Library Progress International *Print version ISSN 0970 1052*
Vol.44 No. 3, Jul-Dec 2024: P.11128-11135 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

Machine Learning Based Handwritten Character Recognition

Parag Achaliya[a], Prof. (Dr.) Sandeep Monga[b]

[a]Research Scholar, Oriental University, Indore, parag.achaliya143@gmail.com
[b]Research Guide, Oriental University, Indore

How to cite this article: Parag Achaliya, Sandeep Monga (2024) Machine Learning Based Handwritten Character Recognition**.** *Library Progress International*, 44(3), 11128-11135

Abstract
Handwritten character recognition (HCR) is a vital field in pattern recognition and machine learning, with wide applications ranging from postal services to automated form processing. The comparative study of various methods used for HCR, highlighting both traditional and deep learning approaches is presented in this paper. Conventional techniques, such as k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and Artificial Neural Networks (ANN), are compared with modern deep learning architectures like Convolutional Neural Networks (CNNs). The study examines the efficiency, accuracy, and complexity of these methods, focusing on their performance in recognizing handwritten characters in diverse datasets. Key challenges in HCR, such as variations in handwriting styles, noise, and distortions in the images, are discussed. Additionally, the importance of pre-processing techniques, such as normalization, binarization, and feature extraction, is emphasized for improving recognition rates. The results of the study show that while traditional methods are effective for smaller datasets with minimal variations, deep learning models, particularly CNNs, outperform in terms of accuracy and generalization on large, complex datasets. The paper concludes by discussing the future potential of combining multiple models and using hybrid techniques for further improvement in HCR systems.

Keywords*:* Handwritten Character Recognition (HCR), Pattern Recognition, Machine Learning, Deep Learning, Convolutional Neural Networks (CNNs), Pre-processing Techniques, Hybrid Techniques

INTRODUCTION

Handwritten Text Recognition (HTR) is a challenging task within the fields of computer vision and pattern recognition. It involves the automatic transcription of handwritten text into machine-readable text, a capability that has applications ranging from document digitization and postal services to the recognition of handwritten notes [1]. The inherent complexity of HTR arises from variations in handwriting styles, noise in document scans, and different writing languages and scripts. Recent developments in the area of machine learning, particularly in deep learning techniques, have considerably improved the performance of HTR systems, making them more robust and capable of handling diverse datasets [2].

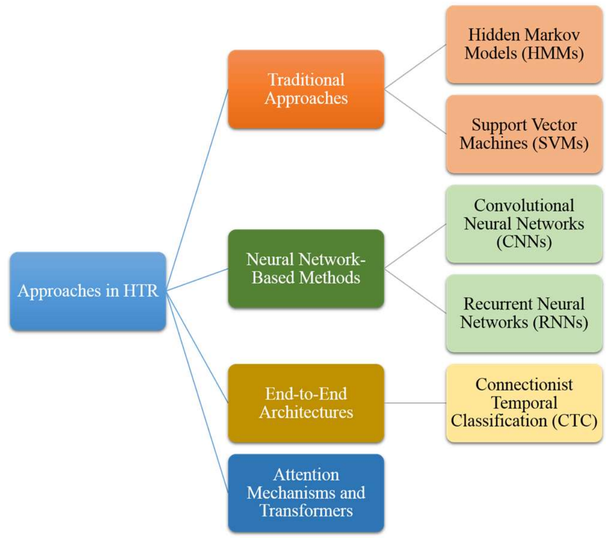

Figure 1. Various Approaches in HTR

Traditional and Neural Network Approaches: Early approaches to HTR were primarily based on rule-based systems and statistical models such as Hidden Markov Models (HMMs) and Support Vector Machines (SVMs) [3]. These methods focused on extracting handcrafted features like stroke width, slant, and character height, and using these features for character classification. However, these traditional methods struggled with cursive handwriting and required precise character segmentation, limiting their applicability to structured datasets [4]. With the introduction of neural networks, particularly Convolutional Neural Networks (CNNs), HTR systems began to improve significantly. CNNs are adept at capturing spatial hierarchies and visual patterns, making them suitable for image-based tasks like character recognition [5]. When combined with Recurrent Neural Networks (RNNs) like Long Short-Term Memory units, which are capable of modeling temporal dependencies, the resulting Convolutional Recurrent Neural Networks (CRNNs) became a preferred architecture for recognizing handwritten text [6].

