INTERNATIONAL JOURNAL OF Telecommunications AIR DEFENSE COLLEGE

A Comprehensive Approach to Arabic Handwriting Recognition: Deep Convolutional Networks and Bidirectional Recurrent Models for Arabic Scripts

Citation: Saber, A.; Taha ,A.; Abd El Salam , K.
Title. Inter. Jour. of Telecommunications, ***IJT'2024****, Vol. 04, Issue 02, pp. 01-11, 2024.*
Editor-in-Chief: Youssef Fayed.
Received: 21/05/2024.
Accepted: 15/07/2024.
Published: 15/07/2024.

Ayman Saber*[1], Ahmed Taha[2] and Khalid Abd El Salam [3].
Electrical Engineering Department, Suez Canal University, Ismailia, Egypt [1]; ayman.saber@eng.suez.edu.eg.
Electrical Engineering Department, Suez Canal University, Ismailia, Egypt [2]; ugs.161750@eng.suez.edu.eg
Electrical Engineering Department, Suez Canal University, Ismailia, Egypt. Department of Information System, College of Information Technology, Misr University for Science and Technology (MUST), 6th of October City 12566, Egypt [3]; khaled.abdelsalam@eng.suez.edu.eg.

Abstract: Arabic handwriting recognition presents unique challenges due to the complexities of Arabic calligraphy and variations in writing styles. Proposing a novel approach to address these challenges by leveraging advanced deep learning techniques. This focus is on Convolutional Neural Networks (CNNs) and Bidirectional Long Short-Term Memory (Bi-LSTM) networks, which are tailored specifically for recognizing handwritten Arabic text. Utilizing the KHATT dataset for comprehensive training and evaluation, implementing rigorous preprocessing steps to enhance data quality. Central to this methodology is the Res-Net152 architecture for feature extraction, which has proven highly effective. This approach achieved remarkable results, with a character error rate of approximately 2.96% and an accuracy of 97.04% on the testing dataset. These results significantly outperform the previous method, representing a substantial advancement in the field of Arabic handwriting recognition. The study demonstrates the potential of deep learning models in overcoming the unique challenges posed by Arabic script, paving the way for further improvements and applications.

Keywords: Optical Character Recognition (OCR), Artificial Neural Networks, Text segmentation, Document Digitization, KHATT Dataset.

1. Introduction

Arabic handwriting recognition has attracted a lot of attention lately because of its useful applications in several fields, such as postal services, automatic text recognition, and document analysis. Given the complexity of Arabic calligraphy and the inherent differences in writing styles, handwritten Arabic text recognition is a challenging task. As a result, it is now imperative to create strong and efficient methods for Arabic handwriting recognition.

Arabic handwriting recognition is more difficult than other scripts. Arabic is a linked language, meaning that several forms can be assigned to the same letter based on where it appears in a word. Furthermore, there is a great deal of individual variety in handwritten Arabic because of things like letter arrangement, spacing, and writing style. These elements add to the challenge of correctly identifying handwritten Arabic text.

There are still a lot of problems and gaps in Arabic handwriting recognition, despite the growing interest in the field. Achieving satisfactory levels of accuracy and robustness is a challenge for current recognition algorithms, particularly when handling unconstrained and noisy handwritten Arabic text. Furthermore, it is more difficult to build and assess novel algorithms and approaches due to the absence of publicly accessible standardized datasets created especially for the handwritten Arabic language.

The main objective of this research paper is to propose a new approach for recognizing handwritten Arabic text that tackles the mentioned challenges. Utilizing the KHATT dataset [1] and extracted a portion of it by selecting fixed paragraphs and performing segmentation. The Approach's advancements include the development of deep learning models, specifically Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), and attention mechanisms designed to recognize handwritten Arabic. Furthermore, the availability of comprehensive annotated datasets, such as KHATT, has facilitated the training and evaluation of these models, leading to improved recognition rates and reduced error rates.
This paper aims to develop a system that can effectively handle the complexities of Arabic handwriting and the inherent variations in handwriting styles. By leveraging recent advances in machine learning and deep neural networks, intending to enhance the accuracy and efficiency of recognizing handwritten Arabic text.
Arabic handwriting recognition technology has diverse applications, including the digital conversion of documents written in the Arabic script, which can accelerate the digital entry of government paperwork. It is also used in machine translation, text analysis, machine learning, and educational tools, where it can be utilized for correcting handwritten exams. Additionally, Arabic handwriting recognition helps preserve cultural heritage and historical manuscripts by converting fragile and deteriorating manuscripts into digital formats, ensuring their preservation, and providing access to researchers, historians, and the public.
This paper is organized into several sections to present a comprehensive analysis of this research topic. In the second section, the paper discusses the related works considered with the same idea and implementation, and the approaches they used, and the third section explains the materials and the preprocessing steps applied and used in this work, along with the methodology implemented, like the architecture and learning algorithm, then in the fourth section the training process is explained, the system the process was tried on, the evaluation metrics, and the results alongside the model summary, finally concluding and summarizing the work in the fifth section, then lastly in the sixth section, discussing the future work and our passionate steps and the goals set for the upcoming advancements. Finally, the list of the references and all the sources cited in this paper.

