International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-8 Issue-6, March 2020

Transfer Learning with Pretrained Neural Network Between Unrelated Tasks for Machine Health Diagnosis

Youssef Maher, Boujemaa Danouj

Abstract: ***Deep Learning (DL) has contributed a lot in the field of industrial maintenance, in particular predictive maintenance by detecting potential failures and breakdowns before their appearance. Unfortunately, the DL has some limitations like the need for a large amount of data to produce an effective prediction model and also the fragility of the model in the face of changes in operating conditions. Another approach, the Transfer Learning (TL), had demonstrated in the literature that he can overcome these weaknesses. In this article, we will be using this technique with the pretrained neural network, AlexNet, which had been previously trained with the ImageNet database. Our method doesn't require a high amount of input data and thus saves a lot of time in retraining the network in another task, which can be related or unrelated to the source task. In fact, the prediction model was successfully adapted to the bearings diagnosis case. It showed also high degree of robustness against changes of functioning conditions.***

Keywords : ***AlexNet, Machine health diagnosis, Pretrained neural network, Transfer learning.***

I. INTRODUCTION

The new industry revolution, industry 4.0, is now taking place. The Machine Condition Monitoring System (MCMS) plays a key role in maintaining an optimal functioning health for the machines.

The predictive-type maintenance is now progressively implemented to all aspects of the industry 4.0. The detection of any imminent defect before its appearance ensures a minimum or no impact costs and operation on the production lines. In addition, it allows the manufacturer to reduce direct and indirect costs of maintenance and of repair by avoiding the untimely unavailability of the production tool by removing premature shutdown for inspections and visits, planning at best the interventions and finally reducing their duration and their extent. This type of maintenance has become increasingly efficient thanks to its data-driven mode of operation; in fact more data is used and more powerful and reliable is the prediction. Currently, the way to get the most out of the data is by using the deep learning, which extracts the necessary information from the raw data to create a powerful health predictive model that is done quickly and efficiently without the need for any human intervention.

However, the deep learning still suffers from two major weaknesses:

- The need for a large amount of data. If the labeled data is unavailable or in very small quantity, the prediction model won't be efficient and thus unreliable.
- Difficulty to adjust to operating condition changes. This problem arises when the labeled target data (for a specific operating condition like for a rotation speed) is unavailable. So if the source data domain is different from the target data domain, the prediction model will no longer be correct.

The industrial environments are highly changing. That's why the new data analysis techniques in the real industrial world must take into account non-stationary environments in order to carry accurate predictive diagnostics.

One widely used method to enable the adaptation to these new changes is transfer learning. This latter allows to extract the learned characteristics of a domain A and to adapt them to the new domain B and thus perform the results with a significant time saving and minimal data in the target domain.

After identifying the source and the target domain, the TL adapts in record time to the new operating conditions like the changing load or speed conditions or even the aging effects on sensors. This process requires much smaller input data, time and energy than traditional approaches. With traditional machine learning, the knowledge is not retained or accumulated. With the TL, the learning of the new tasks relies on the previous learned tasks. And so the learning process is faster, more accurate and needs less training input data.

The TL learns invariant representations from source samples and adapts to the new distribution in the target domain. The methods used are feature transfer, fine-tuning and freezing the first layers.

There are several TL methods depending on the similarity between the source and the target.

Manuscript received on February 10, 2020.
Revised Manuscript received on February 20, 2020.
Manuscript published on March 30, 2020.
* Correspondence Author
Youssef Maher*, Department of Electrical and Mechanical Engineering, Univ. Hassan 1, Settat, Morocco. Email: maher-youssef@hotmail.com
Boujemaa Danouj, Department of Electrical and Mechanical Engineering, Univ. Hassan 1, Settat, Morocco. Email: bdanouj@yahoo.fr

