Convolutional Neural Network Model for Arabic Handwritten Characters Recognition

Murtada Khalafallah Elbashir[1,2], Mohamed Elhafiz Mustafa[1,3]

College of Computer and Information Sciences, Jouf University, Sakaka, Kingdom of Saudi Arabia[1]

Faculty of Mathematical and Computer Sciences, University of Gezira, wad Madani, Sudan[2]

Faculty of Computer Science, Sudan University of Science and Technology, Khartoum, Sudan[3]

Abstract: In this paper, we presented a Convolutional Neural Network (CNN) model for off-line Arabic handwritten character recognition. The proposed CNN model used the dataset which prepared by Sudan University of Science and Technology- Arabic Language Technology group. The dataset is pre-processed before feeding it to the CNN model. In the pre-processing, all the characters images are size normalized to fit in a 20 by 20 pixel and then centred in a scaled images of size 28×28 pixel using the centre of mass then all the images are converted to be having a black background and white foreground colours. The pre-processed images are fed to the CNN model, which is constructed using the sequential model of the Keras library under tensorflow environment. The accuracy obtained varied from 93.5% as test accuracy to 97.5% as training accuracy showing better results than other methods that used the same dataset.

Keywords: Convolutional Neural Network (CNN), Handwritten Characters, Keras, pre-processing

I. INTRODUCTION

There are many scripts and old Arabic manuscripts that are written with the traditional methods such as pen and paper, which makes it difficult to store and retrieve them in an efficient manners. This leads to the loss of a lot of knowledge and historical information due to the fact that these manuscripts are not stored in digital format. Thus Optical Character Recognition (OCR), which is the method that can be used to convert typed, handwritten or printed text into machine-encoded text become very important in the recent decades. OCR can minimizes the time needed to extract the important information from a file since it eases the conversion from a hard copy document to digital files. While it is easy to recognize the printed characters, recognizing handwritten characters is a challenging task due to the fact that the handwritten characters are written by different writers each with his own writing style. Moreover, writers do not write in same way always as in printed text.

234 million people speak Arabic and its characters are used in writing many other languages such as Persian and Urdu [1]. The Arabic language consists of 28 basic letters and many letters look alike but are differentiated from one another by dots. These dots can be under or above the letter and they vary in the number. The number of the dots can be one, two, or three. In addition to the 28 basic letters there are 6 additional isolated letters that are not counted as Arabic alphabet but they have very important usage in Arabic grammar, lexicon, and spelling. The existence of dots and other strokes that differentiate similar characters make Arabic character recognition even more challenging.

The methods that are applied in character recognition use classifiers such as Artificial Neural Network (ANN), Support vector Machines (SVM), Hidden Markov Model (HMM), k-Nearest Neighbours (k-NN) etc. These classifiers use the features of the characters as an input, but before the features of the characters can be fed to the classifiers some pre-processing steps need to be done. These pre-processing steps use many techniques to enhance the character's image. The pre-processing techniques involve normalization, skew correction, noise removal, thinning and slant. These pre-processing techniques are applied in order to remove any element in the character's image that may affect the recognition process in a negative way. The state of the art methods that are used for Arabic character recognition use the deep learning approaches such as Convolutional Neural Networks (CNN) [2, 5], Deep Bidirectional Long Short-Term Memory (DBLSTM) [3], Deep Belief Network (DBN) [4], Deep Support Vector Machine [6], The deep learning architectures show great performance for Chinese and Latin handwritten character recognition and also can perform very well in recognizing Arabic characters.

In this paper, we constructed a CNN model using the sequential model of the Keras library under tensorflow environment for off-line Arabic handwritten character recognition. The constructed CNN model used the dataset which prepared by Sudan University of Science and Technology- Arabic Language Technology group. All the characters images are size normalized to fit in a 20 by 20 pixel and then centred in a scaled images of size 28×28 pixel using the centred of mass before feeding it to the CNN model.

II.   MATERIAL AND METHODS

A. Data Set: SUST ALT (Sudan University of Science and Technology- Arabic Language Technology group) has built many datasets to be used in studying, investigating, training, and testing Arabic recognizers [7, 8]. Isolated Arabic letters is one of these data set. Even though Arabic letters are 28, the data set contains 34 classes. This is one of the main important features of this dataset. Since almost all other data set neglect this important component of Arabic handwriting. These additional 6 classes are not considered Arabic letters (the first 6 columns from the left in fig 1 and 2). However, five of them represent variations for the letter Alif "أ" and the sixth (لا) is a combination for two letters " ل" and "ا " it is also important word in Arabic which means 'no'.