2. Related Works

Document Image Quality Assessment (DIQA) plays a crucial role in document analysis, Optical Character Recognition (OCR), and document management systems.
[2] surveys the field of Document Image Quality Assessment (DIQA), focusing on the impact of portable capturing technologies that generate diverse document images for various uses. The authors present a thorough analysis of both subjective and objective DIQA methods, with subjective methods relying on human opinions through ratings and pair-wise comparisons, and objective methods employing quantitative measurements, including document modeling and human perception-based techniques. They categorize the types and sources of document degradations and review techniques for modeling these degradations. The discussion includes two standard measures for assessing document image quality: OCR-based and objective human perception-based metrics. The survey concludes by identifying open challenges in the development of DIQA methods and offering future research directions, aiming to serve as a valuable resource for the document analysis research community.
DIQA methods aim to evaluate the quality and integrity of document images, which can be affected by factors like scanning artefacts, noise, blur, and degradation during image acquisition and processing. Accurate assessment of document image quality is essential for ensuring reliable and efficient document analysis and processing. In the following sections, exploring different DIQA methods, including traditional and deep learning-based approaches, highlighting their key principles, advantages, and limitations.

2.1. Traditional DIQA Methods

Traditional DIQA methods have been widely used in the past and often incorporate handcrafted features and expert knowledge to assess document image quality. These methods focus on various aspects such as character quality, sharpness, distortions, and geometric attributes.
[3] proposed a method for evaluating children's handwriting by analyzing symbols and letters represented in various geometric forms. Their approach utilized fuzzy inference systems, incorporating both generative and discriminative functionalities to assess handwritten symbols, particularly emphasizing geometric attributes

such as shape and direction. The study explored three distinct feature sets to evaluate symbol morphology, sequence, and direction, highlighting a substantial dependency on input data for accurate assessment. However, despite its detailed feature analysis, this method faced limitations in capturing the nuanced complexities of children's handwriting, potentially restricting its applicability in comprehensive handwriting assessment.
[4] Discussed the use of Adaptive Neuro-Fuzzy Inference System (ANFIS) and LTSM for offline OCR for English script, they discussed the utilization of OCR and Named Entity Recognition (NER) techniques, specifically LSTM for OCR and ANFIS for NER, in the context of English text documents. However, when considering Arabic handwritten OCR, several adjustments and considerations must be made due to the specific characteristics of the Arabic script.
[5] concentrated on evaluating the quality of Arabic letter handwriting. Their methodology integrated the Beta elliptical character segmentation model to meticulously examine individual characters and the Cartesian Fourier Descriptor model for character shape and boundary analysis. Data acquisition utilized a tablet, necessitating preprocessing procedures. Notably, the system was trained on a modest dataset comprising only 20 correctly written samples for each letter. Such a limited dataset size may prompt concerns regarding the reliability of the obtained results.
[6] introduced a model leveraging low-level attributes such as aspect ratio, zero-crossing distributions, and width distributions across the character's height. These characteristics were subsequently converted into high-level feature vectors. The evaluation and scoring of individual handwritten letters were conducted through a fusion of artificial neural network and expert system methodologies. However, it is pertinent to acknowledge that the implementation of these techniques demands a certain level of complexity in comprehension and application.
[7] employed no-reference techniques and metric-based approaches for assessing document quality. Their evaluation criteria encompassed sharpness quality and character quality metrics. Sharpness was evaluated using the LPC Sharpness Index and Log-Gabor filters of various scales. Character quality metrics included considerations such as black and white noise, as well as the presence of overlapping or touching characters. The primary focus of the study was on the specific detection and evaluation of distortions in scanned and mobile-captured images of documents, such as bills and receipts. Additionally, the analysis involved 175 images from the Tobacco dataset, with results compared against human visual perception.
[8] proposed an Arabic OCR system that tackles multiple fonts. It achieves this through a multi-stage pipeline: pre-processing, word-level feature extraction, character segmentation, character recognition, and post-processing. Notably, the authors use a novel character segmentation method that surpasses existing techniques, specifically the baseline segmentation approach used for making the system font independent. The system demonstrates high accuracy in character and word recognition on standard Arabic datasets.
[9] conducted a comprehensive review of various strategies for handling Arabic handwriting recognition. This study details general handwriting recognition, specific challenges in Arabic recognition, and the distinctions between online and offline recognition methods. Additionally, it highlights efforts related to Arabic datasets, such as the first online Quranic handwritten word dataset, and addresses other relevant efforts like estimating dates of historical Arabic documents. The paper emphasizes the complexity of Arabic handwriting recognition due to its ligatures, cursive nature, diacritics, and overlapping characters.
[10] provides a thorough review of the current state of Arabic OCR, focusing on the methodologies and techniques used in the OCR process, including preprocessing, segmentation, recognition, and postprocessing. The authors analyze existing literature using keyword-search methodology and backward and forward citation reviews to identify the most effective approaches. They highlight the challenges unique to Arabic OCR, such as complex morphology, contextual variations, cursive writing, diacritic marks, and the limited availability of labeled datasets. The survey emphasizes the superior performance of segmentation-based approaches and underscores the necessity for improved datasets and postprocessing techniques. By identifying research gaps, the paper aims to guide future research and development, contributing to the creation of more accurate and efficient Arabic OCR systems.

