

Arabic Handwritten Recognition Using Deep Capsule Networks

Adwaa Bukhari

Artificial Intelligence, Midocean University

Project 1 : Paper Proposal

Abstract Recognizing Arabic handwritten characters is more complicated compared to English and other languages, due to its nature cursive, and connected letters. According to the recent researches with different approaches the accuracy and size of the dataset is very challenging. In this paper, I propose a new architecture for Arabic handwritten recognition based on Capsule Network (CapsNet), which proved to generate more than the expected accurate result in different languages. The proposed model is to apply CapsNet in Arabic handwritten recognition and trained the dataset to achieve up to 98% accuracy for unseen characters\words.

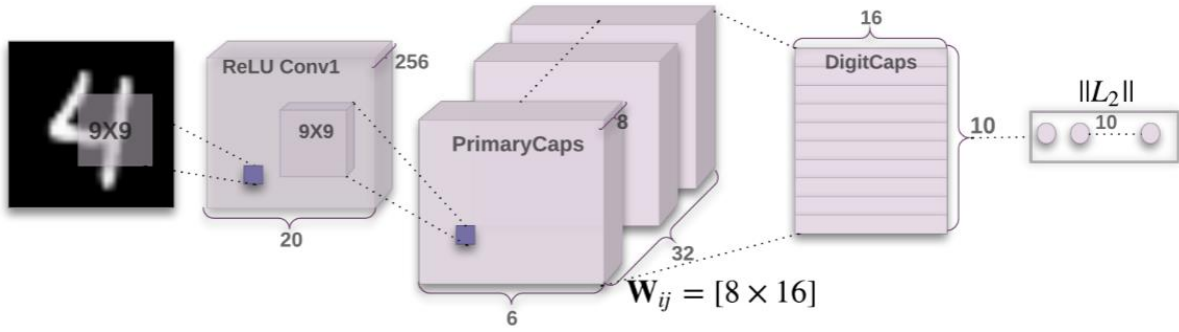
1. INTRODUCTION

Arabic language is the most used language but still a lot of technology everyday going have the weaknesses in Arabic. It is really challenging language based on different shapes of letters within its word depending on the position, as well as the dots which play important role in the letter. Also handwritten characters style is different from person to person. Applying a model to recognize all the differences is still challenging. Fewer efforts are seen using CapsNet for handwritten recognition. On the other hand, Convolution Neural Networks (CNN) happened to be very successful in computer vision application, but they are still weak in understanding the rotation or change of proportion of the input images and the high amount of dataset (Ghofrani & Toroghi, 2019 [1]). However, Capsule Networks has the advantages of predict the parameters of any object at any location with less data.

2. BACKGROUND

Convolutional Neural Networks (CNNs) had an important role in deep learning history and it is one of the models that has an excellent performance in handwritten recognition usage. CNNs are based on neurons layers that are capable of learning by using weights and biases, neurons in the layers are connected together and each layer are extracting the features by filtering the images and taking the maximum value in the pooling layer to extract features from data and sometimes this causing data loss. Capsule Network (CapsNet) is rectifying the problem by considering the feature as a vector and replacing the pooling layer to routing-by-agreement. Capsule Network composed of many layers, each one divided into a group of neurons called “capsules” (Sabour et al., n.d. [11]).

Figure [1] is a simple CapsNet Layers with Convolution layer, PrimaryCaps layer, and DigitCaps (Sabour et al., n.d. [11]).



3. PROBLEM STATEMENT

Manual work of documenting and digitalizing the paper work is consuming time and cost. Therefore, handwritten recognition system can be employed to read the extract the document data very efficiently however, the accuracy of the output still a biggest challenge specially in Arabic Language. Conventional Neural Network (CNN) is used commonly which has multiple connected layers to recognize pattern of the images due to the ability by using conventional filters and non-linearity units. According to the recent researches, the most gap and limitation of the model is the cursive nature of Arabic, and the limited dataset, which still might lead to manual work to clarify the output work, therefore, this needs to enhance the performance of the model to achieve the minimum percentage of errors.