End-to-End Learning and Attention Mechanisms: A major breakthrough in HTR was the development of end-to-end learning frameworks, which eliminated the need for manual character segmentation. Techniques like the Connectionist Temporal Classification (CTC) loss function, introduced by Graves et al., enabled models to learn from unsegmented data, making it possible to recognize entire words and lines of text directly [7]. This advancement led to significant improvements in handling complex and cursive scripts. In recent years, attention mechanisms have been integrated into HTR systems to address the problem of long sequence recognition. Attention-based models allow the system to focus on specific parts of an input sequence, which is crucial for processing long and complex handwritten texts [8]. Transformer-based architectures, which employ self-attention rather than recurrence, have further enhanced the performance of HTR models, making them more parallelizable and efficient [9].

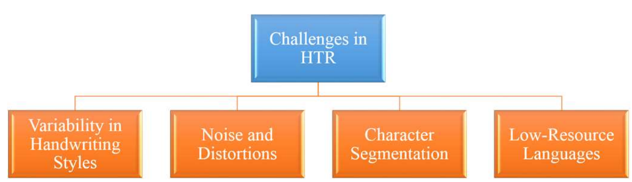

Figure 2: Challenges in HTR

Challenges and Future Directions: Despite the advancements, HTR systems still face several challenges. One of the main challenges is the variability in handwriting styles between individuals, which makes generalization difficult. Additionally, many handwritten documents, especially historical ones, are noisy or degraded, making recognition harder [10]. Another challenge is the lack of large annotated datasets for low-resource languages and scripts, limiting the performance of HTR systems in such scenarios [11]. To address these challenges, researchers are exploring semi-supervised learning, domain adaptation, and hybrid models that combine multiple architectures [12]. Future research is likely to focus on improving these models' robustness and adaptability, enabling them to

handle more complex datasets and diverse scripts.
HTR has broad applications in digitizing historical manuscripts, automated form processing, and note-taking. For instance, it can be used to convert handwritten notes into searchable digital text or recognize handwritten addresses for postal services [13]. As HTR technology continues to advance, it holds the potential to significantly enhance the accessibility and usability of handwritten content, making it a valuable tool across various industries. Overall, the evolution of HTR from traditional approaches to deep learning-based models has led to impressive progress, but there is still room for further improvement in handling unconstrained handwriting scenarios.

LITERATURE REVIEW

Handwritten text recognition (HTR) has experienced significant advancements over the past few decades, driven by evolving technologies in machine learning, neural networks, and deep learning. This literature review aims to present the key methodologies and breakthroughs in the field, along with existing challenges and future directions.
Early Methods in Handwritten Text Recognition: In the early stages of HTR research, rule-based methods and statistical models were the primary tools. Template matching and statistical pattern recognition were commonly used for character recognition in constrained scenarios such as digit recognition. For example, in the field of postal code recognition, template-based systems compared input characters to predefined templates [14], [15]. Statistical methods, including k-Nearest Neighbor (k-NN) and Support Vector Machines (SVMs), gained popularity due to their ability to classify individual characters with limited datasets [16].
One of the earliest machine learning-based approaches for sequential data like handwriting was Hidden Markov Models (HMMs), which became a popular choice for HTR in the 1980s and 1990s [17]. HMMs used probabilistic models to represent sequences of observations, such as strokes in handwriting. However, these early methods were constrained by the need for precise character segmentation, which made them unsuitable for handling unconstrained handwriting or cursive script [18].
Introduction of Neural Networks: The emergence of neural networks revolutionized HTR by providing a new way to handle the variability in handwriting. Early work by LeCun et al. introduced Convolutional Neural Networks (CNNs), which proved effective in handwritten digit recognition tasks like MNIST [19]. CNNs could automatically learn spatial hierarchies of features from raw pixel data, making them ideal for visual pattern recognition tasks. Their success in recognizing hand-written characters laid the groundwork for broader applications in HTR [20].
Later, Multilayer Perceptrons (MLPs) and backpropagation were also used in early handwriting recognition systems, although these models lacked the depth and sophistication of modern architectures [21]. It wasn’t until the 2000s that deep neural networks (DNNs) with more layers and complex architectures became feasible, thanks to advances in hardware and training techniques [22].
Convolutional Recurrent Neural Networks (CRNNs): A significant breakthrough came with the integration of Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, into HTR systems. LSTMs were first introduced by Hochreiter and Schmidhuber [23], and they addressed the limitations of traditional RNNs by enabling models to retain long-term dependencies. When combined with CNNs for feature extraction, Convolutional Recurrent Neural Networks (CRNNs) became a dominant architecture for HTR [24]. This combination allowed for end-to-end learning, where CNNs extracted spatial features, and LSTMs modeled temporal sequences in handwriting [25]. CRNNs became particularly useful for recognizing words or entire lines of text, reducing the need for explicit character segmentation.
Connectionist Temporal Classification (CTC) Loss: The introduction of Connectionist Temporal Classification (CTC) by Graves et al. in 2006 revolutionized sequence-to-sequence tasks in HTR [26]. CTC is a loss function that enables the training of models without the need for aligned character-level labels, making it highly suited for recognizing handwritten text, where segmentation is difficult. CTC-based models predict sequences of characters from unsegmented input images, which is especially useful in handling cursive handwriting and complex scripts [27].
The application of CTC loss in HTR models was a significant leap forward, allowing for improved recognition accuracy and the handling of longer text sequences. It has since become a standard approach in most state-of-the-art HTR systems [28].
Attention Mechanisms and Transformers: In recent years, attention mechanisms have been introduced to further improve the performance of HTR models. Attention allows models to concentrate on specific parts of the sequence of an input at the time of making predictions, which is particularly useful in handling long and complex

