ENHANCING DEGRADED HANDWRITTEN DOCUMENT RECOGNITION USING RESNET BASED CRNN MODEL WITH OPTIMIZED IMAGE ENHANCEMENT

M. Antony Sheela, J. Jencewin and S.L. Rinu
Department Computer Science and Engineering, St. Xavier's Catholic College of Engineering, India

Abstract

Handwritten text recognition in degraded documents remains a major challenge in document image analysis due to factors such as noise, handwriting variability, uneven lighting, faded ink, and physical distortions in historical or low-quality scans. Traditional OCR methods often perform poorly under these circumstances. To improve recognition accuracy, this research proposes a strong architecture that combines a fixed Convolutional Recurrent Neural Network (CRNN)with refined image prose. The preprocessing pipeline includes grayscale normalization, adaptive thresholding, noise filtering (e.g., median and Gaussian smoothing), and morphological operations like dilation and erosion to improve image clarity while preserving critical handwriting features. These refined images are then processed by a CRNN architecture, comprising convolutional layers for spatial feature extraction, bidirectional recurrent layers (LSTM) for sequence modelling, and a Connectionist Temporal Classification (CTC) loss for transcription without character-level segmentation. The addition of preprocessing models reduces the rate of transmitter rate (CER) and word error speed (WER), increasing training stability and flexibility. Our study forms a base line to detect handwriting, such as creating old manuscripts digital and analysing multilingual documents under complex, real -world conditions, which are both effective and expandable.

Keywords:

Handwriting Recognition, CRNN, Preprocessing, Degraded Documents, CTC Loss, Document Analysis, OCR Enhancement, Historical Digitization

1. INTRODUCTION

1.1 BACKGROUND

Handwritten text recognition reflects various applications such as digitization of old manuscripts, preserving administrative or legal documents, and greater access to handwritten materials. Despite the impressive progress achieved by Contemporary Optical Character Recognition (OCR) system, handwritten text is still a challenging task, especially when faced with a fall in the real world. These problems are made more complex in historical manuscripts, where ink appears down, the paper is aged, and structural distortions such as folds, smoothies and noise tend to cover the written text.

1.2 CHALLENGES

Handwritten text recognition from degraded text documents is a unique case outside the usual OCR scope. Handwriting style inconsistency attributable to multiple writers and periods, yellow ink, bleeding, stains and extruding background details, poor scanning or lighting conditions, as a result, is difficult for segmenting. These contribute greatly to the accuracy of classic recognition systems and require strong solutions that can perform optimally in noisy and uncontrolled conditions.

1.3 PROBLEM DEFINITION

Even with the development of deep learning-based OCR systems, the systems are very poor when used to identify degraded handwritten inputs. The classic OCR technique depends very much on clean and structured visual input and is far away from being enough against real deformation. In addition, most of the current systems are ignoring or not sufficiently consider the importance of preprocessing in such a scenario, resulting in poor performance. Therefore, there is an important requirement for a strong and adaptive system that integrates better image preprocessing with deep learning techniques to obtain maximum recognition performance on degraded handwritten documents.

1.4 OBJECTIVE

The current research targets overcoming the limitations of current OCR system to handle poor handwritten documents. The objective is to develop a complete end-to-end image enhancement technique that improves the quality and legibility of degraded handwritten texts and integrate the preprocessing module into a Residual Network (ResNet) based Convolutional Recurrent Neural Network (CRNN) architecture for secure handwriting recognition. Then compare the proposed system on the reference database with different levels and declare its superiority compared to the reference model and then shows how preprocessing of noisy environment improves the model performance and training efficiency.

