INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

IJARIIT

ISSN: 2454-132X

Impact Factor: 6.078

(Volume 8, Issue 2 - V8I2-1319)

Available online at: https://www.ijariit.com

Leveraging a convolutional neural network for developing a hand-written text recognition system

Likith Pulluri
pullurilikith1234@gmail.com
Sharda University, Greater Noida, Uttar Pradesh

Yashwanth Konakanchi
2018004614.konakanchi@ug.sharda.ac.in
Sharda University, Greater Noida, Uttar Pradesh

Divya Prabha
2018005685.divyaprabha@ug.sharda.ac.in
Sharda University, Greater Noida, Uttar Pradesh

Asima Nayak
2018016430.asima@ug.sharda.ac.in
Sharda University, Greater Noida, Uttar Pradesh

Pradeep Kumar Mishra
pradeepkumar.mishra@sharda.ac.in
Sharda University, Greater Noida, Uttar Pradesh

ABSTRACT

The explication revolving around the digital outlook of Hand-written Text has been carried out widely with the upsurge of Technological Advancements in the field of Artificial Intelligence. The automation in daily chores has made human life more consistent and yet stable. Machine Learning and Deep Learning modus operandi has been actuated for contemplating the automated outlook. Hand-written Text Recognition has been a well-scrutinized domain wherein the incorporation of different paradigms of learning has been elucidated. The escalation into digitization of almost everything, necessitates the prevalence of the same in assuaging the Text Recognition System as well. In this articulation, we've consummated Convolutional Neural Networks (CNNs) i.e., 03-Layered, 04-Layered and 05-Layered Convolutional Neural Networks (CNNs) which is trained over Hand-written Characters which includes Alphabets (A-Z), Digits (0-9) and widely utilized Symbols and thereby detecting the Text from a Dataset of diverse Hand-written words consisting of Names and Surnames. The proposed Hand-written Text Recognition System is structured in Offline method, as scarcity of proficient outcome revolves around it, when compared to Online prospect. The proposition also gave us a comparative insight over the impact of layers in models. The efficacy we achieved was phenomenal, which gave us a potent resolution for contemplating it into real-time.

Keywords *— Convolutional Neural Network (CNN), Hand-Written Text Recognition, Optical Character Recognition (OCR), Deep Learning.*

1. INTRODUCTION

The ability of a Computation System to recognize the Hand-written Input of an Individual can be contemplated as an Automated Hand-written Text Recognition System. There can be diverse Hand-Written prospects i.e., Documents in Different Formats, Text over Images (Screen-shots), etc. which can be targeted for Text Recognition [1]. Moreover, Offline and Online are the two methodological outlooks for Hand-written Text Recognition [2]. In the Online Method, the data is collated through sensors when some text is written, thereby formulating the data into the information to be dynamically available as per the fences of the System. Contrarily, in the Offline Method, the Text of an individual is made available through an Image or the scanned documents i.e., pdfs, screenshots, etc. and thus the hand-written Text is detected through those references. To summarize the whole context, the Online Method proposes a method where the Text is recognized "During the Writing Process", whereas the Offline Method proposes the recognition of Text "After the Writing Process" [3]. The Online Method has been explicated well in the past years, but there revolves a scantiness in the Offline Method modus operandi [4]. As a result, we approached for the Offline Method of Hand-written Text Recognition. Developing a Hand-written Text Recognition System commemorates varied shortcomings such as Difference in Shapes, Sizes and Patterns of Letters in the Handwriting of diverse Individuals [5], thereby resulting in the deterioration of the overall Hand-written Text Recognition efficacy. Thus, it's prominent to explore the potent systems which extracts the prime features [6]. Deep Learning has made its mark in the field of Feature Extraction widely and as a result it's utilized eminently in diverse problem statements [7].

Thus, we instantiated with the process of developing an efficient Hand-written Text Recognition System pursuing the Offline methodological outlook and contemplating Convolutional Neural Network (CNN) over profound Dataset. We utilized the Hand-written Characters Dataset consisting of Alphabets (A-Z), Digits (0-9), and Symbols for training our model and further utilized Hand-written Words Dataset which consisted of Names and Surnames of varied individuals for Recognition of Text and Post-Processing. Furthermore, we inculcated 03-Layer, 04-Layer, and 05-Layer Convolutional Neural Network for explicating the prowess of modular depth over system's efficiency. To be more precise, the ultimate motive of comparing differently layered network was to understand the impact of depth over a problem outlook. The output we got were promising and it provided us a well elucidated prospect contemplating Hand-written Text Recognition System.