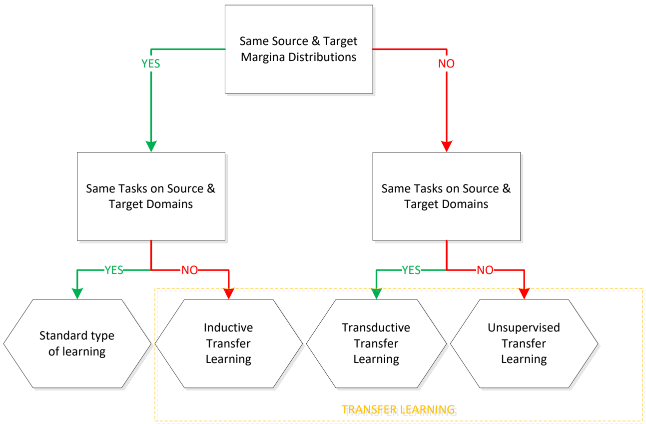

Fig. 1.Different types of transfer learning

Based on the availability of labeled samples, transfer learning can be divided into three categories as seen in Fig. 1 above:

- Inductive transfer learning which requires the availability of labeled target data.
- Transductive transfer learning is used when labeled source data are available while labeled target data do not exist.
- And unsupervised transfer learning remains the most appropriate approach when labeled target and source data are not available.

In our chosen approach, we will use the inductive transfer learning method because the labels for source and target data are available.

The transfer learning had been extensively studied in "in press" [1] who did a survey on the subject of knowledge transfer by showing its opportunities and benefits for manufacturing data and also the state of the art achieved in this subject. They showed how the knowledge transfer allows a fast adaptation to a new environment. They pointed out also its flexibility because it offers several TL methods that differ according to the availability of labels and the type of changes in data properties. These methods are categorized into two types: using knowledge from data and using knowledge from models.

But transfer learning has drawbacks which is the negative transfer that emerges when the source and target data are not closely related, leading to a performance deterioration of the classifier.

Yao and Doretto, in "in press" [2], tackled this problem by proposing a new algorithm, MultisourceTrAdaBoost, that exploits and extracts the knowledge from multiple sources in order to increase the chances to find one with a beneficial knowledge closely related to the target domain.

Traditionally, transfer learning is used primarily to enable the training without overfitting on small target datasets but this new finding shows that transferring features boost generalization performance even if the target dataset is large.

In "in press" [3], the authors used a deep transfer learning method based on three transfer strategies, namely weight transfer, hidden function transfer learning and weight update, demonstrating its effectiveness in predicting the RUL.

"In press" [4] proposes a new fault diagnosis method by using a digital assisted body production line using deep transfer learning. The concept of digital twin is fully compatible with the corresponding physical entity.

The previously trained diagnostic model can be transferred by DL from the virtual space to a real physical space for real-time monitoring and predictive maintenance.

In the "in press" [5], they present a 2-step transfer learning approach based on a deep convolutional neural network. The first part is built with a pre-trained deep neural network used to automatically extract input features: ImageNet's large image data. And the second part is a fully connected layer used for classifying the features to be trained using experimental data on gear defects.

In our study, we will use a pretrained neural network AlexNet with data from general object categories database and we will adapt it to our health diagnosis topic by using only very little training data.

This paper is organized into three chapters as follows:

Chapter 2 will describe our approach,

Chapter 3 will describe the results,

Chapter 4 will conclude, at the end of this article, on the multiple benefits of our method by using the transfer learning in achieving high performance results and time savings.

II. PROPOSED METHOD

To exploit the potential of transfer learning, we will be using the pretrained neural network AlexNet.

The Fig. 2 below shows a simplified illustration of the considered network model. It is a type of convolutional neural network containing eight layers; the first five are convolutional layers, some of them followed by Max-pooling layers, and the last three were fully connected layers. It uses the non-saturating ReLU activation function.

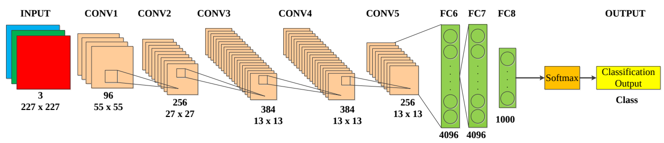

Fig. 2.Simplified illustration of AlexNet architecture

The Table- I. below displays some of the characteristics of the network.