1.5 NOVELTY

The novelty of this research is in the integrated approach where adaptive preprocessing methods are used in combination with a Convolutional Recurrent Neural Network (CRNN) architecture, that is a network designed to handle degraded handwritten documents. Although CRNN is implemented for the recognition of handwriting, their performance in handling compromised input has not been much explored. The novelty of this research is to promote the important role of specially designed preprocessing steps like grayscale normalization, adaptive thresholding, noise filtering, and morphological transformations - not as additional procedures but as intrinsic processes that enable accurate recognition under poor conditions. In addition, the model is trained with Connectionist Temporal Classification (CTC) loss function, thus obviating the need for explicit character segmentation, an additional significant challenge when dealing with degraded texts

2. RELATED WORKS

Handwritten text recognition (HTR) and detection have seen significant advancements through the integration of deep learning techniques and large-scale datasets. Abroad study on text

detection and recognition was presented [1] in natural scene, laying the foundation for understanding the complexities in text extraction, localization, and recognition across diverse image types.

Efforts to build language-specific datasets have enabled better model training for underrepresented scripts. Hussain [2] introduced the PHTI dataset for Pashto handwritten text, catering to deep learning applications and demonstrating the need for localized datasets. Similarly, [4] addresses Tibetan handwritten character recognition using a spatiotemporal network with data augmentation.

Segmentation-free methods have shown promise in handling complex document layouts. Coquenet [3] proposed the Document Attention Network (DAN), which leverages attention mechanisms for end-to-end recognition of handwritten documents without explicit segmentation. Likewise, a residual-attention FCNN for offline handwritten Chinese text recognition [5], demonstrating improvements in robustness and accuracy.

Several approaches have adopted hybrid models for improved performance and also implemented CNN-RNN and CRNN-based methods for cursive and printed handwritten text, respectively, offering better handling of sequential dependencies [14] [20]. A CNN-BiLSTM model on the IAM dataset for English handwriting provided a comprehensive evaluation [16].

Handwriting generation and augmentation also play vital roles in enhancing recognition systems. AugGPT explores the use of ChatGPT for data augmentation in text recognition tasks [7]. Handwritten text generation based on visual archetypes was proposed to synthesize realistic handwriting samples [8]. A method called TextStyleBrush was introduced to transfer text aesthetics from a single example image to new text, supporting style-specific recognition [9].

Language-specific studies include Tamil character recognition [11], Nandinagari recognition [12], and Arabic handwriting classification using deep rule-based systems [13] [15]. Convolutional autoencoder-based keyword spotting was explored for historical Devanagari documents [21].

Foundational works such as gradient-based learning for document recognition and a connectionist system for unconstrained handwriting laid the groundwork for modern architectures [22] [24]. These influenced the development of models like Hutnet for Tibetan character recognition [27] and DropSample, which enhances deep CNNs with a new training technique [28].

Recent end-to-end systems include a unified neural model for text localization and named entity recognition in full document pages, emphasizing integrated solutions [26]. A comprehensive literature review systematically categorized techniques in handwritten OCR, offering insights into the evolution and performance of classical and deep learning-based methods [18].

3. RESNET BASED CRNN WITH IMAGE ENHANCEMENT

3.1 IMAGE ENHANCEMENT

Image enhancement is the process of making an image more suitable than the original image for a specific application. It emphasizes the image features such as edges, boundaries, or contrast to make a graphic display more helpful for display and analysis. It increases the dynamic range of the chosen features so that they can be detected easily. The Fig.1 illustrate the various techniques use to increase the quality of the degraded image.

This work aims to enhance the degraded handwritten input while preserving text features. A sequence of four different techniques is followed to enhance the quality of the handwritten image. The Fig.1 shows the preprocessing steps to improve the image quality.

