation itself) [25]. In this section, different models are provided that show how this information inflow may be potentially integrated and what influence it has on weight perception under various conditions (i.e. exposed to illusions).

Hirsiger et al. [10] state, that the nervous system merges different sources of information to determine an output which is the perceived weight in this context. It is assumed that there are two representations of an object within one action: one that is derived from sensory perceptions, such as vision, and one that is derived from the motor system. Both of the information inflows are continuously adapted and integrated to determine human's action [10]. This could be for example the determination of the required grip force that, with an increase on repetitions, becomes more accurate [16]. These aspects have been particularly examined in the context of the SWI.

Barnard and Teasdale [26] presented the Interactive Cognitive Subsystem (ICS) Model, that aims to explain the effect of multiple sensations on the perception. This model covers the entire processing chain from multisensory data to a final perception. It comprises "nine equally organized subsystems" [27] which are receiving information from each other whilst storing, processing and sending them to another higher subsystem. In case of multiple inputs of consistent information, the information can be blended. In pseudo-haptics, for instance, different inputs would be typically the visual and the haptic stimuli. Memory plays an important role again, as data can be replaced by previous knowledge on any level, which may lead to feedback loops between some of the subsystems [26].

28.3.4.3 Importance of different information inflows

The aim of this section is to determine which information inflows are more important for a human's statement on weight perception. For example, Freeman et al. [28] showed in an experiment, where both cognitive and sensory input has been manipulated, that cognitive perceptions do have a lower impact on the SWI experienced than sensorimotor information. In the same year, this was confirmed by Saccone et al. [15].

Comparing visual and haptic aspects (VHI), a "visual dominance" [15] can often be observed when there is a conflict between the visual and haptic sensors. In an experiment performed by Rock & Victor [29], test subjects were shown a cube through a lens which distorted the image along the horizontal axis resulting in representation of a rectangular cuboid. They were then asked to grasp the object under a silk cloth preventing to see their hands while touching it. When interrogated about their impression the subjects perceived the boxes in their hands a rectangular – an impression that is very likely derived from the visual stimuli to which they were exposed at the same time. In literature, various concepts exist to explain the interaction and influence of simultaneous stimuli such as the ICS and concepts that will be presented below [27, 30]. In the case of pseudo-haptics at least two stimuli are considered as input, the visual and the haptic one, which are fed into a multisensory processing system [27].

In order to express the connection between two inputs, Ernst & Banks [30] presented a statistical model based on the Maximum Likelihood Estimate (MLE) to determine the dominance of one of the inputs (i.e. visual or haptic stimuli). The multisensorial information about an environment "is integrated by summing in the contributing sensory estimate's weighted average" [27] and being fused in a second step. Thus, the CNS is more likely to favor information from the modality with higher effectiveness in measures, for example low variance associated with visual estimation leads to a higher dominance of the visual stimulant.

However, this theory was revised and developed further to the Bayesian Multimodal Cue Integration (BMI) due to inconsistencies observed between prediction and testing results in further experiments [27, 30]. BMI uses elements from the Bayesian Decision Theory (BDT) which considers the influence of knowledge and experience (referred to as coupling prior) [27,30]. The BMI also integrates the intention of the task when attributing the impact of multimodal stimuli. For example, in ambiguous situations information retrieved from the prior and the task goal may compensate for the lacking data and influence the multisensory integration process through giving more importance to one of the sensorial stimuli. This has been studied in the context of VHI. In 2011, Pusch and Lécuyer [27] suggested a synthesized system which comprises the two models of BMI and ICS (Figure 2) to explain information processing and the role of different data inputs, memory and task intention.

Information from the environment is transformed into a final perception. In a first step multimodal raw data from vision, audition and the body state (incl. haptics) are processed to a higher level of information. A first fusion of non-redundant multimodal data then occurs to reach the sensory combination. For the sensory integration a fusion of redundant information is done. In the case of ambiguities such as missing data for the information transformations or for both information fusions, information can be retrieved from the memory and the task goal for compensation. This may result in a higher influence from one of the input modes (e.g. vision dominating body state input). Based on the sensory integration a final perception is developed. The system was validated by [27] with the example of the pseudo-haptic feedback in Hand-displacEMent-based Pseudo- haptics (HEMP).

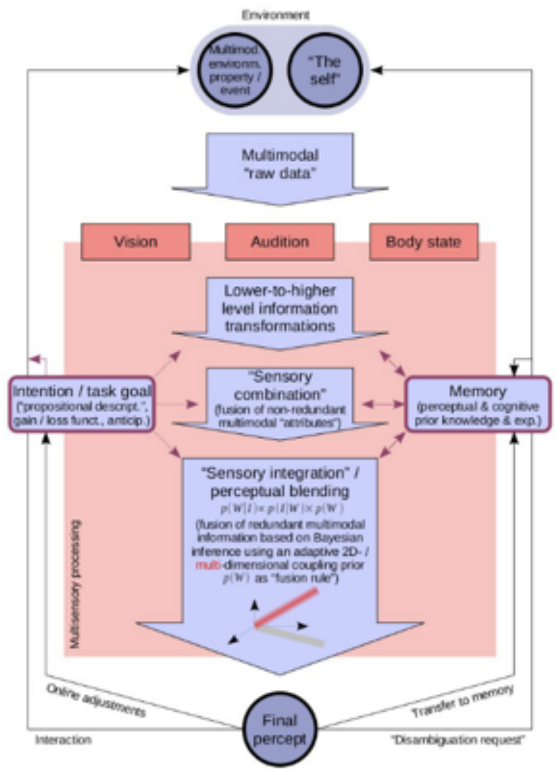


Fig. 28.2 Synthesized system for human perception comprising BMI and ICS [27].

28.3.4.4 Modification and change of input

Since the weight of an object is not always perceived correctly by subjects a certain modification of input streams is assumed. As it has been stated previously, the information obtained by the subject is continuously updated. However, as it has been stated by Brayanov and Smith [8], the adaption remains "a bias that drives weight estimation towards previous experiences". Nonetheless, as soon as an object is lifted, "prior expectations are integrated with real time proprioception information" [8] for example in order to determine the correct amount of force required.

Here we briefly discuss how sensory inflows are manipulated on purpose. Above, the manipulation of one channel of inflow has already been presented. Another illusion that is of interest to consider is the so-called somatosensory illusion [9], that has been shown by an experiment. For this purpose, Augmented Reality (AR) has been used to modify the position of an object on a person's hand or better, to slightly lift it. As a result, the weight perceived by the subject increases [9]. However, from a motor system perspective no measurable adaptions are performed, and the illusion created solely impacts the person's perception.

28.3.4.5 Measurable body effects

Considering the SWI, it has been notified that with an increasing number of trials forces are adapted, even though a derivation in weight perception remains [16]. In addition, it seems that the sensorimotor system assumes small disturbances, which are evened out automatically for example during a grasping process [9].

In order to allow a better understanding of the role of the nervous systems, studies with deafferented subjects were done as well. Usually, healthy subjects can make what has been considered as a "state-estimation" by Grafton [24], that is determined by "object position, vision, haptic feedback, proprioception and efference copy" [24]. Deafferented people are not able to make this estimation to full extent, what leads to the fact that they may only experience the SWI to a very limited extent.

28.3.4.6 Biological foundations

To provide a better understanding, some biological aspects were examined as well. In the human brain all information obtained is collected and it has been found that for various actions various areas of the brain are involved. Chouinard et al. [17], examined which are these regions that show actions during the modification of size, weight and density of an object. These actions were measured using fMRI (function magnetic resonance imaging). The results of Chouinard et al. [17] obtained for this examination are shown in table 2 (active brain areas for object modification).

Table 28.2 Active brain regions for modification of size, weight and density

The temporal cortex is assigned "a strong role in visual perception" [7], p. 963, whilst the parietal cortex is "primarily concerned with visuospatial processing" [7], p. 959. Since it is assumed that the modification of size is primarily noticed by vision (and haptics), a reaction of the brain regions for visual processing is expectable.

