MACHINE LEARNING MODELS AND DATASETS FOR BLURRED TEXT DETECTION IN WILD SCENE: A Review

First Author (Corresponding Author):

NAME: NWUFOH, Chinonyelum Vivian

AFFILIATION: Computer Science Department, Lead City University, Ibadan. Oyo State. Nigeria

E-MAIL: chinonyelum.tabansi@yahoo.com

ORCID: 0000-0002-1663-5137

Second Author:

NAME: SAKPERE, Wilson (Ph.D)

AFFILIATION : Computer Science Department, Lead City University, Ibadan. Nigeria

E-MAIL: sakpere.wilson@lcu.edu.ng

ABSTRACT

Text recognition and natural image recognition are key problems in machine learning with broad applications such as sports video analysis, autonomous driving, and industrial automation. Recognizing and understanding scene text has received much attention in recent years because scene text contains rich semantic information that can be used in a variety of visual applications, but they face common challenges that affect how this is expressed and be influenced by different environments. , especially when rendering text with multiple resolutions and multiple orientations. Current scene text recognition and/or scene text recognition methods leverage advances in machine learning architectures and report superior accuracy on reference data sets. However, there are still some challenges affecting the text-in-the-wild scene; the model's inability to generalize to invisible data and the lack of labeled data, resulting in poor performance of existing methods. In this study, we examined advances in scene text recognition and recognition, as well as classification by traditional methods. These methods are illustrated with key issues and techniques in commonly used benchmarking datasets and test protocols that base text recognition performance and recognition methods on representative scenes.

Keywords: Text Detection and Recognition, Scene Texts, Machine learning, Scene Text Recognition and Detection, Datasets, Classification Algorithms and Dimension (Dimensionality) Reduction, Clusters and Clustering Algorithms

Statement and Declaration

I want to certify that there are no conflicts of interest related to this publication or financial support that could affect the results of this study. I certify that the authors accept all financial contributions to this study.
I certify that it has been read and approved by the named author(s) and that there are no others who meet the author criteria but are not listed. I have fully considered the copyright protection of this work and certify that there are no copyright issues, including the timing of publication. This is a confirmation that we adhere to the standards of the Institute of Intellectual Property.
The second author understands that the corresponding author is the sole point of contact in the editorial process (including direct communication with the editor-in-chief and the secretariat). I will be responsible for communicating with the other author about progress, submitting revisions, and finally approving the proofs.

Signed by

Nwufoh, Chinonyelum Vivian

1. INTRODUCTION

The detection of blurry text in wild scenes is gaining popularity and has become a major topic in machine learning research with significant advances. Data analysis is a rich field of research that has been extensively studied and successfully applied in various fields over the last few decades. Since the advent of big data, the number of digital images has skyrock eted, exposing vast amounts of multimedia data. [55].

Image analysis is a prolific research area that has been extensively studied and successfully applied in various fields over the last few decades. Not only do we need to process an increasing number of images, but we also need to know what features will result, and feature selection can help in this scenario. Image analysis, which has been extensively studied in recent years, is one of the most productive areas of research. Part of its success is due to the fact that it can be applied to a wide variety of fields with satisfactory results [21].

Digital images contain a large amount of information that is used and extracted for various purposes. Machine learning has the ability to effectively classify remote sensing images. The strengths of machine learning include the ability to deal with multidimensional data and map classes with very complex properties. However, implementing machine learning classifiers is not straightforward, and the literature offers conflicting guidance on many important issues. [39].

Working on data mining or machine learning (ML) tasks requires more efficient techniques to achieve the desired results due to the dramatically increased data size. Therefore, in recent years, researchers have proposed and developed a large number of methods and techniques to reduce the large data size while maintaining the required precision through dimensionality reduction. Dimension reduction is used as a pre-processing step to improve the accuracy of learning functions and reduce training time. You can remove superfluous data, noise and redundant functions[57].

Dimension Reduction (DR) is achieved through two main methods: Feature Selection (FS) and Feature Extraction (FE). As data is generated continuously and with increasing speed, FS is considered an important method; This method can reduce some serious size problems, hence effectively reducing redundancy, removing superfluous data and improving the understandability of the results. On the other hand, FE solves the problem of identifying the most limited, informative, and sophisticated set of functions to improve data processing and storage efficiency.

Classification is an important and difficult task in the field of image processing. Description, texture or similarity of elements or objects used to classify them, tag images into one of many predefined categories called image classifications. Pixels are the smallest representation in an image. Image classification classifies pixels into different groups. Image acquisition, image preprocessing and image segmentation are part of image classification [4].

This study provides an overview of efficient machine learning from an applied standpoint for blurred text detection in wild scenes. All these techniques reviewed can be applied to scene text detection specially to blurred text with a little novel modification.

2. RELATED WORKS

Under this section, we delve into various works of scholars with innovative methods that can be experimented on text recognition. These reviews are done in paragraphs as follows;

2DPCA fractal features and genetic algorithms are proposed for efficient face representation and recognition using automatic face recognition systems [25]. They showed that fractal features derived from iterative function systems enable successful face recognition and outperform traditional approaches. To speed up the feature extraction step, they proposed a new fractal feature extraction algorithm based on genetic algorithms. They used two-dimensional principal components analysis to extract another important pieces of information contained in faces with different fractal features. Using the experimental results from two databases, we show that the optimal recognition rate and recognition time make the system a useful tool for automatic face recognition. Using principal component analysis helps extract key text image features from the scene, allowing this method to be used to detect blurry text.

Detection of blurred image areas based on twenty-one (21) functions has been proposed. Twenty of them were based on discrete wavelet transforms and one was based on grayscale co-occurrence matrix relationships [9]. The features described are introduced to blur segments regardless of the blur type such as background blur, motion blur or blur. In addition, the proposed function is relatively easy to calculate and parallelize. Multilayer perceptron's are trained on the above properties using a backpropagation algorithm. Three approaches are presented and tested. Added non-overlapping fixed-size windows, non-overlapping recursively split windows, and morphological closure operators.

An image fusion method based on PCA and genetic algorithm has been proposed [26]. The art of combining multiple images to create a well-developed image is well established. Various fusion methods have been proposed in the literature but the current research is based on PCA image fusion and genetic algorithms. Images of the same size are considered for the experiment. Genetic algorithms can be used in combination with PCA (Principal Component Analysis) techniques to overcome the limitations of traditional techniques. If image fusion requires parameter optimization, a genetic algorithm can be used. A genetic algorithm is also used to optimize the weight values. Root mean square error, entropy, mean, bit error rate, average and maximum signal to noise ratio are some of the parameters used to evaluate the capabilities of image fusion techniques. They found from previous experiments that this method works well and the output image quality is much better than previous methods.

