



Faculty of Engineering and Materials Science
German University in Cairo

Title

Bachelor Thesis

by

Name

Supervised by

Prof. Dr. Gamal Abdel Shafy

Date

Declaration

This is to certify that:

- (i) the thesis comprises only my original work toward the Bachelor of Science (B.Sc.)
at the German University in Cairo (GUC),
- (ii) due acknowledgement has been made in the text to all other material used

Name

Date

Acknowledgement

Contents

1	Introduction	8
2	Literature Review	11
2.1	Advances in Arabic Text Recognition: OCRs	11
2.2	Reviewing Arabic Optical Character Recognition	13
2.3	Transformers for Scalable Image Recognition	14
2.4	Student Outcomes Assessments Using Deep Learning	14
2.5	Optimized DenseNet121 for Arabic Sign Language Recognition Using Harris Hawks Algorithm	15
2.6	Segmentation-based, omnifont printed Arabic character recognition with- out font identification	15
2.7	Automatic recognition of handwritten Arabic characters:a comprehen- sive review	17
2.8	Attention-Based CNN-RNN Arabic Text Recognition from Natural Scene Images	17
2.9	Novel Deep Convolutional Neural Network-Based Contextual Recogni- tion of Arabic Handwritten Scripts	18
2.10	Enhanced Arabic Handwriting Recognition with Standalone and Hybrid CNN Models	19
2.11	Automated Classification of Arabic Script Styles Using Enhanced CNN Frameworks	20
2.12	Adaptive Offline Arabic Handwriting Recognition with Deep Learning .	21
2.13	Arabic Handwritten Character Recognition Using Deep CNN	21
3	Extra section	22

4	Methodology	23
4.1	Model	23
4.2	the Data Set	23
4.3	How is the data processed?	23
4.4	Output	25
4.5	comparison	25
5	Results	26
6	Conclusion	27

List of Figures

2.1	Types of OCR [1]	12
2.2	intermediate steps during OCR main stages. [8]	16

List of Tables

Abstract

write here

Chapter 1

Introduction

The classification of Arabic script styles is an essential yet challenging task due to the unique characteristics of the Arabic language, including its cursive nature, contextual letter shaping, and diacritics. Arabic script has various styles, such as Naskh, Thuluth, and Kufic, each with distinct visual features. Manually classifying these styles is time-consuming and prone to errors, making automation a necessity. With advancements in deep learning, particularly Convolutional Neural Networks (CNNs), automated Arabic script classification has become increasingly feasible. CNNs excel at pattern recognition and have been successfully applied in font classification, achieving high accuracy. Recent studies have leveraged large Arabic text datasets and models like AlexNet and ResNet to improve classification accuracy. Additionally, transformer-based models, originally designed for natural language processing, have shown promising results in image recognition, further enhancing script classification techniques.

The broader field of Optical Character Recognition (OCR) has played a significant role in digitizing printed and handwritten Arabic text. While OCR has advanced significantly for many languages, Arabic remains a challenge due to its complex script and diacritical marks. Research on Arabic OCR has focused on improving recognition accuracy through deep learning techniques, such as fine-tuning the Tesseract 5.0 engine, feature extraction, and data augmentation. Despite notable progress, challenges persist, especially in handling diacritics and complex ligatures. Various deep learning models, including CNNs, EfficientNet, and ResNet, have been explored in different applications, such as Arabic sign language recognition, air-writing recognition, and traffic sign detection. Studies have demonstrated the effectiveness of these models in rec-

ognizing Arabic characters and styles, highlighting the importance of dataset quality, preprocessing techniques, and transfer learning.

The growing demand for accurate Arabic script classification is driven by applications in historical text preservation, automated document processing, and digital typography. By developing an enhanced classification system using deep learning, this research aims to bridge existing gaps in Arabic script recognition. Leveraging CNN-based models, feature extraction techniques, and pre-trained architectures like ResNet, this study seeks to improve classification accuracy while addressing challenges related to font variations and script complexity. Through these advancements, automated Arabic script classification can contribute to preserving Arabic cultural heritage and improving digital text processing systems.

