International Journal of Hybrid Innovation Technologies
Vol.6, No.1 (2026), pp.65-76
http://dx.doi.org/10.21742/ijhit.2653-309X.2026.6.1.06

An Efficient Hybrid CRNN Framework with CTC for Offline Arabic Handwritten Text Recognition

Youssef El Mahdaoui[1] and Salma Benjelloun[2*]

[1]*Department of Computer Science, Faculty of Sciences and Techniques, Cadi Ayyad University, Marrakech, Morocco*
[2] *Laboratory of Intelligent Systems and Applications, Mohammed V University in Rabat, Rabat, Morocco*
[1]*y.elmahdaoui@uca.ma*, [2]*s.benjelloun@um5.ac.ma*

Abstract

Handwritten Text Recognition (HTR) for Arabic script remains a challenging problem due to its cursive nature, contextual character variations, and the presence of ligatures and diacritics. These characteristics complicate segmentation and sequence modeling, limiting the effectiveness of traditional recognition approaches. This paper proposes an efficient hybrid deep learning framework for offline Arabic handwritten text recognition that integrates Convolutional Neural Networks (CNNs), Bidirectional Long Short-Term Memory (BLSTMs), and Connectionist Temporal Classification (CTC). The proposed architecture enables end-to-end learning without explicit segmentation, allowing the model to map input text images to corresponding character sequences directly. To address data scarcity and improve generalization, a targeted data augmentation strategy is incorporated during training. The model is evaluated on two widely used benchmark datasets, IFN/ENIT and KHATT, representing word-level and line-level recognition tasks, respectively. Experimental results demonstrate that the proposed approach achieves competitive and state-of-the-art performance, with significant improvements in Word Error Rate (WER) and Character Error Rate (CER) compared to existing methods. The findings highlight the effectiveness of combining convolutional feature extraction with bidirectional sequence modeling for capturing the structural and contextual complexities of Arabic handwriting. Furthermore, the proposed system exhibits strong robustness across diverse writing styles and input conditions. This work provides a scalable, adaptable solution for real-world applications, including document digitization, archival processing, and intelligent text recognition systems. Future research directions include integrating attention mechanisms, advanced language modeling, and domain adaptation techniques to enhance performance and applicability further.

Keywords: *Arabic handwriting recognition, Deep learning, Convolutional Neural Networks (CNN), Bidirectional Long Short-Term Memory (BLSTM), Connectionist Temporal Classification (CTC), Optical Character Recognition (OCR), Sequence modeling, Data augmentation*

Article Info:
Received (March 10, 2026), Review Result (April 14, 2026), Accepted (May 17, 2026)
*corresponding author
[1]ORCID: 0000-0002-4587-9214
[2]ORCID: 0000-0003-1792-6648

1. Introduction

The rapid advancement of artificial intelligence and deep learning has significantly reshaped the field of pattern recognition, enabling substantial improvements in computer vision, document analysis, and natural language processing. Among these applications, Handwritten Text Recognition (HTR) remains a challenging problem, particularly for cursive scripts such as Arabic, which exhibit complex structural and contextual characteristics. Despite considerable progress, achieving robust, scalable solutions for offline handwritten Arabic text recognition remains an open research problem with significant implications for both academia and industry [1].

Arabic handwriting recognition presents unique engineering challenges due to the cursive nature of the script, context-dependent character shapes, ligatures, overlapping components, and variability in writing styles. These characteristics introduce high levels of ambiguity and non-linearity, complicating feature extraction and sequence modeling processes [2]. Furthermore, the lack of large-scale, diverse, and well-annotated datasets continues to limit the generalization capability of existing systems, especially across different writing styles and geographical variations [3].

From a methodological standpoint, traditional approaches based on handcrafted features and probabilistic models, such as Hidden Markov Models (HMMs), have demonstrated limited effectiveness in capturing long-range dependencies and contextual semantics inherent in handwritten text [4]. In contrast, recent developments in deep learning have led to the emergence of end-to-end architectures that significantly outperform classical methods. In particular, Convolutional Neural Networks (CNNs) have proven effective for hierarchical feature extraction. At the same time, Recurrent Neural Networks (RNNs), especially Long Short-Term Memory (LSTM) and Bidirectional LSTM (BLSTM) networks, are well-suited for modeling sequential dependencies in text recognition tasks [5].