A classification of blurred images based on deep learning has been proposed. In the field of image processing, the detection of blurring is essential for the reconstruction of blind images [48]. In this article, they developed an accurate classification system based on convolutional neural networks (CNN) to identify four types of blurry images; Blur Blur, Gaussian Blur, Haze Blur, Motion Blur. A Simplified-Fast-Alexnet (SFA) (abbreviated and modified version of Alexnet) supervised learning model is built to map the input image into a high-dimensional feature space that can accurately classify blur. SFA succeeded in simplifying Alexnet by proportionally compressing the output counts of each Alexnet convolution layer by a ratio of 0.5 and removing the two fully connected layers at the beginning of Alexnet, reducing parameter redundancy. Overcome the fatal error of sexuality. Also, a stack normalization layer is added to the specified classifier to replace the dropout method. This reduces internal covariate changes and alleviates overfitting problems, speeding up the convergence of the
deep network during the training phase. Experiments show that the proposed approach outperforms the original Alexnet and the widely used modern 2007 Berkeley and Pascal VOC datasets.

A dimensionality reduction technique based on random projections was investigated [50]. Dimension reduction techniques are very important in big data analysis. Traditional dimensionality reduction techniques, such as principal component analysis (PCA) and linear discriminant analysis (LDA), have been extensively studied over the last few decades. However, as the dimensionality of the data increases, the computational burden of traditional dimensionality reduction methods increases exponentially and becomes very difficult to compute. These shortcomings led to the development of random projection (RP) techniques to map high-dimensional data to low-dimensional subspaces in a very short time. RP transformation matrices, on the other hand, are generated without regard to the inherent structure of the original data and therefore typically exhibit a relatively high degree of distortion. As a result, RP-based methods have been proposed in recent years to address this problem. A general description of the methods used in various situations to help practitioners use the appropriate techniques for their application. They listed the pros and cons of different methods and provided additional resources for researchers to develop new RP-based approaches. Dimension reduction techniques are useful for extracting textual image features from a scene. Extract the blurred image from the normal image.

Adaptive Filtering Algorithm
A fuzzy filtering algorithm based on the Genetic Algorithm-Backpropagation Network has been proposed [22]. The GA-BPN algorithm uses a Genetic Algorithm (GA)-(BPN) to determine the weights of a backpropagation neural network. It outperforms traditional optimal algorithms in terms of global optimization. Examine image noise filters using GA-BPN. First, this white paper uses a training example to train GA-BPN as a noise detector. Then use a well-trained GA-BPN

to detect noisy pixels in the target image. Finally, an adaptive weighted averaging algorithm is used to recover the noise pixels detected by GA-BPN. Experimental data shows that this algorithm outperforms other filters.

The main purpose of detecting and classifying blurry images with DNN has been proposed using TensorFlow and Keras networks [35]. Its purpose is to detect and classify naturally blurred, artificially blurred, or distorted images. Since this document is an overview and algorithms are proposed and implemented, it should be properly recognized and classified. The proposed algorithm has been implemented and its accuracy has been improved with respect to existing image classification models. The proposed method is simple and efficient. Experiments show that this method works on different types of images and can be applied to different media analysis applications, such as deep reconstruction, information retrieval and segmentation.

A genetic algorithm has been proposed to blur the motion of individual image frames. One of the most difficult aspects of reconstructing a degraded image from motion blur is estimating an unknown blur filter from a single input blur image [19]. Many blind deconvolution techniques use frequency domain constraints on images, simplified parametric shapes of motion paths during camera movement, or multiple input images. In this article, an algorithm to remove motion blur from a single blurred input image using genetic algorithms was proposed. In science and engineering, genetic algorithms are used as adaptive algorithms to optimize real-world problems. The authors also did a recent study on natural image statistics to show that photographs of natural scenes tend to follow a highly curved distribution. Motion blur is handled in both linear and non-linear modes. Experiments on a large data set of degraded standard images with different grains of different sizes show that the proposed approach is particularly effective for short blur lengths.

The technical problems, methods and achievements of color image text recognition and recognition research have been studied through analysis, comparison, and contrast [55]. It summarizes the basic problems and lists the factors to consider when solving these problems. Existing techniques are classified as incremental or integrated, emphasizing sub-problems such as localization, verification, segmentation, and text recognition. Specific issues related to degraded text, teletext processing, multi-orientation, perceptually warped text, and multilingual text enhancement are also addressed. Various text categories and subcategories are presented, reference data sets are listed, and the performance of the most representative approaches is compared. This overview provides a basic comparison and analysis of what has been considered and remains in the field as such a problem.

A state-of-the-art deep learning approach has been proposed for wildlife surveillance that automatically identifies and isolates species-specific activities from still image and video data

[16]. Using their dataset of 8,368 images of wild and domestic animals in farm buildings, how to distinguish badgers from other species (binary classification) and how to distinguish between each of the six animal species. (Multiple classification) developed. The authors started with a binary taxonomy of badgers because such a tool would be useful to monitor the transmission of Mycobacterium bovis (the cause of bovine tuberculosis) between badgers and cattle. The algorithms developed here find wide application in wildlife surveillance, where large amounts of visual data need to be searched for specific species. An automated deep learning detection method to identify and isolate species-specific activity from still image and video data. If this technique is applicable to this area, you can also perform it on text images in your scene to detect blurry text in those images.

The recognition of text data from text images of scenes is an intriguing topic in the field of computer graphics and visualization [54]. The challenge becomes even more difficult when smart edge devices are involved in the process. Text recognition and classification is complicated by low quality images with challenges such as blur, low resolution, and contrast. The blurred image is pre-processed after synthetic blur, and the blur process is used to restore the image. Standard Maximum Stable Boundary Region (MSER) technique was used to recognize and localize the text. Then, using K-Means, they took three different groups from the query image, separated foreground and background, and integrated the character-level grouping. Finally, the segmented text is classified into text and non-text regions using a novel Convolutional Neural Network (CNN) framework. The goal of this task is to eliminate false positives. The proposed method was tested on three large data sets including SVT, IIIT5K and ICDAR 2003.The classification results for the SVT dataset were 90%, the IIIT5K dataset was 96%, and the ICDAR 2003 dataset was 94%. They demonstrate the superiority of the proposed methodology by showing that it works well for an excellent learning model. Finally, the proposed methodology is validated by comparison with previous reference text recognition techniques.