Degraded Image

↓

Gray scale Normalization

↓

Adaptive Thresholding

↓

Noise Filtering
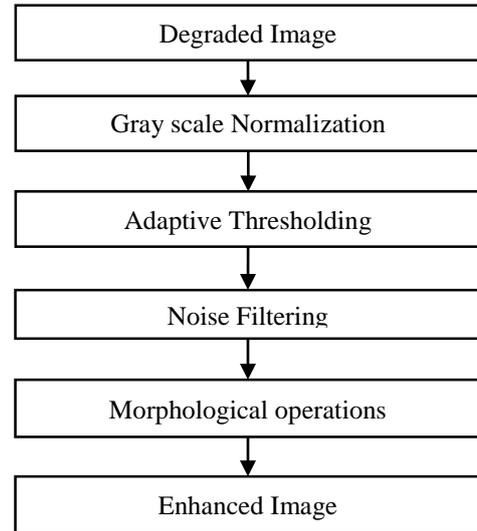
↓

Morphological operations

↓

Enhanced Image

Fig.1. Sequence of image enhancement techniques

First one is converting the image into gray representation, omitting other colours. The image has one channel.

$$I_{Gray}(x,y)=0.299R+0.587G+0.114B \qquad (1)$$

Grayscale normalization scales the intensity values of an image to a specific range (commonly [0, 1] or [0, 255]) to standardize brightness and contrast across images. This improves stability in later processing stages. The second phase is Adaptive Thresholding which Converts a grayscale image into a binary image by calculating a local threshold for each pixel based on the neighbourhood, making it very effective in the condition of non-uniform lighting conditions. The third step is Noise Filtering that removes unwanted objects such as salt-and-pepper noise or Gaussian noise that can interfere with the text structure. Median filtration replaces each pixel with a median of neighbouring pixel. The Gaussian filtering uses a Gaussian kernel to smooth the image. It preserves edge details, such as handwriting strokes, while lowering background and sensor noise. The Morphological Operation, the last stage, is used to tidy up the binary images by filling in the gaps, removing tiny objects, and joining disparate parts. Erosion eliminates pixels from object boundaries, while dilation adds new pixels.

3.2 RESNET BASED CRNN

The Reset-CRNN architecture is a hybrid model that combines the powerful feature extraction capabilities of ResNet (Residual Network) with the sequence modeling strength of CRNN (Convolutional Recurrent Neural Network). This combination is especially effective for tasks such as handwritten name

recognition, where both spatial pattern and sequential dependencies are important.

ResNet addresses the problem of vanishing gradients in deep networks by using skip connections or residual connections, which allow gradients to flow directly through the network. Instead of learning a direct mapping $H(x)$,

$$F(x)=H(x)-x \rightarrow H(x)=F(x)+x \tag{2}$$

The Eq.(2) depicts the residual function through which ResNet learns. ResNet serves as an excellent feature extractor, capturing complex patterns like curves, edges, and letter structures in the input image. It transforms the raw input into a high-level feature map, which can later be fed to a sequence model for temporary learning.

CRNN introduces Recurrent Neural Networks (RNNs), such as LSTM or GRU, on top of the convolutional features to model the temporal sequence of the input, ideal for handwriting, where letters are sequentially connected. In a CRNN, after convolutional feature extraction, the spatial features (typically along the width) are reshaped into a time-series-like format and passed into Bidirectional LSTMs.

$$h_t=BiLSTM(x_t)=LSTM_{fw}(x_t)+LSTM_{bw}(x_t) \tag{3}$$

the Eq.(3) allows the model to learn dependencies from both past and future positions. This bidirectional context improves character-level understanding in handwritten words, capturing the flow of writing more accurately.

3.3 RESNET BASED CRNN ALGORITHM

Image enhancement is the first step in the process of identifying handwritten names from deteriorated grayscale images. The Fig.2 provides the ResNet-based CNN algorithm.

To guarantee uniformity in brightness and contrast, the image must be resized and the pixel values must be normalized. After that, the image is reshaped to incorporate a channel dimension, readying it for convolutional layer processing. Deep spatial features are extracted from the improved image using a ResNet model. ResNet's residual block structure helps it identify complex features even in low-quality images by capturing significant patterns while maintaining the original information.

Following feature extraction, the output is reshaped to represent the time steps of a sequence by reshaping the data into a sequential format. After that, the sequence is sent to a bidirectional LSTM network, which reads it both ways to improve context comprehension.