Table- I: Characteristics of the Alexnet network

Network	Depth (Number of hidden layers)	Parameters (Millions)	Image Input Size
AlexNet	8	61.0	227-by-227

AlexNet is trained on over a million images and can classify images into 1000 object categories (such as keyboard, coffee mug, pencil, and many animals). The network has learned rich feature representations for a wide range of images. The network takes an image as input and outputs a label for the object in the image together with the probabilities for each of the object categories.

So we will be exploiting and using this accumulated knowledge in our work and in particular the initial learned features. In our work, we used the open dataset from the Case Western Reserve University Bearing Data Repository [6].

The Fig. 3 below presents the experimental rig. The procedure was conducted using:

- 2 hp Reliance Electric motor (left),
- a torque transducer/encoder (center): it collects the speed and horsepower data,
- a dynamometer (right),
- and control electronics (not shown).
- Bearings: SKF bearings were used for the 7, 14 and 21 mils diameter faults, and NTN equivalent bearings were used for the 28 mil and 40 mil faults.
- Accelerometers: They are attached to the housing with magnetic bases and were placed at both the drive end and fan end of the motor housing.

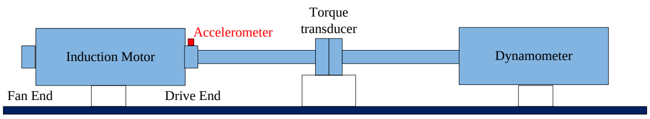

Fig. 3.Schematic of experimental setup

The actual test conditions of the motor as well as the bearing fault status have been carefully documented for each experiment: The faults in the motor bearings were generated by using electro-discharge machining (EDM). Faults ranging from 0.007 inches in diameter to 0.040 inches in diameter were introduced separately at the inner raceway, rolling element (i.e. ball) and outer raceway. The vibration data was recorded for motor loads of 0 to 3 horsepower (motor speeds of 1797 to 1720 RPM).

Outer raceway faults are stationary faults, therefore placement of the fault relative to the load zone of the bearing has a direct impact on the vibration response of the motor/bearing system. In order to quantify this effect, experiments were conducted for both fan and drive end bearings with outer raceway faults located at 3 o'clock (directly in the load zone), at 6 o'clock (orthogonal to the load zone), and at 12 o'clock.

All data files are in Matlab (*.mat) format. Digital data was collected at 12,000 samples per second, and data was also collected at 48,000 samples per second for drive end bearing faults. In our experiment, we will be using the 48,000 samples per second one.

In our work, we will be using the data signals for the 4 different conditions: Normal, Inner Race fault, Ball fault and Outer Race fault (sensor centred at 6:00 only) in 4 different motor loads (0, 1, 2 and 3 HP) corresponding respectively to the following motor speeds (rpm) (1797, 1772, 1750, 1730). We will use only a fault diameter of 0.007" inches.

For each of the 4 conditions signals and for 4 different speeds, we generated several spectral representation images: spectrogram and scalogram, as presented in the two figures, Fig. 4 and Fig. 5, below. These images give the possibility to know whether there is more or less energy but also how energy levels vary over time. The amplitude or energy of a particular frequency at a particular time is represented by the third dimension, color, with dark blues corresponding to low amplitudes and brighter colors up through red corresponding to progressively stronger amplitudes. And so, the images offer an energy map of the signal with high amplitudes for specific frequencies.