If the weight of an object is changed, this requires an adaption of the force that is applied to lift it. The primary motor area is responsible to steer movements and to send orders to the motor neurons [31]. The primary motor area and the ventral premotor area are both "crucial for transforming an object's geometrical properties into a motor command" [32]. They point out, that the left ventral premotor area that "did not adapt to either the size or the weight of objects, adapted instead to the density of objects, and responded more when subjects falsely perceived differences in weight between objects in the SW illusion" [17]. In a study involving monkeys, it has been found that the ventral motor cortex was also active when observing other monkeys lifting objects [17].

28.4 Discussion and future research

The connection amongst the four weight illusions (SWI, MWI, BWI and VHI) studies in this review is shown in figure 3. The initial frame of this work was the assembly/disassembly operation of a product and the resulting muscle fatigue. The topic then was developed towards research considering the peripheral nervous system (muscles). During the research process, it has been observed that there is a lot of content treating the various information channels of the body. These were briefly presented here. The concepts found in literature when looking for publications in the context of the four illusions presented were then structured and a frame for future research was created.

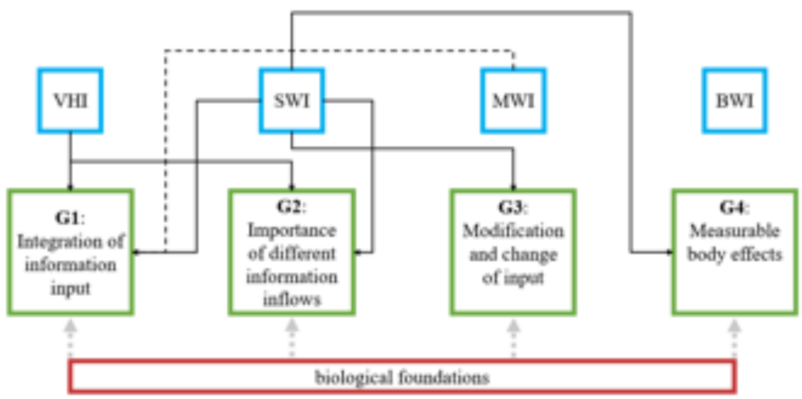


Fig. 28.3 Structuration and interactions of selected weight illusions and suggested groups

Each one of the groups presented in Figure 3 (G1, G2, G3, G4) could function as a potential area for research. Whatever research is made, it must be highly correlated to biological aspects. It is seen that the BWI does not have a direct connection to any of the groups. Thus, it could be of further interest to examine the connection of BWI with the other groups in the way for further research. In addition, the MWI is frequently treated in combination with the SWI [15] and might deserve further attention. Furthermore, the biological aspects were only considered to a very limited extent and it could be of interest to find further information on the brain region of the left ventral premotor (cf. Section 3.4.6) and its involvement in weight perception.

28.5 Conclusion

Considerable work remains to be done to build the necessary ethical framework for integrating AI technology in health care safely and effectively. The ethical challenges presented here assist AI and healthcare stakeholders inappropriately approaching the ethical issues in future discourses to establish responsible AI in healthcare techniques. The study anticipates that this research will spur additional rigorous theoretical and empirical research into the most ethical use of AI systems.

As literature was very scarce for the original research question treating the impact of weight illusions on the PNS, the research scope was enlarged. The general processing of environmental inputs to form perception and the measurable effects on the human body were thus considered. With the findings, a clustering of the concepts attempting to explain the phenomena as well as their impact was structured, which also shows the relationship between illusions and concepts. However, the question of how visual impressions in the form of object property changes might be represented in the PNS and how this could be used in an assembly/disassembly operation context, remains open to discussion.

Acknowledgements

This work was supported by European Project Cyb-SPEED, H2020-MSCA-RISE-2017, GMN 777720, http://www.ehu.eus/ccwintco/cybSPEED/#about-1, https://cordis.europa.eu/project/rcn/212970_fr.h

References

1. O'Neill, A. (2019) Germany: Life expectancy 1875-2020 | Statista [Online]. Available at https://www.statista.com/statistics/1041098/life-expectancy-germany-all-time/ (Accessed 20 April 2020).
2. Puthenveetil, S. C., Daphalapurkar, C. P., Zhu, W., Leu, M. C., Liu, X. F., Gilpin-Mcminn, J. K. and Snodgrass, S. D. (2015) 'Computer-automated ergonomic analysis based on motion capture and assembly simulation', Virtual Reality, vol. 19, no. 2, pp. 119–128.
3. Ban, Y., Narumi, T., Fujii, T., Sakurai, S., Imura, J., Tanikawa, T. and Hirose, M. (2013) 'Augmented Endurance: Controlling Fatigue while Handling Objects by Affecting Weight Perception using Augmented Reality', CHI 2013 Changing Perspectives.
4. Taima, Y., Ban, Y., Narumi, T., Tanikawa, T. and Hirose, M. (2014) 'Controlling fatigue while lifting objects using Pseudo-haptics in a mixed reality space', 2014 IEEE Haptics Symposium (HAPTICS). Houston, TX, USA, 23.02.2014 - 26.02.2014, IEEE, pp. 175–180.
5. Potvin, J. R. and Fuglevand, A. J. (2017) 'A motor unit-based model of muscle fatigue', PLoS computational biology, vol. 13, no. 6, e1005581.
6. Nicolas S., Ross H.E., Murray D.J, Charpentier's Papers of 1886 and 1891 on Weight Perception and the Size-Weight Illusion, August 1,, 2012, https://doi.org/10.2466/24.22.27.PMS.115.4.120-141

7. Bell, A. H., Pessoa, L., Tootell, R. B.H. and Ungerleider, L. G. (2013) 'Visual Perception of Objects', in Squire, L. R. (ed) Fundamental neuroscience, 4th edn, Amsterdam, Elsevier Acad. Press, pp. 947–968.
8. Brayanov, J. B. and Smith, M. A. (2010) 'Bayesian and "anti-Bayesian" biases in sensory integration for action and perception in the size-weight illusion', Journal of neurophysiology, vol. 103, no. 3, pp. 1518–1531
9. Diedrichsen, J., Verstynen, T., Hon, A., Zhang, Y. and Ivry, R. B. (2007) 'Illusions of force perception: The role of sensori-motor predictions, visual information, and motor errors', Journal of neurophysiology, vol. 97, no. 5, pp. 3305–3313.
10. 10. Hirsiger, S., Pickett, K. and Konczak, J. (2012) 'The integration of size and weight cues for perception and action: Evidence for a weight-size illusion', Experimental Brain Research, vol. 223, no. 1, pp. 137–147.
11. Hutchison, D., Kanade, T., Kittler, J., Kleinberg, J. M., Mattern, F., Mitchell, J. C., Naor, M., Nierstrasz, O., Pandu Rangan, C., Steffen, B., Sudan, M., Terzopoulos, D., Tygar, D., Vardi, M. Y., Weikum, G., Isokoski, P. and Springare, J. (2012) Haptics: Perception, Devices, Mobility, and Communication, Berlin, Heidelberg, Springer Berlin Heidelberg
12. Tanaka, M., Misra, A., Oshima, K., Hashiguchi, S., Mori, S., Shibata, F., Kimura, A. and Tamura, H. (2017) '[POSTER] Further Experiments and Considerations on Weight Perception Caused by Visual Diminishing of Real Objects', 2017 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct). Nantes, France, 09.10.2017 - 13.10.2017, IEEE, pp. 160–161
13. Van Polanen, V. and Davare, M. (2015) 'Sensorimotor Memory Biases Weight Perception During Object Lifting', Frontiers in human neuroscience, vol. 9, p. 700.
14. Taylor, B., & Hardcastle, C. (2019). Encyclopedia of Gerontology and Population Aging. In Encyclopedia of Gerontology and Population Aging (pp. 0–4). https://doi.org/10.1007/978-3-319-69892-2
15. Saccone, E. J., Landry, O. and Chouinard, P. A. (2019) 'A meta-analysis of the size-weight and material-weight illusions', Psychonomic bulletin & review, vol. 26, no. 4, pp. 1195–1212.
16. Buckingham, G., Cant, J. S. and Goodale, M. A. (2009) 'Living in a material world: How visual cues to material properties affect the way that we lift objects and perceive their weight', Journal of neurophysiology, vol. 102, no. 6, pp. 3111–3118.
17. Chouinard, P. A., Large, M.-E., Chang, E. C. and Goodale, M. A. (2009) 'Dissociable neural mechanisms for determining the perceived heaviness of objects and the predicted weight of objects during lifting: An fMRI investigation of the size-weight illusion', NeuroImage, vol. 44, no. 1, pp. 200–212
18. Guerraz, M., Provost, S., Narison, R., Brugnon, A., Virolle, S. and Bresciani, J.-P. (2012) 'Integration of visual and proprioceptive afferents in kinesthesia', Neuroscience, vol. 223, pp. 258–268.
19. Luu, B. L., Day, B. L., Cole, J. D. and Fitzpatrick, R. C. (2011) 'The fusimotor and reafferent origin of the sense of force and weight', The Journal of physiology, vol. 589, Pt 13, pp. 3135–3147.