Furthermore, the Paper flows through the Literature Review in Section-II, with detailed Methodological Outlook in Section-III along with the Implementation Prospect in Section-IV, Experimental Results and Analysis in Section-V and finally the Conclusion and Future Scope of the modus operandi in Section VI.

2. RELATED WORK

Before initiating with the proposed modus operandi, we scrutinized diverse workflows induced for creating Hand-written Text Recognition System.

Bi-Directional Long Short-Term Memory (Bi-LSTM) technique along with Convolution Neural Network (CNN) was imposed over 6000 unidentical Hand-written Texts for assuaging Cursive Text Recognition System. This proposition gained an efficacy of 83.69% which is poor in case of real-time applicability [1]. An explication revolving around detection of Doctor's Hand-written Text utilizing Deep Convolutional Recurrent Neural Network (CRNN) was also carried out exhibiting a Training and Validation Accuracy of 76% and 72% respectively. The proposition lacked the availability of training dataset which thereby deteriorated the overall modular efficacy [2]. Singular Approach for Character Recognition utilizing MNIST Hand-written Dataset, trained over Convolutional Neural Network (CNN) was also implored for detecting Hand-written Text in Random Dataset. The accuracy for Character Recognition was not promising as the graphical outlook demonstrated the condition of Overfitting [3].

Gujarati Text Recognition System was also proposed utilizing Image Processing and Deep Learning Prospects but the accuracy was trivial due to complex nature of Gujarati Characters [4]. Going on to the next-level of Text Recognition proposition, Hand-written Text Recognition in the availability of struck-out text was also elucidated utilizing Convolutional Recurrent Neural Network (CRNN). Thus, with this inculcation the Character Error Rate (CER) incremented from 0.09 to 0.11 giving a poor outlook for the inducted modus operandi [5]. The elucidation over Text Block, on the basis of Manuscripts from the Thai Archives was also contemplated utilizing the Hybrid approach amalgamating Convolutional Neural Network (CNN) for Training and Connectionist Temporal Classification (CTC) for expelling Loss. Here, we got the Character Error Rate of 11.9 which is not promising and more scrutiny can be done for higher efficacy [7]. Kannada Hand-written Character Recognizer modus operandi was also proposed inducing the Offline Method of Text Recognition. The approach got different accuracies for different dataset i.e., 93.2% for one and 78.73% for the other. These efficacies exhibited inconsistent modular outlook for real-time application [8]. Moreover, a survey on different Machine Learning and Deep Learning prospects was carried out demonstrating the existing modus operandi's for developing Hand-written Text Recognition System [10].

Furthermore, an elucidation based on the Online and Offline methodological prospects for Character Recognition and thereby Hand-written Text Recognition was explicated demonstrating the structural outlook for the Optical Character Recognition (OCR) [12]. Style-Conditioned Generative Adversarial Network (SC-GAN) was also implored for escalating Hand-written Optical Character Recognition (OCR) leveraging the Augmented Text Line Images. The proposed modus operandi managed to get Character Error Rate (CER) of 4.2, which can be improved with fine tuning the architectural outlook [13]. Moreover, Japanese Text Recognition System was also proposed utilizing Semantic Segmentation-based algorithm, clocking an average accuracy of 92% overall [14].

The above explicated modus operandi gave us an ideation for formulating our proposed approach diminishing all the shortcomings i.e., Modular, Dataset and Complexity Limitations.

3. METHODOLOGY

The exegesis over the existing Hand-written Text Recognition System led us to the formulation of well-structured modus operandi. Initiating with methodological outlook, we constricted ourselves to Convolutional Neural Network (CNN), as it works phenomenally well with Image Dataset.

As a result, we went on training different layered CNNs and found that before 03-Layered CNN Architecture, the model was Underfitting and above 05-Layered CNN Architecture, the model was Overfitting. We explicated this modular outlook for varied epochs along with tuning the hyperparameters, but the results were same. Thus, we restricted ourselves with 03, 04 and 05-Layered CNN Architectures. Let's explore each CNN Architecture in detail.

A. CNN Architecture with 03-Layers:

A Convolutional Neural Network with 03-Layers comprising of different filter sizes i.e., (32, (5,5)), (64, 3,3)), (128, (3,3)) collated with Max-Pooling Layer for prime feature extraction ending with Fully-Connected (FC) Layers, formulating and generating the output. Figure.1, represents the diagrammatic outlook for 03-Layered CNN Architecture.

