Writer Identification via CNN Features and SVM

Mahmoud rahhal

Aleepo, Syria, Computer Engineering Department, College of Electrical and Electronic Engineering, University of Aleppo

---***---

Abstract – *Writer recognition is the process of identifying the author of a document based on his or her handwriting. Recent advances in computational engineering, artificial intelligence, data mining, image processing, pattern recognition, and machine learning have demonstrated that it is possible to automate writer identification. This paper proposes a writer recognition model based on Arabic handwriting.*

This research presents the study and implementation of the stages of writer identification, starting from data acquisition by scanned images of handwritten text, and then augmented the data through study of Maaz et al. [1] that generate a large number of texts from the set of texts available within the database, and then building a convolutional Neural Network (CNN) Which is usually useful for extracting features information and then classification the data, but in this research it is used for feature extraction, finally support vector machine is used for classification.

The experiments in this study were conducted on images of Arabic handwritten documents from ICFHR2012 dataset of 202 writer, and each writer have 3 text. The proposed method achieved a classification accuracy of 97.20%.

Key Words: Arabic handwriting, data augmentation, writer identification, deep learning, convolutional Neural Network, support vector machine.

1. INTRODUCTION

Identification of persons is mainly through the physiological characteristics like fingerprints, face, iris, retina, and hand geometry and the behavioral characteristics like a voice, signature, and handwriting. Writer identification is a behavioral approach of handwriting recognition, which offers with matching unknown handwriting with a database of samples with recognized authorship.

Writer identification has been a lively discipline of lookup over the previous few years and it used in many applications as in biometrics, forensics and historic report analysis.

Identifying people involve a fields of Artificial intelligence (AI), image processing and pattern recognition, and it contributes in many applications: biometrics, forensics, security, and legal matters, and financial field etc. [2]. All methods of writer identification and verification can be classified into two categories [3]: text-dependent method and text-independent method. Text-dependent methods are methods where the text in test samples must be the same as in training samples. Scale to extract features from text image. Text-independent methods are methods in which the text in test samples does not have to be the same in training samples and in which structural methods for extracting features from a character shape are involved.

Writer identification is classified into on-line writer and off-line writer identification [3], in the case of on-line writer identification, features are collected directly from signals sent from digital devices while writing, and in the off-line case, and attributes are collected from written text Handwriting obtained from the scanned image. . It is considered as more complex than on-line case due to many dynamic features of handwriting are missing, for instance pen-pressure, order of strokes, and writing quickness [4].

The task of classification in pattern recognition is to identify the pattern to a class of the known group of classes, in which features will be extracted from images of scanned texts handwritten and trained it by using a classification algorithm. This study proposed a system for identifying people through Arabic handwritten text by extracting features by deep learning, which required data augmentation. Because of limited number of handwritten text samples, it is impossible to gather a wide number of samples from different writers. The text is divided into parts and then Gathering all text structures randomly to obtain several lines, which enables us to generate many texts from them, and then using a convolutional neural network [1] for feature extraction , finally using support vector machine (svm) for classification which proves its ability to achieve high accuracy in writer identification. This research includes a number of previous studies on handwriting features extraction, person identification techniques, and some data augmentation techniques.

2. RELATED WORKS

In this paragraph, the studies will be discussed that have been done on writer identification and the techniques used so far, with an overview of writer identification systems in several languages, such as Chinese, English, Arabic ... etc.