20. Flanagan, J. R., Bittner, J. P. and Johansson, R. S. (2008) ‘Experience can change distinct size-weight priors engaged in lifting objects and judging their weights’, Current biology : CB, vol. 18, no. 22, pp. 1742–1747
21. Miall, R. C., Ingram, H. A., Cole, J. D. and Gauthier, G. M. (2000) ‘Weight estimation in a "deafferented" man and in control subjects: Are judgements influenced by peripheral or central signals?’, Experimental Brain Research, vol. 133, no. 4, pp. 491–500
22. Smith D. (2007) ‘Understanding Muscle Contraction’, pp. 15–38.
23. Sugaiwa, T., Fujii, G., Iwata, H. and Sugano, S. (2011) ‘A methodology for setting grasping force for picking up an object with unknown weight, friction, and stiffness’, pp. 288–293.
24. Grafton, S. T. (2010) ‘The cognitive neuroscience of prehension: Recent developments’, Experimental Brain Research, vol. 204, no. 4, pp. 475–491
25. 25. Lecuyer, A., Coquillart, S., Kheddar, A., Richard, P. and Coiffet, P. (2000) ‘Pseudo-haptic feedback: Can isometric input devices simulate force feedback?’, Proceedings IEEE Virtual Reality 2000 (Cat. No.00CB37048). New Brunswick, NJ, USA, 18-22 March 2000, IEEE Comput. Soc, pp. 83–90.
26. Barnard, P. J. and Teasdale, J. D. (1991) ‘Interacting cognitive subsystems: A systemic approach to cognitive-affective interaction and change’, Cognition & Emotion, vol. 5, no. 1, pp. 1–39.
27. Pusch, A. and Lécuyer, A. (2011) ‘Pseudo-haptics’, Proceedings of the 13th international conference on multimodal interfaces - ICMI ’11. Alicante, Spain, 14.11.2011 - 18.11.2011. New York, New York, USA, ACM Press, p. 57.
28. Freeman, C. G., Saccone, E. J. and Chouinard, P. A. (2019) ‘Low-level sensory processes play a more crucial role than high-level cognitive ones in the size-weight illusion’, PloS one, vol. 14, no. 9, e0222564.
29. Rock, I., Victor, J., Vision and Touch: An Experimentally Created Conflict between the Two Senses, Science, 7 Feb 1964, Vol 143, Issue 3606, pp. 594-596, DOI: 10.1126/science.143.3606.594
30. Ernst, M., Banks, M. Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415, 429–433 (2002). https://doi.org/10.1038/415429a
31. Schwerin, S. (2013) The Anatomy of Movement [Online]. Available at https://brainconnection.brainhq.com/2013/03/05 anatomy-of-movement/ (Accessed 30 April 2020). 32. Davare, M., Montague, K., Olivier, E., Rothwell, J. C. and Lemon, R. N. (2009) ‘Ventral premotor to primary motor cortical interactions during object-driven grasp in humans’, Cortex, vol. 45, no. 9, pp. 1050–1057.

Chapter 29
Design of Mobile Service Robots like Cyber Physical Systems for Pedagogical Rehabilitation, in Special Education

Roman Zahariev, Nina Valchkova, Georgi Angelov, Jasen Paunski

Abstract The article presents the created original method of conduct research into the adoption of robotic technologies and “AnRI” and “BOT” family robots in special education by disabled children and young people, in the same time and pedagogical specialists for the implementation of the EC-FUNDED MARIE CURIE PROJECT H2020-MSCA-RISE-2017, No 77720 "CybSPEED: Cyber-Physical Systems For Pedagogical Rehabilitation In Special Education. The design of the robots created by the “AnRI” and “BOT” families and the requirements to be ready to perform the role of lecturers in the process of educating children with mental disabilities are briefly described. The experiments conducted with representatives of the mentioned families of robots with children from the Day Center St. Sunday, Sandanski are also shown as an illustration. Based on the obtained results, conclusions and summaries are made.

29.1 Introduction

Pedagogy is known to be based on several scientific fields. But in this article, the word "pedagogy" everywhere means the science of the art of raising children. Many types of pedagogical approaches are known, for example: mental, moral, physical, ideological, patriotic, aesthetic, artistic, musical, legal and others. The more developed a society is, the more important the role of Pedagogy in it. This article talks about educating children with mental disabilities with the application of innovative technology through the use of a robot lecturer, but this technology is fully applicable and can be useful in the learning process of young people. But in every approach in the process of education there is something in common, something useful

all authors
Bulgarian Academy of Sciences, Institute of Robotics, “Acad. G. Bonchev” Str. Bl.2, Sofia, Bulgaria roman.zahariev@abv.bg roman.zahariev@abv.bg; nvalchkova@abv.bg; george@robotics.bg; yasen@robotics.bg;

and we can approach it based on human nature, which is the meaning of the concept of education. According to the science of Cybernetics, man is a purposeful system in which his behavior is determined by the set of tasks. The process of successfully pursuing and achieving the set goals, without affecting the previously set limits, is largely determined by the desire and interests of the learner. These interests of the learner are a powerful stimulus in the processes of learning and assimilation of information by pursuing the set strategic task. At the present stage of development of human society, with the introduction of new modern technologies, including Robotics, there is a tendency for increased interest of students, thus illustrating the perception of information in the educational process. In the extremely complex process of educating children with mental disabilities, an important point is the definition of clear activities in the pursuit of tactical goals. The teacher must be aware of the pre-set goals so that he can create interest and desire in the learners to absorb the information based on the set key elements. These elements are usually key words or concepts that make it easier for students to perceive the issues under consideration and create a prerequisite for higher efficiency of the learning process. With the introduction and use of the Cyber physical system as a robot lecturer, it must meet in his behavior the requirements of modern pedagogy. Especially in the process of special education of children with mental disabilities, it is especially important to seek new directions in the application of pedagogical approaches to overcome certain difficulties in the process of learning the information. For example, according to Marvin Minsky's Frame Theory, information can be organized by the robot into separate frames, with each subsequent frame inserted into the previous one clarifying and refining it, and thus the learning process progresses to a deep insight into the problem. The information derived and offered in this way by the robot lecturer will arouse a certain interest in the learners, which will be a powerful stimulus in the process of acquiring knowledge on the way to achieving the set goals. On this basis, the robot lecturer should be designed so that he has the opportunity to communicate appropriately with the learner, to reach his consciousness and to arouse some interest in order to apply a pre-selected approach to implementing the strategy for learning.

29.2 Formulation of Specific Requirements for Mobile Service Robot – Lecturer

This article presents the results of the activities of the CyBSPEED project, which aims after the creation of cyberphysical systems as robot lecturers, to systematically explore the peculiarities of the acquisition of robotic technologies in special education by young learners and teachers, especially in the following aspects:

- The socializing role of robotic technology - there is evidence that robots do not enhance isolation (as might be expected) but instead play the role of social

mediator - robot communication is easy, accessible and fun and helps to strengthen relevant skills.
- Three dimensional objects as a service robot with a certain degree of autonomous behavior and skills for verbal or nonverbal communication with young people is generally a new impetus, the perception of which compared to traditional objects deserves more careful and systematic study in a new theoretical framework.
- The theoretical framework of this project is the system of special school education, in which specialists, teachers, young people and robot designers work together to overcome the main difficulties accompanying the learning process such as distraction, functional illiteracy, lack of cognitive motivation, inadequate social motivation. - both in the regular school and in the special education. For these cases, the Mobile Service Robots for Training series, the ARI series (Anthropomorphic Robots with Intelligence) and the BOT series have been designed and implemented.