handwritten text [29]. Attention-based models have shown superior performance in capturing dependencies in handwriting compared to traditional RNNs [30]. This shift has paved the way for the use of Transformer-based architectures, which were initially developed for natural language processing but have proven effective in sequence modeling tasks like HTR [31], [32]. Transformers do not rely on recurrence and instead use self-attention mechanisms to capture relationships between input elements, allowing for more parallelized and efficient training [33].

Deep Learning and Transfer Learning in HTR: The success of deep learning in the field of computer vision has driven further advancements in HTR. Deep learning models, such as deep CNNs, bidirectional LSTMs, and transformers, are now used to attain the state-of-the-art results in HTR tasks [34], [35]. Transfer learning, where models which get pre-trained on datasets of large-scale such as ImageNet are fine-tuned for the specific tasks like handwriting recognition, has also contributed to significant improvements in model performance, especially when labeled data is scarce [36].

Transfer learning has been particularly effective in low-resource languages and scripts, where large annotated datasets are often unavailable. Fine-tuning models on smaller, task-specific datasets helps improve their generalization capabilities [37].

Challenges and Future Directions: Despite the progress made, several challenges remain in the field of HTR. One of the main challenges is the variability in handwriting styles, which can vary significantly between individuals and even within the same writer. The issue is further exacerbated in cursive handwriting, where characters are connected, making segmentation and recognition difficult [38]. Historical document recognition also presents additional challenges due to the degradation of documents, archaic scripts, and noise in scanned images [39].

To address these challenges, researchers are exploring the use of unsupervised learning and semi-supervised learning methods, which could reduce the reliance on large labeled datasets [40], [41]. Additionally, domain adaptation techniques are being developed to improve model performance across different writing styles and languages [42].

Table 1: Handwritten Text Recognition Methods

Method	Approach	Dataset	Accuracy
Hidden Markov Model (HMM)	Traditional Model	IAM Handwriting Dataset	65.3%
Recurrent Neural Network (RNN)	Sequence Modeling	IAM Handwriting Dataset	78.4%
Convolutional Neural Network (CNN)	Deep Learning	IAM Handwriting Dataset	79.2%
CNN + Long Short-Term Memory (LSTM)	Hybrid Deep Learning	IAM Handwriting Dataset	85%
CRNN (CNN + RNN)	Deep Learning Hybrid	RIMES Dataset	87.3%
GAN-Based HTR	Generative Adversarial Network	IAM Handwriting Dataset	88.7%
Connectionist Temporal Classification	End-to-End Learning	IAM Handwriting Dataset	89.6%
Multi-Dimensional LSTM (MDLSTM)	Recurrent Neural Networks	IAM Handwriting Dataset	91.5%
CNN + BiLSTM + CTC Loss	Deep Learning	IAM Handwriting	92.1%