Animal Species Recognition from Camera Trap Images with the use of machine learning and deep learning models has been proposed [44]. Monitoring and timing wildlife movement is critical to saving animal lives. Camera traps are the most widely used animal surveillance technology as they automatically trigger the cameras when an animal is present and collect a large amount of data. This current work aims to explore different machine learning algorithms like Support Vector Machines (SVM), Random Forests (RF) and deep learning models like his Alexnet and Inception V3 for animal species classification. Deep learning models outperform machine learning algorithms in this respect. In this article, the overall accuracy comparison between machine learning and deep learning models was considered and explained. Experimental results show that InceptionV3 achieves higher accuracy than SVM, Random Forest and AlexNet, and sufficient data and accurate methods lead to high-precision classification. This experiment uses the KTH data set consisting of 19 different animal categories, 12 of which were chosen to assess model performance.

Detection of curved text in blurred/blurry scenes/video images has been realized [52]. Video/closed text recognition is difficult due to blurring and multiple blurring caused by motion. In this article, a new method to detect text in blurry and non-blurry images was introduced. In contrast to existing methods that use blurring or classifiers, the proposed method uses contrast changes of neighboring pixels and a low-pass filter to estimate the degree of blurring of the image and generate candidate pixels for the blurring. It uses the gradient value of each pixel as the blur weight. Then the proposed method uses K-means binning on the weighted values of the candidate pixels to find the candidate text regardless of the blur type. It then uses the Bhattacharyya distance to extract text symmetry, remove spurious text candidates, and provide text components. In addition, the method sets the bounding box of each text component based on the nearest neighbor criteria and the alignment of the text component. The proposed method outperforms existing methods in terms of defocus, motion, blur, and standard curved text datasets.

Noise removal and image blur classification have been proposed [23]. Blur is a general term for image degradation that occurs when shooting with a low-quality camera or to intentionally emphasize moving or prominent objects. However, most blur classifiers only classify images as blurred or sharp and cannot distinguish between intentionally and unintentionally blurred images. Some unintentionally blurry images are too valuable to throw away. In this article, a robust
image blur classifier that classifies images into sharp intentional blur and unintentional blur was developed. The basic idea for determining whether pixel blurring is intentional or not is whether the blurring occurs on a prominent and semantically significant object by using blur hints, highlighting, and semantic targeting. The authors used spatial pyramid clustering to extract global features. Then perform classification using Random Forest. To detect more unintended blurry pixels one can, achieve that by embedding signals in the CRF. Also one can intentionally create a blurred image by adding the unwanted blurred area back into the blurred image. The UBICD dataset performs blur classification and unintended blur removal for different types of unintended blur. The experimental results of our method show excellent performance in image blur classification and promising results in unintentional blur removal.

The deep learning era has been used in the study of text recognition and scene recognition [28]. With the advent and development of deep learning, our view of computing has changed and reshaped drastically. A major research area in computer vision, scene recognition and text recognition are inevitably influenced by this revolutionary wave that ushered in the deep learning era. In recent years, the community has made great strides in terms of thinking, methodology, and performance. The purpose of this research was to analyze together the most important changes and significant advances in text scene recognition and recognition in the age of deep learning.

A text extraction algorithm was validated in this article for scene text and document images [38].

One of the most common uses for retrieving text from images is to extract information from text and recognize its characters. This helps index the images on your storage media. One no longer have to search through dozens of images when looking for a specific image or document. Search only groups of indexed images to make it easier to find a specific image. Extracting lines of text from scanned document images is a major problem in optical character recognition. This is because distorted lines of text complicate the process. This problem is compounded by the different alignments of the text lines. These are called multi-slope lines. These curved lines are found in both printed and handwritten documents. Developing a real-time system that maintains high recognition rates and accuracy regardless of document type or character set is a difficult task. This article analyzed and categorized different text extraction schemes for document images and scene text, compared approaches to these images based on common problems, and discussed their strengths and weaknesses.

Character recognition based on deep learning of isolated Arabic scenes was proposed [1]. As technology advances and cameras and devices become more sophisticated, researchers are turning to image analysis and text comprehension. As has been reported in recent years, deep learning techniques have worked well to assess the potential for classifying text from images of natural scenes. There are many deep learning approaches that promise to recognize text from images. This article demonstrated text recognition of Arabic scenes using Convolutional Neural Networks (ConvNets) as deep learning classifiers. As the text data becomes increasingly distorted in the scene, the authors used five (5) alignments for the appearance of a single character to account for maximum deviations. The training was designed with filter sizes of 3x3 and 5x5 and stride values of 1 and 2. During the text classification phase, they trained the network with different learning rates. This method gave promising results in recognizing Arabic characters from images of segmented Arabic scenes.

It has been proposed to detect and classify text in low quality natural images [54]. The recognition of text data from text images of scenes is an interesting topic in the field of computer graphics and visualization. The challenge becomes even more difficult when smart edge devices are involved in the process. Text recognition and classification is complicated by low quality images with challenges such as blur, low resolution and contrast. This takes account of such an urgent topic in research. The proposed technique consists of three main contributions. The blurred image is pre-processed after synthetic blur and the blur process is used to restore the image. We then use the standard Maximum Stable Extreme Region (MSER) technique to recognize and localize the text. Then, using K-Means, they took three different groups from the query image, separate foreground and background, and integrated the character-level grouping. Finally, the segmented text is classified into text and non-text regions using a novel Convolutional Neural Network (CNN) framework. The goal of this task was to eliminate false positives. The method was evaluated on three main datasets including SVT, IIIT5K and ICDAR 2003. The classification results were 90.3% for the SVT dataset, 95.8% for the IIIT5K dataset and 94 .0% for ICDAR 2003 records. They demonstrated the superiority of the this methodology by showing that it works well for an excellent learning model. Finally, the proposed methodology is validated by comparison with previous reference text recognition techniques. Note that blurred text is classified here as text in a natural, low-quality image.