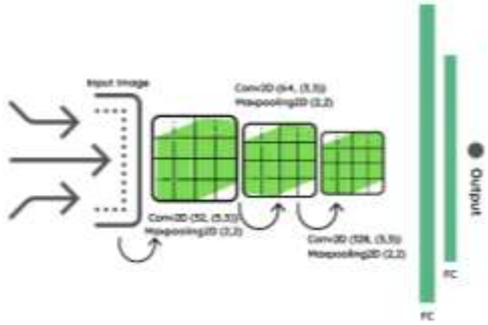

Fig.1. 03-Layered CNN Architecture.

B. CNN Architecture with 04-Layers:

A Convolutional Neural Network with 04-Layers comprising of different filter sizes i.e., (32, (5,5)), (64, 3,3)), (64, (3,3)), (128, (3,3)) collated with Max-Pooling Layer for prime feature extraction ending with Fully-Connected (FC) Layers, formulating and generating the output. Figure.2, represents the diagrammatic outlook for 04-Layered CNN Architecture.

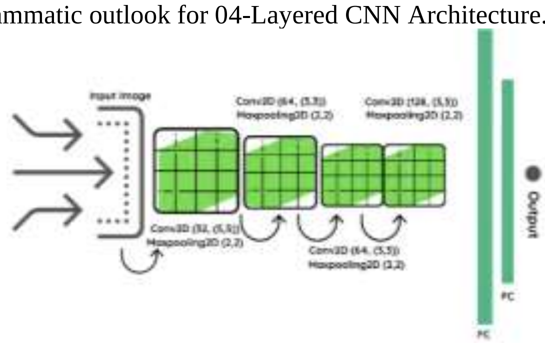

Fig.2. 04-Layered CNN Architecture.

C. CNN Architecture with 05-Layers:
A Convolutional Neural Network with 05-Layers comprising of different filter sizes i.e., (32, (5,5)), (64, 3,3)), (64, (3,3)), (128, (3,3)), (256, (3,3)) collated with Max-Pooling Layer for prime feature extraction ending with Fully-Connected (FC) Layers, formulating and generating the output. Figure.3, represents the diagrammatic outlook for 05-Layered CNN Architecture.

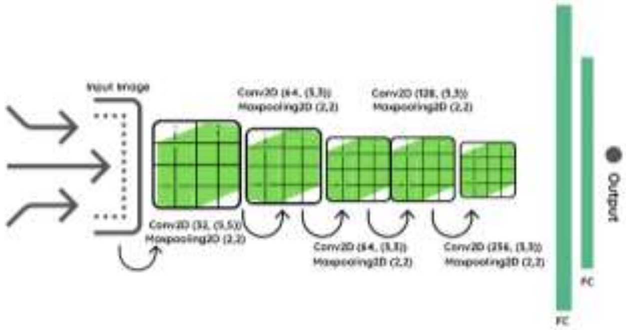

Fig.3. 05-Layered CNN Architecture.

Actuating the Hand-written Text Recognition System over a particular Image encapsulates varied parameters which directly or indirectly impacts the overall performance of an algorithm i.e., the Image Clarity [15], the Character Segmentation process [16], and the Algorithms utilized for Recognition [17].

Thus, the ultimate proposed modus operandi is summarized as,
Step-1: Building a Character Classifier System utilizing Convolutional Neural Network.
Step-2: Applying Character Segmentation for the Image consisting Hand-written Text.
Step-3: Getting the Final word in the Image, by classifying each segmented character.

The diagrammatic outlook for the proposed modus operandi has been elucidated in Figure. With the clarification of our proposed approach, we moved forward with the actuation of the same.

4. IMPLEMENTATION AND TOOLS

The structural analysis provided us a potent visualization of our outcome, deducting varied limitations implicitly seen in related work for Hand-written Text Recognition System. Let's see the Implementation process in detail to get a more elucidated prospect of our modus operandi.

A. The Dataset:
Here, we've utilized several Open-Source Dataset available over Web for our explication.

i] Hand-written Character Dataset: Formulating this Dataset, we utilized MNIST Handwritten Data [18] for Alphabets (A-Z) and Digits (0-9). Furthermore, we processed those Images and converted it into 32x32 pixel [19] in Dimension. We also collated the Dataset consisting of widely utilized Symbols i.e., (@, #, $, &).