Contemporary robots can be characterized by four central features

- mobility, which is important to function in human environments like hospitals and offices;
- interactivity, made possible by sensors and actuators, which gather relevant information from the environment and enable a robot to act upon this environment;
- communication, made possible by computer interfaces or voice recognition and speechsynthesis systems;
- autonomy, in the sense of an ability to 'think' for themselves and make their own decisions to act upon the environment, without direct external control.

29.2.1 Classification of Mobile Service Robots

Mobile service robots for training and servicing disadvantaged people, according to their purpose, can be classified as follows:

- Mobile service robots to serve people with mental health problems.
- Mobile service robots to serve disadvantaged people with physical disabilities.
- Universal Mobile service robots with a high degree of intelligence to serve disadvantaged people.

In turn, mobile service robots to serve people with mental health problems can be divided into the following types:

According to its geometric dimensions:

- Anthropoid type with geometric dimensions close to human.

- Miniature robots with minimal geometric dimensions.

According to the type of mechanisms of movement in space:

- Anthropoid robots with a walking mechanism for movement.
- Robots with a mechanism for moving wheels.

According to the way of contact with the students:

- With eye contact:
  - With information visualization capabilities.
  - With the ability to recognize images using a camcorder and elements of artificial intelligence.
- With speech contact:
  - With speech recognition with elements of artificial intelligence (Voice feedback).
  - With the ability to synthesize speech.

According to the possibilities of controlling the movement of the robot by the servicer or the trainee in a dialog mode.

During the implementation of Cyber physical systems such as Mobile Service Robots for Training at the Institute of Robotics, a series of such robots was created, such as the Anthropomorphic Robots with Intelligence series, and the BOT series, choosing a mechanism for moving wheels.

29.3 Design of Cyber Physical Systems as Service Robots for Training

The general idea when creating a service robot is to use the advantages of modular design. Various ready-made, standard modules are combined to build the whole system [1]. The designer must select the most appropriate components and make the connections between them and organize their communication so as to ensure the operation of the created structure of the robotic system [4]. Each component strongly affects the performance of the system and therefore the choice of individual building components is very important. The price of the total product must also be taken into account according to the set of restrictions imposed. The designer must design the robot's communication with the operator. An example of a graphical user interface has been developed and carefully considered, as it is a decisive factor in the control of the robot and the successful completion of its task. The robot is designed as follows:

- The main components of the robot are defined so as to perform the assigned tasks.
- Design considerations for each part of the robot are taken into account and a ready-made component is selected. For example, the CCD camera, the TV transmitter, the mobile platform, etc. [2].
- The secure connection between the components is created to ensure their interaction and the construction for mounting the camera and the platform manipulator is made. [3].
- The graphical user interface is created based on the capabilities of the selected platform, which will be controlled through it.

The "AnRI-1" and "AnRI-2" robots (Fig. 2 and Fig. 3) are built on the basis of mobile platforms with four wheels, two of which are driven and two have independent movement around the vertical axis so that it can a general movement with a high degree of mobility is realized. In the robot "AnRI-1" the wheels are arranged in the form of a "trapezoid". The driven wheels are on both sides of the platform at the front and the "free" wheels are at the rear. The wheels are arranged in the form of a "cross" in the robot "AnRI-2", and the driven wheels are on both sides of the platform in the middle, and the "free" wheels are one on the back and one on the front. In the robot "AnRI-2", with the help of this arrangement of the wheels, it is possible to realize the movements of the robots around the vertical axis in the geometric center of the structure [10]. In the hub of the driven wheels are built-in electric motors with gearboxes, powered by direct current from a rechargeable battery. The AnRI-1 robot is equipped with an anthropomorphic-type manipulator located on the platform with three regional and three local degrees of mobility and a separate three-finger gripper. The robot "AnRI -2" has the same construction and the difference is in the position of the wheels.

29.3.1 AnRI Service Robots Hardware Platform

The drive of the robots is realized on the basis of "servo" controllers with feedback from incremental sensors located in each degree of mobility of the manipulator. For the manipulators, the regional levels are equipped with electromagnetic brakes, and the drive wheels of the platform are equipped with worm gearboxes that do not allow reversing using their braking effect. The robot has a Control System, which is hierarchical, distributed, microprocessor type and includes different levels, different devices and systems and the corresponding software modules. The connection between all devices is made via the serial interface RS 232 of the control. The control module is based on a 32-bit microprocessor built into the processor module. 32-bit microcontrollers have been widely used in recent years for robot control applications. The most popular and widespread architecture of the Cortex-M processor family is used. It is a range of energy-efficient, scalable and easy-to-use processors and is currently available in six variants: Cortex-M0, Cortex-M0 +, Cortex-M1, Cortex-M3, Cortex-M4 and Cortex-M7 applications.

Fig. 29.1 General View of the AnRI 1 and AnRI 2 Robots.

For example, the AnRI-2 is a service robot that consists of the following components: A mobile robot platform with a differential drive using two DC motors and two free-spinning, centrally oriented wheels built on a modular basis. The robot's electronics use modern 32bit Arm MCU architecture I / O boards, internally connected via CAN bass and connected to an Ethernet network interface. The robot is controlled by a standard computer with Ubuntu Linux OS. The network is implemented using a standard WiFi 802.11 router. The router's integrated switch is used for the internal network between the computer and the I/O boards. A dual-channel amplified audio subsystem is connected to the computer for speech synthesis and multimedia playback. A color LED display attached to the level of the robot's head allows the robot to display a variety of media and emotional states. The robot is powered by a lithium-ion battery with control and charging station.

The Cortex-M4 family is very suitable for robot control. Integrated digital signal processing unit (DSP) with floating point support for fast and energy efficient algorithm. Cortex-M4 can be used in digital control applications, such as sensor information merging, motor control and power management [10]. A control system based on a universal I / O board Cortex-M4 for service robots is being implemented, which provides a set of digital and analog inputs, 10/100 Ethernet interface, PWM control and CAN, SPI and I2C buses. This configuration allows a flexible solution for connecting various sensors and actuators to the robot, using the modern Cortex-M4 architecture, simplifying the design and reducing the number of external components. [4]. The platform has independent control using a standard TCP / IP control [6, 9] network to communicate with the motherboard. Various standard libraries and diagnostic tools can be used for rapid deployment. At the same time, another serial interface module, GPS navigation module and GPRS for Internet communication are presented. It is intended to be used to control the visual module and the laser interferometer module, together with a module for integration, reading and loading data from sensor systems using the ROS (Robots Operating System) operating system [7, 10].

29.3.2 AnRI Service Robot Sensor System

“AnRI" robots are equipped with several types of sensors to monitor the presence of obstacles in the working area of the robot:

- Infrared sensors - 6 pieces located on the periphery of the base with a working range of 10-80 cm.
- Ultrasonic sensors - 2 pieces located in front of the base with a working range of 0.5-3.0 m.
- Tactile "limit switch" sensors located on the periphery of the platform, which are triggered by an obstacle and stop the movement of the robot [8].

Power supply: The robot is powered by a rechargeable LiFePO4 battery with a nominal voltage of 12 V and a capacity of 17 Ah. The battery has a built-in combination protection. Communication between each module is built through a LAN connection (10/100 Mbit).

29.3.3 Implementation of TCP/IP Communication with User Interface for Service Robot in ROS Environment

The development and implementation of service robots often raises the question of how to implement management that is flexible enough to allow the user interface to work remotely and without compromise on reliability, efficiency and safety. The combination of these serious requirements requires the development and use of sufficiently reliable means in the communication layer to provide two-way interaction in real time. The implementation of the communication layer must be based on established standards and practices and meet the requirements for safety in robotics. The main goal is to provide an intuitive, user-friendly interface, secure and reliable connection with minimal network load, optimized for robot control.