An evolutionary approach to improve and classify dermatological images using deep learning models has been proposed [12]. Dermatological diseases are the most common skin diseases in humans. They can be contagious and chronic, and in some cases lead to serious health problems like skin cancer. Images and other relevant information should be shared based on expert analysis that can be accessed remotely. In such cases, the deterioration in image quality caused by the recorder leads to misdiagnosis. In this study, a genetic algorithm (GA)-based approach was used as an image enhancement technique to enhance poor-quality dermatological images acquired in rural clinics. Disease identification is performed on the enhanced images using a convolutional neural network (CNN) classifier. The scope of this document is limited to motion blur, the most common problem in image capture, especially when one of the two (device or object) moves unexpectedly. ResNet-152 was used to examine types of skin diseases including melanoma, melanocytic nevus, basal cell carcinoma, actinic keratosis, benign keratosis, vascular lesions, and squamous cell carcinoma, resulting in blurry images. The overall accuracy was 87.40%. Accuracy improved to 95.85% when using GA enhanced images. Statistical studies based on confusion matrices and t-tests were used to further analyze the results. The proposed method has the advantage of reducing analysis time and manual diagnostic errors.

It has been proposed to detect motion blur using support vector machines. Motion blur is the most common factor affecting the quality of captured images [53]. Users will only notice the blur after viewing the image on a high-resolution screen. Object movement, camera shake, or the relative speed between the object and the camera all contribute to motion blur. Many anti-shake or image stabilization techniques have been developed to solve this problem. However, there is still no mechanism to detect motion blur. For this reason, this white paper describes several possible solutions and evaluates their effectiveness. The purpose of the motion blur detector is to classify and inform the user if a digital image is blurry or sharp. This feature can provide information to help the user decide whether to recapture the image immediately rather than going into playback mode to review it. It uses machine learning techniques to achieve greater fault tolerance and adaptation to different imaging environments. Explore different digital image processing schemes to identify the most discriminating features.Support Vector Machines (SVM) is one of many machine techniques that have been implemented. To get the best performance from your SVM, it is important to extract unique information from motion blur images. Therefore, several signal transforms have been studied, such as the discrete Fourier transform, the discrete cosine transform, and the Radon transform. The document also compares the performance of different feature vectors, functions, and kernel parameters.

For robust scene text recognition, the use of multifit twin support vector machines has been proposed [20]. Text recognition in images of natural scenes remains a challenging problem as their appearance is highly variable in unconstrained environments. The proposed work improved the generalization of Twin Support Vector Machines (T-SVM) through diverse regularization and was later extended to check and recognize text in images of natural scenes, thus this type of is the first of Se that a multiclass T-SVM added with intrinsic and environmental regularization terms that help form a smooth functioning of the model. Understanding the text of nature scenes requires finding, recognizing, and reconstructing the text. This work also includes a revalidation module that discards the false alarms of the text objects recognized during the localization phase.500 and 86% at SVT. The results show that the model could recognize most of the characters in images with high precision.

Aiming to develop an automated algorithm for detecting moths from captured images under real-world conditions, we used an adapted Support Vector Machine (SVM) detection approach to establish contact with individual moths in captured images. I counted the number of moths on the ground [7]. This method builds on previous detection work and adds an improved classification step. More specifically, the SVM classifier is trained using a multi-scale descriptor called Histogram Of Curviness Saliency (HCS). This light-resistant descriptor can recognize and describe the outer and inner contours of target insects at multiple scales. The proposed classification method can be trained on small sets of images. The quantitative evaluation shows that the proposed method outperforms current approaches in terms of insect classification accuracy (95.8% rate).

Several texture descriptors and classifiers for face detection have been proposed [30]. A novel facial recognition system that works well in practice, based on a set of descriptors using various pre-processing techniques. Angular distance is used in the FERET dataset where the goal is identification. Validate specific matches using support vector machines on the LFW dataset and a learned similarity metric. Our proposed system works well on both datasets, and to our knowledge,
achieves one of the highest throughput rates published in the literature for the FERET dataset. It's worth noting that we got these good results on both sets of data without using any additional training patterns.

The use of deep learning-based object detection and severe weather scene detection has been suggested [41]. Population growth in big cities causes traffic jams. The city's road network must be constantly monitored, expanded and modernized. With the development of autonomous vehicles, intelligent vehicle sensor solutions are needed to solve traffic problems. Intelligent road traffic monitoring includes identifying and tracking vehicles on roads and highways. They demonstrated how the v5 You Only Look Once (YOLO) model can be used to identify cars, traffic lights and pedestrians in various weather conditions, enabling real-time identification in a typical vehicle environment. Object detection in normal or autonomous environments can be hampered by adverse weather conditions. Bad weather can make driving dangerous in a number of ways, including: the illusion of icy roads or dense fog. In this study, they use the YOLOv5 model to extract objects from traffic logs for wet and normal scenarios for 11 different vehicle classes (cars, trucks, bikes), pedestrians, and road signs (red, green, yellow). To train the proposed system, a publicly available Roboflow dataset was used. In addition, they evaluated the performance of the proposed system using real-time traffic video recordings. The results of the conducted research showed that the proposed method allows detecting cars, trucks and other roadside objects with acceptable results in different situations.

A combination of machine learning and human knowledge is proposed. Machine learning has been extensively studied and applied in various fields [18]. However, achieving high accuracy requires large amounts of data that can be difficult, expensive, or impractical to obtain. Integrating human knowledge with machine learning can significantly reduce data requirements, making machine learning more robust and reliable, and creating understandable machine learning systems. Not only does this leverage extensive human knowledge and machine learning capabilities to achieve previously unavailable functionality and performance, but it also facilitates the interaction between humans and machine learning systems, making learning decisions understandable to humans. This whitepaper provides an overview of knowledge and its representation and explains how it can be integrated into machine learning and methods.

A study on face recognition in nature was conducted; past, present and future.[56]. Face recognition is one of the most studied topics in the computer vision literature, not only because of the difficulty of faces as objects, but also because of the large number of applications that require face recognition in the first place. The availability of data on the Internet with “in the wild” searches, community efforts to develop publicly available references, and advances in the development of robust computer vision algorithms have resulted in significant advances in recent years. This article discusses the latest advances in real-world face recognition technology, beginning with Viola Jones' ground-breaking face recognition method. These techniques can be roughly divided into two categories: Fixed models learned primarily with methods based on reinforcement or deep neural networks, and deformable models describing parts of the face. A representative method and other effective methods are listed.