2.2. Deep Learning-based DIQA Methods

With the advancements in deep learning techniques, deep learning-based DIQA methods have gained significant attention in recent years. These methods utilize neural networks, such as CNNs [11] and RNNs [12], to automatically learn features and patterns from document images. Some common deep learning-based DIQA techniques include:

2.2.1. Convolutional Neural Networks (CNN):

CNNs have been widely used in DIQA to extract discriminative features from document images. These networks leverage multiple convolutional layers to capture hierarchical features and learn representations directly from the data. CNN-based DIQA methods often involve training models on large-scale annotated datasets and optimizing them using techniques like backpropagation and gradient descent.

2.2.2. Recurrent Neural Networks (RNN):

RNNs, particularly with Long Short-Term Memory [13] units have been applied to DIQA for capturing sequential dependencies and contextual information within document images. RNN-based models can effectively model the spatial and temporal relationships in document images, enabling improved quality assessment [14].

2.3. Integration of OCR with Existing DIQA Approaches

The paper titled "Recent advances of ML and DL approaches for Arabic handwriting recognition: A review" [15] provides a comprehensive bibliographic study of various techniques used for Arabic handwriting recognition. The authors review and compare different machine learning (ML) and deep learning (DL) methods, discussing holistic, analytical, and segmentation-free approaches. They describe the overall process of Arabic handwriting recognition, including pre-processing, feature extraction, and segmentation, and offer a detailed synthesis of the main techniques utilized in the field. The paper aims to highlight the differences between ML and DL approaches and motivate further research to develop more advanced recognition systems.
The paper "An End-to-End OCR Framework for Robust Arabic-Handwriting Recognition" [16] introduces a novel approach using Vision Transformers as encoders and vanilla Transformers as decoders, eliminating the need for CNNs. This system integrates OCR, deep learning, and comprehensive pre- and post-processing to improve the accuracy of Arabic handwritten text recognition. Utilizing a dataset of 270 million words and 30.5 million images, the approach achieved a 4.46% Character Error Rate (CER), marking a significant advancement in OCR for complex scripts like Arabic, although they needed a huge amount of Graphics Processing Unit (GPU) power and data to achieve this mark.
The recent “Handwritten Arabic Bills Reader and Recognizer” [17] paper contributes to the landscape DIQA. Specifically, propose an approach similar to [18] for recognizing and interpreting Arabic handwritten bills. This system integrates OCR, NER, and Deep Learning techniques, addressing the challenges associated with diverse handwriting styles and document content. Their approach on the KHATT dataset reached a CER of 13%, meaning an accuracy of 87%.
Going forward, that approach will be referred to as the previous method as this approach has improved upon their work, significantly increasing its accuracy.

3. Materials and Methods

The research endeavors to enhance the automation of grading Arabic exams through a refined approach in OCR and sequence modelling techniques. Presenting a detailed methodology leveraging the ResNet152 [19] architecture for feature extraction, as well as Connectionist temporal classification for LSTM, focusing on improving the accuracy and efficiency of automated grading systems for Arabic text-based exams.

3.1. Dataset

This research is aimed to take an open and accessible approach to Arabic OCR. This is why from the few datasets available, the choice landed on the KHATT dataset, which provides a diverse collection of handwritten paragraphs in Arabic script as well as being a publicly available dataset [20]. Before model training, the dataset was subjected to rigorous preprocessing steps to enhance its suitability for OCR tasks.

Using histogram projections; specifically, gradients of intensity histogram [21], for text line segmentation, employing a vertical projection profile to identify peaks corresponding to lines of text. By setting appropriate thresholds and utilizing peak detection algorithms, isolating individual text lines effectively, an example is presented in Figure 1.