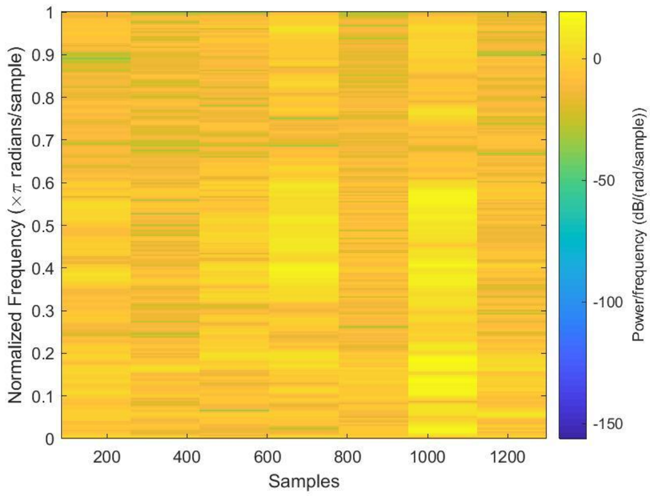

Fig. 4.Spectogram of the signal part

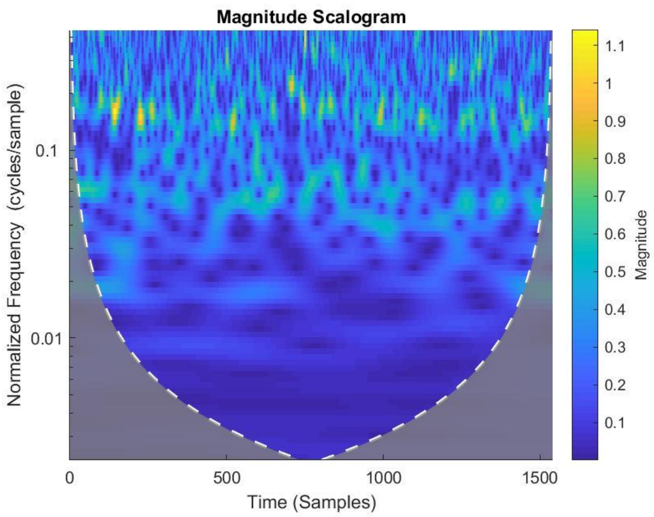

Fig. 5.Scalogram of the signal part

Here below in The Table- II, we summarized in a table the number of images generated for each health condition and motor loads that we used in our work.

Table- II: Number of images used for the experience

Motor speed	Bearing condition	Spectogram	Scalogram
Motor speed 1797 rpm (Load 0)	*Normal*	121	121
	Inner Race Fault	121	121
	Ball Fault	122	122
	Outer Race Fault (Centered @6:00)	121	121
Motor speed 1772 rpm (Load 1)	*Normal*	241	241
	Inner Race Fault	243	243
	Ball Fault	243	243
	Outer Race Fault (Centered @6:00)	243	243
Motor speed	*Normal*	242	242

Retrieval Number: F8489038620/2020©BEIESP
DOI:10.35940/ijrte.F8489.038620
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

1750 rpm (Load 2)	*Inner Race Fault*	242	242
	Ball Fault	243	243
	Outer Race Fault (Centered @6:00)	243	243
Motor speed 1730 rpm (Load 3)	*Normal*	242	242
	Inner Race Fault	242	242
	Ball Fault	244	244
	Outer Race Fault (Centered @6:00)	243	243

All these images will be used as inputs in order to fine-tune the pretrained neural network AlexNet.

We will take this pretrained network and use it to learn a new task which is differentiating between the 4 different health conditions. Fine-tuning a network with transfer learning is faster and easier than training a network with random initialized weights from the beginning because it doesn't require a high number of training images. Just a smal number of training images is sufficient to quickly fine-tune the network into the new task. In addition, in the "in press" [7], the authors showed that initializing with transferred features from a source distant from the target and then fine-tuning them to the new task gives much better results than starting with random weights.

Secondly, we made changes on the image inputs. The network requires input images of size 227-by-227-by-3, but our images have a different size. Thus, we resized these images into this format.

Thirdly, we replaced the last three layers, configured for 1000 classes from the previous pretraining, of the network by new ones in order to adapt to our new classification problem. We replace the last three layers with a fully connected layer, a softmax layer, and a classification output layer. The fully connected layer is set to have the same size as the number of classes in the new data which 4 conditions: Normal, Inner Race fault, Ball fault and Outer Race fault.

The first layers contain very generic variables (For example: detection of angles or colors) and therefore relevant for many different tasks while the last layers are specific to the details of the classes contained in the target database. Thus, the early layers of AlexNet possess the learned low-level features from previous training and the last 3 layers have the task specific features.

For transfer learning, we need to keep the features from the early layers of the pretrained network (the transferred layer weights). So we slow down the learning rate in the transferred layers, set the initial learning rate to a small value and we increase it for the fully connected layer to speed up learning in the new final layers.

The Fig. 6 below provides a summary of all the steps that are used in the approach.