4. PROJECT OBJECTIVE

To improve and analyze the accuracy of the system, we need to develop a more robustness Arabic handwriting recognition. Capsule Networks (CapsNets) have been introduced as an alternative to Convolutional Neural Networks (CNN). CapsNets are group of capsules composed of neurons representing the presence of an entity and the vector indicating the relationship between features by encoding instantiation parameters (Goceri, 2021 [2]). Unlike the CNN, CapsNets needs less data to train, and faster training time. This makes the accuracy of the network recognition higher.

5. RESEARCH SCOPE

This study will develop a new model of Arabic handwritten recognition using Capsule Networks. This research will involve collecting data to be trained and tested. The goal is to achieve the higher accuracy with unseen data and determine the effectiveness of Capsule Network.

6. LITRATURE REVIEW

Handwriting recognition has gained a lot of attention and many researchers have proposed different technologies and approaches. In this review section we present previous works on Arabic handwritten recognition methods.

Elsawy et al., n.d. [3] claim that the performance of Arabic recognition is affected by the limited dataset. He uses Convolution Neural Networks (CNN) which has the best performance in images and big data. A CNN consists of a number of layers (convolutional layers, pooling layers, fully connected layers). Dataset collected and made by 60 participants between 19 to 40 years. Each participants wrote each character ten times on different form. This dataset composed approximately 16K. The accuracy of the experimental result is 94.9%. Therefore, Ahmad et al., 2020 [4] deployed a couple of approach using Multi-Dimensional Long Short-Term Memory (MDLSTM) networks and Connectionist Temporal Classification (CTC) with KHATT dataset. MDLSTM is great with scanning the Arabic text in all directions and covering the dots, strokes, and diacritics. They implemented data augmentation methods in the pre-processing stage to enhance the performance of the MDLSTM such as blurring, contouring, edge enhancing, and bleeding through effect. The CTC layer is used as the final layer for prediction and alignment. The model achieved 80.02% recognition rate on KHATT dataset.

Alaasam et al., 2017 [5] claimed that the best approach to use with the limitation of the historical handwritten text that might have loosely layout format, poor quality and overlapping components in a historical text is Convolution Neural Networks (CNN). Since CNN required very large dataset for training, they performed two methods to extend the labeled dataset (automatic synthesis of Arabic sub-words, and generating printed of Arabic sub-words). They perform the experiment with three dataset (historical handwritten text image (HD), synthesized text image (SD), and printed text image (PD)). The best result of the three experiments was 85% accuracy of combined handwritten text image and printed text image dataset.

In this paper Elleuch et al., 2016 [6] proposed a new design based-Support Vector Machine (SVM) of Convolution Neural Network (CNN) classifier architecture with offline Arabic handwritten text recognition which CNN as an automatic feature extractor from raw images and it let SVM do classification by analyzing the error classification rate on the Arabic handwritten character. The target is to mix the CNN respective capacities and the SVM to obtain a new effectual recognition system inspired by the two formalisms and using HACDB database and IFN/ENIT database. The result of the experimental achieve a classification error rate of 6.95%. So, in order to improve this architecture, they added a dropout technique on the CNN based-SVM model to the only fully connected hidden layers, reduces the error to 5.83% with HACDB database.

Amrouch et al., 2018 [7] deployed system to integrate Convolution Neural Networks (CNN) and Hidden Markov Models HMM classifier, to use HMM model the dynamic of Arabic handwriting and CNN to extract salient features. The Convolution layers works as a feature extractor, and HMM as a recognizer. They use the IFN/ENIT database that consists of 946 handwritten Tunisian city names and their corresponding post codes, which achieve a rate of 88.95%.

El-Melegy et al., 2019 [8] proposed high performance of the Convolution Neural Networks (CNN) recognition system in the recognition of Arabic handwritten literal amounts that are used in bank cheques. They use AHDB standard dataset which consists 50 different Arabic literal amounts. They used 80% of the dataset for training and 20% for testing. The model achieved 96.58%, Therefore, they improved the model performance by applying the data augmentation that achieved 97.8% accuracy.

Yi et al., [9] experimented a Hidden Markov Model (HMM) using database IFN/ENIT with segmentation-free, due to the cursive handwritten and connected character in Arabic handwritten. So, they avoided the

segmentation stage and directly extracting the features from the word image using a sliding window then process it to HMM classifier that is mainly for automatic embedded training especially in non-segmented words. The result of the experiment leads to average recognition rate of 84.09%.

