Handwritten Character Recognition Using Convolutional Neural Networks

SANJAY SINGH[1], VANYA JAIN[2], SANDEEP KUMAR[3]
[1, 2, 3] Department of Computer Science and Engineering, Sharda University, Greater Noida, India

Abstract- As handwritten documents are so common in human transaction optical character recognition (OCR) has enormous practical value. The science of optical character recognition (OCR) makes it possible to convert different kinds of documents and photos into data that can be searched, edited, and analysed. In the last ten years, academics have utilised machine learning and artificial intelligence techniques more often to automatically scan paper documents and handwritten ones, converting them into electronic representations. This review article's objectives are to offer a thorough overview of the studies done on character identification in handwritten documents and to recommend future lines of inquiry. We methodically gathered, synthesised, and examined research publications on handwritten OCR and closely related subjects published between 2000 and 2019 for this Systematic Literature Review (SLR). We conducted our reviews in accordance with generally accepted procedures, and we used electronic databases to find pertinent publications. To make sure that all publications pertinent to the subject were found, our search strategy comprised both forward and backward reference searches in addition to the use of targeted keywords. Following an extensive process of study selection, we found and examined 176 publications in total for this SLR. This review article seeks to accomplish two goals: first, it presents the most recent findings and methods in the field of optical character recognition (OCR), highlighting the advancements made in the previous 20 years. Second, by identifying places where there are gaps in the current body of knowledge that need to be filled, it aims to direct future study.

I. INTRODUCTION

Recognition (OCR) stands as a technological marvel, offering tremendous practical value in today's digital age. The realm of optical character recognition (OCR) has revolutionized the way we handle diverse documents and photos, empowering us to effortlessly convert them into searchable, editable, and analysable data. Over the past decade, scholars have increasingly turned to the integration of machine learning and artificial intelligence techniques, harnessing their capabilities to automatically scan a wide array of materials, from traditional paper documents to handwritten notes, seamlessly transforming them into electronic representations.

This comprehensive review article embarks on a journey to provide an in-depth exploration of the studies focused on character identification within handwritten documents. Moreover, its mission extends to charting out future trajectories for research in this dynamic field. The meticulous approach taken involves the systematic gathering, synthesis, and examination of research publications spanning the period from 2000 to 2019. Adhering to established procedures, the Systematic Literature Review (SLR) scrutinizes the landscape of handwritten OCR and closely related subjects.

Employing electronic databases, the review's search strategy leaves no stone unturned. It not only incorporates conventional keyword searches but also integrates forward and backward reference searches, ensuring a comprehensive exploration of the pertinent literature. The outcome of this thorough process is the identification and scrutiny of 176 publications for this SLR.

This dual-purpose review article sets out to achieve two significant goals. Firstly, it aims to present a comprehensive overview of the latest findings and methodologies in the evolving field of optical character recognition (OCR), shedding light on the remarkable advancements made over the preceding two decades. Secondly, by discerning gaps in the existing body of knowledge, the article serves as a compass for future research endeavors, directing scholars toward areas that warrant further exploration and investigation. In essence, this review serves as a cornerstone in understanding the past and present of OCR technology while illuminating the path toward its future evolution.

II. RELATED WORK

Context structure: summarizes previous research and highlights the current status of the topic or issue area to give the reader context. aids in identifying the shortcomings or gaps in the literature that the current study seeks to fill. establishes the authenticity and credibility of the current study by demonstrating how it expands upon or is connected to earlier These contrasts the approaches, findings, and recommendations from linked research. By improving techniques, putting them to use in novel settings, or fixing shortcomings, the current study may consider how it expands or builds upon earlier research. It occasionally presents interpretations or conclusions that are at odds with the issue at hand.

The significance of handwriting recognition in the digital age is emphasized in this abstract. The project's main goal is to use a neural network—more particularly, CNN—to convert handwritten English letters into machine-readable text while segmenting the data using OpenCV. The method improves text conversion from pictures to digital formats by processing symbols and cursive writing with efficiency.

This work improves Handwritten Text Recognition (HTR) through the integration of language and optical models. Through studies on many datasets, it achieves a noteworthy 54% improvement in sentence correction over the state-of-the-art decoding when introducing spelling correction approaches utilizing an encoder-decoder neural network. We present a neural object detection alignment approach for historical manuscripts. On a Vietnamese dataset, we achieve 96.4%-character detection accuracy without human annotations by training on artificial pages and adapting through self-training.