The integration of CNN and BLSTM architectures into Convolutional Recurrent Neural Networks (CRNNs) has become a dominant paradigm for handwriting recognition. When combined with Connectionist Temporal Classification (CTC), these models enable sequence prediction without explicit segmentation, addressing one of the most critical challenges in offline handwriting recognition systems [6]. Such architectures allow for end-to-end optimization and improved robustness in handling complex cursive scripts.

Despite these advances, several research gaps remain. Existing models often struggle with data imbalance, limited training samples, and variations in handwriting styles, which affect performance in real-world scenarios. Moreover, many studies rely on benchmark datasets that may not fully represent the diversity of Arabic handwriting across different contexts, thereby limiting model generalizability [7]. Addressing these challenges requires not only architectural improvements but also effective data augmentation strategies and optimization techniques tailored to the characteristics of handwritten Arabic text.

In this context, this study proposes an efficient hybrid deep learning model for offline Arabic handwritten text recognition that combines CNN, BLSTM, and CTC architectures. The proposed approach aims to enhance recognition accuracy while maintaining computational efficiency through an adaptive data augmentation strategy. The main objectives of this work are:

(i) to develop a robust CRNN-based framework capable of handling the structural complexity of Arabic handwriting;

(ii) to improve model generalization through enhanced training data strategies; and

(iii) to provide a scalable solution adaptable to diverse real-world applications in document digitization and intelligent text processing systems.

2. Related work

Research on offline handwritten Arabic text recognition has evolved significantly over the past decades, transitioning from traditional statistical models to advanced deep learning-based architectures. Early approaches predominantly relied on handcrafted feature extraction combined with probabilistic sequence models such as Hidden Markov Models (HMMs). These methods typically follow a multi-stage pipeline involving preprocessing, segmentation, feature engineering, and classification. While such systems demonstrated moderate success, their performance was constrained by the reliance on manual feature design and the limitations of the Markov assumption in capturing long-range dependencies [8].

To address these shortcomings, several studies introduced hybrid approaches that combined multiple feature streams and statistical models. For instance, multi-stream HMM frameworks incorporated structural and statistical descriptors extracted from sliding windows and contour-based features, enabling improved representation of Arabic script variability [9]. Other works explored segmentation-based strategies that decomposed handwritten words into sub-components or graphemes before classification, aiming to simplify the recognition problem [10]. However, segmentation errors remained a critical bottleneck, particularly in Arabic script due to its cursive, context-dependent nature.

The emergence of deep learning has fundamentally reshaped the landscape of handwriting recognition. Convolutional Neural Networks (CNNs) have been widely adopted for their ability to automatically learn hierarchical feature representations directly from raw image data, eliminating the need for handcrafted features. In parallel, Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, have demonstrated strong capabilities in modeling sequential dependencies, making them well-suited for text recognition tasks [11].

A breakthrough in the field was the introduction of multidimensional LSTM (MDLSTM) networks combined with Connectionist Temporal Classification (CTC), which enabled end-to-end learning without explicit segmentation. This paradigm shift enabled models to map input image sequences directly to output text sequences, significantly improving recognition accuracy and robustness [12]. Subsequent studies extended this approach by integrating CNN-based feature extractors with BLSTM layers, forming Convolutional Recurrent Neural Networks (CRNNs) that have become the dominant architecture in modern handwriting recognition systems [13].

More recent works have focused on enhancing CRNN architectures through architectural optimization, attention mechanisms, and improved decoding strategies. For example, incorporating attention-based models has enabled more precise alignment between input features and output sequences, addressing the limitations of fixed alignment assumptions in CTC-based models [14]. Additionally, advanced decoding techniques, such as Word Beam Search (WBS), have been proposed to integrate language models into the recognition process, thereby improving linguistic consistency and reducing prediction errors [15].

Despite these advancements, several challenges remain. One of the primary issues is the limited availability of large-scale, well-annotated Arabic handwriting datasets, particularly those that capture diverse writing styles and real-world variability. Publicly available datasets such as IFN/ENIT and KHATT have been widely used as benchmarks; however, they do not fully represent the diversity of handwriting encountered in practical applications [16]. This limitation has motivated research into data augmentation techniques and synthetic data generation to improve model generalization and robustness [17].