[12pt]report fontspec polyglossia english arabic [Script=Arabic]Amiri graphicx,float

Chapter 2

Literature Review

Recognition of handwriting had been one of the hardest problems in pattern recognition and Artificial Intelligence. A lot of systems have been developed over the years to boost the level of accuracy and flexibility especially for complicated scripts like Arabic. Old approaches betrayed, like the Hidden Markov Models (HMMs), faced difficulties in dealing with long sequences of texts and variations of handwriting. Recent developments in deep learning with the advent of Convolutional Neural Networks (CNNs) together with Recurrent Neural Networks (RNNs) changed the game because they made data features' extraction and the sequences' modeling automation possible. This paragraph summarizes major achievements in offline recognition of Arabic handwriting with special regard to CNNs, RNNs, and domain adaptation techniques to increase writer-independent recognition.

2.1 Advances in Arabic Text Recognition: OCRs

Optical Character Recognition (OCR) extracts text from images or scanned documents and converts it into a machine-readable format. It is widely used for digitizing documents, making them searchable, editable, and more accessible. This paper surveys Arabic OCR, its applications, techniques, and challenges.

Key Points: Types of OCR: Unilingual OCR (supports one language, e.g., Arabic). Multilingual OCR (supports multiple languages). Offline OCR (processes scanned or printed documents). Online OCR (real-time text recognition, e.g., number plate recognition).

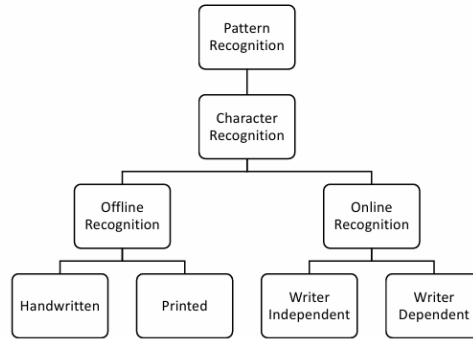


Figure 2.1: Types of OCR [1]

Arabic OCR Challenges: Arabic is written right to left and has 28 letters, including vowels ا , إ ,

Many letters change shape depending on their position in a word. Diacritics and ligatures make recognition complex. Lack of large datasets for deep-learning models. Future Directions:

Improving recognition of diacritics and complex letter shapes. Developing more accurate deep-learning models. Enhancing OCR performance for handwritten Arabic text.[1]

the identification of air-written Arabic letters using machine learning techniques, OCR, and convolutional neural networks is performed. Air writing is a technology that has been researched extensively for English and Chinese languages but remains unsolved for Arabic. Introduced are deep learning models based on VGG16, VGG19, and SqueezeNet and combined them with ML approaches based on neural networks, random forest, and SVM to increase accuracy. The AHAWP dataset was used, which contains many styles of writing, and preprocessing steps were taken to minimize bias. Different OCR algorithms are used to segment individual letters from strokes. VGG16 with neural network gave the best results with accuracy of 88.8 percent. This research fulfills an important need in the recognition of Arabic air-writing and has possible uses in education (teaching Arabic), assistive technology (for users with speech disabilities), and human computer interaction [6]

Optical Character Recognition (OCR) is the process of converting scanned or printed

text and images into a machine-readable format. OCR technology has evolved significantly for different written languages and scripts, but Arabic language and script offers unique challenges owing to its cursive style, letter form, diacritics shapes, and context sensitive letter joining. This study aims to enhance Arabic OCR by customizing Tesseract 5.0 engine. The study collects and trains a tailored dataset to improve the recognition accuracy denoted in terms of character error rate (CER) and word error rate (WER). The outcome showcases significant advancement in the recognition of Arabic scripts which facilitates the digitization of historical documents and automated extraction of texts. Nonetheless, challenges persist, particularly concerning diacritical marks and intricate ligature combinations, which restricts predictive precision. The research work enhances the Tesseract 5.0 engine to better Arabic Optical Character Recognition (OCR) and integrate it into practical applications across different disciplines.[4]

2.2 Reviewing Arabic Optical Character Recognition

The progress made in Arabic Optical Character Recognition (OCR) over the past decade. OCR is used to convert printed or handwritten text into editable text, helping in tasks like searching, analyzing, and modifying documents. While OCR systems for languages like English, Chinese, and Japanese have advanced, Arabic OCR still faces challenges due to its cursive nature, diacritics, and letter similarities.