Researcher Bangy Li et al in 2009 proposed a method based on combining the static feature and dynamic feature, and then the Nearest Neighbor Classifier was used as a classification methodology [5]. In 2009, researcher Zhu et al presented a study in this field using the Shape Codebook and Multi Class SVM Classifier, and their proposed method was evaluated by collecting images of texts written in eight languages (Arabic, Chinese, English, Hindi, Japanese, Korean, Russian and Thai). The Chinese identification rate was much lower (55.1%) compared to the eight other languages [6]. In 2013, researcher Djeddi and his colleagues proposed a way to identify the identity of the text writer based on a set of features extracted through the Gray Level Run Length (GLRL) matrices method and comparing them with the latest known features at the time, and good results were shown in identifying the writer's identity, and K-Nearest was also used. Neighbors and Support Vector Machines and their experiments were conducted on handwritten Greek and English texts of 126 writers and each writer 4 samples, achieving an interesting performance in author identification and verification in a multitext environment [7]. A study presented by the researcher Alaei et al in 2014 proposing a method for identifying the writer based on Histogram, using two different sets of handwritten data written in Canadian and English languages. Their experiments were conducted on 228 Canadian-language texts for 57 writers with 92.79% accuracy for F-measure and 330 texts in English for 55 writers with 26.67% accuracy for F-measure [8]. Then dral et al conducted a 2015 study of person-identification using images of Tamil handwriting of 300 writers and each writer 100 samples using a discriminant model to categorize with features aggregated from handwriting and use it in the Support Vector Machine, using three kernel function: Linear The RBF and Polynomial and together with Polynomial and RBF gave the highest prediction accuracy [9]. In 2015, researcher Fiel and Sablatnig [10] presented a study on the databases ICDAR2013, ICDAR2011 and CVL, in which they propose to use the Convolutional Neural Network (CNN) to generate a ray of traits for each writer that is compared with the rays of traits stored in the database. The method gave 98.6% on ICDAR-2011, 97.6% on CVL and 40.5% on ICDAR-2013 might be because of missing Greek training data and wrong segmentation.

The Arabic writer identification was not covered as widely as English or Chinese writer identification until the last few years, and the first study goes back to the proposal of the researcher AL-Zoubeidy et al in 2005 using multi -channel Gabor filtering and gray scale co-occurrence matrices to distinguish writer from his writing style [11]. Fouad and Volke in 2014 proposed a system for Arabic writer identifition using only 21 feature, and Gaussian Mixture Models (GMMs) were also used as the kernel of this system, where GMMs provide a strong representation of the distribution of the extracted features using a fixed-length sliding window It is constructed and trained using images of words and texts for writer [12]. Djeddi et al in 2014 conducted their study on handwritten texts of 1,000 writers using three methods to extract features: run-length distribution, edge-hinge distribution, and edge-direction features. They used Multiclass SVM (Support Vector Machine), and then had an identification rate of 84.10% [13]. In 2016, researcher Elleuch and his colleagues presented a study based on Convolutional Neural Network (CNN) and SVM (Support Vector Machine) for handwriting recognition. The performance was evaluated with HACDB and IFN/ENIT database for character images. The results showed the effectiveness of the proposed method compared to CNN Standard Classifier [14]. In 2018, the researcher Amar and his colleagues proposed a deep learning methodology for slicing and recognizing printed and handwritten characters from words. An algorithm was proposed to slice the word into letters, and then extract the features using a wavelet transform, and these extracted features are exploited as connection weights to build a convolutional neural network for each letter shape, The proposed method was tested on APTI and IESKarDB database, and their results showed the effectiveness and speed of the proposed method for each of the databases [15]. Dengel et al presented the Arabic text recognition system using the (KHATT) database. This study mainly contributes to three aspects: preprocessing, deep learning, and data augmentation by creating copies of the data. The first version is the blurred version of the data, the second version defines the contours, the third version is enhanced edges, and the fourth version is making effects and where Data augmentation with a deep learning approach has been shown to produce a better and more promising improvement in outcomes. This study achieved a recognition accuracy of 80.02%, which is better compared to other studies conducted on the same database [16]. In 2019, Aref and colleagues presented a study based on integrating OBI features with the character codebook using IAM databases for English texts and ICFHR-2012 for Arabic texts. The classification accuracy with the KNN classifier was 96% [17]. In 2019, Rehman and his colleagues applied their study to identify the writer's identity on the QUWI database that contains Arabic texts and English texts for 1017 writers, and each writer has 4 texts, using techniques to increase the data to improve performance, by obtaining from each original image an image containing the boundaries surrounding the words (Contours), as well as obtaining new samples from the original images by increasing their sharpness and the image unit obtained in the first stage of data augmentation when obtaining the bounding boundaries of words (Sharped Contours), and finally applying Negatives to the original images and images resulting from the previous steps to augment the data, And using the AlexNet structure to extract attributes to then apply the Support Vector Machine as a classification method, the accuracy was 92.78% in English texts, 92.20% in Arabic and 88.11% when combining Arabic and English, respectively [18].

3. PROPOSED SYSTEM

In this paragraph, the Arabic writer identification is presented. First, Image Acquisition of handwritten texts in Arabic, Secondly, the data-augmentation technique for increasing the input data images, third, finally, training process and the experimental are presented.

3.1. Image Acquisition

The study was conducted on an ICFHR2012 database downloaded from Kaggle for 202 writers, each writer had three different handwritten texts in Arabic, and the images were at a resolution of 600 DPP binary [19].