Techniques have been proposed to reduce the dimensionality of face recognition [37]. Various linear and non-linear dimensionality reduction techniques for face recognition are discussed. Face images contain different global and local features, so using one technique may not improve detection accuracy. It can be advantageous to combine basis vectors from different approaches to achieve higher accuracy. The face subspace can be obtained from the basic structure of the collector in image space using the LPP and ICA methods. LDA-based algorithms can be used to extract more salient features from the space. Random and PCA projections essentially give a global appearance model. The future of face recognition may include advanced dimensions such as depth information for recognition purposes. Algorithmic models should strive to deal with scale-invariant feature vectors. This makes it possible to solve recognition tasks even when the images are extremely fluctuating. Inspired by information theory, face recognition is based on a small set of image features that are more like a set of images of familiar faces, rather than our intuitive notions of facial parts and features. This gave rise to the idea of carrying out face recognition. This technique offers a practical solution to the face recognition problem, is relatively simple, and has proven effective in limited environments. Case studies with captured image sets have shown that the size, skin tone, ambient light and viewing angle of the profile play an important role in image recognition.

A new machine learning approach has been proposed to extract text from the scene [3].In recent years, extracting text from images has become a popular and challenging research area in computer vision. Due to chaotic backgrounds, unstructured scenes, orientations, ambiguities and other factors, this article explores such an urgent aspect as natural scene identification and text extraction. Contrast enhancement is performed on the input image to identify text by applying a LUV channel to obtain perfectly stable areas. Then, using the standard MSER segmentation technique, the L-channel is selected for region segmentation. This work also takes into account various geometric properties to distinguish between textual and non-textual regions. Furthermore, we obtain a segmented picture by ranking the combined components using a combination of two feature descriptors, LBP and T-HOG. First, the two feature descriptors are classified separately using linear SVM. The two results are combined using the weighted sum technique and classified into textual and non-textual parts. Text areas are recognized and labeled with the new CNN in text recognition. To create text words, save the CNN output to a text file. Finally, the text file is searched in the lexicon for matches to optimized scene text words using Hamming distance (error correction) techniques where appropriate.

The learning of
visual properties of video objects and the recognition of human actions were examined [27]. Mot ion detection of video objects and people is used in various applications like CCTV, face recognition and more. Video object detection includes not only the position of objects in the image, but also their classification. Tracking human behavior is referred to as human behavior tracking. Video detection is more difficult than image detection because video frames tend to be blurrier than frames. Additionally, video tracking is often hampered by issues like blurry video, motion blur, and partial blackout. At present, video detection technology not only offers real-time detection, but also high-accuracy detection of blurry video images. This article examines and describes the different methods of detecting video objects and human behavior. Many of them deliver ground-breaking results. They reviewed and discussed traditional video recognition methods using supervised learning. They also looked at the most used datasets for video object

detection and human activity detection. Finally, they gave an overview of video detection methods. Example: frame-by-frame detection (frame-based), keyframe detection and use of timing information. Methods that use the timing information of adjacent video frames mainly include optical flow, long-term storage, and convolution between adjacent frames methods.

Image signals were detected and classified with the refined PSO-SVM model [58]. For recognition and classification, PCA and SVM are combined and applied to a set of handwritten MNIST digits. At the data level, dimensionality reduction is used to compress high-dimensional image data. This greatly improves the performance of the algorithm. The recognition accuracy reaches 98% and the execution time is reduced by about 90%. This model first processes the original image data and then uses coarse set theory to select features. This reduces the dimensionality of the AI network input, speeds up learning and recognition, and further improves recognition accuracy. We have applied this model to the recognition of handwritten digital images of works, and the experimental results show that it is efficient and feasible. The system has the advantage of being easy to implement, maintain and integrate. Experiments show that the system has good synchronization performance during the fusion parallel image processing process using multiple algorithms.

All these articles reviewed above be it video and image signaling, objection recognition, video object recognition and image fusion all have direct relationship to study of blurry text of objects because all these study areas delas with challenges of irregularities that arise from camera and lens technology. Also, with the aid of machine learning, deep learning and dimensionality reduction scholars have been able to come up with methods that have solved the issues of blurriness in images or videos.

3. MATERIALS AND METHODS

Scene text recognition is classified taxonomically in general object recognition and divided into one and two-stage processes. In fact, many scene text detection algorithms are heavily modeled on generic object detectors. On the other hand, scene text recognition has its own characteristics and challenges that require unique methods and solutions. To solve these important problems, many methods rely on special representations of scene text.

Consequently, the development of scene recognition algorithms goes through three main stages. Learning-based methods have multi-stage pipelines but are still slow and complex. Therefore, general concepts and object detection methods are well integrated into this task. In the third stage, the researchers create specialized representations based on partial text components to solve the problem of long and irregular texts.

In this article, we will introduce a trending method to detect blurry and sharp areas in dense, blurry and noisy images. All model training details are designed for experimentation with models and datasets.

High quality data and text annotations are required for both training and model evaluation. The most frequently used comparative data sets are listed below [34]; [28]; [11]; [10]; [59]; [60].

2003 ICDAR (IC03). The ICDAR Robust Reading Competition has published its first scene text recognition and recognition benchmark. There are 258 natural images for training and 251 natural images for testing. All instances of text in this data set are in English and arranged horizontally.

Text for Street View (SVT). This dataset contains 350 Google Street View images annotated at the word level with
bounding boxes aligned to axes. It always has lower-resolution text and not all instances of the text are annotated.

KAIST. This dataset contains 3,000 photos taken in a variety of environments, including indoor and outdoor scenes, and under different lighting conditions (bright day, night, strong artificial lighting, etc.).). The photos were taken with a high-resolution digital camera or a low-resolution mobile phone camera. All images have been reduced to a size of 640,480 pixels.

2011 ICAR. This dataset is based on ICDAR 2003 with some modifications.There are 229 training scene images and 255 test scene images.++
MSRA-TD500 (M500). This dataset contains a total of 500 images of nature scenes, 300 for training and 200 for testing. It supports text line annotations and polygon fields for text box annotations. Contains example texts in English and Chinese.

2013 ICDAR (IC13). It is also a modified version of ICDAR 2003. There are 229 natural images for training and 233 natural images for testing++
USTB-SV1k. Contains a total of 1,000 Google Street View images and 2,955 text instances. Considers only English words and provides word-level annotations.

2015 ICDAR (IC15). Contains 1500 scene images, 1000 for training and 500 for validation/testing. Because they were captured without prior user preference or intent, text instances (described with four square vertices) tend to be distorted or blurry. Contains a total of 1 7,548 text instances. Provides word-level annotations. The first random scene text record is IC15 and is in English.

Text COCONUT. It is the largest text recognition and recognition model to date. The original images are from the Microsoft COCO dataset and the 173,589 text instances are annotated from 63,686 images. There are 43,686 training images and 20,000 test images with handwritten and printed text, clear and blurry images, and English and non-English text.