Finally, we collated everything, thereby creating a Dataset consisting of Alphabets, Digits and Symbols comprising of 39 Categories. The total number of Instances in this Dataset stands to be 857,000.

ii] Hand-Written Text Dataset: The Hand-written Text Dataset consists of approximately more than 400,000 Hand-written Name and Surname Images. The segregation of the Dataset is such that, there are 206,799 Instances of First Names and 207,024 Instances of Surnames. Figure.4 demonstrates the outlook of both the Datasets i.e., (i) & (ii).

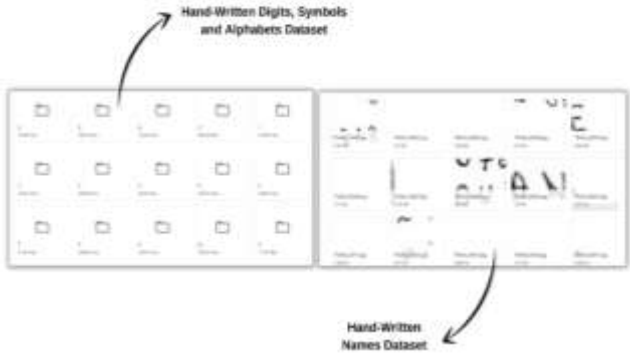

Fig.4. Diagrammatic Outlook for Utilized Dataset.

B. Tools Used:
We utilized the Frameworks such as TensorFlow and Keras. Image Data Generator was also used for Data Augmentation. The Training of our Model was contemplated through free GPU available over Kaggle Notebook IDE.

C. Metrics Used:
As we trained our model for Character Recognition, the Dataset was segregated into Train, Validation and Test Dataset. Furthermore, the Output is shown based upon the Character Recognition modus operandi performance. Thus, for gauging our efficacy we explicated Train, Validation and Test Accuracy for juxtaposing between utilized architectures.

D. Implementation Flow:
Thus, after collating every resource required, we carried out our modus operandi. Figure.5, depicts the applied workflow for Hand-written Text Recognition System. In a nutshell, we Trained each CNN Model with Hand-written Character Dataset. Then, we applied Character Segmentation over the Hand-written Text Dataset. Finally, the segmented letters in an Image are depicted as an Output.

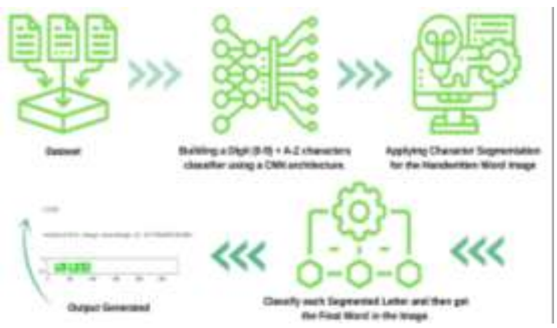

Fig.5. Demonstration of Proposed Modus Operandi.

5. EXPERIMENTAL RESULTS AND ANALYSIS

The implementation gave us potent results demonstrating the Train and Validation Efficacy along with Train and Validation Loss for each CNN Architecture in Figure.6, Figure.7 and Figure.8. through Graphical explication.

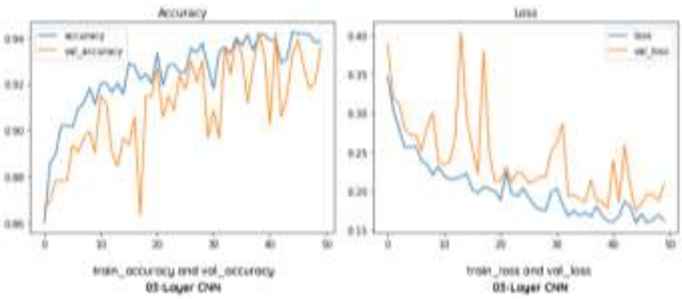

Fig.6. Graphical Demonstration of 03-Layered CNN Architecture Accuracy and Loss.

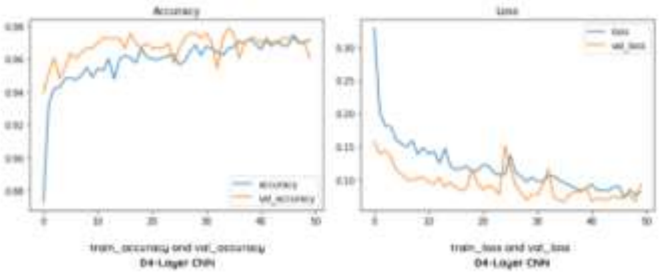

Fig.7. Graphical Demonstration of 04-Layered CNN Architecture Accuracy and Loss.