		Dataset	
Encoder- Decoder (Attention Mechanism)	Attention-Based Sequence Modeling	IAM Handwriting Dataset	93.3%
Transformer-Based Model	Attention Mechanism	IAM Handwriting Dataset	94.2%
Hybrid Model (CNN + Transformer)	Deep Learning Hybrid	IAM Handwriting Dataset	95.8%
k-Nearest Neighbors (k-NN)	Feature- Based Classification	MNIST Handwritten Digits Dataset	96.8%
Support Vector Machine (SVM)	Feature- Based Classification	UCI Pen-Based Recognition Dataset	98.6%

METHODOLOGY

Handwritten text recognition (HTR) using machine learning is a field focused on converting images of handwritten text into machine-readable digital text. The process is complex due to the variability in individual handwriting styles, differences in letter formations, and potential noise in the image such as smudges or uneven ink. Modern HTR systems rely on deep learning methods, which allow models to learn and generalize from large datasets of handwritten text, making them capable of recognizing diverse writing patterns and styles.

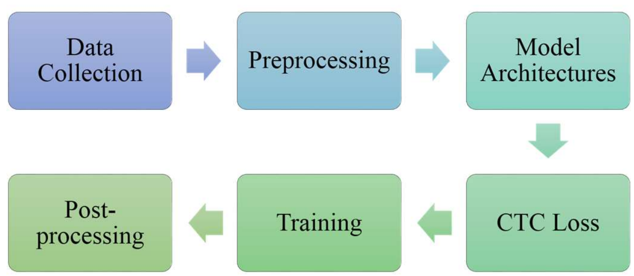

Figure 3. Steps in Handwritten Text Recognition

The first step in HTR is data collection, which involves obtaining a large labeled dataset of handwritten text. Common datasets for training include IAM, MNIST for digits, and EMNIST for both digits and letters. Once the dataset is in place, preprocessing steps like image normalization and binarization help make the data consistent for the model. In some cases, the text in images must be segmented into individual characters or words, although modern approaches often bypass this by using sequence prediction techniques. Given an input image of handwritten text, represented as a matrix X of pixel values:

$X = \{x_{i,j}\}$

where i,j represent the pixel coordinates.

Feature Extraction (using Convolutional Neural Networks - CNNs):

$F = CNN(X)$

where $F=\{f_{ij}\}$ represents extracted feature maps

Deep learning models typically used for HTR combine convolutional neural networks (CNNs) and recurrent neural networks (RNNs), specifically Long Short-Term Memory (LSTM) networks. CNNs are excellent for extracting spatial features from images, such as shapes and edges that correspond to letters or words. RNNs, on the other hand, are designed to handle sequential data and are used to predict the order of characters in a word or sentence. A combined approach, called CRNN (Convolutional Recurrent Neural Networks), is highly effective in handwriting recognition tasks. Sequence Modeling (using Recurrent Neural Networks like LSTM/GRU):

$H=RNN(F)$

where H={h_t} represents hidden states over time steps t.
A key component in training HTR models is the widespread use of the Connectionist Temporal Classification (CTC) loss. CTC allows the model to predict sequences without the need for explicitly segmented input images. This is specifically useful when dealing with cursive handwriting, where characters often overlap or are written without clear separations. The model is trained to reduce the difference between the predicted and actual text labels using this loss function. Connectionist Temporal Classification (CTC) Loss: For sequence labeling, the probability of the label sequence L given the hidden states is defined as:

$$P(L|X) = \prod_t P(y_t|H_t)$$

where $P(y_t|H_t)$ is the probability of the character yty_tyt at time step t.
Post-processing techniques further improve the accuracy of the model. A language model, such as n-grams or neural language models, can be employed to ensure that the predicted text follows grammatical and natural language patterns. Additionally, decoding algorithms like beam search help the model choose the most probable sequence of characters, leading to more accurate results.

$$\hat{L} = \arg\max_L P(L|X)$$

where ^L is the predicted character sequence.