The 5,000-word (IIIT) IIIT dataset is from Google image search. Question words include signs, house signs, movie posters, etc. Also, all words in the images are manually annotated with bounding boxes and simple real images.

The Street View Text (SVT) dataset is based on Google Street View and consists of exterior shots annotated with a list of text. Texts in SVT come from business tags that are easily sourced from geographic search companies.

The StreetViewText-Perspective (SP) data set contains texts in street images from different perspectives. SP is based on the original SVT data set and was specially developed for perspective text recognition.

The data set CUTE80 (CT) is offered for the recognition of curved writing. With 288 images, CT word images are collected from natural scenes++ short text.It contains 858,750 synthetic images in which text is accurately reproduced on set designs with randomized colors, fonts, scales, and orientations for a realistic look. The text in this record is annotated at the character, word, and line level.

In kind, Chinese text (CTW). This dataset contains 32,285 high-resolution Street View images, annotated at the character level, including basic font, border, and detailed attributes such as use of text graphics. The data set is the largest so far and the only one with detailed annotations. However, it only annotates Chinese text and ignores other scripts such as English. Contains 32,285 high-resolution images of Street View Chinese text for a total of 1,018,402-character occurrences. Each image is annotated at the character level, including the base font, bounding box, and six other attributes. These features indicate whether the background is complex, embossed, handwritten or printed, obscured and distorted, or whether graphic images are used.

RCTW-17. This dataset contains images of various types, such as street views, posters, menus, interior scenes, and screenshots, for a contest to recognize and recognize Chinese text in images. The data set contains about 8000 training images and 4000 test images with annotations of type ICDAR2015.

Total Text (ToT). Compared to several instances in previous datasets, this dataset has a relatively high percentage of warped text. These images were mostly taken from street billboards and described as polygons with varying numbers of vertices.

SCUT-CTW1500. This dataset contains a total of 1,500 images, 1,000 for training and 500 for testing, and 10,751 clipped word images. CTW-1500 annotations are 14-vertex polygons. The data set consists mainly of Chinese and English words.

MLT ICDAR 2017 (MLT17). This is a large multilingual text dataset containing a total of 10,000 nature scene images, including 7,200 training images, 1,800 validation images, and 9,000 test images. Provides word-level annotations.

ICDAR 2019 Free Format Text (ArT19). ArT consists of 10,166 images, 5,603 for training and 4,563 for testing.They were collected to preserve the variety of text shapes, and all text shapes (horizontal, multidirectional, and curved) are very common.

MLT ICDAR 2019 (MLT19). This dataset contains 18,000 annotated word-level images. Compared to MLT, this data set contains ten languages. This is a more realistic and complex dataset for scene text detection.

Text from ICDAR 2019 Large Scale Street View (LSVT19). This dataset contains 20,000 test images, 30,000 fully annotated training images, and 400,000 lightly annotated training images, also known as partial labels. can be used to test omnidirectional text recognition algorithms in natural scenes.

The NEOCR dataset contains images of nature scenes with multidirectional text. Contains 659 real images with annotations in 5 x 238 pixel frames.It is a multilingual dataset as the texts it contains are in several languages, including English, Hungarian, Russian, Turkish and Czech

The Chars74K dataset was made available to evaluate individual character recognition algorithms in natural images. This dataset contains the symbols for English and Kannada. The GoodImg subset contains 636 images where commenters have highlighted latin letters and arabic numbers.

The
Street View House Numbers (SVHN) dataset is a large real-world database of over 600,000 numbers in natural scenes.The numbers are house numbers cut out of Google Street View images. This benchmark is mainly used for developing and testing digit recognition algorithms.

The IIIT 5K Word dataset is the largest and most demanding benchmark in this field to date. This database contains 5,000 images with text in natural and digital environments. This is difficult due to differences in font, color, size, and layout, as well as noise, blur, distortion, and variable lighting. 2000 images are used for training and 3000 images for testing.

ImageNet, Princeton and Stanford Universities: An image database organized in the WordNet hierarchy that can be used for a variety of visual recognition tasks.

Original images, image URLs, SIFT functions, object boundaries and object attributes are available.

PASCAL VOC, University of Leeds et al: Standardization of an image database for object class recognition. Original images, annotated objects, segmentations and a set of tools to interact with them.

MNIST Google Labs and Microsoft: An image database with manuscript numbers by different authors, useful for image classification. normalized size and centered images (no preprocessing or formatting required).

Caltech256 is the abbreviation for California Institute of Technology. A database of images that can be used to classify objects. original images organized in a taxonomy, annotations of objects in the form of outlines, and a Matlab script for their display.

LabelMe: MIT: An external image database useful for object detection and image segmentation. Original images and partially annotated objects.

NUS-WIDE stands for National University of Singapore. An online image database that can be used to annotate and search for images. Image URLs, original images, tags and six different types of low-level functions are supported.

Zurich, Food 101. Useful image database for food classification, showing different food categories and noise in images and labels. Food categories, original images.

Faces in the Wild: University of Massachusetts A database of face images that can be used for face recognition. Original images, images with deep paths and annotations with the person in the photo.

Tightness-IoU: Tightness-IoU takes into account that scene text recognition is sensitive to missing and redundant parts in the recognition results.

Non-recoverable fields produce missing characters in the recognition results, while redundant fields produce unexpected characters. The proposed indicators penalize the IoU by reducing them by the percentage of missing areas and redundant areas that overlap with other texts.

4. MACHINE LEARNING

Data is typically used and processed by machine learning algorithms to learn relevant patterns about people, business operations, transactional data, events, etc. Below we discuss different types of real data and their classification according to machine learning algorithms [39]. There are four types of machine learning algorithms: Supervised learning, unsupervised learning, semi-supervised learning, reinforcement learning.

Supervised Learning is a machine learning activity that learns a function that maps inputs to outputs based on sample input and output pairs. Derive the function using labeled training data and a set of training examples. In supervised learning, a specific goal is set to be achieved with a specific set of inputs. It is task- based approach. The most monitored activities are 'classification', where the data is split, and 'regression', where the data is fitted. For example, predict a label or class sentiment for a text such as: B. Tweet or product review, an example of supervised learning [8].

Unsupervised Learning is a data-driven process for analyzing unlabeled data sets without requiring human intervention. It is commonly used to identify significant patterns and patterns, group results, and research goals. Clustering, density estimation, feature learning, dimension reduction, association rule detection, anomaly detection, and other unsupervised learning tasks are common [31].