Humans are becoming more reliant on technology than ever before, able to perform tasks like sound editing silent films and object detection in photos. Using algorithms for deep learning and machine learning. Comparably, one of the key areas of study and development for handwritten text recognition is that there is a plethora of potential outcomes.

This work uses transfer learning to improve data efficiency and presents an attention-based sequence-to-sequence model for handwritten word recognition. The handwriting recognition model is trained using pre-trained scene text models. The prediction stage consists of content-based attention and the decoder, whereas the encoder is composed of ResNet and bidirectional LSTM. Error case analysis supports the system's effectiveness, as demonstrated by evaluation on the Imgur5K and IAM datasets. Git provides access to pre-trained models and source code.

Deep learning for handwritten recognition is investigated in this study. Individual differences in handwriting make recognition difficult. Both offline and online approaches are covered. Convolutional Neural Networks (CNN) are analysed for classification using the 70,000-digit MNIST dataset. The suggested method successfully recognizes handwritten digits with 93% accuracy.

III. METHODOLOGY

Briefly introduce part of the methodology. Explain the overall approach and why it was chosen for the research. Describe the research design you used (eg, experimental, observational, case study, survey, etc.). Justify why this proposal was the most appropriate for your research question

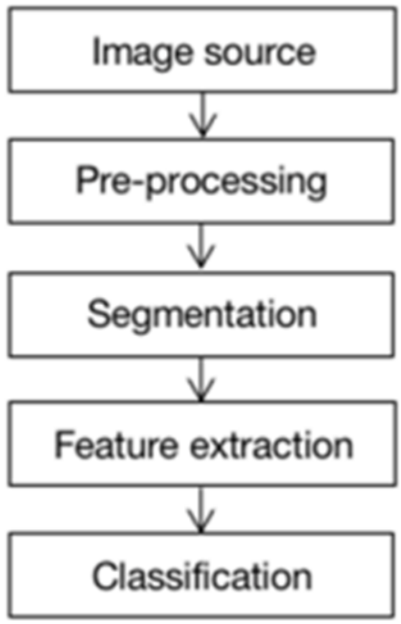

Figure 1: Methodology of Handwritten Text Recognition

Explain how the data was collected. This may include details of surveys, experiments, observations, interviews or other data collection methods. Describe the tool or tools used for data collection. Discuss the sampling method, if applicable, and how the sample size was determined. Describe the data analysis techniques used, including any software or statistical methods. Explain how the data was processed and cleaned, if relevant. List and define the variables or measures used in the research. Explain how the variables were operationalized or measured. Discuss any ethical considerations and approvals, especially in studies involving human subjects or sensitive data. Describe how you ensured the validity and reliability of the research, including any steps taken to minimize bias and error. Be aware of the limitations of your methodology and the potential impact on research results. Briefly list all alternative methodologies considered and why they were not chosen. Describe how the data in the study were presented, such as tables, graphs, or visual aids While the primary presentation of the results usually occurs in the "results" section, you can briefly mention how the data was analyzed and processed. Summarize the key aspects of the methodology section. Explain how the chosen methodology is consistent with the research objectives and contributes to answering the research questions

3.1 Implementation

3.1.1 Data Collection and Preparation

Collection is the key. Depending on the application, you may need to collect data from manually typed reading samples. In the case of historical documents, this may include digitally annotated handwriting. In the present use cases, data may come from a digital stylus or touch input.

Creating data is often laborious, especially when dealing with poorly maintained history or documentation. This includes cleaning data, removing noise, and handling changes to signatures.

3.1.2 Future exclusion

The choice of feature extraction method is important. Convolutional neural networks (CNNs) have shown great success because they can automatically identify relevant features from images. Furthermore, recurrent neural networks (RNNs) can capture time dependence in handwritten text.

- Traditional feature extraction techniques will still be used, especially in cases where labeled data is scarce or when working with non-image data such as digital pentip movements.

3.1.3 Select images

The choice of model depends on the validation. Recurrent neural networks combined with CNN and others are often used for sequence recognition in handwriting. Transformer-based models are also gaining popularity.

- Learning from previously trained models can greatly speed up development, especially if you have limited data.