Altwaijry & Al-Turaiki, 2021 [10] presented a Convolution Neural Network (CNN) to recognize an Arabic letter written by children. They trained their model by Hijja dataset and AHCD dataset to compare the results. Hajja dataset is collected data from children between the age of 7-12. Children handwritten caused many challenging such as missing letters, wrong positioning of letters, and duplicated letters. They trained and compared both dataset with Convolution neural network. The result of the average accuracy of Hajja dataset is 88%, while AHCD achieved an average accuracy of 97%.

Butt et al., 2021 [12] present a Convolution neural network (CNN) and Recurrent neural network (RNN) approaches with an attention mechanism. they used CNN to take the input image and extract the features, then generate the output from an attention mechanism over the RNN. During the second stage after extracting the feature, they used bidirectional Long Short-Term Memory (LSTM) networks to solve the problem for vanishing gradient. They used two dataset ACTIV and ALIF, and the result achieved 87% accuracy.

The table [\[1\]](#) summarize the literature reviewed on Arabic handwriting recognition. From the literature review, I collected a several paper that uses Arabic handwritten in many different approaches such as Convolution neural network (CNN), Support Vector (SVM), and Hidden Markov Model (HMM). The Convolution neural network (CNN) hit very well in the handwritten and test recognition in general and achieved excellent result. It is very efficient to abstract features and identify object. However, (Ramesh et al., 2019) [13] proposed a Capsule Network approach for Kannada languages using manual collecting of dataset from 47 users with total 23500 images, out of 18800 images have been used for the training and the remaining for testing. Kannada language have a complex format of characters with different curve position for each character. So, the images go through a series of convolution layers that including the primary capsule layer. Each layers perform kernal convolution over the image with different kernal size to detect the basic features of the image. The result of using the Capsule Network for Kannada handwritten character obtained 98.7% accuracy.

CapsNet approach especially with complex language than English and with its curve complexity, that nearly same with Arabic complexity. This will give us the hope of increasing the accuracy of recognizing the Arabic handwritten characters by using Capsule Network and achieve an excellent result.

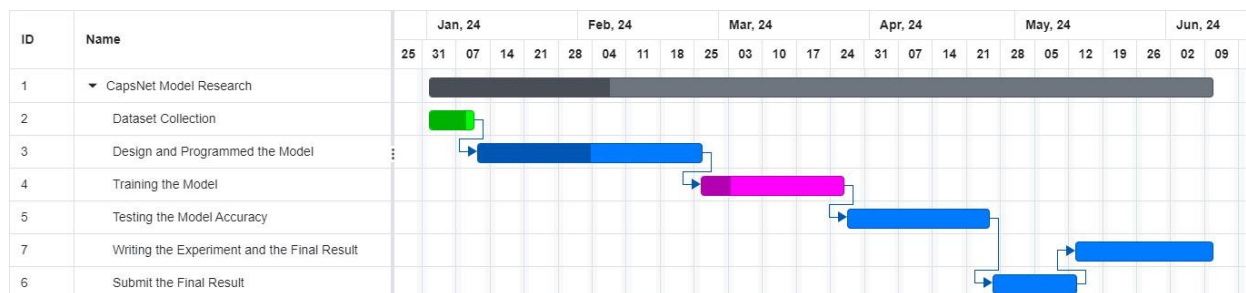
Paper	Method	Dataset	Result
(Elsawy et al., n.d. 2017) [3]	CNN	Collect and made by participant	Accuracy 94.9%
(Ahmad et al., 2020) [4]	MDLSTM + CTC	KHATT	Accuracy 80.02%
(Alaasam et al., 2017) [5]	CNN	HAHPT	Accuracy 85%

(Elleuch et al., 2016) [6]	CNN based-SVM with Dropout	HACDB + IFN/ENIT	Error rate 5.83% Error rate 7.05%
(Amrouch et al., 2018) [7]	CNN Based HMM	IFN/ENIT	Accuracy 88.95%
(El-Melegy et al., 2019) [8]	CNN	AHDB	Accuracy 97.8%
(Yi et al.,) [9]	Hidden Markov Model (HMM)	IFNIENIT	Accuracy 84.09%
(Altwaijry&Al-Turaiki, 2021) [10]	CNN	Hajja + AHCD	Accuracy 88% Accuracy 97%
(Butt et al., 2021) [12]	Attention-Based CNN-RNN	ALIF+ ACTIV	Accuracy 87%
(Ramesh et al., 2019) [13]	Capsule Network	47 participants 23500 images	Accuracy 98.7%

7. EXPECTED OUTCOME

The research is expected to deliver best performance and accurate result of average 98% in Arabic handwritten recognition, in addition to determine their suitability for recognizing characters with dots and the different shapes of letters depending on the handwriting style.