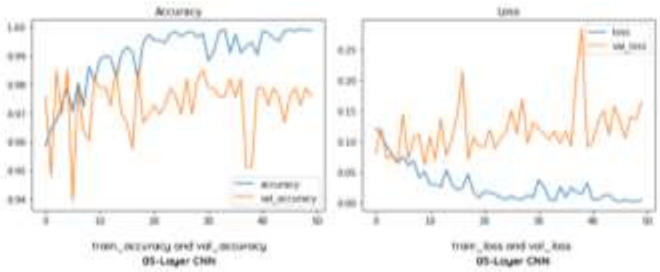

Fig.8. Graphical Demonstration of 05-Layered CNN Architecture Accuracy and Loss.

Moreover, Table – I demonstrates the accuracies for our 03, 04, and 05 Layered CNN Architecture for Training, Validation and Test.

Table 1. Comparative analysis of different models accuracies.

Model	Training Accuracy (%)	Validation Accuracy (%)	Test Accuracy (%)
03-Layer CNN	95.06	92.30	92.05
04-Layer CNN	96.50	95.89	93.15
05-Layer CNN	97.22	96.05	95.11

Contrasting our modus operandi with already existing techniques, it gave us an add-on edge in terms of Efficacy, in terms of Feasibility, in terms of Reduced Computational Complexity, and in terms of Real-time Applicability.

Figure.9 demonstrates the Output we procured through our elucidated modus operandi.

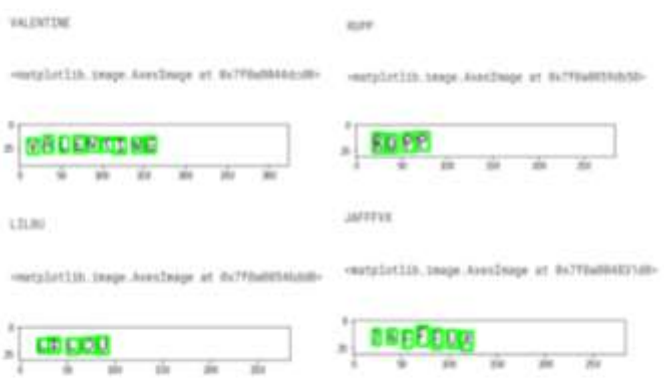

Fig.9. Demonstration of Procured Output.

6. CONCLUSION AND FUTURE SCOPE

The proposed approach exhibited the efficacy of Convolutional Neural Networks (CNNs) over the Image Dataset. The Highest Accuracy aggregated, was through 05-Layered CNN Model as shown in Table-I, giving us great results over our exegesis. The comparative analysis over these different layered CNN Architectures gave us an elucidation about the impact of Layers in a Network. We also got an insight that tuning of the Hyper-parameters play a major role in contemplating potent efficacies. Thus, through this explication we extracted the most efficient outlook for real-time Hand-written Text Recognition System.

Adding to the proposed modus operandi, more data could be accumulated for added potent outlook. Transfer Learning utilization along with addition of certain Layers can also be formulated for understanding the behavioural outlook of the architectures over this particular problem statement.

7. REFERENCES

[1] S. Hassan, A. Irfan, A. Mirza and I. Siddiqi, "Cursive Handwritten Text Recognition using Bi-Directional LSTMs: A Case Study on Urdu Handwriting," 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), 2019, pp. 67-72, doi: 10.1109/Deep-ML.2019.00021.

[2] L. J. Fajardo et al., "Doctor’s Cursive Handwriting Recognition System Using Deep Learning," 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), 2019, pp. 1-6, doi: 10.1109/HNICEM48295.2019.9073521.

[3] A. Garg, D. Gupta, S. Saxena and P. P. Sahadev, "Validation of Random Dataset Using an Efficient CNN Model Trained on MNIST Handwritten Dataset," 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), 2019, pp. 602-606, doi: 10.1109/SPIN.2019.8711703.

[4] S. Aniket, R. Atharva, C. Prabha, D. Rupali and P. Shubham, "Handwritten Gujarati script recognition with image processing and deep learning," 2019 International Conference on Nascent Technologies in Engineering (ICNTE), 2019, pp. 1-4, doi: 10.1109/ICNTE44896.2019.8946074.