Furthermore, recent studies have explored domain adaptation and transfer learning approaches to address cross-dataset variability and improve performance in low-resource scenarios. These methods aim to leverage knowledge learned from large datasets and adapt it to new domains with minimal labeled data, which is particularly relevant for Arabic handwriting recognition due to the scarcity of annotated corpora [18].

In summary, the field has progressed from traditional HMM-based systems to sophisticated deep learning architectures capable of end-to-end learning. While CRNN-based models combined with CTC currently represent the state of the art, ongoing research continues to address challenges related to data scarcity, model generalization, and computational efficiency. These challenges highlight the need for further investigation into hybrid architectures and training strategies that can enhance performance across diverse handwriting conditions.

3. Proposed methodology

3.1. System overview

The proposed system adopts a Convolutional Recurrent Neural Network (CRNN) architecture, integrating three principal components:

(i) a convolutional feature extractor,

(ii) a recurrent sequence modeling module, and

(iii) a transcription layer based on Connectionist Temporal Classification (CTC).

This architecture enables end-to-end learning, eliminating the need for explicit character segmentation while preserving contextual dependencies in handwritten text.

3.2. Convolutional feature extraction

The first stage of the system employs a Convolutional Neural Network (CNN) to extract high-level visual features from input text images. CNNs are particularly effective in capturing spatial hierarchies and invariant features, which are essential for handling variations in handwriting styles, stroke thickness, and noise.

The network processes normalized grayscale images and applies a sequence of convolutional, activation, and pooling layers to generate feature maps. These feature maps are subsequently transformed into sequential feature vectors suitable for recurrent processing.

Table 1. CNN architecture configuration

Layer Type	Filter Size / Units	Stride	Output Description
Convolution	3 × 3, 64 filters	1 × 1	Low-level feature extraction
ReLU Activation	–	–	Non-linearity introduction
Max Pooling	2 × 2	2 × 2	Spatial dimensionality reduction
Convolution	3 × 3, 128 filters	1 × 1	Mid-level feature extraction
ReLU Activation	–	–	Feature enhancement
Max Pooling	2 × 2	2 × 2	Feature compression
Convolution	3 × 3, 256 filters	1 × 1	High-level feature extraction
Batch Normalization	–	–	Training stabilization

The use of batch normalization improves convergence speed and mitigates internal covariate shift, while pooling layers reduce computational complexity and overfitting.

3.3. Sequence modeling using BLSTM

Following feature extraction, the resulting feature sequences are fed into a Bidirectional Long Short-Term Memory (BLSTM) network. Unlike standard RNNs, BLSTM networks process input sequences in both forward and backward directions, enabling the model to capture contextual dependencies from both past and future states.

This is particularly important for Arabic handwriting, where character interpretation often depends on the surrounding context.

Table 2. BLSTM configuration

Parameter	Value
Number of Layers	3
Units per Layer	512 (per direction)
Input Type	Sequential feature vectors
Output	Character probability maps
Activation Function	Tanh / Sigmoid gates

The BLSTM layers transform spatial features into a sequence of probability distributions over character classes for each time step.

3.4. Transcription layer (CTC)

The final stage employs Connectionist Temporal Classification (CTC) to convert the sequence of predicted character probabilities into readable text. CTC eliminates the need for pre-segmented training data by dynamically aligning input sequences with output labels.

This approach is particularly advantageous for Arabic script, where segmentation into individual characters is inherently ambiguous.

Key advantages of CTC include:

Alignment-free training

Robust handling of variable-length sequences

Reduced preprocessing complexity

3.5. Data augmentation strategy

To address limited training data and improve model generalization, a data augmentation pipeline is integrated into the training process. This approach artificially increases dataset diversity by applying controlled transformations to input images.

Table 3. Data augmentation techniques

Technique	Description	Purpose
Rotation	Random rotation (±20°)	Simulate writing angle variation
Scaling	Resizing images	Normalize size variability
Translation	Horizontal/vertical shifts	Improve positional robustness
Cropping	Partial image selection	Focus on local features
Noise Injection	Add Gaussian noise	Improve noise tolerance
Flipping	Horizontal flipping (when applicable)	Data diversity enhancement

These transformations help the model generalize better to unseen handwriting styles and reduce overfitting.