To accomplish this task, various options have been considered, some of which have been rejected as a direct connection to the ROS server because they do not meet some of the many requirements. As part of the communication solution it was decided to use ROS Bridge, ROS package. ROS Bridge allows access to ROS through JSON formatted messages with one socket and also through the so-called WEB socket. It uses TCP port (9090) for communication, which is very convenient for remote access, as it can be easily forwarded through network protections (Fig. 3), [1, 3]. WEB socket support is a new feature in the new HTML 5 browsers, which is also an advantage for the WEB-based version of the interface. WEB sockets provide a way to communicate with low latency, which is very important in such an application where real-time response is critical. The current version of ROS Bridge supports creating ROS themes, posting messages, calling ROS services, authentication and security features, and launching a ROS startup file. In Fig. 2 is a block diagram of communication between the user interface and the robot using ROS Bridge. Based on the ROS Bridge stack, it allows network communication between ROS and a remote device running UI_PRI and UI_LOC user interfaces. The design concept allows an open interface for different types of devices based on different operating systems. Standard network communication (TCP / IP) and a widely accepted JSON encapsulation format are used [5].

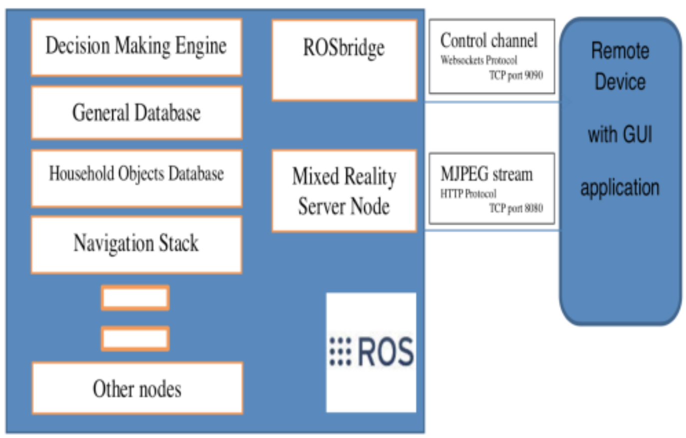


Fig. 29.2 Scheme of The Realized Communication of ROS with User Interface.

29.3.4 Hardware Platform of the Service Robots of "Bot" Family

"MaxiBot" and "BeBot" are Mobile service robots, implemented on a platform for educational robots (Fig. 3). They have a differential drive base with two motor wheels and a 52 mm free-rolling sphere as a third passive wheel. With "MaxiBot" the basic size of the print is 260x260 mm, and the total height of the robot is 540 mm. The platform's control board is a Raspberry Pi 3B + single-board computer with Raspbian Linux and Robot Operating System (ROS) pre-installed. MaxiBot has a built-in LIDAR sensor that allows the robot to create a map and move in space using

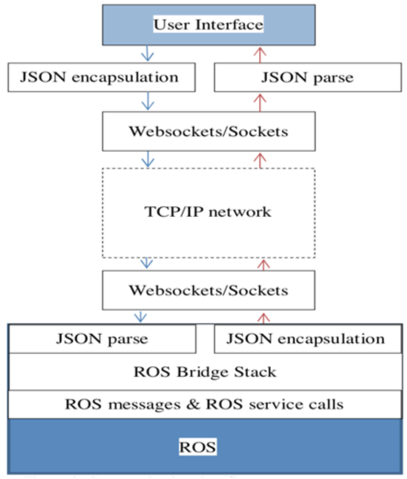


Fig. 29.3 Scheme of the Developed Communication System.

the ROS navigation stack. The robot's head uses a 7-inch touch screen (Raspberry Pi) display for expressing emotions and playing media and an 8MPix CMOS camera. In BeBot, a projector is installed instead of a screen, which allows the information needed for the learning process to be displayed on a non-raised screen installed, for example, on the wall of the room. The platform has a reinforced system of two speakers connected to the control board using a 3.5 mm stereo jack. The power supply of "MaxiBot" is carried out with the help of Li-Ion PD compatible power bank with a capacity of 60Wh, which allows more than 4 hours of autonomy of the robot.

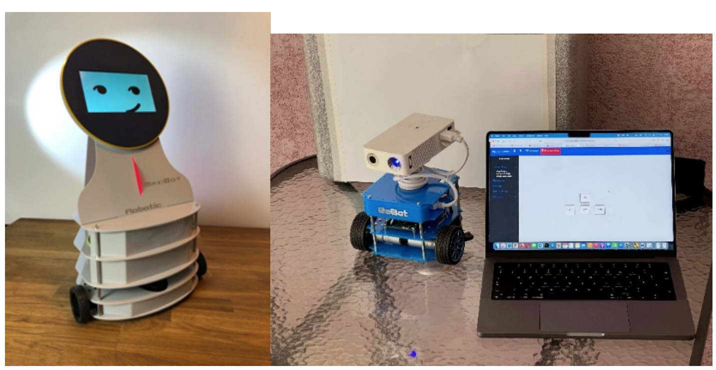


Fig. 29.4 View of Mobile Service Robots "Maxibot" and "Bebot" from The Family "Bot"

29.4 Possible Scenarios for Training Sessions for Children with Disabilities with Mobile Service Robots

All scenarios for the behavior of the Mobile Service Robots during the special training process have been developed in detail together with pedagogical specialists. The Vagatsuma Laboratory at the Kyutech Institute in Japan is a consultant in optimizing the robot's sensor information system so that maximum information can be obtained from and about the learner. This is possible on the basis of the operating system "ROS" used in the robot control system of the families "AnRI" and "BOT". The goal is for the Robot to be able, through the optimized information from the sensory system, to perform movements and functions that meet the desires and needs of both the trainee and the specialist educator. Possible test scenarios were considered and implemented together with educators by improving the good practices of the special learning process of young people, such as:

- Experiment 1
  - Digits Game Learning with BeBot
  - A kid-student teaches summing the digits from 1 to 10 assisted by the robot BeBot.
  - Cyber-physical system consists of Robot BeBot, Choreography Software Framework, Computer Network, with help of Pedagogue.
- Experiment 2
  - Robot Driving Game with BeBot and ANRI2
  - A game that enhance student solving skills, creative thinking, and fine-motor skills Cyber-physical system consists of Robots ANRI2, BeBot, Choreography Software Framework, Computer Network, with help of Pedagogue and Engineer.

Of course, such scenarios can be written at the suggestion of the teachers of the Day Center "St. Sunday", who are working on testing the created Mobile service robots, depending on the specific requirements and desires in the process of special training. During the experiments, the overall behavior of the robot and the trainees was recorded with a video camera for further detailed analysis of the training. Based on it, appropriate conclusions have been made to improve the design of the robot using the achieved results.

29.5 Conclusions

Making service robots accessible to humans is the main challenge of the SybSPEED project, which must be achieved by implementing innovations in all aspects of the operation of the robotic device - materials, sensors, cognitive, communication, drives, energy consumption, etc. The question is to define a clear path of work and tactical

Fig. 29.5 Experiment 1: Robot Driving Game with BeBot and ANRI2.

Fig. 29.6 Experiment 2: Digits Game Learning with BeBot.

goals in the extremely complex learning process with the help of Service Robots. The hardware and software systems for robot control must be further developed on the basis of the behavioral models of robots studied in special conditions. This can provide the necessary basis for further development of service robots in order to better meet the training needs of people with mental disabilities. Advances in modern technology in this regard develop and increase the capabilities of the equipment used to control robots in "real time". It is known that not only the rapid development of digital technologies has led to an unprecedented growth of communication tools in society. The mechanical design of service robots is also a very important aspect. Improving the process of controlling mobile robot systems allows to realize the ability of the robot to move in a stochastic changing environment. The behavior of the robots must be stable and easy to control. The required number of degrees of mobility must be determined to ensure service at each point of the Mobile Service Robot workspace. These two processes of development of the hardware and software part of the service robots must go in parallel and iteratively with the development of one part, which implies a leap in the development of the other part. This in turn sets requirements for new opportunities, which catalyzes the development of Robotics.