It is important to mention here that all Arabic letter are consonants. In Arabic there are three vowel sounds. These vowel sounds considered just variation that occur for the consonant letters. Normally these vowels do not appear in text (printed or handwritten) whoever when needed appear as diacritics connected to one consonant letters. Fig.1 shows one filled table. All subjects (writers) are students or lecturers in SUST. Using this dataset Balola et al, has designed and tested a multi-layer perception for the isolated letters classification [9].   The main result of his work shows that the feature that causes the main challenge for Arabic letters classification is the usage of dots to differentiate similar letters like dal and thal ("ذ" & "د").

Fig. 1: sample for filled table with handwritten characters

B.   Handwritten characters pre-processing: Our handwritten character are obtained using scanner and then the letters are cut from the scanned paper so most of the characters are not in the centre. The originally obtained character are depicted in the Fig. 2.
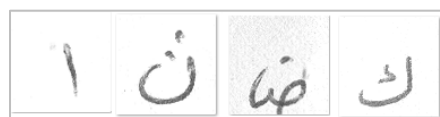

Fig. 2: the originally obtained Characters

We tried to use these characters as the input for Convolutional Neural network and we get random accuracy and it does not go beyond 10%. In the pre-processing we started by size normalizing the images to fit in a 20 by 20 pixel and then cantering them in an scaled images of size 28 by 28 pixel images using the centre of mass. This is done after removing every row and column at the side of the 20 by 20 images that are completely black . Then all the images are converted to be having a black Background and white foreground colours. Fig. 3 shows the characters after the pre-processing step.

Fig. 3 The handwritten characters after the pre-processing step

C.   Deep Neural Networks: In the area of machine learning and computer Vision, Deep Neural Network (DNN) is considered to be an active area [10]. There are three well known architectures for DNN, these architectures are: Convolutional Neural Network (CNN), Deep Belief Net (DBN), and Stacked Auto Encoder (SAE) [11, 12]. Deep Neural net get it is name from the fact that beside its input and output layers it has many intermediate or hidden layers in oppose to shallow Neural network, which has few number of hidden layers. Because of its composition of many layers, DNN can handle the highly varying nonlinear function compared to the shallow net [13]. In visual pattern

recognition the DDN several layers can learn to recognize specific feature of the pattern, for example the first layer's neuron can recognize edges, and the other more complex shapes can be recognized by the neurons in the middle layers, so the multiple layers become an advantages for the DNN in learning complex patterns and this make them intrinsically more powerful than shallow networks. CNN are considered to be very popular models among the deep learning methods and it provides outstanding and state-of-the art performance in image recognition. They were inspired by the biological processes, where the connection between the different neurons mimic the animal visual cortex organization [14]. CNN is very good in classifying images because they use a special architecture, which makes them fast to train. The fundamental components of the CNN are: a stack of convolution layers and max pooling layers followed by fully connected and softmax layer at the end layer [15].

D. The Convolutional Neural Network (CNN): The CNN model that we created has the following architectural components: INPUT, CONV, RELU, POOL, and the fully connected or dense layers. The model is constructed using the sequential model of the Keras library under tensorflow environment using python programming language. The input layer receives an images of size (28, 28, 1), which means that the image is gray scaled and has height and width of 28 pixel. Thus the input layer will hold the raw pixel data of the character image. The convolutional layer can perform a set of operations that result in a single value. We started the width of the convolutional layer by 32 and applied RELU as an activation function for introducing nonlinearities into the model. We used the pooling layer to down sample the 28 by 28 image data to reduce the features' dimension so as to increase the performance of our model. Max pooling algorithm is used in the pooling layer to extract a 2x2-pixel tiles from the feature map and return the maximum values while discarding other values. The task of the fully connected layer or the dense layer of our model is to perform the actual classification using the features that are extracted by the convolutional layer and down sampled by the max pooling algorithm. The number of the classes that we have according to our data set is 34 classes, which means that we have 34 output neurons, each represent one Alphabet characters from the dataset. The architecture of our CNN is depicted in Fig. 4.