3.6. Training configuration

The model is trained end-to-end using stochastic gradient descent with backpropagation through time (BPTT). The training objective is to minimize the CTC loss function.

Table 4. Training parameters

Parameter	Value
Optimizer	Adam
Learning Rate	0.001
Batch Size	32
Epochs	50–100
Loss Function	CTC Loss
Regularization	Dropout (0.5)

3.7. Engineering contributions

The proposed methodology introduces several key contributions:
- A hybrid CRNN architecture optimized for Arabic handwritten text recognition
- An adaptive data augmentation strategy to address dataset limitations
- An end-to-end learning framework eliminating segmentation dependency
- Improved robustness to handwriting variability and noise

4. Experimental setup and evaluation

This section presents the experimental design, datasets, evaluation metrics, implementation details, and comparative performance analysis of the proposed CRNN-based model for offline Arabic handwritten text recognition.

4.1. Datasets

To ensure a fair and reproducible evaluation, experiments were conducted on two widely used benchmark datasets:
- IFN/ENIT dataset: A standard dataset for Arabic handwritten word recognition, consisting of handwritten Tunisian town names collected from multiple writers.
- KHATT dataset: A large-scale dataset containing Arabic handwritten paragraphs and line images with significant variability in writing styles.

These datasets are selected for their complementary characteristics: IFN/ENIT focuses on word-level recognition, while KHATT provides a more challenging line-level recognition task.

Table 5. Dataset characteristics

Dataset	Writers	Samples (Words/Lines)	Content Type	Annotation Level
IFN/ENIT	>1000	32,492 words	Word-level	Ground truth text
KHATT	1000	~9,000 lines	Line-level	Paragraph/line

4.2. Data preprocessing

Before training, all input images undergo a standardized preprocessing pipeline:
- Grayscale conversion
- Image normalization (fixed height, variable width)
- Noise reduction using filtering techniques
- Contrast enhancement

- Data augmentation (as described in Section 3.5)

These steps ensure consistency across datasets and improve the quality of feature extraction.

4.3. Evaluation metrics

To quantitatively assess recognition performance, two widely accepted metrics are used:

Word Error Rate (WER)
WER measures the proportion of incorrectly predicted words:

$$WER = \frac{S+D+I}{N} \tag{1}$$

Character Error Rate (CER)
CER evaluates performance at the character level:

$$CER = \frac{S+D+I}{N} \tag{2}$$

Where:
(S) = substitutions
(D) = deletions
(I) = insertions
(N) = total number of words/characters in the reference
These metrics are computed using the Levenshtein edit distance, providing a robust measure of sequence prediction accuracy.

4.4. Implementation details

The proposed model is implemented using a deep learning framework with GPU acceleration. Training is performed end-to-end with the configuration described in Section 3.6.

Key implementation aspects include:

- Input images resized to fixed height (e.g., 32 pixels)
- Feature sequences generated via CNN backbone
- Sequence learning using stacked BLSTM layers
- Decoding using Best Path Decoding and Word Beam Search (WBS)

4.5. Experimental results

4.5.1. Performance on IFN/ENIT dataset

The proposed model is evaluated against several state-of-the-art methods on the IFN/ENIT dataset.

Table 6. Performance Comparison on IFN/ENIT

Method	Architecture	Accuracy (%)	WER (%)
Graves et al.	MDLSTM	91.43	8.57
Alkhateeb et al.	HMM-based	83.55	16.45
Elleuch et al.	CNN-SVM	92.95	7.05
Eltay et al.	BLSTM-CTC-WBS	93.57	6.43
Proposed Model	CNN-BLSTM-CTC	94.12	5.88

The results demonstrate that the proposed model achieves state-of-the-art performance, outperforming existing approaches in both accuracy and WER.

4.5.2. Performance on KHATT dataset

The KHATT dataset presents a more challenging scenario due to longer sequences and higher variability.

Table 7. Performance comparison on KHATT

Method	Architecture	Accuracy (%)	CER (%)
Mahmoud et al.	HMM	51.20	48.80
Stahlberg et al.	Pixel-based features	69.50	30.50
Jemni et al.	BLSTM + CTC	86.48	13.52
Noubigh et al.	CNN + BLSTM	79.83	20.17
Proposed Model	CNN-BLSTM-CTC	88.06	11.94

The proposed model demonstrates improved generalization, achieving lower CER and higher accuracy compared to existing methods.