Acknowledgement

The authors acknowledge the financial support of the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska – Curie Grant Agreement No 777720 (H2020-MSCA-RISE-2017 for project "Cyber-Physical Systems for Pedagogical Rehabilitation in Special Education" (CybSPEED).

References

1. Angelov, G., et al.: Remote Interface Communication to ROS Based Robotic System. In: Proceedings of the Twenty First International Conference Robotics and Mechatronics; "Invited Session - Austrian-Bulgarian Automation Day"; ISSN 1310-3946; Varna Bulgaria; pp. 22 – 27. (2011). https://www.acin.tuwien.ac.at/fileadmin/acin/ijaa/ijaajg20h2.pdf
2. Cheah C.C., Y. Zhao, J. Slotine : Adaptive Jacobian Motion and Force Tracking Control for Constrained Robots with Uncertainties. In: Proc. IEEE International Conference on Robotics and Automation, 2006, pp. 2226–2231. (2006). LNCS Homepage, https://ieeexplore.ieee.org/abstract/document/1642034
3. Chivarov N, Paunski Y., Ivanova V., Vladimirov V., Angelov G., Radev D., Shivarov N.: Intelligent Modular Service Mobile Robot Controllable via Internet, In: Proceedings IFAC International Conference " SWIIS 2012", June 11th -13th, Waterford, Ireland, pp. 149-53. (2012). LNCS Homepage, https://doi.org/10.3182/20120611-3-IE- 4029.00029
4. Dimitrova M.,Vegt N., & Barakova E.: Designing a System of Interactive Robots for Training Collaborative Skills to Autistic Children. In: Proceedings Int. Conference of Interactive Collaborative Learning ICL2012, IEEE, Xplore, Villah, Austria, pp.1-8. (2012).
5. Kaburlasos, V.G., Dardani, Ch., Dimitrova, M., & Amanatiadis, A.: Multi-Robot Engagement in Special Education: A Preliminary Study In Autism. In: Proceedings of the 36-th IEEE International Conference on Consumer Electronic (ICCE), Las Vegas, USA, pp. 995-996. (2018)., https://ieeexplore.ieee.org/abstract/document/8326267
6. Khatib 0., P. Moore: Sensor Based Control of Robots. Stanford University Publisher. Stanford, USA, (1997). https://ieeexplore.ieee.org/abstract/document/6402179
7. Serpanos, D.: The Cyber-Physical Systems Revolution. J. Computer, V.51(3), 70–73. (2018). LNCS Homepage, https://doi.org/10.1109/MC2018.1731058
8. Valchkova N., R. Zahariev: Optimization of Model Operator for Service Robot, Intended to Service Persons with Disability. In: Proceedings of "TECIS-19" IFAC Conference on International Stability, Technology and Culture, Sozopol, Bulgaria. (2019). LNCS Homepage, https://doi.org/10.1016/j.ifacol.2019.12.468
9. Zahariev R, N. Valchkova, G. Angelov, Y. Paunski, and A. Krastev: Robots for Help in Pedagogy and Rehabilitation. In: Proceedings of 20-th International Conference on Computer Systems and Technologies (CompSysTech'19), ACM, New York, NY, USA, 1-7. (2019). LNCS Homepage, https://doi.org/10.1145/3345252.3345253

10. Zahariev R., N. Valchkova: Existing Robotics Technologies for Implementation of Special Education. In book: M. Dimitrova & H. Wagatsuma (Eds.) Cyber-Physical Systems for Social Applications, IGI-Global, ISBN13:9781522578796, Pennsylvania, USA (2019).

[english]article [T1]fontenc [latin9]inputenc textcomp graphicx babel

Chapter 30
Study of The Ethics in The Implementation of Collaborative Robots in The Training of Disadvantaged People

Nina Valchkova, Roman Zahariev, Georgi Angelov, Jasen Paunski,

Bulgarian Academy of Sciences, Institute of Robotics, “Acad. G. Bonchev” Str. Bl.2, Sofia, Bulgaria nvalchkova@abv.bg; roman.zahariev@abv.bg;george@robotics.bg; yasen@robotics.bg; **Abstract** T

his article analyzes the ethical issues identified in the implementation of collaborative robots (cobots) in the training of disadvantaged people and examines how this is changing the practices of both trainees and staff of social workers and psychologists. With the introduction of cyber-physical systems, such as cobots, changes in human values occur in people’s lives and there are ethical consequences for those involved in this process. Presenting a number of dilemmas that arise during the training raises questions about the extent to which ethics in relationships can be regulated and predetermined in the processes of robot implementation and the resulting reconfiguration of work with robots.

30.1 Introduction

Robotics ethics is a relatively new aspect of analysis focused on the ethical aspects of robotics design and implementation, but Isaac Asimov’s Three Laws of Robotics also delineates its ethical boundaries.

- First Law: A robot may not injure a human being or, through inaction, allow a human being to come to harm.
- Second Law: A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
- Third Law: A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

Robot ethics combines insights from experts in robotics, AI, computer science, and engineering with insights from experts in philosophy, law, psychology and sociology, in an attempt to ensure that automation designs and deployments do not create ethical hazards for individuals and society.

The development of robotics and its implementation through service robotics in the social sphere, in areas outside the industry, cause widespread social and economic changes. This opens up new social and ethical issues for which designers, constructors, manufacturers, end users in the public and private sectors must be prepared. As the field of application for robots (cobots) expands and the robot leaves the factory halls, new challenges are seen and even a paradigm shift is formed due to new human-robot interactions, including social interactions, healthcare, human communication leading to challenging privacy. Robotics deals with the ethical aspects of the design, development and implementation of intelligent machines, such as collaborative robots. Robotic problems are new and not fully understood, their study requires transdisciplinarity to find an original solution, which requires the contribution and efforts of all scientific disciplines, not only those working in the field of robotics, but also in law, sociology, psychology, anthropology and philosophy. Topics related to robotic problems include, but are not limited to, the following:

- Economy (replacement of people in the workplace; use of robotics and labor market). • Psychology (position of people in the hierarchy of management; robots and children).
- Law (robots and responsibility; implementation of autonomously operating robots).
- Health (robotics in surgery; robotics in healthcare, care, prosthetics and therapy).
- Military application of robotics (acceptability, advantages and risks).
- Environment (e.g. cleaning up nuclear and toxic waste).
- Service (social robotics, personal assistants).
- Technical reliability (availability; reliability; safety; security).

In recent years, various robotic systems have been developed to support, use and promote skills in social interaction and mediation of people with and without cognitive and / or physical disabilities. The specialists and disadvantaged participants in the trial had the task to determine the requirements of the users and to formally assess them by the specialists, as well as by the guardians who will support the patient in the process of recovery and social adaptation. The introduction of service mobile collaborative robots for social assistance to disadvantaged people is increasingly used both in medicine and for the purposes of training this category of people.

30.2 Advantages of Collaborative

Collaborative robots are a natural evolutionary branch of robotics. They save valuable space, making them suitable for implementation in even the most spatially limited applications. At the same time, it is becoming easier to program and adapt such a robot to a specific task, even to the most specific processes, without the need for a specially trained programmer [5]. Configuration can be done quickly and easily by anyone via the robot's touch screen.

Among the most serious arguments in favor of collaborative robots is safety. These robots have integrated sensors, soft and rounded surfaces and a number of other ways to minimize the risk of hitting, squeezing or crushing the user. The most important feature of collaborative robots in terms of safety is the limitation of force in the movable joints (hinges, joints). In addition, the robot is usually programmed to stop immediately when in contact with a person or by shortening the distance to him too much to avoid any possible collisions and injuries.