A 0.2 dropout ratio is used for regularizing the two dense layers, which means that the output of each hidden neuron that is less than 0.2 should be set to zero. Setting of neuron output to zero means that this neuron will not contribute to the forward pass thus it will not participate in backpropagation to avoid overfitting.
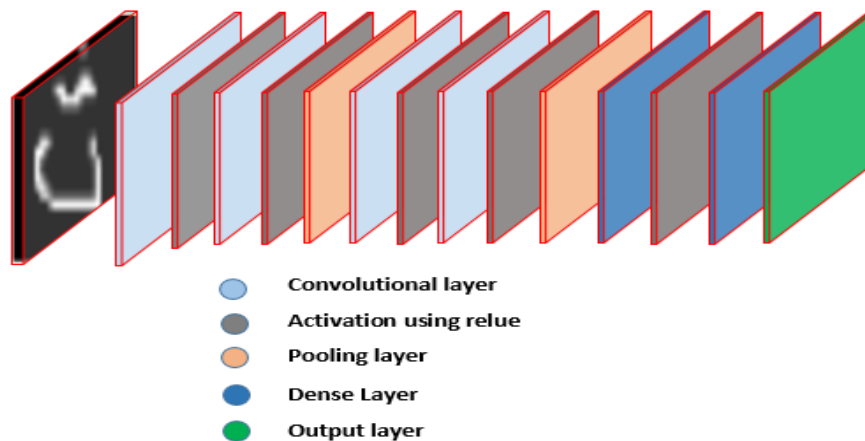

● Convolutional layer
● Activation using relue
● Pooling layer
● Dense Layer
● Output layer

Fig. 4: The architecture of the used convolutional neural network

III. RESULTS AND DISCUSSION

We used keras library under tensorflow environment to construct our CNN model, the whole dataset was divided randomly into groups. These groups are training, validation and testing. The training data represents 70% of the total data while both the testing and validation data represents 30%. The accuracy of prediction is used as performance measure. The test, train, and validation accuracies obtained are 93.5, 97.5, and 93.56 respectively. Fig. 5 shows the accuracy curve of the training and validation process.

The literature survey reveals that there are two papers which describe recognition models using the same data set [9, 16]. The first paper contains one main classifier that classifies characters into groups. Each group contains similar characters. The second phase contains set of classifiers each classifier specialized in classifying one group characters. The accuracy of this paper classifier is 78.8 [9]. The second paper uses the Adaptive Neural Network Fuzzy Inference System (ANFIS). This paper uses the same strategy used by the previous paper in dividing the classification into two

phases. Moreover, the paper reported the classification accuracy of the method when used in one phase which is 91.39 [16]. Although these results are limited however, it indicates that the proposed method result is very good.
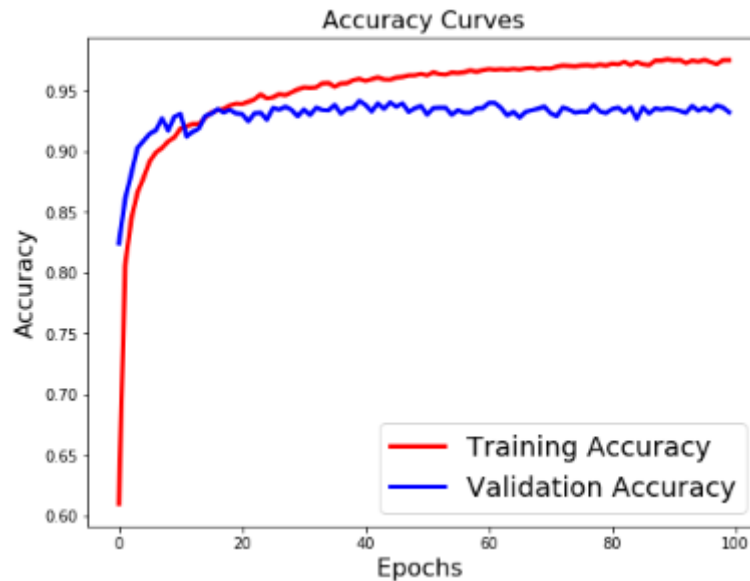

Fig 5: The accuracy curve

From the results we can see that the CNN model that we developed performed very well and it is found that the miss classified characters resemble very well the one that the CNN predicted i.e the character DAD (ض) is predicted as SAD(ص). And sometime it is very difficult to differentiate between these two characters even for human. When we investigated the most important errors to obtain the similarity between the probabilities of real characters and the predicted ones we found that most of these misclassifications occurs in the similar or badly written characters. Fig. 6 shows the predicted label for some of the misclassified characters and the actual or true character.