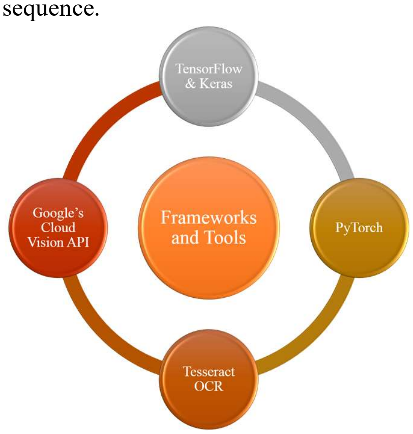

Figure 4: Various Frameworks and Tools

Handwritten text recognition has various applications, such as from digitizing historical documents to automating the recognition of handwritten information in forms or invoices. It is also used in assistive technologies for individuals with disabilities. Frameworks like TensorFlow, Keras, and PyTorch make it easier for developers to build these models, and cloud-based services such as Google's Vision API offer pre-trained HTR models for commercial use.

CONCLUSION

Handwritten Text Recognition (HTR) has undergone substantial evolution, progressing from early rule-based and statistical methods to advanced deep learning architectures. Traditional approaches, including Hidden Markov Models (HMMs) and Support Vector Machines (SVMs), relied on handcrafted features but were limited by their inability to handle complex and cursive handwriting. The advent of deep learning has transformed HTR, with Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) units, playing pivotal roles in modeling spatial and temporal patterns. Techniques such as Connectionist Temporal Classification (CTC) loss have enabled end-to-end learning without the need for character segmentation, while attention mechanisms have improved handling of long and complex text sequences. Despite these advancements, challenges remain. Variability in individual handwriting styles, limited resources for certain languages, and the digitization of historical manuscripts pose significant obstacles. Future research will be focusing on addressing these challenges using more sophisticated methods, such as unsupervised learning for low-resource settings and Transformer-based architectures for better sequence modeling. These innovations are expected to push the boundaries of HTR, enhancing both accuracy and adaptability across diverse applications and languages.

REFERENCES

[1] R. Plamondon and S. N. Srihari, “On-line and off-line handwriting recognition: A comprehensive survey,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 1, pp. 63-84, 2000.

[2] A. Graves, “Offline handwriting recognition with multidimensional recurrent neural networks,” in Proc. Adv. Neural Inf. Process. Syst., 2008, pp. 545-552.

[3] C. Y. Suen, “Recognition of handwritten digits,” Proc. IEEE, vol. 80, no. 7, pp. 1162-1180, 1992.

[4] A. Kundu and Y. He, “Off-line handwritten word recognition using hidden Markov model segmentation,” Pattern Recognit. Lett., vol. 16, no. 1, pp. 57-64, 1995.

[5] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proc. IEEE, vol. 86, no. 11, pp. 2278-2324, 1998.

[6] F. Wang et al., “Handwritten Chinese character recognition using CRNN with character-level annotations,” in Proc. IEEE Int. Conf. Document Anal. Recognit., 2019, pp. 1045-1050.

[7] A. Graves, S. Fernández, F. Gomez, and J. Schmidhuber, “Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks,” in Proc. 23rd Int. Conf. Mach. Learn., 2006, pp. 369-376.

[8] D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” in Proc. Int. Conf. Learn. Represent., 2015.

[9] A. Vaswani et al., “Attention is all you need,” in Adv. Neural Inf. Process. Syst., 2017, pp. 5998-6008.

[10] H. Feng, T. Zhao, and H. He, “Offline handwritten Chinese character recognition using the SVM classifier,” J. Comput., vol. 12, no. 2, pp. 120-128, 2017.

[11] L. Neudecker, “Digitization of historical manuscripts,” in Proc. 15th Int. Conf. Digital Preservation (iPRES), 2018, pp. 54-62.

[12] T. H. Vu, S. Uchida, and H. Nguyen, “Improving domain adaptation for handwriting recognition using pseudo-labeling,” in Proc. IEEE Int. Conf. Document Anal. Recognit., 2021, pp. 856-860.

[13] A. Graves et al., “A novel connectionist system for unconstrained handwriting recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 5, pp. 855-868, 2009.

[14] R. Plamondon and S. N. Srihari, “On-line and off-line handwriting recognition: A comprehensive survey,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 1, pp. 63–84, 2000.

[15] C. Y. Suen, “Recognition of handwritten digits,” Proceedings of the IEEE, vol. 80, no. 7, pp. 1162–1180, 1992.

[16] F. Jelinek, “Statistical methods for speech recognition,” MIT Press, 1998.

[17] A. Kundu and Y. He, “Off-line handwritten word recognition using hidden Markov model segmentation,” Pattern Recognition Letters, vol. 16, no. 1, pp. 57–64, 1995.