A collection of feature vectors containing rich information about the handwritten content is the end result. To create probabilities for every potential class, these vectors are then passed through a dense layer with softmax activation. The predicted name is chosen from the class with the highest probability.

Input: Grayscale image of a handwritten name.

Output: Predicted identity label corresponding to the handwritten name

Step 1: Image Enhancement.
1.1 Resize the input image for processing.
1.2 Normalize the pixel values to the range [0, 1] to ensure uniform intensity distribution

1.3 Expand the image dimensions to include a channel and reshape it and send for compatibility with convolutional networks.

Step 2: Feature Extraction using ResNet
2.1 Pass the preprocessed image through a customized Residual Network (ResNet) to extract deep spatial features.
2.2 Each residual block computes an output given:

$$Output=F(x,\{W_i\})+x$$

where $F(x,\{W_i\})$ is the residual mapping function (typically a stack of convolution, batch normalization, and ReLU layers) and x is the shortcut input.

Step 3: Reshape Feature Map
3.1 Convert the spatial feature map output by ResNet into a sequential format by collapsing the height and channel dimensions, resulting in a sequence of vectors across the width dimension.
3.2 This transformation allows the width dimension to be interpreted as a temporal sequence.

Step 4: Sequence Modeling using BiLSTM
4.1 Feed the sequential feature vectors into a Bidirectional Long Short-Term Memory (BiLSTM) network.
4.2 This layer captures dependencies from both left-to-right and right-to-left across the sequence.
4.3 The BiLSTM outputs a context-enriched feature vector for each position in the sequence

Step 5: Classification

$$y'=Softmax(W.h+b)$$

Pass the final BiLSTM output through a Dense layer with softmax activation to obtain class probabilities.

Step 6: Output Prediction
Select the class with the highest probability as the predicted identity label.

Fig.2. Algorithm for ResNet based CRNN

4. RESULT

A dataset of 30,000 handwritten name images is used to evaluate six deep learning models. The models are calculated for each set of data records and the dataset is split into five categories. CRNN, CNN, ResNet, and BiLSTM are the models that are being compared.

As shown in Fig.3-Fig.7, the algorithms' execution demonstrated that ResNet-CRNN continuously performed better than the others across all important metrics. The Fig.3's loss graph illustrates that ResNet-CRNN is attaining the lowest error. BiLSTM, CRNN, and CNN came next in order of increasing loss, but ResNet alone also demonstrated strong performance, confirming the strength of residual connections.

Accuracy results in Fig.4 reaffirmed ResNet-CRNN's dominance, with ResNet also performing well. Sequence-aware models like BiLSTM outperformed CNN and CRNN due to their ability to process temporal data. Figures 5 and 6, displaying precision and recall, further highlighted ResNet-CRNN's superior ability to make both accurate and comprehensive predictions. While ResNet remained precise, it was outpaced in recall by models incorporating recurrent layers, emphasizing the need for temporal understanding in handwriting tasks.
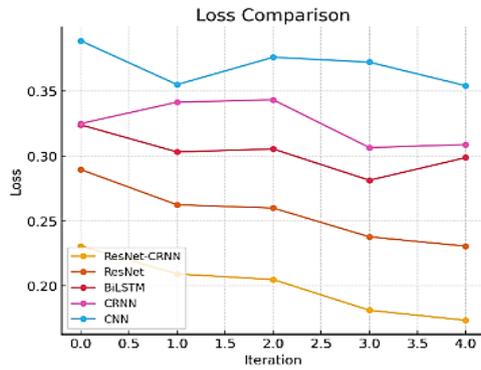
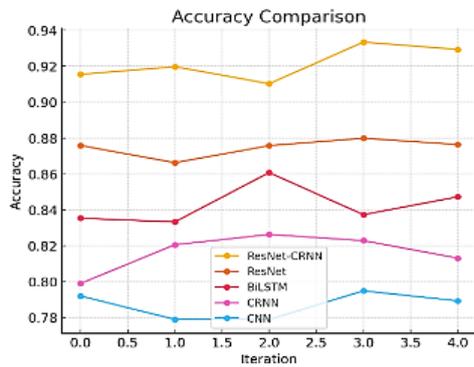Fig.3. Loss comparison between ResNet based CRNN and other models