3.2. Pre-processing

The database did not need to be pre-processed so it was binary images with white background and text in black, this data just need to be increased.

3.3. Data Augmentation

Handwritten documents may be limited, which prompted a number of researchers to increase data techniques, and it was necessary to train the neural network from the presence of a large number of samples. This prompted us to use an algorithm that seeks to increase the number of images for handwritten texts in Arabic [1], especially that the database The data on which the study was conducted (ICFHR2012 dataset) contains only three texts for each writer, and this number is not enough to train the proposed neural network in the research, in addition to the images of large and different sizes, which does not fit with the nature of the proposed network and to avoid the problem of out of memory. To solve this problem, huge local image patches centered on the contour of the handwriting images [20] or inside the slipping windows on text lines [21] are extracted by past approaches, which only consider the local structural information of handwriting.

After reviewing some research in this field, the method of researcher Maaz et al. was applied to increase the data[1], and through the following stages as shown in" Figure1":

1) Consequently, the proposed method for dividing the text into words or word structures was adopted to avoid the problems of using the techniques for dividing the text into lines and then into words, in proportion to the database through the following stages:

   a) Read original pictures and find their Negative and then determining the connected components (CCs: connected-components) from the binary image.

   b) Applying the (LoG: Laplacian of Gaussian) filter to the matrix of connected components, which resulting from phase (a).

   c) Filling the filtered image based on morphological reconstruction.

   d) Determining the connected components (CCs: connected-components) from the output image of phase (c). Then, determining the connected components, which resulting on the original.

   e) Deleting the connected components that are within larger connected components and make the interconnected components a single component that includes them.

2) Collecting all the words or word structures randomly, which resulting from the previous stage in one line and repeat this step to get several different lines.

3) Forming an image of a five-line text generated from stage (c) in the second stage.

4) Cropping each image generated from the third stage into a number of images with a size of 400 * R, where (R) represents the height of the image generated from the third stage, as shown in "Figure 10", and then the size of all images was standardized when entered into the deep learning network .

Hundreds of pictures of handwriting texts for each writer are got from this stage; the first 200 handwriting images resulting from Data Augmentation stage for each writer were taken to obtain a database of 40400 handwriting scripts for 202 writers and for each writer 200 texts.