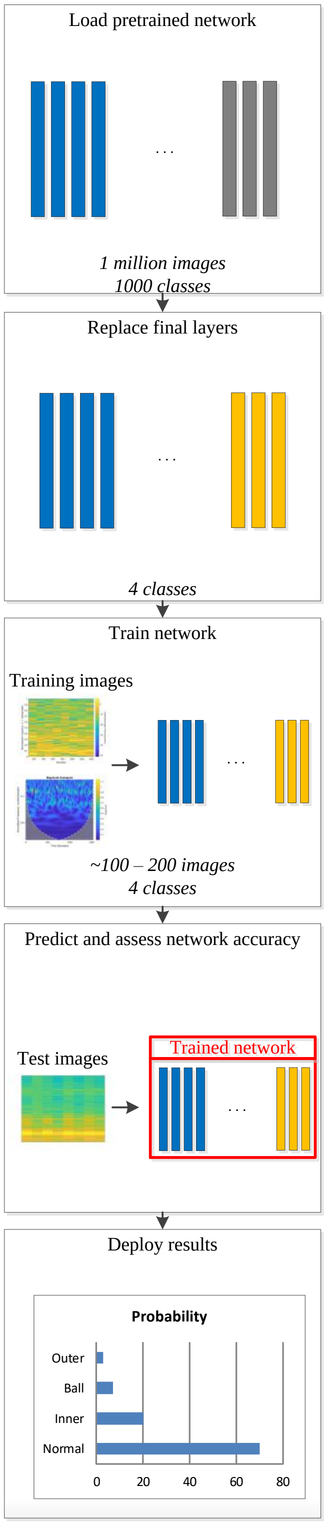

Fig. 6.Simplified diagram of the process

The training was done with a small number of images from each class of the machine health in 4 different speed functioning conditions.

The AlexNet-based extraction method can extract signal trend to failure with high accuracy using the Fourier and the wavelet transforms.

III. RESULTS

We trained the new neural network with source data from a specific motor speed and we tested the diagnosis model on the 4 different cases of target data. We used alternatively spectrogram and scalogram images as inputs.

We summarized the results in the Table- III below.

Table- III: Accuracy results with the 2 types of images

Image type			Motor speed used for target data (rpm)			
			1797	*1772*	*1750*	*1730*
Spectrogram	***Motor speed used for source data (rpm)***	***1797***	100%	99,90%	99,90%	99,90%
Scalogram			100%	100%	100%	99,38%
Spectrogram		***1772***	100%	100%	100%	100%
Scalogram			91,75%	100%	100%	100%
Spectrogram		***1750***	98,97%	100%	100%	100%
Scalogram			97,73%	100%	100%	100%
Specrtogram		***1730***	98,97%	100%	100%	100%
Scalogram			81,65%	100%	99,90%	100%

We can see that the approach gives very positive results even if the speed changes. The diagnostic model built with the newly adjusted network AlexNet stays robust against speed changes.

In general, the results show a high level of accuracy but there is some differences detected between the ones using the spectrogram as input and scalogram images. The precision is higher with spectrogram images in comparison with scalogram ones. The average precision achieved by spectrogram is at least 99% while by scalogram the minimum reached is at 78%. Also, the accuracy stayed stable with spectrogram which is not the case with scalogram. It becomes lower when the source operating speed is different from the target one. With spectrogram, the accuracy fluctuates between 99% and 100% whatever the source operating speed. More there is a gap between source and target speed and lower is the precision with scalogram. Thus spectrogram is a robust and stable spectral representation against speed variations.

In the study of Ferguson et al [8], they transferred the knowledge of different CNN architectures previously trained using images of common objects (for example: a person, a bicycle and a car) to detect and locate defects in metal castings with a small set of X-ray images. They took advantage of the stored knowledge in 2 available pretrained architectures which are VGG-16 (the Visual Geometry Group) and ResNet-101 (Residual Network) and improved the diagnosis system by reducing the total training time and improving the accuracy. They achieved a precision of 86 and 92%. In our work, especially with spectrogram images, we never go under 99.5% in accuracy.

Finally, this finding is very interesting because even if the machine operating parameters change, the model would be able to detect the correct health condition with high accuracy.