8. GANTT CHART



Task	Start Date	End Date	Duration
Research Implementation	Jan 01, 2024	Jun 10, 2024	116 days
Dataset Collection	Jan 01, 2024	Jan 10, 2024	8 days
Design and Programmed the Model	Jan 11, 2024	Feb 26, 2024	33 days
Training the Model	Feb 26, 2024	Mar 26, 2024	22 days
Testing the Model Accuracy	Mar 27, 2024	Apr 25, 2024	22 days
Submit the Final Result	Apr 26, 2024	May 13, 2024	12 days
Writing the Experiment and the Final Result	May 13, 2024	Jun 10, 2024	21 days

REFERENCES:

- Ahmad, R., Naz, S., Afzal, M., Rashid, S., Liwicki, M., & Dengel, A. (2020). A deep learning based arabic script recognition system: Benchmark on khat. *International Arab Journal of Information Technology*, 17(3), 299–305. <https://doi.org/10.34028/iajit/17/3/3>
- Alaasam, R., Kurar, B., Kassis, M., & El-Sana, J. (2017). *Experiment study on utilizing convolutional neural networks to recognize historical Arabic handwritten text*. <https://beratkurar.github.io/papers/Asar2017Experiment.pdf>
- Altwaijry, N., & Al-Turaiki, I. (2021). Arabic handwriting recognition system using convolutional neural network. *Neural Computing and Applications*, 33(7), 2249–2261. <https://doi.org/10.1007/s00521-020-05070-8>
- Amrouch, M., Rabi, M., & Es-Saady, Y. (2018). Convolutional feature learning and CNN based HMM for Arabic handwriting recognition. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10884 LNCS, 265–274. https://doi.org/10.1007/978-3-319-94211-7_29
- Butt, H., Raza, M. R., Ramzan, M. J., Ali, M. J., & Haris, M. (2021). Attention-Based CNN-RNN Arabic Text Recognition from Natural Scene Images. *Forecasting*, 3(3), 520–540. <https://doi.org/10.3390/forecast3030033>
- Elleuch, M., Maalej, R., & Kherallah, M. (2016). A New design based-SVM of the CNN classifier architecture with dropout for offline Arabic handwritten recognition. *Procedia Computer Science*, 80, 1712–1723. <https://doi.org/10.1016/j.procs.2016.05.512>
- El-Melegy, M., Abdelbaset, A., Abdel-Hakim, A., & El-Sayed, G. (2019). Recognition of Arabic Handwritten Literal Amounts Using Deep Convolutional Neural Networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11868 LNCS, 169–176. https://doi.org/10.1007/978-3-030-31321-0_15
- Elsawy, A., Loey, M., El-Bakry, H. M., El-Sawy, A., & El-Bakry, H. (n.d.). *Arabic Handwritten Characters Recognition using Convolutional Neural Network*. <https://www.researchgate.net/publication/313891953>
- Ghofrani, A., & Toroghi, R. M. (2019). *Capsule-Based Persian/Arabic Robust Handwritten Digit Recognition Using EM Routing*. <https://doi.org/10.1109/PRIA.2019.8785981>
- Goceri, E. (2021). *ANALYSIS OF CAPSULE NETWORKS FOR IMAGE CLASSIFICATION*.

Ramesh, G., Manoj Balaji, J., Sharma, G. N., & Champa, H. N. (2019). Recognition of off-line kannada handwritten characters by deep learning using capsule network. *International Journal of Engineering and Advanced Technology*, 8(6), 4767–4778.
<https://doi.org/10.35940/ijeat.F8726.088619>

Sabour, S., Frosst, N., & Hinton, G. E. (n.d.). *Dynamic Routing Between Capsules*.

Yi, Hang., Wen, Desheng., & Sandhu, P. S. (2010). *Proceedings, 2010 3rd IEEE International Conference on Computer Science and Information Technology : July 9-11, 2010 Chengdu, China*. IEEE.