Semi-supervised learning is a hybrid of the above supervised and unsupervised methods as it works on both labeled and unlabeled data. Somewhere between "unsupervised" learning and "supervised" learning. Semi-supervised learning is useful because unlabeled data is plentiful in the real world, while labeled data can be scarce in some contexts. The ultimate goal of a semi-supervised learning model is to provide better prediction results than those produced by a model using only labeled data.
Semi-supervised learning is used in various applications such as machine translation, fraud detection, data labeling and text classification [40].

Reinforcement Learning is a type of machine learning algorithm that allows software agents and machines to automatically evaluate their optimal behavior in a given context or environment to improve performance. It Uses environmental approach. This type of learning is based on reward or punishment, with the ultimate goal of using environmental activists' insights to take action to increase reward or decrease risk. It is a powerful tool for training AI models and helps automate or optimize the operational performance of advanced systems such as robotics, autonomous tasks, manufacturing and supply chain logistics. However, it is not recommended for solving simple or explicit problems [36].

Therefore, depending on the type of data discussed above and the desired outcome, different types of machine learning techniques can play an important role in building effective models in different application domains.

In machine learning, classification is considered a supervised learning technique and is related to the predictive modeling problem of predicting a class label for a given example. We map the function (f) mathematically as a target, label or category from the input variable (X) to the output variable (Y). It can be used with structured or unstructured data to predict the class of a specific data point. For example, spam detection can be a classification problem for email service providers such as "Spam" and "Unspam". The following section summarizes the most common classification problems.

Many classification algorithms have been proposed in the machine learning, data science, and wilderness detection literature. The most common and popular methods, widely used in various application areas and related trends are:

Decision Tree (DT): A well-known non-parametric method of supervised learning is the Decision Tree (DT). The DT learning method is used for both classification and regression tasks.ID3, C4.5 and CART are known DT algorithms [14].

K Nearest Neighbors (KNN): K-Nearest Neighbors (KNN) is an instance-based learning or non-generalized learning algorithm, also known as lazy learning. You are not trying to create a generic internal model. Instead, we store all instances corresponding to the training data in an n-dimensional space [24].++
Logistic regression (LR) is another popular stochastic statistical model used to solve classification problems in machine learning [42].

SVM stands for Support Vector Machine. Support vector machines are another popular machine learning technique that can be used for classification, regression or other (SVM) tasks [5].

Analysis of Clusters

Cluster analysis, also known as clustering, is an unsupervised machine learning technique for identifying and grouping related data points in large data sets, regardless of revenue. It does this by classifying a collection of objects such that objects within the same category, called clusters, are more similar to each other than objects within other groups. It is commonly used as a data analysis technique to identify interesting trends and patterns in data, Example behavioral consumer groups. Clustering can be used in various applications including cybersecurity, e-commerce, mobile computing, health analytics, user modeling, behavioral analytics and more. A brief description and summary of the different types of grouping methods can be found in [29]. This grouping approach classifies data into multiple groups or clusters based on the characteristics and similarities of the data. Depending on the type of target application, the data scientist or data scientist usually decides how many clusters to create dynamically or statically for the clustering method. The most commonly used algorithms are K-Means, K-Mediods, CLARA and other partition-based clustering algorithms [47].

Density Based Method: Uses the concept of clusters in the data space. It is a continuous area of high point density separated from other similar clusters by continuous areas of low point density, identifying distinct clusters or clusters. Noisy spots are spots that are not part of the cluster. DBSCAN, OPTICS and other density-based clustering algorithms are popular. Density-based methods generally struggle with such dense clusters and multidimensional data [49].

Hierarchical method: In general, hierarchical clustering attempts to build a hierarchy of clusters, for example a tree structure. There are two types of hierarchical clustering strategies:

The “bottom-up” approach, where each observation starts with a cluster and pairs of clusters are combined into one, moves up the hierarchy.

Divisible by: A top-down approach, in which all observations start with a cluster and the division is recursive, is derived from the hierarchy [43].

Grid-based method: Grid-based clustering is ideal for processing large amounts of data. The principle of obtaining clusters is to first summarize the data set in a grid representation and then to combine the grid cells. The standard mesh-based clustering algorithms are STING, CLIQUE, etc. [2].

Model-based approach: There are two types of model-based clustering algorithms; One using statistical learning and one using neural network learning methods. For example, GMM is a statistical learning method and SOM is a neural network training method [51].

Constraint-Based Approach: Constraint-based clustering is a semi-supervised approach to data clustering that captures domain knowledge through constraints. Application or user oriented constraints are used to create clusters. The most popular algorithms for this type of clustering are COP K-Means, CMWK-Means, etc. [6].

Many clustering algorithms capable of clustering data have been proposed in the machine learning and data science literature. Common methods that are widely used in different fields of application are summarized below [39].

K-Means Clustering is a fast, robust, and simple algorithm that provides reliable results when datasets are well separated. The K-Means algorithm identifies k centroids and assigns each data point to the cluster with the fewest number of centroids.

Mean-Shift Clustering: Mean-shift clustering is a nonparametric clustering method that requires no prior knowledge of the number of clusters or the shape constraints of the clusters. The purpose of mean lag groupi ng is to find "points" in the uniform distribution or density of the samples.

DBSCAN: Density-Based Spatial Clustering for Noisy Applications (DBSCAN) is a basic density-based clustering algorithm widely used in data mining and machine learning. It is a density-based nonparametric clustering technique used to separate dense clusters from sparse clusters in model building.

GMM Clustering: The Gaussian Mixture Model (GMM) is a distribution-based clustering algorithm commonly used to cluster data. The Gaussian mixture model is a probabilistic model where all data points are generated by combining a finite number of Gaussian distributions with unknown parameters.

Clustering Hierarchical clustering: Cluster clustering is the most common method of hierarchical clustering and is used to group objects into clusters based on similarity. This technique uses a bottom-up approach, where the algorithm first treats each object as a single cluster.

5. TECHNIQUES FOR REDUCING DIMENSIONALITY

Dimension reduction is the process of transforming a high-dimensional representation of data into a low-dimensional representation. With the tremendous increase in the amount of multidimensional data, the use of various dimensionality reduction techniques has become popular in a wide range of applications. In addition, new modern approaches are constantly emerging. Dimension reduction techniques take a high-dimensional dataset and transform it into a low-dimensional dataset, preserving as much of the data's original meaning as possible. A low-dimensional representation of the original data helps to solve the dimensionality problem. Low-dimensional data is easier to analyze, process and visualize [57].