Then we made a classification of the images in the test set by plotting a confusion matrix of the true and predicted labels, presented in the two figures below, Fig. 7 and Fig. 8, in order to verify that the network is accurate at classifying new data.

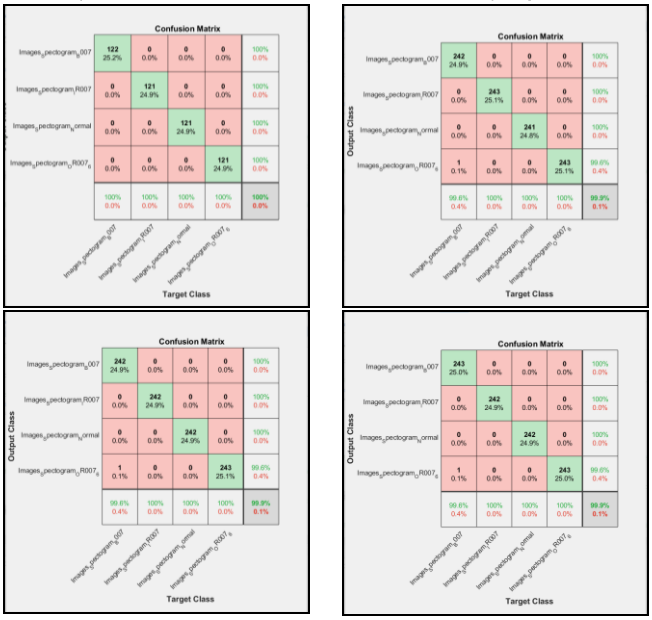

Fig. 7.Confusion matrices of the diagnosis results with 1797 rpm source motor speed and with input data spectogram images

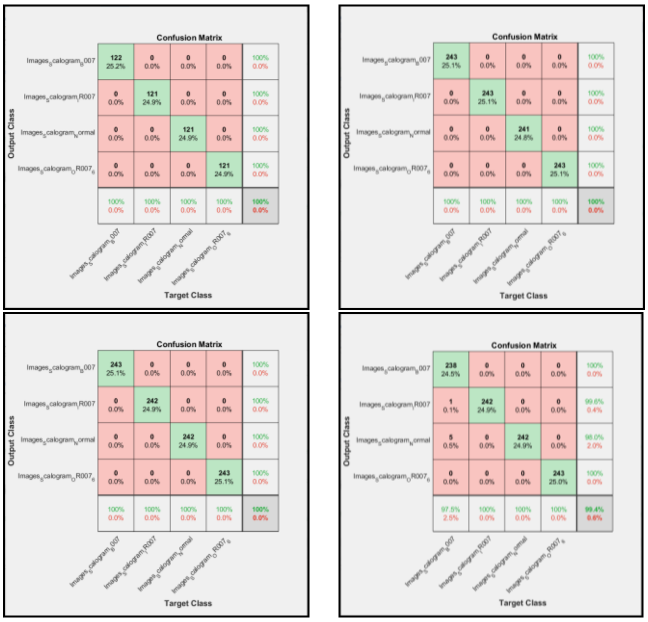

Fig. 8.Confusion matrices of the diagnosis results with 1797 rpm source motor speed and with input data scalogram images

To analyze the network performance further, we computed activations for every observation in the dataset at an early max pooling layer, the final convolutional layer, and the final softmax layer.

We used the t-SNE function to reduce the dimensionality of the activation data. Early layers tend to operate on low-level features such as edges and colors. Deeper layers have learned high-level features.

Retrieval Number: F8489038620/2020©BEIESP
DOI:10.35940/ijrte.F8489.038620
Journal Website: www.ijrte.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Therefore, activations from early layers do not show any clustering by class. Two images that are similar pixelwise (for example, they both contain a lot of green pixels) are near each other in the high-dimensional space of the activations, regardless of their semantic contents. Activations from later layers tend to cluster points from the same class together. This behavior is most pronounced at the softmax layer and is preserved in the two-dimensional t-SNE representation.