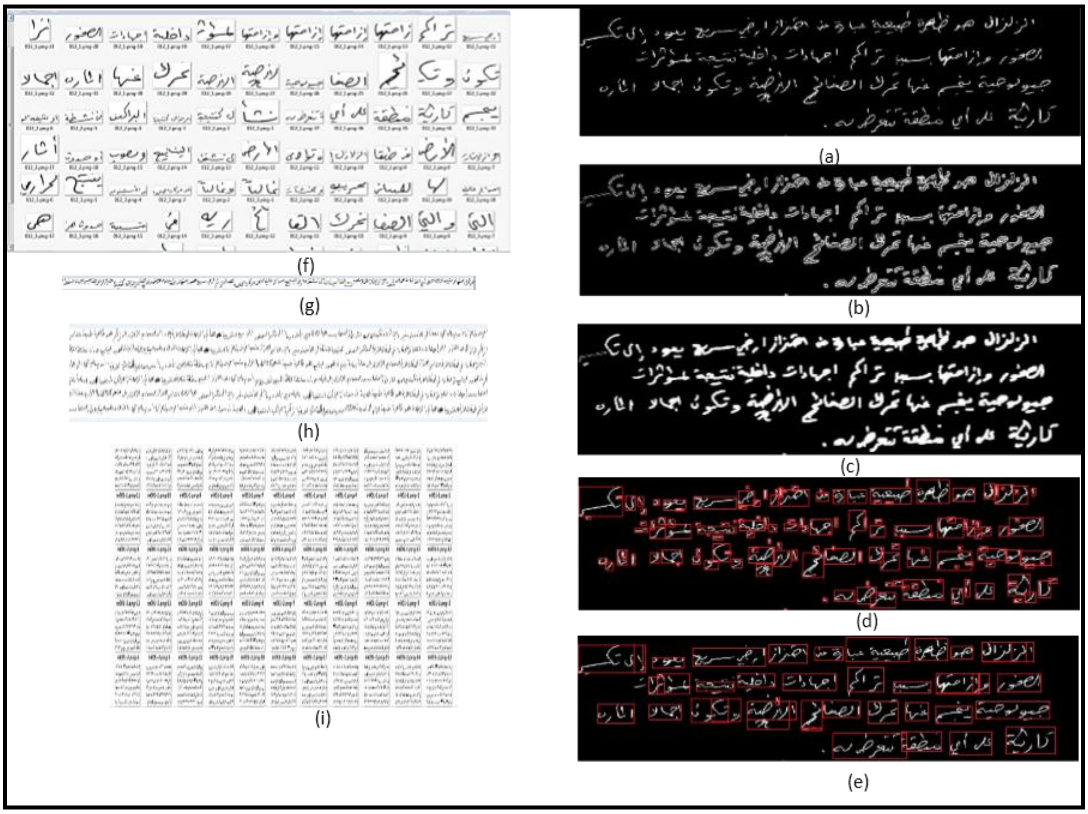

Fig -1: Data augmentation process: (a) Negative original image. (b) Image after filtering. (c) Filling image regions and holes after filtering. (d) Determining connected components on image shown in (a). (e) Delete connected components that are among other components. (f) Sample of word pictures and word structures for a writer. (g) Sample of generated text line image. (h) : sample Image of a 5-line generated text. (i) Sample of the text images after cropping.

3.4 Convolutional neural network

After increase the data, a wide number of images are generated from the training dataset, grounded on which, the coming step is to extract some distinct features to represent their characteristic. As done by Christlein et al and Fiel et al. [20, 21], the Convolutional neural network is employed to reckon the features in this work.

Convolutional neural network's architecture has several famous designs, including: LeNet, AlexNet [22], GoogLeNet [23], and VGGNet [24].

Three types of A Convolutional Neural Network (CNN): Convolutional Layer (CONV), Pooling Layer (POOL), and Fully-Connected Layer (FC).

3.4.1 Convolutional Layer (CONV)

This layer is the most significant in any CNN structure due to the fact filters are utilized to learn the features from the enter image. The 2D function map consists of the fold product between the filter values and the scanned area with the corresponding filter dimension of the enter image, i.e. every neuron of this layer will join to a precise enter location according to for filter size.

In a CONV layer, the input volume is represented as [D1x M1x P1] corresponding to the dimensions of the input image. Four hyperparameters are represented as (N, L, Z, W) corresponding to the number of pollutants, the size of the filter, the stride and the quantum of zero padding. The output volume is represented as (W2 x H2 x D2) corresponding to:

$$D2= (D1 - L + 2W) / Z + 1 \quad (3.4.\ 1.1)$$

$$M2= (M1 - L + 2W) / Z + 1 \quad (3.4.\ 1.2)$$

$$P2 = N \quad (3.4.\ 1.3)$$

3.4.2 Pooling Layer (POOL)

It is an intermediate layer in the network that reduces or compresses the spatial dimensions of the incoming input to it, and this reduces the computations within the network.

Pooling layers decrease the dimensions of data by combining the outputs of neuron clusters at one layer into a single neuron in the subsequent layer. Pooling acts on all the neurons of the feature map, by combining small clusters, tiling sizes such as 2 x 2 are frequently used. There are two common types of pooling in famous use: max and average. Max pooling uses the maximum value of each local cluster of neurons in the feature map, while average pooling takes the average value.

In a POOL layer, the input volume is represented as [D2x M2x P2] corresponding to the dimensions of the input volume, two hyperparameters are represented as [L1, Z1] corresponding to the receptive field or size of the filter and the stride, and the output volume is represented as [D3 x M3 x P3] corresponding to:

$$D3= (D2 - L) / Z + 1 \qquad (3.4.2.1)$$

$$M3= (M2 - L) / Z + 1 \qquad (3.4.2.2)$$

$$P3 = P2 \qquad (3.4.2.3)$$

3.3.3 Fully-Connected Layer (FC)

The FC layer in a CNN like in ANNs has neurons that are connected every neurons in one layer to every neuron in the past layer. This layer is often kept as the final layer of a CNN with “SOFTMAX” as its activation function for multi-class classification problems. The FC layer is responsible to predict the final class. Thus, it has an output dimension of [1x1xD] where D represents the number of classes or labels considered for classification [25].

The network architecture that was used in this research consists of 4 convolutional layers, 4 pooling layers, and 2 Fully Connected Layer (FC). [1] Each convolutional layer is followed by a Batch Normalization Layer (BN) and Relu layer. This network also includes a Softmax layer, which is used for Classification problems.