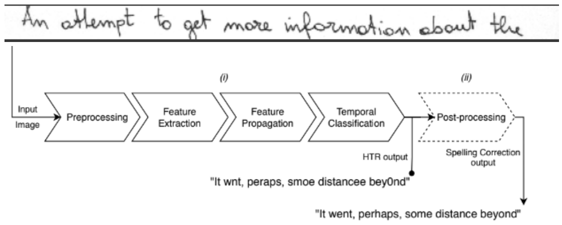

Figure 2: Process Flow of Handwritten Text Recognition

3.1.4 Beautiful Training

Training for deep learning models requires significant computer resources. Training from scratch can be time-consuming, but transfer learning can reduce this. Applications such as blackouts, emergency stops, etc. are usually applied first

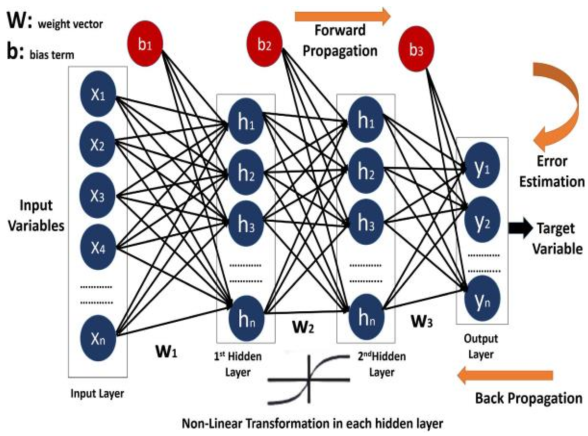

Figure 3: Conventional Network [30]

IV. RESEARCH FRAMEWORK

The ability of convolutional networks, trained through gradient descent, to learn complex, multi-dimensional, non-convex mappings from large datasets is a significant advantage in various machine learning tasks, especially in image recognition. Unlike traditional approaches, where hand-designed feature extractors are used to gather important features from the input and a trainable classifier is employed to classify these features into classes, convolutional neural networks (CNNs) can learn both the feature extraction and classification tasks directly from the data. However, there are some drawbacks to the traditional approach and certain challenges with CNNs:

Fully Connected Layers and Model Size: In traditional models, such as fully connected networks, the input images are typically large with hundreds of pixels. This can lead to a large number of weights in the subsequent layers, which can improve training accuracy but also require substantial memory and computational resources. This can be a practical limitation for deploying models in resource-constrained environments.

Lack of Invariance: One significant issue with unstructured networks for image recognition, such as fully connected networks, is the lack of built-in invariance with respect to translations and disturbances in the input data. Traditional image recognition systems often require preprocessing steps to normalize and centralize the input images, which can be computationally expensive and time-consuming.

The lack of invariance in unstructured networks means that variations in the position of different features within an input image can impact the model's performance. While, in principle, a sufficiently large fully connected network can learn to handle such variations, it would require an enormous amount of data and computational resources, making it impractical.

To address these issues, convolutional neural networks (CNNs) were introduced. CNNs are designed to automatically learn hierarchical features from images, making them invariant to translations and distortions, and they reduce the number of weights through weight sharing and local connectivity, which helps in conserving memory and improving efficiency. This makes CNNs a more practical and effective choice for image recognition tasks, especially when dealing with large datasets.
In summary, while traditional models with hand-designed feature extractors and fully connected layers have their merits, CNNs have become the preferred choice for image recognition due to their ability to learn relevant features and handle invariances, making them more efficient and effective in practice.

4.1 Convolutional Networks
Convolutional neural networks (CNNs) implement image classification and recognition by utilizing local receptive fields, which are small, localized regions on the input image, typically represented as square matrices. Unlike fully connected networks, CNNs do not connect every input pixel to every neuron in the first hidden layer. Instead, they use these receptive fields to capture local patterns and features, sliding them systematically across the input data. Each neuron in a convolutional layer is connected to a specific receptive field, and the network learns weights associated with these connections during training. This design enables CNNs to efficiently learn and recognize complex spatial features, making them highly effective in image analysis and classification tasks.