The undeniable benefits of collaborative robots are:

- Quick installation. Collaborative robots can be installed by untrained personnel and put into operation in just a few hours.
- Easy programming. The innovative technology of collaborative robots allows anyone without prior training or experience to program the system quickly and seamlessly using intuitive tools and logical steps through an easy-to-use software programming wizard.
- Improved safety. Collaborative robots, in turn, can work alongside humans, without safety equipment and without the need for large dedicated spaces for their installation.
- Flexible adaptation to multiple applications. Collaborative robots are light, small, space-saving and suitable for deployment in different areas of technological lines, and can be easily moved. Able to reuse programs for repetitive operations, they require minimal time and effort to reconfigure.

30.3 Cobots and Ethics

30.3.1 Ethical Challenges

Robotic technologies, such as the introduction of cyber-physical learning systems [4, 11], blur the line between human subjects and technological objects. In doing so, they not only have social consequences that can be ethically assessed, but also affect the central categories of ethics: our concepts of freedom and responsibility, and our value frameworks. Given the growing autonomy of robots, the question arises as to who exactly should be ethically and / or legally responsible for the behavior of robots. There usually seems to be a "shared" or "shared" responsibility between robot designers, engineers, programmers, manufacturers, investors, vendors and consumers, but this trend completely blurs the notion of responsibility: if everyone is involved in shared responsibility, no one bears full responsibility. When considering recommendations on robotics ethics, it is important to distinguish between deterministic and cognitive robots. In the determined case, the robot's behavior is determined by the program that controls its actions and the responsibility for the robot's actions is therefore clear and regulation can largely be regulated by law. In cognitive robotics, robot decisions and actions can only be statistically estimated and are therefore unpredictable. The responsibility for the robot's actions is unclear, and its behavior in environments other than those it "experienced" during training, such as essentially "accidental" ones, can be potentially dangerous. The scheme [2]. is illustrated in the table below (Table1). While the proposed structure is simple, its implementation in terms of assigning accountability and regulating use is complex and challenging – for scientists and engineers, policy makers and ethicists alike.

With developments in advanced computing, the concept of computers demonstrating artificial intelligence arose. Although this term may be subject to different interpretations, and some may regard as implying real human-like intelligence, AI-based machines can demonstrate human-like sensory ability, language, and interaction. Moreover, these machines can show human-like learning capability, which is being improved – and will continue to be further improved – by employing deep learning techniques. These developments are leading to what may be termed cognitive robotics, such as mobile service collaborative robots. This is a field of technology involving robots that are based on cognitive computing and therefore can learn from experience, not only from human teachers, but also on their own, thereby developing an ability to deal with their environment on the basis of what has been learned. Compared to 'traditional' or deterministic robots, cognitive robots can make decisions in complex situations, decisions that cannot be predicted by a programmer.

Table 30.1 The responsibility for the robot's actions [2].

Decision by Robot	Human Involvement	Technology	Responsibility	Regulation
Made out of finite set of options, according to preset strict criteria	Criteria implemented in a legal framework	Machine only: deterministic algorithms/ robots	Robots' producer	Legal (standards, national or international legislation)
Out of a range of options, with room for flexibility, according to a preset policy	Decision delegated to robot	Machine only: AI- based algorithms, cognitive robots	Designer, Manufacturer, Seller, User	Codes of practice both for engineers and for users; Precautionary Principle
Decision s made through human machine interaction	Human controls robot's decisions	Ability for human to take control over robot in cases where robot's actions can cause serious harm or death	Human beings	Moral

Collaborative robots have a sensor-information system that can be assigned responsibility for their actions, taking into account the degree of their autonomy in the stochastic environment. Avoiding the potential paralyzing effect of this difficulty in taking on and attributing responsibility is a significant challenge for robotics ethics. Taking responsibility anyway is a decision for which:

- techniques for predicting the impact of robotic development can and should be developed as much as possible [8, 9, 10];
- to work carefully with the inevitable occurrence of unexpected consequences, given that the introduction of robotic technologies is a "social experiment" that must be carried out with great care and attention [6, 7]. Because of their ability to act autonomously, robots also challenge our notion of agency.

The main question is how cobots change human practices and how the quality of the human-robot relationship can guide the design, implementation and use of robots. One way to address the issue of moral freedom is offered by the emerging discipline of machine ethics, which aims to equip machines with ethical principles or procedures for resolving ethical dilemmas so that they can function in an ethically responsible way. Robots eventually become morally valuable, as devices designed to perform specific tasks, and cobots deserve moral respect and immunity for moral rights. Service mobile collaborative care robots can also change people's views on care, for example the use of cobots can have a positive effect on our criteria for good education and social care.

These impacts responsibly require a careful balance between expectation and experimental results, closely following the impact of robotic technologies on value frameworks, in small experimental settings. This is important in order to be able to take into account design decisions, public discussions and policy making for social care and training through service mobile collaboration robots (cobots).

Cognitive Theories for Socially-Competent Robotics in Education [1]. Considers three aspects of learning, namely:

- Ethics in using a robot with cognitive and social abilities as an assistant to the teacher; - ethical issues of the use of robot assistants to the teacher
- Ethics in using a robot with cognitive and social abilities in one-to-one communication; - ethical issues of the use of robots that communicate with the child
- Education of children in ethics towards the robots; - teaching children ethics about robots.

An interesting and innovative approach is Multi-robot engagement in special education for a preliminary study in autism [3], where to avoid ethical issues when the robot communicates with the child will use another robot trained to behave like a child, and tests the robot's behavior towards the child-imitating robot.

30.3.2 Ethical Declaration

In this article, we present a new approach that seeks to help disadvantaged children develop their skills with the help of the cobwebs from the Henry and Bot series. Using a robot to train these children has a clear advantage, as evidenced by the fact that what the robot can do or see can be shown directly to children and affect their emotional and physical condition. Robots also have an emotional impact on social educators and psychologists working with disadvantaged people, who see cobots as a helper in their work to more quickly and easily cope with the difficult task of social inclusion of these adolescents. This study was approved by the Ethics Committee of the Institute of Robotics at the Bulgarian Academy of Sciences for research involving disadvantaged people, written consent was obtained from all participants in the study, as well as from the parents of the children.

30.4 Experiments

The team of robotics specialists has developed and implemented games with service collaborative robots from the "Henri" and "Bot" series, through which to study the impact of collaborative robots in their implementation in the learning process for disadvantaged people. The aim is to improve the skills of this category of people and accelerate the process of socialization in society. Preliminary and ex-post evaluations are described in detail, through inclusion criteria, taking into account the progress in socialization.

Participants in this study differed in age, which led to differences in both their abilities and personal interests. This is a study examining whether robots can be used to help disadvantaged children, so we did not have an inclusion criterion and therefore selected a wide range of children from the age range, but found that selection based on an IQ and ADOS as estimates comparable to the results before and after the experiments. In conclusion, children who could benefit from learning with collaborative mobile robots are in nonverbal mental age ranges from 6 to 14 years, and their level of autism spectrum symptoms is "moderate" to "high." In this study, we found that children who met this criterion had a significant improvement in the results of games that had some beneficial effect on their cognitive abilities in the following experiments.

Experiment 1 (Fig.1)

Digits Game Learning with BeBot.

A kid-student teaches summing the digits from 1 to 10 assisted by the robot BeBot.

Cyber-physical system consists of Robot BeBot, Choreography Software Framework, Computer Network, Pedagogue.

Experiment 2 (Fig.2)

Robot Driving Game with BeBot and ANRI2.

A game that enhance student solving skills, creative thinking, and fine-motor skills

Cyber-physical system consists of Robots ANRI2, BeBot, Choreography Software Framework, Computer Network, Pedagogue, Engineer.

Experiment 3 (Fig.3)

Repeat Game with humanoid Marty

A game that teach student different poses and gestures. The goal is to improve fine-motor skills.

Cyber-physical system consists of Humanoid Robot Marty, Choreography Software Framework, Computer Network, Pedagogue.