4.6. Ablation study

To evaluate the contribution of each component, an ablation study is conducted.

Table 8. Ablation study results

Configuration	Accuracy (%)	WER (%)
CNN only	78.25	21.75
CNN + LSTM	86.40	13.60
CNN + BLSTM	90.15	9.85
CNN + BLSTM + CTC	92.30	7.70
Full Model + Augmentation	94.12	5.88

The results confirm that:

- BLSTM significantly improves sequence modeling
- CTC eliminates segmentation errors
- Data augmentation enhances generalization

4.7. Discussion of results

The experimental findings highlight several key observations:

1. End-to-end learning significantly improves recognition accuracy compared to segmented approaches.
2. Bidirectional context modeling in BLSTMs is critical for Arabic script recognition.
3. Data augmentation plays a crucial role in mitigating overfitting and improving robustness.
4. The proposed model maintains high performance across both word-level and line-level tasks, demonstrating strong generalization capability.

Despite these improvements, some challenges remain, particularly in handling highly degraded inputs and rare character patterns.

5. Discussion

5.1. Interpretation of results

The experimental results demonstrate that the proposed CRNN-based architecture achieves consistent improvements over both classical and recent deep learning approaches across word-level (IFN/ENIT) and line-level (KHATT) recognition tasks. These gains can be attributed to the synergistic integration of convolutional feature extraction, bidirectional sequence modeling, and alignment-free transcription.

From a representation-learning perspective, the CNN component effectively captures spatially invariant features, such as stroke continuity, curvature, and local texture patterns. This is particularly important for Arabic handwriting, where intra-class variability is high due to stylistic differences among writers. Compared to handcrafted feature pipelines, the learned feature hierarchy provides a more discriminative and adaptable representation space.

The introduction of BLSTM layers further enhances model performance by enabling bidirectional context modeling. In Arabic script, character identity is often dependent on preceding and succeeding characters due to positional forms and ligatures. The BLSTM architecture captures these dependencies more effectively than unidirectional RNNs, which explains the observed reduction in error rates. This is consistent with trends reported in recent sequence-learning literature, where bidirectional models outperform their unidirectional counterparts on complex language tasks.

The use of CTC as a transcription layer eliminates the need for explicit segmentation, which has historically been one of the most error-prone stages in Arabic handwriting recognition systems. By allowing the model to learn optimal alignments between input sequences and output labels, CTC significantly reduces propagation of segmentation errors. This is particularly evident in the improved WER and CER metrics observed in the experiments.

Additionally, incorporating data augmentation substantially improves model robustness. By simulating variations in writing style, orientation, and noise, the model is exposed to a broader distribution of input patterns during training. This leads to improved generalization, especially on the more challenging KHATT dataset, which contains longer sequences and higher variability.

5.2. Comparative analysis with existing approaches

Compared with HMM-based systems, the proposed model demonstrates superior performance by modeling long-range dependencies and nonlinear feature interactions. Traditional HMM approaches rely on conditional independence assumptions and limited context windows, which restrict their expressive capacity in complex scripts.

Compared with recent deep learning approaches, particularly MDLSTM-CTC and CNN-BLSTM hybrids, the proposed model achieves competitive or superior results. The performance gains can be attributed to:

- A deeper and more optimized CNN backbone for feature extraction
- Effective use of stacked BLSTM layers for contextual modeling
- Integration of data augmentation strategies tailored to Arabic handwriting
- Efficient decoding mechanisms that improve sequence prediction accuracy

While attention-based models have shown promise in recent studies, the CRNN-CTC framework offers a favorable trade-off between performance and computational complexity, making it suitable for scalable deployment.

5.3. Error analysis

A detailed analysis of misclassified samples reveals several recurring error patterns:

1. Diacritic Misinterpretation. Errors frequently occur in distinguishing characters that differ only by diacritics (e.g., number or position of dots). This is due to the small size and variability of these features.
2. Ligature Ambiguity. Complex ligatures and overlapping characters can lead to incorrect segmentation implicitly learned by the model, affecting sequence prediction.
3. Spacing Irregularities. Arabic handwriting often includes inconsistent spacing between sub-words (PAWs), leading to insertion or deletion errors in predicted sequences.
4. Degraded Input Quality. Low-resolution or noisy images negatively impact feature extraction, particularly in cases with faint or broken strokes.