The description of this architecture convolutional network layers is as follows and illustrated according to "Figure 2":

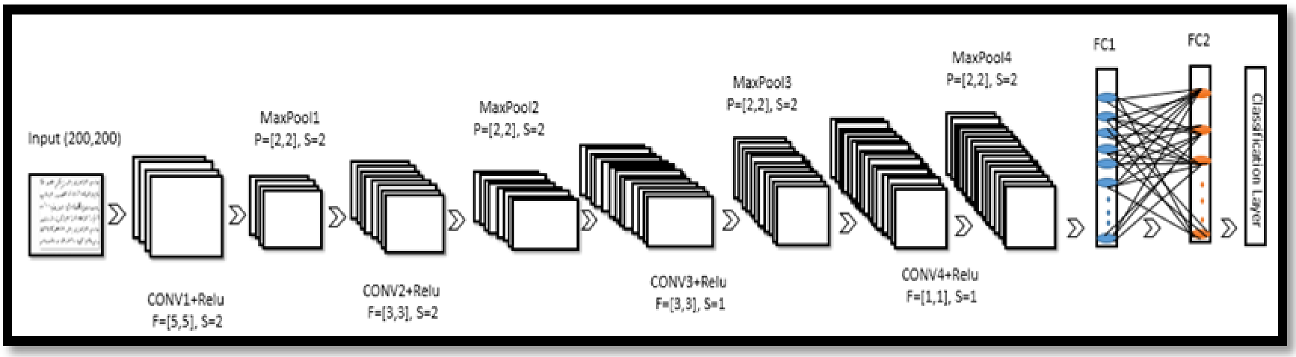

Fig -2: The structure of the convolutional network [1]

1. The input layer receives training samples, which is binary images and size [400,400].

2. Convolutional Layer (CONV1): In this layer, 96 filters of size F = [5, 5] were used, and stride (S = 2) without applying zero padding.

3. Batch Normalization Layer: It normalizes the input via mini batch, speeds up the training of convolutional neural networks, and reduces the sensitivity of network initialization, so batch normalization layers are used between convolutional layers and nonlinear layers, such as ReLU layers.

   The layer first normalizes the activations for its input by subtracting the mean of the mini batch and dividing by the standard deviation of the mini batch. [26]:

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{{\sigma_B}^2 - \epsilon}} \qquad (3.4.1)$$

4. Relu Layer: This layer applies a threshold to all its income items from the previous layer according to the function shown below:

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x \leq 0 \end{cases} \qquad \text{.... (3.4.2)}$$

5. The pooling layer: (Pool1) In this layer Max Pooling has been used in the maximum pool and the size of the area from which the maximum value will be chosen is size P=[2,2], and step (S=2).

6. Convolutional Layer (CONV2): In this layer, 128 filters of size F = [3, 3], and stride (S = 2) without applying zero padding.

7. Batch Normalization Layer.

8. Relu Layer.

9. The pooling layer: (Pool2): The size of the area from which the maximum value will be chosen is of size F = [2,2], and a stride (S = 2).

10. Convolutional Layer (CONV3): In this layer, 256 filters of size F = [3,3], and stride (S = 1) without applying zero padding.

11. Batch Normalization Layer.

12. Relu Layer.

13. The pooling layer: (Pool3): The size of the area from which the maximum value will be chosen is of size F = [2, 2], and a stride (S = 2).

14. Convolutional Layer (CONV4): In this layer, 300 filters of size F = [1, 1], and stride (S = 1) without applying zero padding.

15. Batch Normalization Layer.

16. Relu Layer.

17. The pooling layer: (Pool4): The size of the area from which the maximum value will be chosen is of size F = [2,2], and a stride (S = 2).

18. Fully Connected Layer (FC1): The FC layer contains neurons completely connected to the neurons in the previous layer, the output volume of this layer is defined (4500).

19. Relu Layer.

20. Dropout Layer: This layer effectively changes the underlying network structure between iterators and helps prevent the network from over-fitting by setting the input elements randomly to zero according to the probability that was defined when building the network and the most appropriate probability value for our network is (0.5) [27],[28].

21. Fully Connected Layer (FC2): The output size has been specified for this layer (202), which represents the number of classes.