Hand written character										
Actual	ت	ر	د	و	ن	ض	لا	ت	ا	ز
Predicted	ن	د	ر	هـ	ت	ص	ي	ق	أ	ذ

Fig. 6: Example of the misclassified characters

CONCLUSIONS

In this paper, we constructed a CNN model for Arabic handwritten character recognition. The dataset which prepared by Sudan University of Science and Technology- Arabic Language Technology group is used to evaluate or CNN model. All the characters in this dataset are pre-processed by size normalizing them to fit in a 20 by 20 pixel and then cantering them in a scaled images of size 28×28 pixel using the centre of mass. All the character images are converted to be having a black Background and white foreground colours before feeding them to the CNN model. The test accuracy obtained is 93.5%, which is the best recognition results using Sudan University of Science and Technology-Arabic Language Technology group's dataset.

REFERENCES

[1]. L.M. Lorigo, V. Govindaraju, "Off-Line Arabic Handwriting Recognition: A Survey", IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 28, no. 5, pp. 712-724, May 2006. Amrouch, M.; Rabi, M. The Relevant CNNs Features Based HMM for Arabic Handwriting Recognition. Int.J. Imaging Robot. 2017, 17, 98–109.
[2]. Maalej, R.; Tagougui, N.; Kherallah, M. "Online Arabic handwriting recognition with dropout applied in deep recurrent neural networks". In Proceedings of the 2016 IEEE 12th IAPRWorkshop on Document Analysis Systems (DAS), Santorini, Greece, 11–14 April 2016; pp. 417–421.
[3]. Elleuch, M.; Tagougui, N.; Kherallah, M. "Towards Unsupervised Learning for Arabic Handwritten Recognition Using Deep Architectures". In Proceedings of the International Conference on Neural Information Processing, Istanbul, Turkey, 9–12 November 2015; pp. 363–372.
[4]. Elleuch, M.; Maalej, R.; Kherallah, M. A New Design Based-SVM of the CNN Classifier Architecture with Dropout for Offline Arabic Handwritten Recognition. Procedia Comput. Sci. 2016, 80, 1712–1723. [CrossRef]
[5]. Elleuch, M.; Kherallah, M. An Improved Arabic Handwritten Recognition System using Deep Support Vector Machines. Int. J. Multimed. Data Eng. Manag. 2016, 7, 1–20. [CrossRef]

[6]. Musa, M. E. M. "Towards building Standard Dataset for Arabic Recognition", International Journal of Engineering and advanced Research Technology and Research (IJEART),2016, 2(2),16-18.

[7]. Musa, M. E. M., "Arabic Handwritten Datasets for Pattern Recognition and Machine Learning", Application of Information and Communication Technologies (AICT 2011) – Baku, Azerbaijan.

[8]. Ali, O. B., Shaoat, A., Musa, M. E. M., (2015) "Two Stage Classifier for Handwritten Character Recognition", International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 2015, Volume 4, Issue 12, pages 646 – 650.

[9]. A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: a brief review," Computational Intelligence and Neuroscience, vol. 2018, Article ID 7068349, 13 pages, 2018.

[10]. G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," Neural Computation, vol. 18, no. 7, pp. 1527–1554, 2006.

[11]. P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol,"Extracting and composing robust features with denoising autoencoders," in Proceedings of the 25th International Conference on Machine Learning, pp. 1096–1103, ACM, Helsinki, Finland, July 2008.

[12]. Y. Bengio et al., "Learning deep architectures for AI," Foundations and Trends in Machine Learning, vol. 2, no. 1, pp. 1–127, 2009.

[13]. Matusugu, Masakazu; Katsuhiko Mori; Yusuke Mitari; Yuji Kaneda ,"Subject independent facial expression recognition with robust face detection using a convolutional neural network" . Neural Networks., 2003, 16 (5): 555–559. doi:10.1016/S0893-6080(03)00115-1.

[14]. Md Zahangir Alom et al , Handwritten Bangla Character Recognition Using the State-of-the-Art Deep Convolutional Neural Networks, Computational Intelligence and Neuroscience, volume 18, 2018

[15]. Balola O. A. and Shaout A. "Habird Arabic Handwritten Character Recognition Using PCA and ANFIS" , International Arab Conference on Information Technology (ACIT), '2016