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

3

List of Figures

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

Handwriting recognition has been a fundamental challenge in pattern recognition and artificial intelligence. Over the years, various approaches have been explored to improve accuracy and adaptability, especially for complex scripts like Arabic. Traditional methods, such as Hidden Markov Models (HMMs), struggled with long text sequences and handwriting variations. The emergence of deep learning, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), revolutionized the field by enabling automatic feature extraction and sequence modeling. This section explores key advancements in offline Arabic handwriting recognition, focusing on CNNs, RNNs, and domain adaptation techniques to enhance writer-independent recognition

2.0.1 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 they 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 like 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.

2.0.2   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]

2.0.3 Autonomous Traffic Sign Detection and Recognition Using Deep CNN

This paper talks about automatically detecting and recognizing traffic signs using Deep Learning (a type of artificial intelligence). This technology can help drivers by giving them important road information and can also be used in self-driving cars to improve safety.

The researchers developed a system called Autonomous Traffic and Road Sign (ATRS) Detection and Recognition System, which works in real time to detect and recognize traffic signs from images. They used a Deep Convolutional Neural Network (CNN), a powerful AI model for image processing, to achieve high accuracy in recognizing traffic signs.

One of the major contributions of this research is the creation of a new database of 24 different traffic signs from roads in Saudi Arabia, called SA-TRS-2018. This database contains 2,718 images captured from different angles and conditions to ensure the system works in various real-world situations.

The system was trained using CNN with different settings to achieve the best recognition accuracy. The results showed 100 percent accuracy, meaning it performed better than previous similar studies.[2]

2.0.4 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.[1]

2.0.5 Optimizing an Arabic OCR Model with Tesseract 5.0

Optical Character Recognition (OCR) is essential for converting printed or handwritten text into digital format, making it searchable and editable. While OCR has advanced significantly for many languages, Arabic poses unique challenges due to its cursive nature, contextual letter shaping, and diacritics. This research focuses on improving Arabic OCR by fine-tuning the Tesseract 5.0 engine. By collecting and training a specialized dataset, the study enhances recognition accuracy, measured through character error rate (CER) and word error rate (WER). The results show significant improvements in Arabic text recognition, helping in applications such as digitizing historical texts and automating text extraction. However, challenges remain, particularly in handling diacritics and complex ligatures, which limit the model's accuracy. The study contributes to advancing Arabic OCR and optimizing it for practical use in various field.[4]

2.0.6 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

2.0.7 A Survey of OCR in Arabic Language: Applications, Techniques, and Challenges

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).
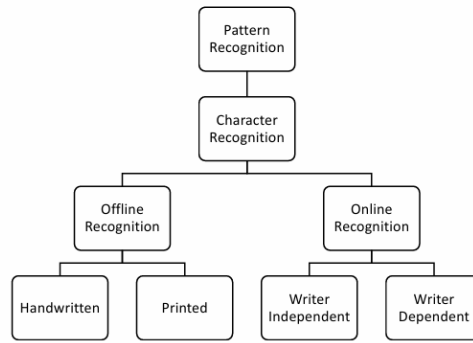

Figure 2.1: Types of OCR [2]

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

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.[2]

2.0.8 Recognition of Arabic Air-Written Letters: Machine Learning, Convolutional Neural Networks, and Optical Character Recognition (OCR) Techniques

This research focuses on air writing, where people write in the air using hand movements. While air-writing recognition is well-studied in English and Chinese, Arabic remains a challenge due to letter connections and handwriting variations. To address this, the study uses deep learning (VGG16, VGG19, SqueezeNet) and machine learning (NNs, RF, KNN, SVM) combined with Optical Character Recognition (OCR) to accurately recognize Arabic letters. The model, trained on the AHAWP dataset, achieved 88.8 percent accuracy with NN and VGG16. This research can enhance Arabic learning, assist people with disabilities, and improve human-computer interaction.[**6**]

2.0.9 Reviewing Arabic Optical Character Recognition

This study reviews 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.[3]

2.0.10 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).
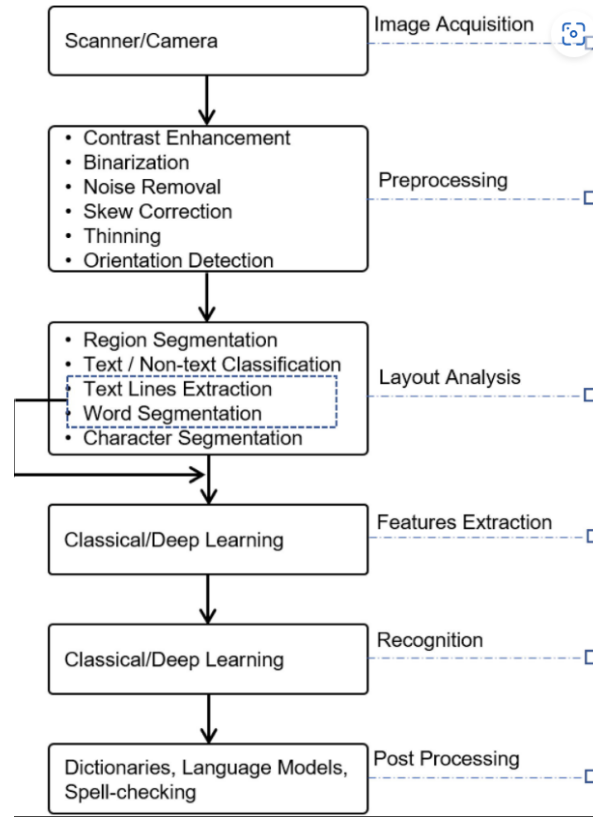

Figure 2.2: intermediate steps during OCR main stages.[8]

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]

2.0.11 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]

2.0.12   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]

2.0.13   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]

2.0.14 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]

2.0.15 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]

Chapter 3

Extra section

write here

Chapter 4

Methodology

Write here

Chapter 5

Results

write here

Chapter 6

Conclusion

write here

Bibliography

[1] I. J. Loay Alzubaidi, "Student outcomes assessments using deep learning," *https://www.researchgate.net/profile/Loay-Alzubaidi/publication/384232231$_student_Outcomes_Assessments_Using_Deep_Learning/links/66ef312$ Outcomes − Assessments − Using − Deep − Learning.pdf*, 2024.

[2] 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*,

[3] S. Alghyaline*, "Arabic optical character recognition: A review," *https://www.researchgate.net/profile/Salah-Alghyaline/publication/366275602$_Arabic_Optical_Character_Recognition_A_Review/links/639a4845e42$ Optical − Character − Recognition − A − Review.pdf*, 2022.

[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.