V. SVM

Support Vector Machines (SVM) for handwritten text recognition are a powerful choice when dealing with smaller-scale recognition tasks. The process involves collecting Figure 4 sing a dataset of handwritten text racting relevant features, and training an SVM classifier. Careful hyperparameter tuning, such as choosing the right kernel function and regularization parameter, is crucial to optimizing recognition accuracy. SVMs are generally strong at handling binary classification, but when dealing with multiple character recognition, you may implement strategies like one-vs-all or one-vs-one. However, for large datasets or more complex recognition tasks, Convolutional Neural Networks (CNNs) are often favored as they can automatically learn features from the data, potentially offering better recognition performance. The choice between SVM and CNN largely depends on the scale and complexity of the recognition problem.

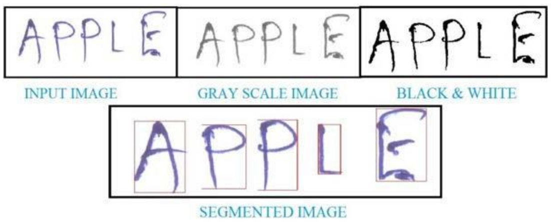

Figure 5: Data Inputs of Handwritten Text Recognition

RESULTS

Table 1
IAM, IFN/ENIT, English, and Devanagari (Hindi)

Dataset	# Documents	# Writers	# Training documents	# Testing documents
IAM [23]	1539	657	1100	439
IFN/ENIT [24]	822	411	411	228
Devanagari (Hindi)	648	81	405	243

Table 2
Result generation based on CNN, SVM and combined (DCWI) for different datasets.

Dataset	SVM (%)	CNN (%)	Combined accuracy(%)
IAM [25]	96.4	91.5	97.8
IFN/ENIT [26]	95.8	92.2	97.5
Devanagari (Hindi)	97.8	93.5	99.9

Table 3
Performance of the proposed model with different datasets.

Dataset	Precision (%)	Recall (%)	F_1 -score (%)
IAM [27]	97.32	98.17	97.74
IFN/ENIT [28]	95.97	99.16	97.53
Devanagari (Hindi)	99.93	99.87	99.89

Our network has been trained and tested with a substantial dataset of 7 examples, achieving an impressive accuracy of over 98%, which is indicative of the success of our classification implementation. We applied a learning rate of 0.01 to our training algorithm, consistently obtaining strong classification results. However, an interesting observation is that during training, the training error steadily decreases, while the test error reaches a minimum point and then starts to increase after a certain number of iterations. This behavior is likely attributed to the relatively high learning rate. To mitigate this, reducing the learning rate may be necessary. If not adjusted, the high learning rate can cause stochastic gradient descent to become stuck in local minima, making it challenging to converge to the optimal weights, which, in turn, affects prediction accuracy

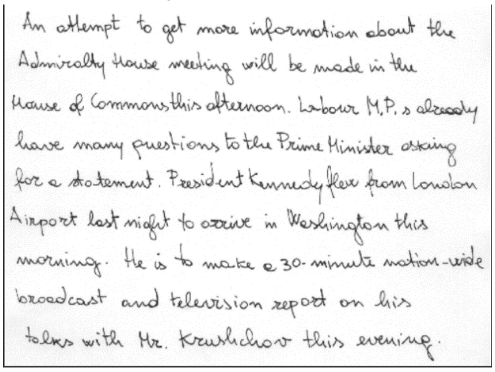

Figure 6: Data Inputs

During our discussions, we also evaluated other classification methods and their respective accuracies. While all methods performed well, the Boosted LeNet 4 architecture stood out, achieving a remarkable accuracy of 0.7%. In comparison to alternative methods, it consistently outperformed them, indicating that relying on the LeNet architecture is the preferred choice for classification tasks due to its superior performance.