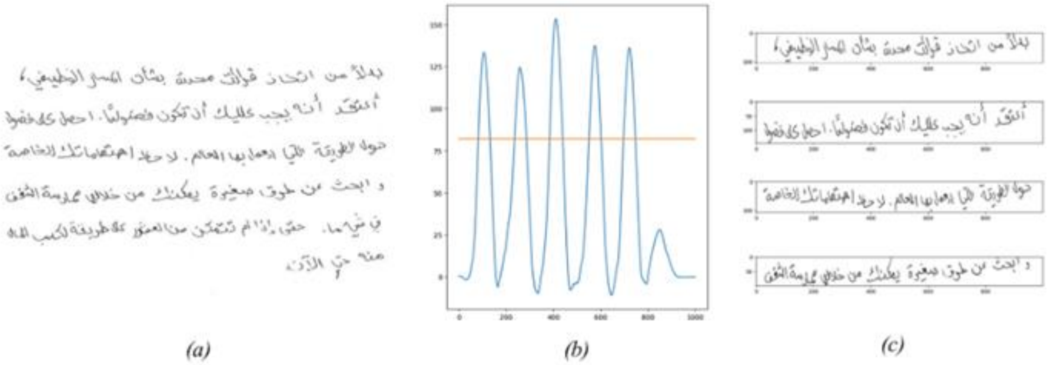

Figure 1. An example from the line segmentation algorithm (a) Original image before segmentation; (b) Histogram projection; and (c) Image after segmentation.

Extending this approach, projection segmentation is applied histogram along the horizontal axis to achieve word/sub-word level segmentation. This involved analyzing peaks in the projection profile to delineate boundaries between words or sub-words within a line. The result is called "text island". Examples are presented in Figure 2.

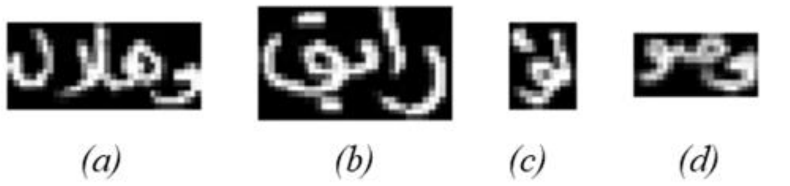

Figure 2. Examples of word/sub-word segmentation algorithm (text islands).

The data is then cropped into content to remove any empty spaces. Image padding was implemented to ensure consistent input dimensions for the model, by adding zero-padding to images, standardizing their size, and facilitating batch processing during training.
Augmenting the dataset by rotating each image twice: once clockwise by 5 degrees and then again counterclockwise by 5 degrees which proved to be beneficial in various deep learning applications [22]. This introduced the desired side-effect of a slight blur. The final image has a height of 32 pixels and a width of 64 pixels.

3.2. Model Architecture

The methodology utilizes the state-of-the-art CNN architecture, ResNet152, known for its deep learning capabilities. It is used to extract features from the input data. Unlike typical transfer learning approaches [23], not freezing any components of ResNet152. Instead, the network incorporates several layers specifically designed for handwritten Arabic text recognition as the flow chart shows in Figure 3.
The first layer is a dense layer with 128 units and Rectified Linear Unit (ReLU) activation, this layer enriches the features by applying a non-linear transformation. To improve training stability and generalization, a batch normalization layer is then applied. Subsequently, a dropout layer with a dropout rate of 0.3 is employed to prevent overfitting by randomly dropping a certain percentage of units during training.
The core layer for sequence modelling is a bidirectional LSTM layer with 512 units. This layer captures contextual information in both forward and backward directions within the sequence, crucial for recognizing

handwritten Arabic text. The output of each step in the LSTM layer is preserved, allowing the model to process the entire sequence. Additionally, a dropout rate of 0.3 is applied within the LSTM layer for further regularization.

The final layer is a dense layer with the number of units equal to the size of the character vocabulary plus two. This layer projects the processed features into a probability distribution over the characters and two additional tokens that are used for representing spaces and new line characters respectively in Connectionist Temporal Classification (CTC) loss. The SoftMax activation function is used to ensure the output probabilities sum to one. Finally, a CTC layer is applied to calculate the connectionist temporal classification loss.