22. Softmax Layer Activation Function: It is used in classification problems of several classes. In classification problems, the last connected layer must be followed by a softmax layer and a classification layer, where the activation function of the output unit is the softmax function

23. Classification Layer.

3.5 Feature extraction using CNN

This section shows how to extract learned image features from a pertained convolutional neural network, and use those features to train an image classifier. Feature extraction is the easiest and fastest way to use the representational power of pertained deep networks. Because feature extraction only requires a single pass through the data, it is a good starting point if you do not have a GPU to accelerate network training with. We showed the network architecture of the CNN based-SVM model in "Figure. 3".

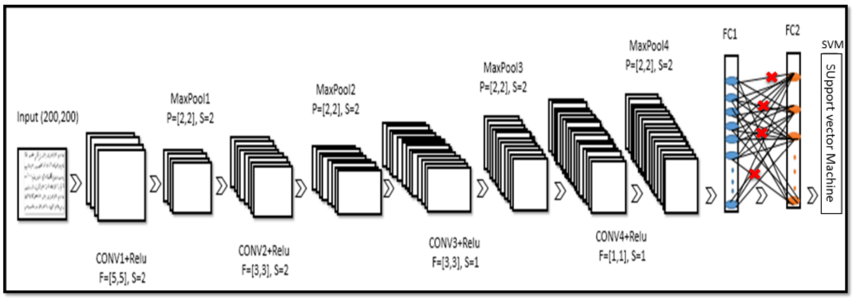

Fig -3: Architecture of the CNN based-SVM model

4 EXPERIMENTS AND RESULTS

It is observed that it looks as follows. First, load the data. Second, load a pre-trained network. Third, the network creates a hierarchical exemplification of the input images. Deeper layers contain higher-level features, which are built using the lower-level features of past layers. To get representations of training features and test images, use activations on the fully connected layer 'FC1'. For a lower level exemplification of the images, use a past layer in the grid. Fourth, from the training images has been used the features extracted as predictor variables and fit a multi-class support vector device (SVM). Finally, from the test images the trained SVM model and features extracted are used to classify the test images

The experiments were conducted on Matlab version R2020a. The database includes 202 writers and each writer has 3 texts to be increased by using the Data Augmentation techniques proposed in this research, so that each writer has 200 texts and the dataset becomes 40400 handwriting images. The dataset was divided into 3 groups 90% as a training set, chose 5% of them as a validation set, and 10% as a test set.

To know the effectiveness of the proposed method, the accuracy was measured using relationship (4.1) to obtain 96.5099% classification accuracy.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{FalsePositives}} * 100 \quad (4.1)$$

Table -1: Table compare result between Proposed System and Maaz et al.

Method	Feature Extraction	Classification Model	Accuracy
Maaz et al.	CNN	CNN	96.51%
Proposed System	CNN	SVM	97.20%

5 CONCLUSION AND FUTURE WORK

Writer identification by handwriting is a relatively new biometric method that has received significant research interest in recent years. Handwriting biometrics can be used in forensic applications to identify individuals based on the characteristics of their writing by comparing unsorted handwritten texts with written samples Handwritten and categorized.

Handwritten texts were used to verify the identity of the author of a document. The font analyzer usually depends on studying some features of the writing style of each writer. This research present using a convolutional neural network, which required the use of a data augmentation method that fits the characteristics of the text written in Arabic to extract the features and then Using support vector machine (SVM).

In the future, experiments will be focused interest on applying this study to a new and more diverse database of texts that contain structures in Arabic and English and numbers, in order to determine the identification accuracy with these cases in the text.

6 REFERENCES

[1] MAAZ, SH., ISSA, H., 2020- Using Deep Learning For Arabic Writer Identification. International Journal of Coputer Applications. Vol. 175, No 25.

[2] SREERAJ, M., IDICULA, S., 2011- A Survey on Writer Identification Schemes. International Journal of Computer Applications, (26).

[3] SHAHABINEJAD, F., RAHMATI, M.,2009 A New Method for Writer Identification of Handwritten Farsi Documents, ICDAR 2009, Spain, pp. 426 – 430

[4] SARANYA, K., VIJAYA, M. S., 2013- Text Dependent Writer Identification using Support Vector Machine. International Journal of Computer Applications, (65).

[5] Li,B., Sun, Z., Tan,T.N., 2009- Hierarchical Shape Primitive Features for Online Text-independent Writer Identification. 2009 10th International Conference on Document Analysis and Recognition, pages 201–210.