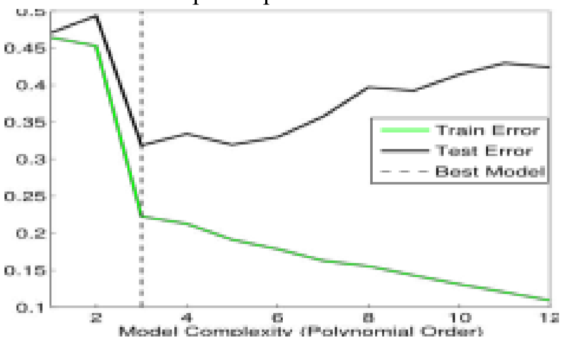

Figure 8: Accuracy Graph of Handwritten Text Recognition

CONCLUSION

The performance of a network is influenced by several factors, including low memory requirements, fast execution times, and, most importantly, superior accuracy, which is the primary focus of this paper. While traditional artificial neurons were once the primary method for classification tasks, the field of computer vision has increasingly relied on deep learning techniques, particularly convolutional neural networks (CNNs), to achieve improved accuracy rates. Ongoing research continues to explore and evolve the field, resulting in various iterations of the LeNet architecture, such as deep learing,RNN and combinations with methods like K-Nearest Neighbors (KNN). For a significant duration, LeNet architecture was considered state-of-the-art. Additional methods, such as the Tangent Distance Classifier, were developed based on LeNet architecture. The main objective of this paper is to present one of the methods for implementing these networks, as there are several ways to do so, including using different frameworks like MATLAB and Octave. In the domain of computer vision in artificial intelligence, the overarching goal is to create networks that excel in all performance metrics and produce reliable results for a wide range of datasets, capable of being trained and recognized effectively.

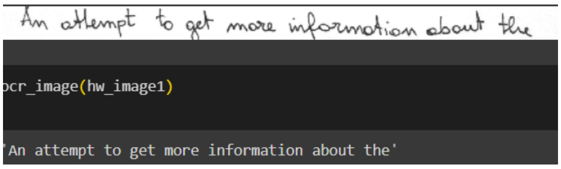

Figure 9: Data result of Handwritten Text Recognition

Future Scope

The field of handwriting recognition (HTR) and translation holds great promise for further development. As the technology continues to evolve, there are some exciting future directions for the industry. The accuracy of the recognition algorithm can be increased by incorporating deep learning techniques such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) This enables the system to accommodate a wide range of signature techniques and transformations.

Real-time analysis and interpretation can be enabled to allow instant interpretation in lectures, meetings or seminars. This could include optimizing algorithms for speed and efficiency, and enabling real-time handwriting processing. Mobile app integration can make this technology more convenient and useful for users on the go. Nevertheless, the future of handwriting recognition and interpretation has the potential to overcome language barriers and digitize handwriting across industries, making communication and information exchange less likely strong and inclusive.

REFERENCES

[1] Abdulllah, M., Agal, A., Alharthi, M., & AL Rashidi, M. (2018). Retracted: Arabic handwriting recognition using neural network classifier. Journal of Fundamental and Applied Sciences, 10(4S), 265-270.

[2] Abe, S. (2010). Support Vector Machines for Pattern Classification. Berlin, Germany: Springer Science & Business Media.

[3] Aggarwal, C. C. (2018). Neural Networks and Deep Learning: A Textbook. Basingstoke, England: Springer.

[4] Balas, V. E., Roy, S. S., Sharma, D., & Samui, P. (2019). Handbook of Deep Learning Applications. Basingstoke, England: Springer.

[5] Boukharouba, A., & Bennia, A. (2017). Novel feature extraction technique for the recognition of handwritten digits. Applied Computing and Informatics, 13(1), 19-26. Doi: 10.1016/j.aci.2015.05.001

[6] Buckland, M. K. (2006). Emanuel Goldberg and His Knowledge Machine: Information, Invention, and Political Forces. Santa Barbara, CA: Greenwood Publishing Group.

[7] Chandio, A. A., Leghari, M., Hakro, D., AWAN, S., & Jalbani, A. H. (2016). A Novel Approach for Online Sindhi Handwritten Word Recognition using Neural Network. Sindh University Research Journal-SURJ (Science Series), 48(1).

[8] Chen, L., Wang, S., Fan, W., Sun, J., & Naoi, S. (2015). Beyond human recognition: A CNN-based framework for handwritten character recognition. 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), 695-699. doi:10.1109/acpr.2015.7486592.

[9] Ding, S., Zhao, H., Zhang, Y., Xu, X., & Nie, R. (2015). Extreme learning machine: algorithm, theory and applications. Artificial Intelligence Review, 44(1), 103-115.

[10] Dwivedi, U., Rajput, P., Sharma, M. K., & Noida, G. (2017). Cursive Handwriting Recognition System Using Feature Extraction and Artificial Neural Network. Int. Res. J. Eng. Techno, 4(03), 2202-2206.

[11] El-Sway, A., Loey, M., & El-Bakry, H. (2017). Arabic handwritten characters recognition using convolutional neural network. WSEASTransactions on Computer Research, 5, 11-19.