Other possible scenarios are:

- n the learning process, the robot shows different geometric shapes (cube, cone, sphere, etc.), and learners recognize and name the figure. If the answer is correct, the robot will sound affirmative by saying the word "true". If the answer is wrong, the robot's answer will be heard with the word "No". The robot can count the positive and negative answers and make an estimate.
- The robot can also perform a popular song by conducting a singing lesson. After the song is over, he will invite the students to sing it with him. In the end, he will let them do it themselves, praising them for their good performance.
- Based on the innovative approach of KyuTech University - Japan for the transmission of information on "Keywords", the AnRI2 robot will offer students a short text, highlighting the keywords in it. He will then go through with invitations for encouragement to retell what has been said, as there is an opportunity to assess the degree of learning information from students and the opportunity to implement their assessment.
- Another possible scenario is the delivery of a newspaper or book from the library shelf. The action involves, after talking with the students, bringing the newspaper or book from the library shelf and placing it on the table in front of the robot so that he can pick it up and use it as intended.

Fig. 30.1 Exp.1: Digits Game Learning with BeBot.

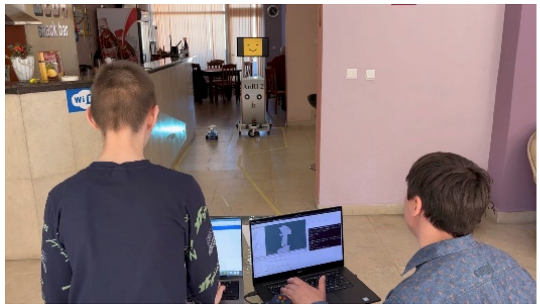


Fig. 30.2 Exp.2: Robot Driving Game with BeBot and ANRI2

Fig. 30.3 Repeat Game with Robot Marty.

Fig. 30.4 Playing with Robots AnRI and Marty.

30.5 Analysis of The Results of The Conducted Training

The following results are achieved with the CybSPEED project:

- Increasing the flexibility and usefulness of robot assistants while respecting the autonomy and privacy of users.
- Supporting the idea and encouraging the future use of robot assistants.
- Research of the requirements and possibilities of the market for robots - assistants in Europe.

As this pilot study is the first of its kind for the team, there are a number of lessons that were learned during the experiments. The project provides an opportunity to exchange experience, knowledge and good practices between its partners, in order to explore the possibilities of using mobile service collaborative robots for the needs of education and social inclusion of disadvantaged people. Reducing social exclusion, removing barriers and barriers and using information technology are among the most important goals in this area. One possibility for this is the introduction of cognitive modular service mobile robots AnRI and Bot (Fig. 5).

Fig. 30.5 Cognitive Mobile Service Robot in Action.

The analysis of the questionnaires from the experimental use of the robot shows the following:

- The conducted training is very well organized, both in terms of the content of the material and in terms of the way it is served.
- The prevailing assessment of the appearance of the robot is very good, as some of the recommendations are related to reducing the size (height) of the manipulation and increasing the range of the robot.
- At this stage, the set functions of the robot are defined as sufficient, but the recommendations indicate new ideas for their expansion.
- A positive evaluation of the experiments was given by the majority of the participants.
- For the possibility to use a robot to help disadvantaged people, a large part of the participants consider it appropriate, giving an excellent assessment.

30.6 Conclusion

The opportunity provided for the study of European experience and sustainable cooperation between the beneficiaries and partners of the project "CybSPEED" will contribute to the introduction of an innovative model for Bulgaria as a whole. Generating knowledge and know-how in the field of pedagogical rehabilitation of people with disabilities will help to improve existing technical tools in this area and identify the needs of the target group of various Mobile Service Collaborative Robots - Cobots. Service mobile collaborative care robots can also change people's views on care, for example the use of cobots can have a positive effect on our criteria for good education and social care.

In view of the specific resource of IR-BAS, which will implement innovative European practices, we believe that the project includes not only an innovative element, the use of robots "AnRI" and "BOT", but also offers innovative, new solutions to social problems , which will be realized with the good cooperation with the professional team of the Institution for social services "Day Centre "St. Sunday", Sandanski, BG. The study showed that some disadvantaged people can benefit from learning through the use of collaborative robots, and this approach to learning needs to be developed, as progress is palpable, especially among adolescents. Research, testing of cobots in disadvantaged people and exchange of knowledge will provide Bulgarian partners with completely new tools and forms of intervention and will contribute to improving existing practices.

The opportunity provided for the study of European experience and sustainable cooperation between the beneficiaries and partners of the project "CybSPEED" will contribute to the introduction of an innovative model for Bulgaria as a whole. Generating knowledge and know-how in the field of pedagogical rehabilitation of people with disabilities will help to improve existing technical tools in this area and identify the needs of the target group of various Mobile Service Collaborative Robots - Cobots.

In view of the specific resource of Institute of Robotics-Bulgarian Academy of Sciences, which will implement innovative European practices, we believe that the project includes not only an innovative element, the use of robots "AnRI" and BOT, but also offers innovative, new solutions to social problems , which will be realized with the good cooperation with the professional team of the Institution for social services "Day Center" St. Sunday ", Sandanski. The study showed that some disadvantaged people can benefit from learning through the use of collaborative robots, and this approach to learning needs to be developed, as progress is palpable, especially among adolescents. Research, testing of cobots in disadvantaged people and exchange of knowledge will provide Bulgarian partners with completely new tools and forms of intervention and will contribute to improving existing practices.

Acknowledgement

The authors acknowledge the financial support of the European Union's Horizon 2020 Researchand Innovation Program under the Marie Skłodowska – Curie Grant Agreement No 777720 (H2020-MSCA-RISE-2017 for project "Cyber-Physical Systems for Pedagogical Rehabilitation in Special Education" (CybSPEED).

References

1. Dimitrova, M. (2016). Cognitive Theories for Socially-Competent Robotics in Education. Lambert Academic Publishing.
2. Corporate author, 2017. Report of COMEST on robotics ethics:World Commission on the Ethics of Scientific Knowledge and Technology [145], Document code:SHS/YES/COMEST-10/17/2 REV.https://doi.org/10.1016/j.ifacol.2019.12.468
3. Kaburlasos, V.G., Dardani, Ch., Dimitrova, M., & Amanatiadis, A.: Multi-Robot Engagement in Special Education: A Preliminary Study In Autism. In: Proceedings of the 36-th IEEE International Conference on Consumer Electronic (ICCE), Las Vegas, USA, pp. 995-996. (2018). LNCS Homepage, https://ieeexplore.ieee.org/abstract/document/8326267
4. Serpanos, D.: The Cyber-Physical Systems Revolution. J. Computer, V.51(3), 70–73. (2018). LNCS Homepage, https://doi.org/10.1109/MC2018.1731058
5. Valchkova N., R. Zahariev: Optimization of Model Operator for Service Robot, Intended to Service Persons with Disability. In: Proceedings of "TECIS-19" IFAC Conference on International Stability, Technology and Culture, Sozopol, Bulgaria. (2019). LNCS Homepage,
6. Van de Poel, I. 2013. Why New Technologies Should Be Conceived as Social Experiments, Ethics, Policy & Environment, Vol. 16, No. 3, pp. 352-55.
7. Van de Poel, I. 2016. An Ethical Framework for Evaluating Experimental Technology,
8. Verbeek, P. P. 2013. Technology Design as Experimental Ethics. In: Van den Burg, S. and Swierstra, Tsj. eds. Ethics on the Laboratory Floor. Basingstoke, Palgrave Macmillan, pp. 83-100
9. Waelbers, K. and Swierstra, Tsj. 2014. The family of the future: how technologies can lead to moral change. In: Hoven, J. V. D., Doorn, N., Swierstra, T., Koops, B-J. and Romijn, H. eds. Innovative Solutions for Global Issues. Dordrecht, Springer, pp. 219-236
10. Zahariev R, N. Valchkova, G. Angelov, Y. Paunski, and A. Krastev: Robots for Help in Pedagogy and Rehabilitation. In: Proceedings of 20-th International Conference on Computer Systems and Technologies (CompSysTech'19), ACM, New York, NY, USA, 1-7. (2019). LNCS Homepage, https://doi.org/10.1145/3345252.3345253
11. Zahariev R., N. Valchkova: Existing Robotics Technologies for Implementation of Special Education. In book: M. Dimitrova & H. Wagatsuma (Eds.) Cyber-Physical Systems for Social Applications, IGI-Global, ISBN13:9781522578796, Pennsylvania, USA (2019).