[6] ZHU, G., YU, X., LI, Y., DOERMANN, D. 2009-Language identification for handwritten document images using a shape codebook. Pattern Recogn.(42), 3184-3191.

[7] DJEDDI, C., SIDDIQI, I., MESLATI, L., ENNAJI, A., 2013- Text-independent writer recognition using multi-script handwritten texts. journal homepage, 1196–1202.

[8] ALAEI,A., ROY, P.P., 2014- A New Method for Writer Identification based on Histogram Symbo lic Representation. 14th International Conference on Frontiers in Handwriting Recognition, 2167-6445.

[9] THENDRAL, T., VIJAYA, M. S., KARPAGAVALLI, S., 2015- Prediction of Writer Using Tamil Handwritten Document Image Based on Pooled Features. International Journal of Computer, Electrical, Automation, Control and Information Engineering, (9).

[10] Fiel, S., Sablatnig, R., Writer identi_cation and retrieval using a convolutional neural network, in Proc. Int. Conf. Comput. Anal. Images Patterns. Cham, Switzerland: Springer, 2015, pp. 26_37.

[11] AL-ZOUBEIDY ,L. M. AL-NAJAR ,H. F., 2005- Arabic writer identification for handwriting images, International Arab Conference on Information Technology, pp. 111-117.

[12] SLIMANE ,F., MARGNER, V., 2014- A New Text-Independent GMM Writer Identification System Applied to Arabic Handwriting. 14th International Conference on Frontiers in Handwriting Recogniti.

[13] DJEDDI, CH., MESLATI, L., SIDDIQI,I., ENNAJI,A., ABED,H., GATTAL,A., 2014- Evaluation of Texture Features for Offline Arabic Writer Identification. 11th IAPR International Workshop on Document Analysis Systems.

[14] 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. Volume 80, 2016, Pages 1712–1723.

[15] ELADEL,A., ZAIED,M., AMAR,CH., 2018 - Trained convolutional neural network based on selected beta filters for Arabic letter recognition. WIREs Data Mining Knowl Discov.2018;e1250.

[16] DENGEL,A., LIWICKI,M., RASHID,S., AFZAL,M., NAZ,S., AHMAD,R. 2018 - A Deep Learning based Arabic Script Recognition System: Benchmark on KHAT. The International Arab Journal of Information Technology, Vol. 17, No. 3, May 2018.

[17] Durou,A., Aref,.I., Al-Maadeed,S., Bouridane,A., 2019 – Writer identification approach based on bag of words with OBI features. ScienceDirect ,Information Processing & Management, vol 56 issue (2) pp. 354-366. ISSN 0306-4573 .

[18] REHMAN,A., NAZ,S., RAZZAK,M., 2019 - Automatic Visual Features for Writer Identification: A Deep Learning Approach. IEEE Volume 7, 2019.

[19] www.kaggle.com/c/awic2012/leaderboard.

[20] Christlein, V., Bernecker, D., Maier, A., Angelopoulou, E. 2015- Offline Writer Identification Using Convolutional Neural Network Activation Features, Proc. German Conference on Pattern Recognition, pp. 540-552.

[21] Fiel, S., Sablatnig, R. 2015- Writer Identification and Retrieval Using a Convolutional Neural Network, Proc. Computer Analysis of Images and Patterns, pp. 26-37.

[22] Krizhevsky, A., Sutskever, I., Hinton, G.E. 2012- Imagenet classification with deep convolutional neural networks, Proc. Neural Information Processing Systems, pp. 1097-1105.

[23] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S. , Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A. 2014- Going deeper with convolutions, arXiv preprint arXiv:1409.4842.

[24] Simonyan, K., Zisserman, A. 2014- Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.

[25] Gogul.,Kumar,S., 2017- Flower Species Recognition System using Convolution Neural Networks and Transfer Learning. 2017 IEEE, 2017 4th International Conference on Signal Processing, Communications and Networking.

[26] IOFFE, S., SZEGEDY, CH., 2015- Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. reprint,arXiv:1502.03167 .

[27] SRIVASTAVA,N., HINTON , G., KRIZHEVSKY, A., SUTSKEVER, I., SALAKHUTDINOV, R., 2014- Dropout: A Simple Way To Prevent Neural Networks From Overfitting. Journal of Machine Learning Research. Vol. 15, pp. 1929-1958, 2014.

[28] KRIZHEVSKY, A., SUTSKEVER, I., HINTON, E., G., 2012- Imagenet Classification With Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems. Vol. 25, 2012.