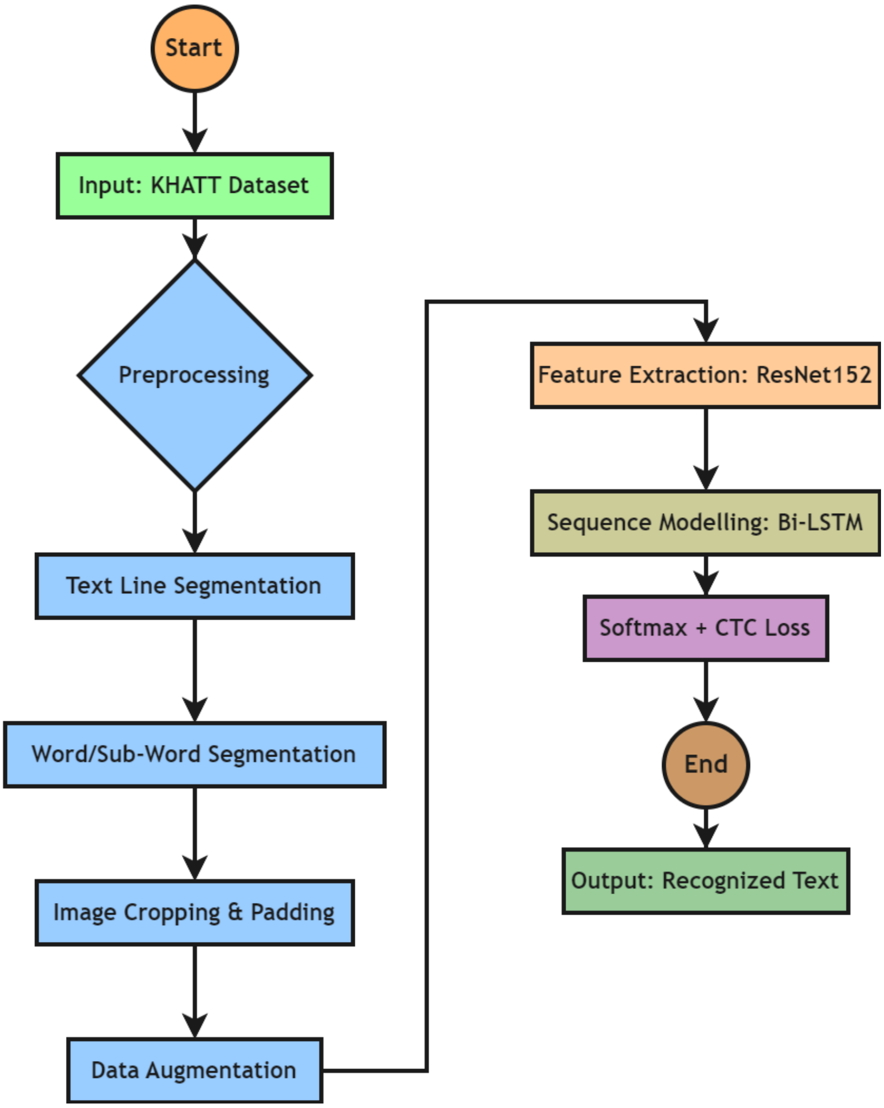

Figure 3. Flowchart of the proposed method.

3.3. Learning algorithm

As mentioned, utilizing a Bidirectional Long Short-Term Memory (Bi-LSTM) layer consisting of 1024 LSTM units in both forward and backward directions, as illustrated in Figure 4. This results in a total of 512 units in each direction when applying equations (1-6) [24]. This bidirectional design enables the model to efficiently recognize relationships between data entries in the sequence.

$$i_t = \sigma\left(x_t U^i + h_{t-1} W^i\right) \tag{1}$$

$$f_t = \sigma\left(x_t U^f + h_{t-1} W^f\right) \tag{2}$$

$$o_t = \sigma(x_t U^o + h_{t-1} W^o) \tag{3}$$

$$C_t = \tanh(x_t U^g + h_{t-1} W^g) \tag{4}$$

$$c_t = \sigma(f_t * C_{t-1} + i_t * C_t) \tag{5}$$

$$h_t = \tanh(c_t) * o_t \tag{6}$$

These equations describe the flow of information, where gates manage the cell state and hidden state, maintaining long-term dependencies and controlling information flow across time steps where i_t is the input gate activation, f_t is the forget gate activation, o_t is the output gate activation, C_t is the candidate cell state, c_t is the cell state, h_t is the hidden state, x_t is the input vector, h_{t-1} is the previous hidden state, U^i, U^f, U^o, U^g are the input weight matrices for the respective gates, W^i, W^f, W^o, W^g are the recurrent weight matrices for the respective gates, σ is the sigmoid function, and $\tanh$ is the hyperbolic tangent function.