The study explains the different stages of Arabic OCR, the techniques used, and the datasets available for research. It also compares handcrafted and deep learning-based OCR methods, showing that deep learning (CNNs, RNNs) has improved accuracy. However, most Arabic OCR systems still struggle, especially with page-level recognition, and commercial systems show low accuracy (under 75 percent).

The study provides a detailed comparison of existing Arabic OCR methods and software, along with a review of evaluation metrics. It also implements preprocessing

and segmentation techniques to enhance Arabic OCR performance. The goal is to help both beginners understand Arabic OCR and researchers improve current techniques.^[2]

2.3 Transformers for Scalable Image Recognition

This paper talks about using Transformers (a type of AI model originally used for text processing) for computer vision (teaching computers to understand images). Normally, Convolutional Neural Networks (CNNs) are used for image tasks, either alone or mixed with attention-based models.

However, the authors show that pure Transformers can work really well for image classification without needing CNNs. Instead of analyzing the entire image at once, they split the image into small patches and process them like words in a sentence.

At first, this Transformer model (called Vision Transformer or ViT) didn't perform as well as CNNs when trained on small datasets. But when trained on large datasets (millions of images), ViT outperformed CNNs while also requiring fewer computing resources.¹

2.4 Student Outcomes Assessments Using Deep Learning

Many universities want their degree programs to be officially recognized (accredited) because it improves teaching quality and helps attract talented students and teachers. Accreditation is now necessary for most universities worldwide. For example, ABET, an organization that accredits engineering and technology programs in the U.S., now accredits programs in other countries too.

To get accredited, universities must collect and report data about how well students meet learning goals. This includes Course Learning Outcomes (CLOs), Program Learning Outcomes (PLOs), and Key Performance Indicators (KPIs). Since there's a

lot of student data available, universities need better ways to analyze it.

This research suggests using an intelligent system based on Artificial Neural Networks (ANN) to measure student performance and accreditation requirements. The model uses deep learning, specifically the Multilayer Perceptron (MLP) method, which includes different layers of processing to classify and evaluate key performance indicators.

The study shows that this method can improve how universities measure student learning, making the accreditation process easier and more accurate.[\[3\]](#)

2.5 Optimized DenseNet121 for Arabic Sign Language Recognition Using Harris Hawks Algorithm

This research presents a new method for recognizing Arabic Sign Language (ArSL) using deep learning. The study compares different models like CNNs, EfficientNet, ResNet50, and DenseNet, and improves their performance using Harris Hawks Optimization (HHO). The best results came from DenseNet121 combined with HHO, achieving an impressive 99.79 percent accuracy on the ArASL2018 dataset (which contains 54,049 images)

2.6 Segmentation-based, omnifont printed Arabic character recognition without font identification

OCR (Optical Character Recognition) is used in many real-life applications like scanning books, recognizing car plate numbers, and processing bank cheques. However, Arabic OCR is difficult to develop because Arabic script is written in a connected (cursive) way, has different fonts, and includes special marks (diacritics).

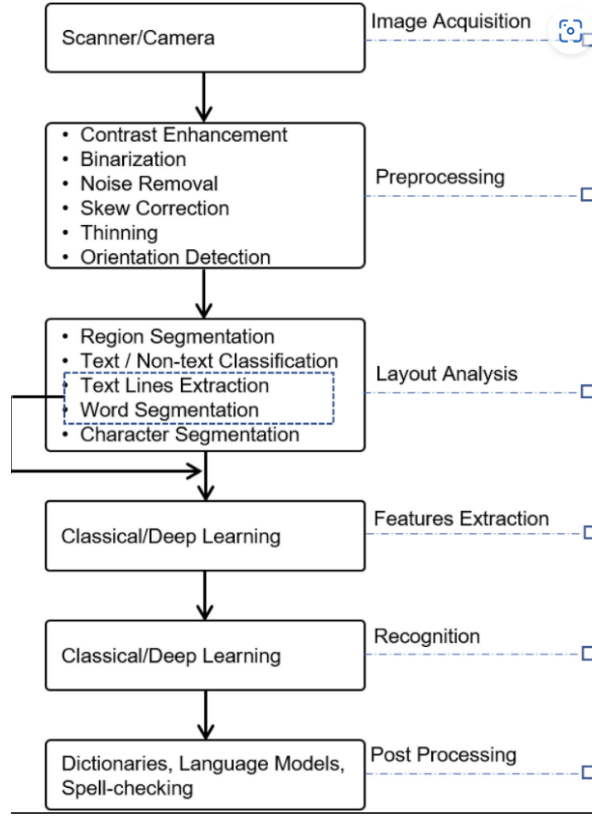


Figure 2.2: intermediate steps during OCR main stages.[8]