Fig.4. Accuracy comparison between ResNet based CRNN and other models
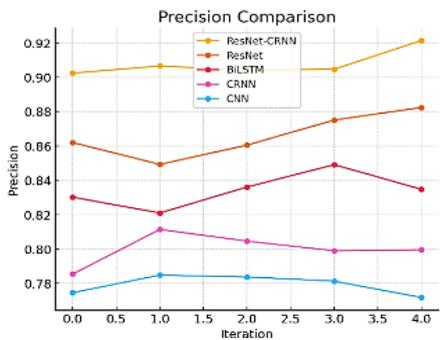

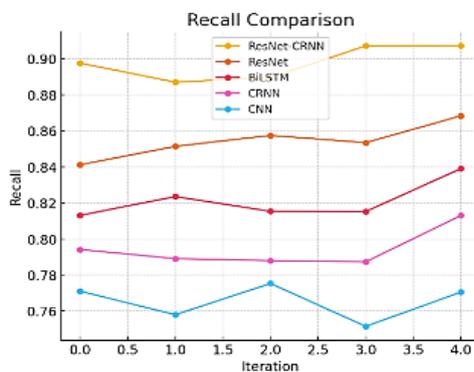Fig.5. Precision between ResNet based CRNN and other models

Fig.6. Recall comparison between ResNet based CRNN and other models

The F1-score as in Fig.7 confirmed the overall balance and robustness of ResNet-CRNN, which excelled in combining precision and recall.
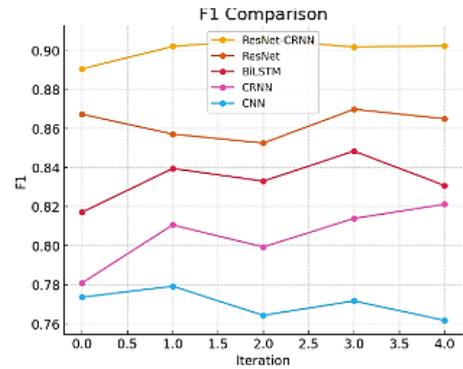

Fig.7. F1 comparison between ResNet based CRNN and other models

From Fig.8, it is inferred that preprocessing's first and performing ResNet-CRNN significantly produced better output than when standardized preprocessing (resizing, grayscale normalization) was applied.
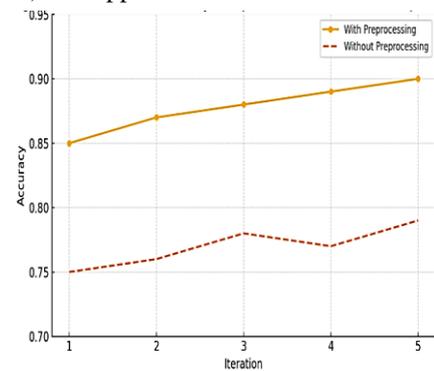

Fig.8. Comparison between the proposed sequence of preprocessing and the scandalized preprocessing

5. CONCLUSION

This paper presents a robust and powerful paradigm that combines ResNet-based CRNN architecture with optimized image preprocessing to improve handwritten document recognition, particularly in degraded scenarios. The quality of deteriorated handwritten inputs is significantly enhanced to allow for improved feature extraction by successively applying grayscale normalization, adaptive thresholding, noise reduction, and morphological processing. Without explicit character segmentation, the CRNN model can effectively handle spatial and sequential variations because it is trained under CTC loss and has residual connections.