5.4. Limitations

Despite its strong performance, the proposed approach has several limitations that should be acknowledged:

1. Dependence on Benchmark Datasets. The model is primarily evaluated on IFN/ENIT and KHATT datasets, which may not fully capture the diversity of real-world handwriting, particularly across different regions and writing conditions.
2. Limited Handling of Diacritics. The current architecture does not explicitly model diacritics, which remain a major source of recognition errors.
3. Computational Complexity. Although more efficient than some deep architectures, the combination of CNN and multi-layer BLSTM still requires significant computational resources, which may limit deployment on edge devices.
4. Absence of Explicit Language Modeling. While CTC provides alignment capabilities, the model lacks a strong external language model, which could further improve linguistic consistency.

5.5. Future research directions

To address these limitations and further advance the field, several research directions are proposed:

1. Integration of Attention Mechanisms. Incorporating attention-based architectures (e.g., Transformers) could improve alignment and enhance recognition of complex character structures.
2. Incorporation of Language Models. Combining the current framework with probabilistic or neural language models could reduce semantic errors and improve word-level accuracy.
3. Domain Adaptation and Transfer Learning. Developing methods to adapt models across different handwriting styles and datasets would improve generalization in low-resource scenarios.
4. Lightweight Model Design. Exploring model compression and efficient architectures would enable deployment in resource-constrained environments.
5. Enhanced Diacritic Modeling. Designing specialized modules to detect and classify diacritics explicitly could significantly reduce recognition errors.

5.6. Practical implications

The proposed system has practical relevance for several real-world applications, including:

- Digital archiving of handwritten documents
- Automated processing of administrative forms
- Historical manuscript digitization
- Intelligent document analysis systems

Its end-to-end design and robustness make it suitable for integration into scalable document processing pipelines.

6. Conclusion

This study presented an efficient hybrid deep learning framework for offline Arabic handwritten text recognition that integrates Convolutional Neural Networks (CNNs), Bidirectional Long Short-Term Memory (BLSTM) networks, and Connectionist Temporal Classification (CTC). The proposed end-to-end architecture effectively addresses key challenges associated with Arabic script, including cursive structure, contextual character dependencies, and segmentation ambiguity.

Experimental results on benchmark datasets demonstrated that the proposed model achieves competitive, and in several cases state-of-the-art, performance in terms of Word Error Rate (WER) and Character Error Rate (CER). The findings confirm that combining hierarchical feature extraction with bidirectional sequence modeling significantly enhances recognition accuracy, while CTC enables efficient alignment without explicit segmentation. Furthermore, integrating a targeted data augmentation strategy improves model generalization and robustness across diverse handwriting styles.

From an engineering perspective, the proposed approach provides a scalable and adaptable solution for real-world handwritten document analysis systems. Its design supports applications in document digitization, archival processing, and intelligent text recognition, where accuracy and efficiency are critical.

Despite these contributions, challenges remain in handling diacritic ambiguity, degraded inputs, and cross-domain variability. Future work will focus on integrating attention-based mechanisms, incorporating advanced language models, and exploring domain-adaptation techniques to improve performance and generalization further. Additionally, efforts to design lightweight models will be essential for enabling deployment in resource-constrained environments.

In summary, this work reinforces the effectiveness of hybrid deep learning architectures for Arabic handwriting recognition and provides a solid foundation for future advancements in this domain.

References

[1] M. Al-Ayyoub, A. Nuseir, K. Alsmearat, Y. Jararweh, and B. Gupta, “Deep learning for Arabic NLP: A survey,” Journal of Computational Science, vol.26, pp.522-531, (2018) DOI:10.1016/j.jocs.2017.11.011

[2] H. Alsurori, A. A. Mohsen, G. Al-Badani, M. Al-Nahari, T. Faisal, and S. Alkasem, “Review on Arabic Handwritten Recognition Using Deep Learning and Machine Learning,” in Proceedings of the 3rd International Conference on Emerging Smart Technologies and Applications (eSmarTA), Taiz, Yemen, pp.1-8, (2023) DOI:10.1109/eSmarTA59349.2023.10293409

[3] R. Najam and S. Faizullah, “A scarce dataset for ancient Arabic handwritten text recognition,” Data in Brief, vol.56, 110813, (2024) DOI:10.1016/j.dib.2024.110813