This research presents a new segmentation-based Arabic OCR system that does not require knowing the font type in advance. It uses a three-step method to separate overlapping characters and ensures high accuracy. The separated characters are then recognized using deep learning a Convolutional Neural Network (CNN). The system was tested using a dataset (APTID-MF) and achieved 95 percent accuracy for segmentation and 99.97 percent accuracy for recognition. The research also compares other existing Arabic OCR systems, explaining their limitations, such as low accuracy and difficulty handling multiple fonts. The study provides a new dataset of segmented Arabic characters and proves that deep learning improves accuracy without needing extra corrections.[8]

2.7 Automatic recognition of handwritten Arabic characters:a comprehensive review

Arabic handwritten text recognition is a growing research field due to the complexity of the Arabic script and its various challenges, such as character connectivity, multiple writing styles, and diacritical marks. Several studies have explored different methods for improving recognition accuracy, leveraging machine learning and deep learning techniques. One of the key contributions in this domain is the review of state-of-the-art approaches, which highlights the challenges and advancements in Arabic handwritten text recognition. Researchers have also developed and utilized several benchmark datasets, such as IFN/ENIT, AHDB, KHATT, AHTID/MW, HADARA, and ARABASE, to train and evaluate recognition models. These datasets provide diverse handwriting samples, supporting the development of more robust systems. Additionally, comparisons between different recognition techniques, including traditional feature extraction methods and modern deep learning architectures, have been extensively studied to enhance system performance. Despite significant progress, open challenges remain, such as improving recognition accuracy for complex handwritten scripts and addressing the lack of high-quality annotated datasets. This review provides a foundation for future advancements by identifying key research gaps and proposing directions for further improvements in Arabic handwritten text recognition.[\[4\]](#)

2.8 Attention-Based CNN-RNN Arabic Text Recognition from Natural Scene Images

Arabic text recognition is a complex challenge due to the cursive nature of the script, variations in fonts, sizes, and environmental conditions. Traditional Optical Character Recognition (OCR) methods have struggled with natural scenes where text appears in complex backgrounds or different orientations. Recent advancements in deep learning,

particularly the use of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have significantly improved accuracy in text recognition tasks. CNNs help in extracting essential features from images, while RNNs process sequences of text, making them effective for recognizing Arabic script. However, RNNs often face limitations in handling long sequences efficiently. To address this, an attention mechanism is introduced, which allows the model to focus on the most relevant parts of an image rather than processing all features equally. This approach improves the accuracy of Arabic text recognition in natural scenes, videos, and printed documents, making it valuable for applications like automated document processing, video analysis, and assisting visually impaired users. Despite these advancements, challenges such as limited annotated datasets and handling complex real-world text conditions remain open research areas.[10]

2.9 Novel Deep Convolutional Neural Network-Based Contextual Recognition of Arabic Handwritten Scripts

Handwritten Arabic recognition remains a challenging problem due to the unique characteristics of the script, including cursiveness, ligatures, diacritics, and variation in handwriting styles. Traditional methods struggle with these challenges, leading to the adoption of deep learning approaches, particularly Convolutional Neural Networks (CNNs). CNNs have proven highly effective in pattern recognition tasks, including handwriting recognition.

This study introduces a Deep Convolutional Neural Network (DCNN) model to address the limitations of previous methods. The model incorporates batch normalization and dropout regularization to enhance generalization and prevent overfitting. It is trained and tested on six benchmark databases of handwritten Arabic text, including digits, characters, and words. The study also applies transfer learning (TL) techniques,

comparing the proposed model to VGGNet-19 and MobileNet, demonstrating superior performance.