According to experimental results, the suggested method performs better than baseline CRNN models and conventional OCR systems, particularly in low contrast and artifact-contaminated situations. This paper highlights the critical role of image enhancement in enhancing deep learning model performance and provides a solid foundation for further research on historical document digitization, multilingual handwriting recognition, and actual handwritten OCR application.

REFERENCES

[1] Q. Ye and D. Doermann, "Text Detection and Recognition in Imagery: A Survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 37, No. 7, pp. 1480-1500, 2015.

[2] I. Hussain, R. Ahmad, S. Muhammad, K. Ullah, H. Shah and A. Namoun, "PHTI: Pashto Handwritten Text Imagebase for Deep Learning Applications", *IEEE Access*, Vol. 10, pp. 113149-113157, 2022.

[3] D. Coquenet, C. Chatelain and T. Paquet, "DAN: A Segmentation-Free Document Attention Network for Handwritten Document Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 45, No. 7, pp. 8227-8243, 2023.

[4] Yuze Li, Lei Li, Zhejun Zhang, Shaoting Guo, Shuo Zhang, Xingyue Zhang, Rinchen Dongrub, Dorje Tashi, Yougbin Yu and Nyima Tashi, "TSTN: Tibetan Spatiotemporal Online Handwritten Character Recognition with Data Augmentation", *IEEE Transactions on Multimedia*, Vol. 14, No. 8, pp. 1-7, 2024.

[5] Y. Wang, Y. Yang, W. Ding and S. Li, "A Residual-Attention Offline Handwritten Chinese Text Recognition based on Fully Convolutional Neural Networks", *IEEE Access*, Vol. 9, pp. 132301-132310, 2021.

[6] N. Saqib, K.F. Haque, V.P. Yanambaka and A. Abdelgawad, "Convolutional-Neural-Network based Handwritten Character Recognition: An Approach with Massive Multisource Data", *Algorithms*, Vol. 15, pp. 1-7, 2022.

[7] H. Dai, "AugGPT: Leveraging ChatGPT for Text Data Augmentation", *IEEE Transactions on Big Data*, Vol. 11, No. 3, pp. 907-918, 2025.

[8] V. Pippi, S. Cascianelli and R. Cucchiara, "Handwritten Text Generation from Visual Archetypes", *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 22458-22467, 2023.

[9] Praveen Krishnan, Rama Kovvuri, Guan Pang, Boris Vassilev and Tal Hassner, "TextStyleBrush: Transfer of Text Aesthetics from a Single Example", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 45, No. 7, pp. 9122-9134, 2023.

[10] Heng Zhao and Huihui Li, "Handwriting Identifcation and Verifcation using Artifcial Intelligence-Assisted Textural Features", *Scientific Reports*, Vol. 2, pp. 1-16, 2023.

[11] S. Sobhana Mari and G. Raju, "Modified View-based Approaches for Handwritten Tamil Character Recognition", *ICTACT Journal on Image and Video Processing*, Vol. 6, No. 1, pp. 1076-1085, 2015.

[12] Prathima Guruprasad and Guruprasad, "An Accurate and Robust Handwritten Nandinagari Recognition System", ICTACT Journal on Image and Video Processing, Vol. 10, No. 3, pp. 2119-2124, 2020.

[13] Soumia Djaghbellou1, Zahid Akhtar, Abderraouf Bouziane and Abdelouahab Attia, "Arabic Handwritten Characters Recognition via Multi-Scale Hog Features and Multi-Layer Deep Rule-based Classification", *ICTACT Journal on Image and Video Processing*, Vol. 10, No. 4, pp. 2195-2200, 2020.

[14] G.R. Hemanth, M. Jayasree, S. Keerthi Venii, P. Akshaya and R. Saranya, "CNN-RNN based Handwritten Text Recognition", *ICTACT Journal on Soft Computing*, Vol. 12, No. 1, pp. 2457-2463, 2021.