[5] H. Nisa, J. A. Thom, V. Ciesielski and R. Tennakoon, "A deep learning approach to handwritten text recognition in the presence of struck-out text," 2019 International Conference on Image and Vision Computing New Zealand (IVCNZ), 2019, pp. 1-6, doi: 10.1109/IVCNZ48456.2019.8961024.

[6] G. Maillette de Buy Wenniger, L. Schomaker and A. Way, "No Padding Please: Efficient Neural Handwriting Recognition," 2019 International Conference on Document Analysis and Recognition (ICDAR), 2019, pp. 355-362, doi: 10.1109/ICDAR.2019.00064.

[7] R. Chamchong, W. Gao and M. D. McDonnell, "Thai Handwritten Recognition on Text Block-Based from Thai Archive Manuscripts," 2019 International Conference on Document Analysis and Recognition (ICDAR), 2019, pp. 1346-1351, doi: 10.1109/ICDAR.2019.00217.

[8] G. Ramesh, G. N. Sharma, J. M. Balaji and H. N. Champa, "Offline Kannada Handwritten Character Recognition Using Convolutional Neural Networks," 2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE), 2019, pp. 1-5, doi: 10.1109/WIECON-ECE48653.2019.9019914.

[9] N. Rahal, M. Tounsi, T. M. Hamdani and A. M. Alimi, "Handwritten Words and Digits Recognition using Deep Learning Based Bag of Features Framework," 2019 International Conference on Document Analysis and Recognition (ICDAR), 2019, pp. 701-706, doi: 10.1109/ICDAR.2019.00117.

[10] R. Sharma, B. Kaushik and N. Gondhi, "Character Recognition using Machine Learning and Deep Learning - A Survey," 2020 International Conference on Emerging Smart Computing and Informatics (ESCI), 2020, pp. 341-345, doi: 10.1109/ESCI48226.2020.9167649.

[11] M. M. Al-Taee, S. B. H. Neji and M. Frikha, "Handwritten Recognition: A survey," 2020 IEEE 4th International Conference on Image Processing, Applications and Systems (IPAS), 2020, pp. 199-205, doi: 10.1109/IPAS50080.2020.9334936.

[12] B. Türkmen and Y. Genç, "Handwriting Recognition via Visual Observation of Pen Movements," 2020 5th International Conference on Computer Science and Engineering (UBMK), 2020, pp. 58-61, doi: 10.1109/UBMK50275.2020.9219381.

[13] M. Guan, H. Ding, K. Chen and Q. Huo, "Improving Handwritten OCR with Augmented Text Line Images Synthesized from Online Handwriting Samples by Style-Conditioned GAN," 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2020, pp. 151-156, doi: 10.1109/ICFHR2020.2020.00037.

[14] K. C. Nguyen, C. T. Nguyen and M. Nakagawa, "A Semantic Segmentation-based Method for Handwritten Japanese Text Recognition," 2020 17th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2020, pp. 127-132, doi: 10.1109/ICFHR2020.2020.00033.

[15] A. Nikitha, J. Geetha and D. S. JayaLakshmi, "Handwritten Text Recognition using Deep Learning," 2020 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), 2020, pp. 388-392, doi: 10.1109/RTEICT49044.2020.9315679.

[16] J. H. Alkhateeb, A. A. Turani and A. A. Alsewari, "Performance of Machine Learning and Deep Learning on Arabic Handwritten Text Recognition," 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE), 2020, pp. 1-7, doi: 10.1109/ETCCE51779.2020.9350863.

[17] N. Darapaneni et al., "Handwritten Form Recognition Using Artificial Neural Network," 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), 2020, pp. 420-424, doi: 10.1109/ICIIS51140.2020.9342638.

[18] V. K. Vaisakh and L. B. Das, "Handwritten Malayalam Character Recognition System using Artificial Neural Networks," 2020 IEEE International Students' Conference on Electrical,Electronics and Computer Science (SCEECS), 2020, pp. 1-4, doi: 10.1109/SCEECS48394.2020.101.

[19] M. S, C. N. Vanitha, N. Narayan, R. Kumar and G. R, "An Enhanced Handwritten Digit Recognition Using Convolutional Neural Network," 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), 2021, pp. 1724-1727, doi: 10.1109/ICIRCA51532.2021.9544669.