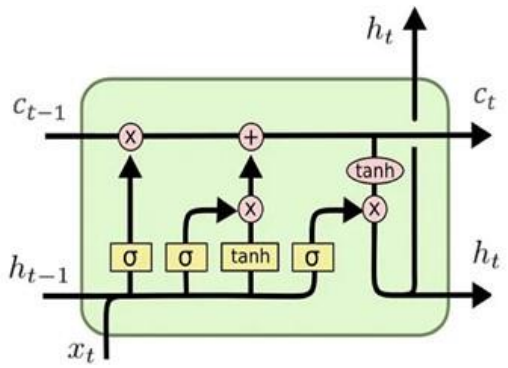

Figure 4. Diagram showing an LSTM unit.

4. Training and Results

4.1. System Configuration

This method is implemented using the following hardware and software configurations:

- OS: Windows 11 Pro.
- CPU: 12th Gen Intel (R) Core (TM) i3-12100F.
- Installed RAM: 16.0 GB DDR4 @ 3200MHZ.
- GPU: NVIDIA GeForce RTX 3060.

4.2. Performance Metric

Evaluating the model's performance on the test dataset by employing the Character Error Rate (CER) metric, it's used to evaluate the accuracy of text recognition particularly in OCR, CER is commonly used in fields such as computer vision, natural language processing, and audio processing, by comparing training and validation losses. The CER calculation is measured as the equation (7):

$$CER = (i + s + d) / n \tag{7}$$

Where (n) is the overall count of characters in ground truth, including spaces, (i) is the minimum number of insertions, representing the additional characters in the recognized text that are not present in the reference text, (s) is the minimum number of substitutions, indicating the characters in the recognized text that differ from those in the reference text, and (d) is the minimum number of deletions, referring to the characters in the reference text that are not recognized by the system.

Insertions occur when the model predicts additional characters that are not present in the ground truth label (e.g., truth = 'رؤوف', predicted = 'رؤووف'). Deletions, on the other hand, involve missing characters that the model fails to predict but are part of the ground truth label (e.g., truth = 'رؤوف', predicted = 'رؤف'). Substitutions refer to characters inaccurately predicted by the model (e.g., truth = 'رؤوف', predicted = 'زؤوف').

4.3. Training

The original set of 5,159 images driven from the "KHATT" dataset underwent augmentation as described previously to expand its size to 15,372. Following this, the dataset was divided into a training-testing set, with 80% designated for training and 20% for testing and validation purposes. Training multiple models to explore which model would get the highest accuracy. The training landed on three different choices as shown in Table 1.
Visual Geometry Group 19 architecture (VGG19): A total of 70 epochs were executed, considering the choice of optimizer, and learning rate, The ADAM optimizer, characterized by its parameters $\beta1=0.9$, and $\beta2=0.999$, and $\varepsilon = 1\times10^{-8}$, were employed, coupled with an exponential decay scheduler for the learning rate, starting at a maximum value of 0.0001. A smaller initial learning rate was chosen to decrease the risk of triggering overfitting prematurely, before achieving acceptable accuracy. These specific parameters for the ADAM optimizer contribute to the efficient optimization of the model's weights during training, enhancing convergence and overall performance. EfficientNetB1 was trained for 150 epochs using the ADAM optimizer with parameters $\beta1=0.9$, $\beta2=0.999$, and $\varepsilon = 1\times10^{-8}$, alongside an exponential decay scheduler for the learning rate starting at 0.001. This choice aimed to mitigate overfitting risk early on. Where ResNet152 underwent 150 epochs of training using the NADAM optimizer with parameters $\beta1=0.9$, $\beta2=0.999$, and $\varepsilon = 1\times10^{-8}$, coupled with an exponential decay scheduler for the learning rate starting at 0.001. This choice aimed to prevent premature overfitting.

Table 1. Comparison of Hyperparameter Tuning for Different CNN Architectures.

Parameters	VGG19	EfficientNetB1	ResNet152
Optimizer	ADAM	NADAM	NADAM
Learning Rate	0.0001	0.001	0.001
β_1	0.9	0.9	0.9
β_2	0.999	0.999	0.999
ε	1×10^{-8}	1×10^{-8}	1×10^{-8}
Batch Size	128	128	128

4.4. Results

Table 2 shows the results of each model in terms of validation loss, CER, and accuracy. Figure 5 shows a graph of the CER calculated after each epoch along with validation loss and the training loss for Resnet152. These visualizations help in understanding the model's performance trends over the training period.

Table 2. Summary of the experimental training.