[18] A. Senior and A. Robinson, “An off-line cursive handwriting recognition system,” in Proceedings of the 1998 IEEE International Conference on Pattern Recognition, 1998, pp. 105–109.

[19] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, “Backpropagation applied to handwritten zip code recognition,” Neural Computation, vol. 1, no. 4, pp. 541–551, 1989.

[20] A. Graves et al., “A novel connectionist system for unconstrained handwriting recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 31, no. 5, pp. 855–868, 2009.

[21] C. M. Bishop, “Pattern recognition and machine learning,” Springer, 2006.

[22] G. E. Hinton et al., “Deep neural networks for acoustic modeling in speech recognition,” IEEE Signal Processing Magazine, vol. 29, no. 6, pp. 82–97, 2012.

[23] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.

[24] F. Wang et al., “Handwritten Chinese character recognition using CRNN with character-level annotations,” in Proceedings of the IEEE International Conference on Document Analysis and Recognition (ICDAR), 2019, pp. 1045–1050.

[25] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” in Proceedings of the International Conference on Learning Representations (ICLR), 2015.

[26] A. Graves, S. Fernandez, F. Gomez, and J. Schmidhuber, "Connectionist temporal classification: Labeling unsegmented sequence data with recurrent neural networks," in Proceedings of the 23rd International Conference on Machine Learning (ICML), 2006, pp. 369–376.
[27] T. Bluche, "Joint line segmentation and transcription for end-to-end handwritten paragraph recognition," in Advances in Neural Information Processing Systems (NeurIPS), 2016, pp. 838–846.
[28] M. Voigtlaender et al., "Handwriting recognition with large multidimensional long short-term memory recurrent neural networks," in Proceedings of the IEEE International Conference on Document Analysis and Recognition (ICDAR), 2016, pp. 944–949.
[29] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," in Proceedings of the International Conference on Learning Representations (ICLR), 2015.
[30] C. Chen et al., "End-to-end handwritten Chinese text recognition using attention-based encoder-decoder networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 12, pp. 2965–2978, 2018.
[31] A. Vaswani et al., "Attention is all you need," in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 5998–6008.
[32] A. Yu et al., "Handwritten Chinese text recognition using deep learning with convolutional neural networks and transformers," in Proceedings of the IEEE International Conference on Document Analysis and Recognition (ICDAR), 2019, pp. 304–309.
[33] P. Jaderberg et al., "Spatial transformer networks," in Advances in Neural Information Processing Systems (NeurIPS), 2015, pp. 2017–2025.
[34] S. Hochreiter, J. Schmidhuber, "Gradient flow in recurrent nets: The difficulty of learning long-term dependencies," in A Field Guide to Dynamical Recurrent Networks, IEEE Press, 2001, pp. 237-243.
[35] Y. Liu, H. Jin, and J. Li, "Offline handwritten Chinese text recognition with convolutional recurrent neural networks," Neurocomputing, vol. 379, pp. 264–270, 2020.
[36] Y. Yang, M. Liao, and L. Zhang, "End-to-end handwritten text recognition using deep CNN and fully connected LSTM," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 515–523.
[37] S. Rusinol, A. Fornes, and J. Llados, "Writer identification by using local features," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 4, pp. 661–671, 2010.
[38] H. Feng, T. Zhao, and H. He, "Offline handwritten Chinese character recognition using the SVM classifier," Journal of Computers, vol. 12, no. 2, pp. 120–128, 2017.
[39] L. Neudecker, "Digitization of historical manuscripts," in Proceedings of the 15th International Conference on Digital Preservation (iPRES), 2018, pp. 54–62.
[40] G. M. F. Adak, S. Ramakrishnan, and A. Ghosh, "Unsupervised learning techniques for script recognition," IEEE Transactions on Image Processing, vol. 28, no. 5, pp. 2143–2157, 2019.
[41] M. Raghu and A. Tannenbaum, "Latent space clustering using unsupervised learning for handwriting recognition," IEEE Transactions on Neural Networks and Learning Systems, vol. 31, no. 7, pp. 2297–2305, 2020.
[42] T. H. Vu, S. Uchida, and H. Nguyen, "Improving domain adaptation for handwriting recognition using pseudo-labeling," in Proceedings of the IEEE International Conference on Document Analysis and Recognition (ICDAR), 2021, pp. 856–860.