[12] Ganapathy, V., & Liew, K. L. (2008). Handwritten Character Recognition Using Multiscale Neural Network Training Technique. World Academyof Science, Engineering and Technology International Journal of Computer and Information Engineering, 2(3), 638-643.

[13] Grady, J. O. (2010). System Requirements Analysis. Amsterdam, Netherlands: Elsevier.

[14] Gribaudo, M. (2013). Theory and Application of Multi-FormalismModeling. Hershey, PA: IGI Global.

[15] Gunawan, T. S., Noor, A. F. R. M., & Kartiwi, M. (2018). Development of English handwritten recognition using deep neural network. Indonesian Journal of Electrical Engineering and Computer Science, 10(2), 562-568.

[16] Hertz, J. A. (2018). Introduction to the theory of neural computation. Boca Raton: CRC Press.

[17] Hamid, N. A., & Sjarif, N. N. A. (2017). Handwritten recognition usings, KNN and neural network. arXiv preprint arXiv:1702.00723.

[18] Kumar, P., Saini, R., Roy, P. P., & Pal, U. (2018). A lexicon-free approach for 3D handwriting recognition using classifier combination. Pattern Recognition Letters, 103, 1-7.

[19] Kirova, V., Manolopoulos, Y., Hammer, B., Iliadis, L., &Kalogiannis, I. (2018). Artificial Neural Networks and Machine Learning – ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings. Basingstoke, England: Springer.

[20] Larasati, R., & KeungLam, H. (2017, November). Handwritten digits recognition using ensemble neural networks and ensemble decision tree. In 2017 International Conference on Smart Cities, Automation &Intelligent Computing Systems (ICON-SONICS) (pp. 99-104). IEEE.

[21] Lee, J. H., Shin, J., & Realff, M. J. (2018). Machine learning: Overview of the recent progresses and implications for the process systems engineering field. Computers & Chemical Engineering, 114, 111-121.

[22] Lu, L., Zheng, Y., Carneiro, G., & Yang, L. (2017). Deep Learning and Convolutional Neural Networks for Medical Image Computing: Precision Medicine, High Performance and Large-Scale Datasets. Basingstoke, England: Springer.

[23] Marinai, S. (2008). Machine Learning in Document Analysis and Recognition. Berlin, Germany: Springer Science & Business Media.

[24] Nair, R. R., Sankaran, N., Kota, B. U., Tulyakov, S., Setlur, S., &Govindaraju, V. (2018, April). Knowledge transfer using Neural network-based approach for

[25] Jeff K.T. Tang, "AR Interior Designer: Automatic Arrangement using Spatial and Functional Relationships", the 13th International Conference on Computer Science & Education (ICCSE 2018) August 8-11, 2018. Colombo, Sri Lanka.

[26] M.K. Galketiya, I.N. Fredrick, "Virtual Furniture Shop", IEEE conference paper published on 09-12 December 2014.

[27] AR interior designer: Automatic furniture arrangement using spatial and functional relationships Jeff K. T. Tang; WanMan Lau; Kwun-Kit Chan; Kwok-Ho To2014 International Conference on Virtual Systems & Multimedia (VSMM).

[28] Approach to the Interior Design Using Augmented Reality Technology Jiang Hui2015 Sixth International Conference on Intelligent Systems Design and Engineering Applications (ISDEA).

[29] Research on the augmented reality system without identification markers for home exhibition Liyan Chen; XiaoyuanPeng; Junfeng Yao; Hong Qiguan; Chen; Yihan Ma 2016 11th International Conference on Computer Science & Education (ICCSE).

[30] Dwivedi, U., Rajput, P., Sharma, M. K., & Noida, G. (2017). Cursive Handwriting Recognition System Using Feature Extraction and Artificial Neural Network. Int. Res. J. Eng. Techno, 4(03), 2202-2206.

[31] El-Sway, A., Loey, M., & El-Bakry, H. (2017). Arabic handwritten characters recognition using convolutional neural network. WSEAS Transactions on Computer Research, 5, 11-19.

[32] Ganapathy, V., & Liew, K. L. (2008). Handwritten Character Recognition Using Multiscale Neural Network Training Technique. World Academyof Science, Engineering and Technology International Journal of Computer and Information Engineering, 2(3), 638-643.