Model	Validation Loss	Character Error Rate	Accuracy
VGG19	0.6	5.4%	94.6%
EfficientNetB1	0.78	7.3%	92.7%
ResNet152	0.3	2.96%	97.04%

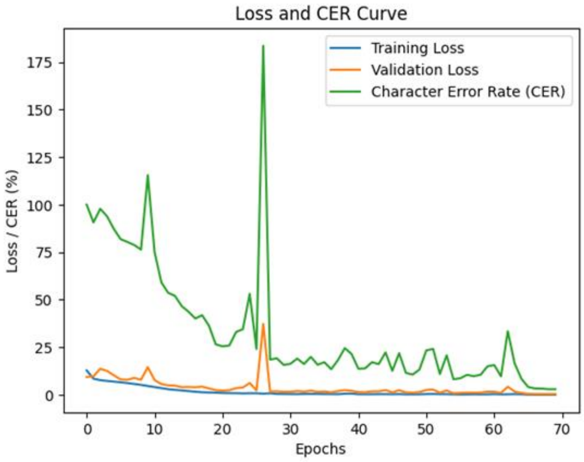

Figure 5. Graph showing CER, Training loss, and Validation loss after each epoch in Resnet152 approach for the first 70 epochs.

4.5. Summary

The chosen model's architecture is summarized in Table 3, and Table 4 shows a comparison between this method and the previous method.

Table 3. Summary of the chosen model architecture.

Layer (type)	Output Shape	Parameters	Connected to
image (InputLayer)	[(None, 32, 64, 3)]	0	[]
resnet152 (Functional)	(None, 1, 2, 2048)	58370944	['image[0][0]']
reshape (Reshape)	(None, 32, 128)	0	['resnet152[0][0]']
dense2 (Dense)	(None, 32, 128)	16512	['reshape[0][0]']
batch_normalization	(None, 32, 128)	512	['dense2[0][0]']
dropout (Dropout)	(None, 32, 128)	0	['batch_normalization[0][0]']
Bidirectional	(None, 32, 1024)	2625536	['dropout[0][0]']
label (InputLayer)	[(None, None)]	0	[]
dense3 (Dense)	(None, 32, 42)	43050	['bidirectional[0][0]']
ctc_loss (CTCLayer)	(None, 32, 42)	0	['label[0][0]','dense3[0][0]']
Total params: 61,056,554 Trainable params: 60,904,874 Non-trainable params: 151,680			

Table 4. Comparison of this approach and the previous method.

Parameters	Previous method	The Approach's method
No.of epochs	450	**150**
Val loss	1.1795	**0.3**
CER	13 %	**2.96 %**
Accuracy	87 %	**97.04 %**

5. Conclusions

Driven by the passion for developing advanced text-processing techniques in Arabic, recognizing their potential to bolster various areas mentioned earlier (education, government, data analysis, and information retrieval). This dedication led us to closely monitor advancements in Arabic handwriting recognition.

The recent research paper "Handwritten Arabic Bills Reader and Recognizer" was impressive. Which gave motivation to push this research, culminating in noteworthy results.

The models, particularly the impressive ResNet152, yielded significantly better performance compared to the referenced research.

This achievement, with its parameters of ADAM optimizer, 0.001 learning rate, 150 epochs, 0.3 validation loss, and 2.96% CER, represents a substantial leap in accuracy (97.04%).

This accomplishment fuels the commitment to further research in this domain, to achieve even more advanced levels of Arabic handwriting recognition.

6. Acknowledgment

Special thanks the following individuals for their significant contributions to this research: Ahmed Nagah Ahmed, Mohamed Fathi Mohamed, Reem Mohamed Fouad, Rawan Gamal Mohamed, Nada Hussain Asran, Kerollos Samir Fathi, Abanoub Aied, Mohamed Abd-El Fattah, and Nada Mahmoud Mohamed. Their guidance, dataset collection, tool development, preprocessing research, and financial support have been crucial to the project's success.

Funding: Collective financial contributions from all team members facilitated hardware provision (RTX 3060 with 12GB of VRAM), highlighting the shared commitment to research excellence.

As a team, each member's unique contribution has been instrumental, and we are looking forward to continued collaboration in future endeavors.