Furthermore, the model's ability to generalize is tested on the MNIST English digits dataset, confirming its robustness across different languages. The findings suggest that deep learning, particularly CNN-based architectures, provides a powerful solution for offline Arabic handwriting recognition, outperforming traditional methods and state-of-the-art pre-trained models.[11]

2.10 Enhanced Arabic Handwriting Recognition with Standalone and Hybrid CNN Models

Handwritten character recognition (HCR) is a well-researched area in pattern recognition and computer vision. Many studies have focused on recognizing English handwriting, achieving high accuracy using deep-learning techniques such as Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs). However, Arabic handwriting recognition has received less attention due to the complexity of the script and the limited availability of high-quality datasets. Several machine-learning methods, such as Support Vector Machines (SVMs), have been used for Arabic handwriting recognition but have not performed as well as CNNs. Research shows that CNN models can automatically extract features from handwritten text images and outperform classical machine-learning techniques, especially when trained on large datasets. However, existing Arabic handwriting datasets are often too small, leading to challenges in training deep-learning models effectively. Studies have also explored various deep-learning architectures such as Recurrent Neural Networks (RNNs) and Generative Adversarial Networks (GANs) to enhance handwriting recognition accuracy. Despite these advancements, Arabic handwriting recognition still faces challenges such as variations in writing styles, dataset noise, and limited dataset size. Future research should focus on creating larger, more diverse Arabic handwriting datasets and exploring advanced deep-learning

techniques to improve recognition performance.[\[5\]](#)

2.11 Automated Classification of Arabic Script Styles Using Enhanced CNN Frameworks

. Arabic script has many different styles, such as Naskh, Thuluth, and Kufic. These styles look different, making it hard to classify them manually, as it takes a lot of time and can lead to mistakes. Using deep learning, especially Convolutional Neural Networks (CNNs), can help automate this process and improve accuracy.

CNNs are good at recognizing patterns in images and have already been used successfully to classify Arabic fonts. Some studies have used large datasets with thousands of Arabic text samples to train CNN models such as AlexNet and ResNet, achieving high accuracy in distinguishing between different fonts.

Besides CNNs, newer AI models called transformers, originally designed for language processing, have also been used for image recognition. These models have shown strong results in complex classification tasks. Other deep learning techniques, such as object detection and segmentation, have also contributed to better script recognition.

This study aims to improve Arabic script classification by using advanced methods like:

Feature extraction – Identifying key structural features of Arabic letters. Data augmentation – Expanding the training data to improve model performance. Transfer learning – Using pre-trained CNN models (like ResNet) and adapting them to Arabic script. These improvements will help classify both printed and handwritten Arabic scripts more accurately. This research will support the automation of script recognition and help preserve Arabic cultural heritage.

2.12 Adaptive Offline Arabic Handwriting Recognition with Deep Learning

Early handwriting recognition relied on Hidden Markov Models (HMMs) but struggled with long text sequences. Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) improved accuracy by remembering past characters. Bidirectional LSTM (BLSTM) with Connectionist Temporal Classification (CTC) further enhanced recognition without requiring precise segmentation. Recently, Convolutional Neural Networks (CNNs) have shown competitive results by recognizing letter shapes. Some studies integrated attention mechanisms for better accuracy. A key challenge is domain adaptation, as handwriting varies across writers. Prior studies explored feature mapping and style transfer, often focusing on individual characters. This research extends domain adaptation to full Arabic handwriting lines using CNN-BLSTM-CTC, evaluating both writer-dependent and writer-independent strategies to improve accuracy across diverse datasets.[\[6\]](#)

2.13 Arabic Handwritten Character Recognition Using Deep CNN

Recognizing handwritten Arabic letters from images automatically is a key challenge in many fields. Although optical character recognition (OCR) has improved a lot, Arabic characters still pose difficulties. Deep Neural Networks (DNNs) offer a powerful way to tackle these issues. In this work, we use a Convolutional Neural Network (CNN) with techniques like batch normalization to avoid overfitting. We tested our model on two datasets, AIA9k and AHCD, achieving high accuracy rates of 94.8 percent and 97.6 percent, respectively. We also analyzed its performance on EMNIST and a form-based AHCD dataset. Our results show that deep learning can effectively recognize handwritten Arabic characters.[\[7\]](#)

Chapter 3

Extra section

write here

Chapter 4

Methodology

4.1 Model

4.2 the Data Set

Handwritten Arabic character recognition systems face a lot of challenges. Because everyone's handwriting is different so there are large public datasets to help researchers understand it. The CNN was trained and tested by 16800 images that have different handwriting Arabic language .