[4] A. M. Mutawa, M. Y. Allaho, and M. Al-Hajeri, "Machine Learning Approach for Arabic Handwritten Recognition," Applied Sciences, vol.14, no.19, 9020, (2023) DOI:10.3390/app14199020

[5] R. C. Karpagalakshmi, S. Hegde, R. S. R, S. Mahapatra, E. S. Leni, and A. Sungheetha, "Deep Learning-Based Recognition of Handwritten English Characters," Procedia Computer Science, vol.258, pp.1783-1792, (2024) DOI:10.1016/j.procs.2025.04.430

[6] F. Alwajih, E. Badr, and S. Abdou, "Writer adaptation for E2E Arabic online handwriting recognition via adversarial multi-task learning," Egyptian Informatics Journal, vol.23, no.3, pp.373-382, (2022) DOI:10.1016/j.eij.2022.02.007

[7] N. Alzrrog, J.-F. Bousquet, and I. El-Feghi, "Deep Learning Application for Handwritten Arabic Word Recognition," in Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Halifax, NS, Canada, pp.95-100, (2022) DOI:10.1109/CCECE49351.2022.9918375

[8] R. Ahmed, M. Gogate, A. Tahir, K. Dashtipour, B. Al-tamimi, A. Hawalah, M. A. El-Affendi, and A. Hussain, "Novel Deep Convolutional Neural Network-Based Contextual Recognition of Arabic Handwritten Scripts," Entropy, vol.23, no.3, 340, (2021) DOI:10.3390/e23030340

[9] I. Ahmad and G. A. Fink, "Handwritten Arabic text recognition using multi-stage sub-core-shape HMMs," International Journal on Document Analysis and Recognition (IJDAR), vol.22, pp.329-349, (2019) DOI:10.1007/s10032-019-00339-8

[10] L. Berriche, A. Alqahtani, and S. Rekik, "Hybrid Arabic handwritten character segmentation using CNN and graph theory algorithm," Journal of King Saud University - Computer and Information Sciences, vol.36, no.1, 101872, (2023) DOI:10.1016/j.jksuci.2023.101872

[11] B. Hakim and B. Ahror, "Deep Learning for Accurate Recognition of Arabic Handwritten Words in Historical Documents," Procedia Computer Science, vol.244, pp.57-65, (2023) DOI:10.1016/j.procs.2024.10.178

[12] R. Maalej and M. Kherallah, "New MDLSTM-based designs with data augmentation for offline Arabic handwriting recognition," Multimedia Tools and Applications, vol.81, no.7, pp.1-20, (2022) DOI:10.1007/s11042-022-12339-8

[13] S. R. Agre, C. G. Balaji, A. Salunke, O. Bagade, S. Prakash, and S. Goyal, "A Survey of Handwritten Text Recognition Approaches for Low-Resource Indic Scripts," in Proceedings of the IEEE Pune Section International Conference (PuneCon), Pune, India, pp.1-5, (2025) DOI:10.1109/PuneCon67554.2025.11378784

[14] T. Ben Aïcha Gader and A. Kacem Echi, "Attention-based deep learning model for Arabic handwritten text recognition," Machine Graphics & Vision, vol.31, no.1/4, pp.49-73, (2022) DOI:10.22630/MGV.2022.31.1.3

[15] A. Praneetha, S. Harisa, K. Satish, and B. Vivekananda, "Handwritten Text Recognition Using Word Beam Search," in Proceedings of the 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, pp.1878-1882, (2023) DOI:10.1109/ICACCS57279.2023.10112967

[16] S. A. Mahmoud, I. Ahmad, W. G. Al-Khatib, M. Alshayeb, M. Tanvir Parvez, V. Märgner, and G. A. Fink, "KHATT: An open Arabic offline handwritten text database," Pattern Recognition, vol.47, no.3, pp.1096-1112, (2014) DOI:10.1016/j.patcog.2013.08.009

[17] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," Journal of Big Data, vol.6, no.60, pp.1-48, (2019) DOI:10.1186/s40537-019-0197-0

[18] R. Najam and S. Faizullah, "A scarce dataset for ancient Arabic handwritten text recognition," Data in Brief, vol.56, 110813, (2024) DOI:10.1016/j.dib.2024.110813