References

1. Mahmoud, S.A.; Ahmad, I.; Alshayeb, M.; Al-Khatib, W.G.; Parvez, M.T.; Fink, G.A.; Märgner, V.; Abed, H. El KHATT: Arabic Offline Handwritten Text Database. In Proceedings of the 2012 International Conference on Frontiers in Handwriting Recognition; 2012; pp. 449–454.
2. Alaei, A.; Bui, V.; Doermann, D.; Pal, U. Document Image Quality Assessment: A Survey. *ACM Comput. Surv.* **2023**, *56*, doi:10.1145/3606692.
3. Bouillon, M.; Anquetil, E. Handwriting Analysis with Online Fuzzy Models. In Proceedings of the 17th Biennial Conference of the International Graphonomics Society; Rémi, C., Prévost, L., Anquetil, E., Eds.; Pointe-à-Pitre, Guadeloupe, June 2015.
4. Suganthi, M.; Arun Prakash, R. An Offline English Optical Character Recognition and NER Using LSTM and Adaptive Neuro-Fuzzy Inference System. *Journal of Intelligent & Fuzzy Systems* **2023**, *44*, 3877–3890, doi:10.3233/JIFS-221486.
5. Akouaydi, H.; Hamdi, Y.; Boubaker, H.; Zaied, M.; Alaya Cheikh, F.; Alimi, A. *Children's Online Handwriting Quality Analysis.*; 2021;
6. Kulesh, V.; Schaffer, K.; Sethi, I.; Schwartz, M. Handwriting Quality Evaluation. In Proceedings of the International Conference on Advances in Pattern Recognition; Berlin, Heidelberg, March 2001; pp. 157–165.
7. Nayef, N.; Ogier, J.M. Metric-Based No-Reference Quality Assessment of Heterogeneous Document Images. In Proceedings of the Document Recognition and Retrieval XXII; February 2015; Vol. 9402, p. 94020L.
8. Osman, H.; Zaghw, K.; Hazem, M.; Elsehely, S. An Efficient Language-Independent Multi-Font OCR for Arabic Script 2020.
9. Youssef, N.I.; Abd-Alsabour, N. A Review on Arabic Handwriting Recognition. *Journal of Southwest Jiaotong University* **2022**, *57*, 746–756, doi:10.35741/issn.0258-2724.57.6.66.

10. Kasem, M.S.; Mahmoud, M.; Kang, H.-S. Advancements and Challenges in Arabic Optical Character Recognition: A Comprehensive Survey. *ArXiv* **2023**, *2312.11812v1*.
11. O'Shea, K.; Nash, R. An Introduction to Convolutional Neural Networks. *arXiv preprint arXiv:1511.08458* **2015**.
12. Jordan, M.I. *Serial Order: A Parallel Distributed Processing Approach. Technical Report, June 1985-March 1986*; United States, 1986;
13. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput* **1997**, *9*, 1735–1780, doi:10.1162/neco.1997.9.8.1735.
14. Sherstinsky, A. Fundamentals of Recurrent Neural Network (Rnn) and Long Short-Term Memory (Lstm) Network. *Physica D* **2020**, *404*, 132306, doi:10.1016/j.physd.2019.132306.
15. Mezghani, A.; Maalej, R.; Elleuch, M.; Kherallah, M. Recent Advances of ML and DL Approaches for Arabic Handwriting Recognition: A Review. *Int. J. Hybrid Intell. Syst.* **2023**, *19*, 61–78, doi:10.3233/HIS-230005.
16. Mostafa, A.; Ahmed, O.; Ashraf, A.; Elbehery, A.; Jamal, S.; Salah, A.; Ghoneim, A. *An End-to-End OCR Framework for Robust Arabic-Handwriting Recognition Using a Novel Transformers-Based Model and an Innovative 270 Million-Words Multi-Font Corpus of Classical Arabic with Diacritics*; 2022;
17. Sweidan, S.; Hammam, M. Handwritten Arabic Bills Reader and Recognizer. *Journal of Computing and Communication* **2024**, *3*, 44–54, doi:10.21608/jocc.2024.339920.
18. Liwicki, M.; Graves, A.; Bunke, H.; Schmidhuber, J. A Novel Approach to On-Line Handwriting Recognition Based on Bidirectional Long Short-Term Memory Networks.; 2007.
19. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016; pp. 770–778.
20. Faizullah, S.; Ayub, M.S.; Hussain, S.; Khan, M.A. A Survey of OCR in Arabic Language: Applications, Techniques, and Challenges. *Applied Sciences* **2023**, *13*, doi:10.3390/app13074584.
21. Santos, R.; Clemente, G.; Ing Ren, T.; Cavalcanti, G. Text Line Segmentation Based on Morphology and Histogram Projection. *Document Analysis and Recognition, International Conference on* **2009**, *0*, 651–655, doi:10.1109/ICDAR.2009.183.
22. Shorten, C.; Khoshgoftaar, T.M. A Survey on Image Data Augmentation for Deep Learning. *J Big Data* **2019**, *6*, 60, doi:10.1186/s40537-019-0197-0.
23. Bozinovski, S.; Fulgosi, A. The Influence of Pattern Similarity and Transfer Learning upon the Training of a Base Perceptron B2. (Original in Croatian). In Proceedings of the Proceedings of Symposium Informatica; Bled, 1976; Vol. 3, pp. 121–125.
24. O'Shea, T.; Hitefield, S.; Corgan, J. End-to-End Radio Traffic Sequence Recognition with Deep Recurrent Neural Networks. **2016**.