[15] Maram Saleh Alwagdani and Emad Sami Jaha, "Deep Learning-based Child Handwritten Arabic Character Recognition and Handwriting Discrimination", *Sensors*, Vol. 23, No. 15, pp. 1-7, 2023.

[16] FiratKizilirmak and Berrin Yanikoglu, "CNN-BiLSTM Model for English Handwriting Recognition: Comprehensive Evaluation on the IAM Dataset", *Nature*, Vol. 98, No. 2, pp. 1-21, 2021.

[17] PriushaNarwariya and Susheel kumarTiwari, "Hybrid Curvelet Transform with K-Nearest Neighbor for Enhanced Satellite Resource Classification", *ICTACT Journal on Image and Video Processing*, Vol. 15, No. 3, pp. 3469-3475, 2025.

[18] J. Memon, M. Sami, R.A. Khan and M. Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)", *IEEE Access*, Vol. 8, pp. 142642-142668, 2020.

[19] P.P. Kumar, C. Bhagvati, A. Negi, A. Agarwal and B.L. Deekshatulu, "Towards Improving the Accuracy of Telugu OCR Systems", *Proceedings of International Conference on Document Analysis and Recognition*, Vol. 72, pp. 910-914, 2011.

[20] A.A. Idris and D.B. Taha, "Handwritten Text Recognition using CRNN", *Proceedings of International Conference on Contemporary Information Technology and Mathematics*, Vol. 33, pp. 329-334, 2022.

[21] S.N. Sushma and B. Sharada, "A Convolutional Autoencoder based Keyword Spotting in Historical Handwritten Devanagari Documents", *Proceedings of International Conference on Inventive Computation Technologies*, pp. 356-362, 2022.

[22] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based Learning Applied to Document Recognition", *Proceedings of the IEEE*, Vol. 86, No. 11, pp. 2278-2324, 1998.

[23] B.L.D. Bezerra, C. Zanchettin and V. Braga de Andrade, "A Hybrid RNN Model for Cursive Offline Handwriting Recognition", *Proceedings of International Symposium on Neural Networks*, pp. 113-118, 2012.

[24] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke and J. Schmidhuber, "A Novel Connectionist System for Unconstrained Handwriting Recognition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 31, No. 5, pp. 855-868, 2009.

[25] Rakesh Kumar Mandal and N.R. Manna, "Hand Written English Character Recognition using Row-wise Segmentation Technique", *Proceedings of International Symposium on Devices MEMS, Intelligent Systems and Communication*, Vol. 29, No. 2, pp. 5-9, 2011.

[26] M. Carbonell, A. Fornés, M. Villegas and J. Llados, "A Neural Model for Text Localization, Transcription and Named Entity Recognition in Full Pages", *Pattern Recognition Letters*, Vol. 136, pp. 219-227, 2020.

[27] G. Zhang, W. Wang, C. Zhang, P. Zhao and M. Zhang, "Hutnet: An Efficient Convolutional Neural Network for Handwritten Uchentibetan Character Recognition", *Big Data*, Vol. 11, No. 5, pp. 387-398, 2023.

[28] W. Yang, L. Jin, D. Tao, Z. Xie and Z. Feng, "Dropsample: A New Training Method to Enhance Deep Convolutional Neural Networks for Large Scale Unconstrained Handwritten Chinese Character Recognition", *Pattern Recognition*, Vol. 58, pp. 190-203, 2016.

[29] A. Goyal, K. Khandelwal and P. Keshri, "Optical Character Recognition for Handwritten Hindi", *CS229 Notes-Machine Learning*, Vol. 97, pp. 1-5, 2010.

[30] P. Shivakumara, D. Tang, M. Asadzadehkaljahi, T. Lu, U. Pal and M. Hossein Anisi, "CNN-RNN based Method for License Plate Recognition", *CAAI Transactions on Intelligence Technology*, Vol. 3, No. 3, pp. 169-175, 2018.