Applying dimensionality reduction techniques to datasets has several advantages. Data storage space can be reduced as the number of dimensions decreases. Calculations take very little time. Redundant, irrelevant or noisy data can be removed. Data quality can be improved. Some algorithms don't work well when the dimensionality increases. Consequently, reducing this size makes the algorithm more efficient and accurate. Seeing big data is difficult. Consequently, the smaller dimensions allow for more precise design and study of the model. It simplifies the classification process while increasing efficiency [46].

ISSN 1947-5500

In general, dimensionality reduction techniques fall into two categories. In other words, dimensionality reduction can be achieved using two different techniques. Select and extract features. Some features may need to be excluded during the feature subset selection process, so information may be lost during feature selection. However, feature extraction can reduce dimensionality without significantly reducing the original feature dataset [33].

Multidimensional data processing is a challenging task for machine learning, data science and application developers. This is why dimensionality reduction, an unsupervised learning technique, is important. This is intended to improve human interpretation, reduce computational costs, and avoid overfitting and redundancy by simplifying the model. Dimension reduction can be achieved using feature selection methods and feature extraction methods. The main difference between feature selection and feature extraction is that feature selection keeps a subset of the original features, while feature extraction creates entirely new features. These techniques are briefly described in the following section [45].

Feature Selection: The process of selecting a subset of unique features (variables, predictors) to use in building models for machine learning and data analysis is called feature selection, also known as variable or data attribute selection. Reduces model complexity by removing non-essential or non-essential features, allowing for faster training of machine learning algorithms. The correct and optimal subset of selected features in the problem domain can reduce overfitting by simplifying and generalizing the model while increasing its accuracy. As a result, “feature selection” is recognized as one of the key machine learning concepts that has a significant impact on the efficiency and effectiveness of objective machine learning models. Common feature selection techniques include chi-square test, analysis of variance (ANOVA), genetic algorithms, Pearson correlation coefficients, and recursive feature removal [17].

Feature Extraction: Feature extraction techniques in machine learning-based models or systems typically provide insight into the data, improve prediction accuracy, and reduce computational costs or training time. The goal of "feature mining" is to reduce the number of features in a dataset by creating new ones from the old ones and deleting the old ones. Much of the information contained in the original feature set can be aggregated using new feature sets such as PCA, PSO, ACO, LSA, LDA, ICA, , and PLS. For example, principal component analysis (PCA) is often used as a dimensionality reduction technique to extract a low-dimensional space with brand new components from existing features in a data set [32].

Many algorithms have been proposed in the machine learning and data science literature to reduce the dimensionality of data. Machine learning is gaining popularity in various application domains due to its ability to learn from the past and make smarter decisions in the current Fourth Industrial Revolution (4IR) era. Image recognition is a well-known and widely used machine learning practice that can identify objects as digital images. Image recognition is widely used to label cancer X-rays, recognize letters and faces in images, and provide label suggestions on social media platforms like Facebook. Pattern recognition is the automatic recognition of patterns and patterns in data such as: image analysis. Various machine learning techniques are used in this field, such as: Classification, feature selection, clustering and sequence labeling. Models based on machine learning can also be used in many other areas. The nature and characteristics of the data and the performance of learning algorithms determine the effectiveness and efficiency of machine learning solutions. Collecting data in the related field “Machine Learning Applications” is not easy, but today's cyberspace can regularly generate large amounts of data. Therefore, collecting and managing useful data for your target machine learning based application is an important factor for further analysis, and when it comes to real data, you need to delve deeper into how the data is collected must be taken into account. In addition, historical data can contain a large amount of ambiguous, missing, anomalous, and nonsensical data.

There are many machine learning algorithms that analyze data and generate insights. However, choosing the right learning algorithm for the target application is difficult. Because different learning algorithms can lead to different results depending on the nature of the data. Choosing the wrong learning algorithm can lead to unexpected results and loss of model effort, efficiency, and accuracy. The ultimate success of machine learning solutions and related applications depends heavily on data and learning algorithms. If the data is unsuitable for training, unrepresentative, poor quality, irrelevant features, or insufficient training can render a machine learning model ineffective or less accurate. Efficient data processing and handling of different learning algorithms is therefore essential for machine learning solutions and ultimately for building intelligent applications.

6. CONCLUSION

The main goal of this article is to describe the basic concepts of machine learning models that can work for blurry text detection in wild scenes in experimentation. Machine learning has attracted the attention of scientists around the world who are working to develop advanced models. This research will help scientists improve their research. Also, we talk about several datasets for scene text recognition. On the other hand, we delve into dimension reduction and clustering algorithms used in image analysis. Further research in the future suggests the development of methods that can be accurately predicted and evaluated by related evolutionary algorithms. Additionally, a combination of two or more of the techniques described here could help researchers develop new techniques for detecting blurry text in the real world.

ISSN 1947-5500

7. ABBREVIATIONS

ML - Machine Learning

DR - Dimensionality Reduction

FS - Feature Selection

FE - Feature Extraction

2DPCA

PCA - Principal Component Analysis

GA - Genetic Algorithm

CNN - Convolutional Neural Network

LDA - Linear Discriminant Analysis

RP - Random Projection

SFA - Simplified Fast Alexnet

GA-BPN- Genetic Algorithm Backpropagation Network

MSER - Maximal Stable Extreme Regions

SVM - Support Vector Machine

RF - Random Forest

T-SVM - Twin SVM

LST-SVM- Least Square Twin SVM

Hcs - Histogram of curviness Saliency

LFV - Labelled Face in the Wild

YOLO - You only look once

SVT - Text for Street View

SP - Structured Text Perspective

CT - Curved Text

CTW - Chinese Text in the Wild

ToT - Total text

MLT - Multilingual Text

SVHN - Street View House Number

DT - Decision Tree

KNN - K-Nearest Neighbor

LR - Logic Regression

8. DECLARATIONS

Availability of data and material: Not Applicable because this is a review paper.

Competing interests: No Competing Interest.

Funding: Not applicable. No Funding received.

Authors' contributions: Nwufoh Chinonyelum Vivian (NCV) and Wilson Sakpere (WS) are the two authors for this paper. NCV and WS both contributed to all sections of this paper (Abstract, Introduction, Related works, Material and Methods, Machine Learning, Techniques for Dimensionality Reduction and Conclusion)

Acknowledgements: I sincerely appreciate and acknowledge the support of Dr. Olaolu Arowolo who assisted with the conceptualization of this study.