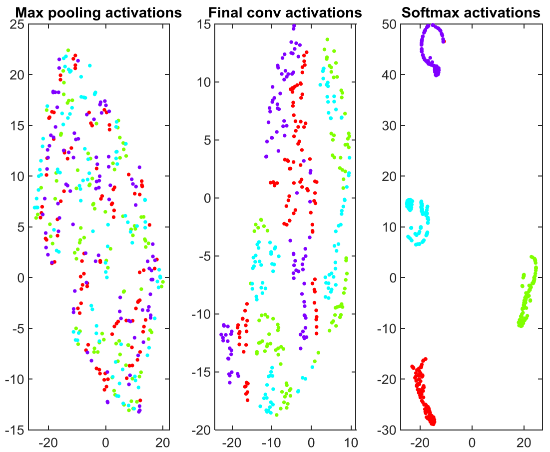

Fig. 9.Ft-SNE representations at the early max pooling layer, the final convolutional layer and the final softmax layer

From the Fig. 9 above, we can observe that the early max pooling activations do not exhibit any clustering between images of the same class. Activations of the final convolutional layer are clustered by class to some extent, but less so than the softmax activations. Different colors correspond to observations of different classes.

It shows that the prediction model generated by the neural network was able to separate successfully the 4 different categories of the machine status without any problem of overfitting.

IV. CONCLUSION

In conclusion, the use of pretrained network with the transfer learning approach made it possible to adapt it to our fault diagnosis case very effectively even with very little data used for training. The approach made the learning process quicker with less energy and less data required.

This work shows also that using a neural network pretrained with data unrelated to the target dataset and retraining it for a new classification task gives an astounding result. By retraining AlexNet with a new set of spectral representation images to classify the condition of the machine according to 4 categories, the prediction model was able to achieve 99% accuracy. In addition, the generated model remained robust in the face of changes in operating conditions, in particular the rotation speed as previously demonstrated.

REFERENCES

1. Bang, S.H., R. Ak, A. Narayanan, Y.T. Lee, and H. Cho, "A survey on knowledge transfer for manufacturing data analytics". Computers in Industry, 2019. 104: p. 116-130.
2. Yao, Y. and G. Doretto. "Boosting for transfer learning with multiple sources". in 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2010.
3. Sun, C., M. Ma, Z. Zhao, S. Tian, R. Yan, and X. Chen, "Deep Transfer Learning Based on Sparse Auto-encoder for Remaining Useful Life Prediction of Tool in Manufacturing". IEEE Transactions on Industrial Informatics, 2018: p. 1-1.
4. Xu, Y., Y. Sun, X. Liu, and Y. Zheng, "A Digital-Twin-Assisted Fault Diagnosis Using Deep Transfer Learning". IEEE Access, 2019. 7: p. 19990-19999.
5. Cao, P., S. Zhang, and J. Tang, "Preprocessing-Free Gear Fault Diagnosis Using Small Datasets With Deep Convolutional Neural Network-Based Transfer Learning". IEEE Access, 2018. 6: p. 26241-26253.
6. University, C.W.R. "Bearing Data". Data files]. Available from: https://csegroups.case.edu/bearingdatacenter/home.
7. Yosinski, J., J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?", in Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2. 2014, MIT Press: Montreal, Canada. p. 3320-3328.
8. Ferguson, M., R. Ak, Y.T. Lee, and K.H. Law. "Automatic localization of casting defects with convolutional neural networks". in 2017 IEEE International Conference on Big Data (Big Data). 2017.

AUTHORS PROFILE

Youssef Maher, is a research scholar in the Department of Electrical and Mechanical Engineering, from the University Hassan 1, Settat, Morocco. He has a degree in industrial and mechanical engineering. His research interests include Data Mining, Big data analytics, Deep Learning, Neural Networks Predictive analytics.

Boujemaa Danouj, Doctor Engineer in mechanical engineering. Graduate of the National School of Electricity and Mechanics Casablanca, Morocco. Master of Philosophy (M.Phil) from Middlesex University, UK. Ph.D. from ÉTS (École de Technologie Supérieure), Montréal, Canada. Expertises: Free form curves and surfaces modeling, partial discharges in high voltage machines, deep learning and predictive maintenance, Metrology and tolerancing.