The code uses CSV files that contain: • csvTrainImages: 13,440 training images (each 32x32 pixels)

- csvTrainLabel: Labels for these training images
- csvTestImages: 3,360 test images
- csvTestLabel: Labels for test images

Each image is a handwritten Arabic letter (28 letters in Arabic alphabet)

4.3 How is the data processed?

Having 'loadahcdcsv()' function that reads a CSV file, then reshapes the images to 32x32 pixels to 32x32x1 channel (grayscale) and then normalizes pixel values from 0-255 to 0-1 (by dividing by 255) to helps the model learn better, so it returns images and their labels.and then using some layers.

- Conv2D(32): First convolution layer with 32 filters (3x3 size) help to detects simple

features like edges and uses ReLU activation to add non-linearity (makes it able to learn complex patterns)

- MaxPooling2D to Reduces image size by taking maximum value in 2x2 areas and makes the network faster and helps detect features better

- Conv2D(64) (Second convolution with 64 filters) to detects more complex patterns

- MaxPooling2D this is another pooling layer

- Flatten layer to converts 2D features to 1D for the dense layers

- Dense(128 neurons) this is a fully connected layer with 128 neurons This is layer help in combine all features for the final decision

- Dropout(0.5) This is randomly turns off 50 percent of neurons during training to prevent overfitting (memorizing instead of learning)

- Dense(numClasses, softmax) this is the output layer with neurons = number of letters ,uses softmax to give probability for each letter

and then training model with 2 function `model.compile()` and `model.fit()` . `model.compile` this function takes `optimizer='adam'` and Adam is an algorithm that adjusts weights to minimize error , `loss='categorical_crossentropy'` and this is to know How wrong the model is (lower is better) and `metrics=['accuracy']` for tracking and `model.fit` this is function for trainig and takes `epochs=10` this is to Full passes through the training data, `batchsize=32` this is the number of samples processed before updating weights and validationdata that used to check performance on unseen data during training

4.4 Output

In terminal appear the training accuracy/loss and validation accuracy/loss and numbers will improve over epochs if learning is working for each epoch then after finishing the all the epochs then the accuracy percentage will appear

4.5 comparson

Chapter 5

Results

write here

Chapter 6

Conclusion

write here

Bibliography

- [1] S. H. Safiullah Faizullah Muhammad Sohaib Ayub and M. A. Khan, “A survey of ocr in arabic language: Applications, techniques, and challenges,” <https://www.mdpi.com/2076-3417/13/7/4584>,
- [2] S. Alghyaline*, “Arabic optical character recognition: A review,” https://www.researchgate.net/profile/Salah-Alghyaline/publication/366275602_Arabic_Optical_Character_Recognition_A_Review/links/639a4845e2Optical-Character-Recognition-A-Review.pdf, 2022.
- [3] I. J. Loay Alzubaidi, “Student outcomes assessments using deep learning,” https://www.researchgate.net/profile/Loay-Alzubaidi/publication/384232231_student_Outcomes_Assessments_Using_Deep_Learning/links/66ef312Outcomes-Assessments-Using-Deep-Learning.pdf, 2024.
- [4] H. A. A. M. B. Hossam Magdy Balaha, “Automatic recognition of handwritten arabic characters: A comprehensive review,” <https://link.springer.com/article/10.1007/s00521-020-05137-6>, 2020.
- [5] W. Albattah and S. Albahl, “Intelligent arabic handwriting recognition using different standalone and hybrid cnn architectures,” <https://www.mdpi.com/2076-3417/12/19/10155>, 2022.
- [6] S. A. Y. K. Sana Khamekhem Jemni, “Domain and writer adaptation of offline arabic handwriting recognition using deep neural networks,” <https://link.springer.com/article/10.1007/s00521-021-06520-7>, 2021.
- [7] K. S. Younis, “Arabic handwritten character recognition based on deep convolutional neural networks,” <https://www.mdpi.com/1424-8220/23/23/9475>, 2017.