Recognizing Arabic handwritten literal amount using Convolutional Neural Networks

Aicha KORICHI · Sihem SLATNIA · Najiba TAGOUGUI · Ramzi ZOUARI · Monji KHERALLAH · Oussama AIADI

Abstract Currently, deep learning techniques have become the core of recent research in pattern recognition domain and especially for the handwriting recognition field where the challenges for the Arabic language are stilling. Despite their high importance and performances, for the best of our acknowledge, deep learning techniques have not been investigated in the context of Arabic handwritten literal amount recognition. The main aim of this paper is to investigate the effect of several Convolutional Neural Networks CNNs based on the proposed architecture with regularization parameters for such context. To achieve this aim, the AHDB database was used where very promising results were obtained outperforming the previous works on this database.

Keywords Arabic handwriting · Literal Amount Recognition · Offline recognition · deep learning · Resnet · VGG

1 Introduction

Handwriting recognition has received a growing interest by researchers and it has become a very active field of research in recent years due to its important applications including automatic postal mail sorting, historical handwritten documents digitization, automatic checks recognition… etc. The handwriting recognition systems are divided into online and offline branches according to the data acquisition mode [1–3]. In the online mode, the input data is acquired from a digitized tactile screen and both static and dynamic information about the handwriting trajectory are available like the trajectory coordinates, temporal order, speed and acceleration [4]. In the offline mode, the input data is captured from a scanned image of the text, and therefore only static information representing the pixel values (0 or 1) is available. The lack of dynamic information makes offline handwriting recognition a very challenging task compared to the online one. Despite the many solutions proposed to deal with offline issues, it is still a challenging task because of several inherent characteristics that are related to the writing style and the writing language itself. In fact, for the Arabic script which is written from right to left, it contains 28 letters where each letter has more than three shapes according to its position within the world. Furthermore, additional strokes can be written above and below the letters like dots, chadda, fatha, etc. Moreover, Arabic is the 4^{th} spoken language with more than 400 million speakers [5], it is one of the six United Nations official languages.

In the literature, existing approaches can roughly be classified into two categories, namely holistic and segmentation approach [6]. As its name indicates, the segmentation approach consists of the segmentation of

A.KORICHI
LINFI Laboratory University of Mohammed Khider Biskra, Algeria
E-mail: aicha.korichi@univ-biskra.dz

S. SLATNIA
LINFI Laboratory University of Mohammed Khider Biskra, Algeria
E-mail: sihem.slatnia@univ-biskra.dz

N. TAGOUGUI
The Higher Institute of Computer Science and Multimedia of Sfax, Tunisia
E-mail: tag.najiba@gmail.com

R. ZOUARI
Faculty of Science of Sfax, Tunisia
E-mail: ramzi.zouari@gmail.com

R. ZOUARI
Faculty of Science of Sfax, Tunisia
E-mail: monji.kherallah@fss.usf.tn

O. AIADI
Depertment of Computer Science, University of Kasdi Merbah Ouargla, ALgeria
E-mail: o_aiadi@hotmail.com

the word into characters or sub words, meanwhile, the holistic approach considers the whole image as an entity that must be recognized without decomposition in smaller units and it is generally used with limited lexical datasets.

For the actual research purpose, we proposed in this paper to deal with the recognition of the Arabic handwriting literal amount that is still highly used in checks. Although this is not a new research area especially that many digitization attempts where realized in the early 1990s, our motivation to revive this domain has been consolidated by two major facts, the first one is that unlike it was expected in the digitization area, online electronic payment methods haven't replaced the use of checks and these ones are still spread used nowadays. The second reason is that according to recent statistical studies, about 100 billion checks are treated manually each year [7] [8]. Therefore, checks are a fundamental payment tool used in many countries and there are still processed thanks to the effort of human agents. There is no real need to stop using them, especially that they offer a trusted biometric measure within the personal handwritten style that differs from one person to another and which is more secure than laying on magnetic cards where the security property could be violated.

For all those reasons, and convinced by the power of the deep learning for the field of handwriting recognition, we tried to investigate in a real solution that enhances the actual automating checks reading methods so as to bring several benefits to ensure a better solution that is efficient, robust, and fast.

Indeed, interest for deep learning has emerged in the last few years [9] due to its effectiveness with the ability to learn high-level abstractions automatically in the data, it has replaced the handcrafted ways where the features were extracted manually. Convolutional Neural Networks (CNN) are among the most famous neural network models. It has proved its efficiency for many pattern recognition tasks. CNNs are present in the literature with several architectures that we decide to test in this experimental study to choose the most relevant one that matches our main motivation of designing a robust and efficient automatic checks reading system using the recognition of Arabic handwriting literal amounts.

To achieve this goal, several CNNs architectures starting from simple architecture and then going to complex ones namely Visual Geometry Group (VGG) and Residual Network (Resnet) were used with regularization parameters for the recognition of Arabic handwriting literal amounts.

The rest of the paper is organized as follows: In section 2, we present some related works, which have dealt with the Arabic handwritten recognition issue. Section 3 gives details of the proposed CNNs architectures and an overview of the system. Section 4 presents the AHDB database of Arabic handwritten literal amounts. The experimental evaluation results are given in section 5. Finally, we finalize the paper by outlining some conclusions.

2 Related work

Nowadays, Deep Learning techniques become the state of the art of the majority of research, they proved their efficiency for many pattern recognition systems [10, 11]. Despite their good performances, little attention has been devoted to dealing with them in the context of Arabic handwriting recognition. Almaageg et al [12] proposed a new system for Arabic handwriting recognition based on two deep neural network techniques. The first one is CNN, and it is used for feature extraction, the second one is bidirectional Long Short-Term Memory (BLSTM) followed by Connectionist Temporal Classification layer (CTC) for classification purposes.. In [13], the proposed model is based on the combination of CNN with Support Vector Machines (SVM) classifier and using raw pixel data. This system was tested on both HACDB and IFN/ENIT Arabic handwriting letters databases. The same system was reproduced in [14] with the application of the Dropout technique. Authors in [15] proposed a handwriting recognition system based on hybrid CNN architectures applied on several databases.

El-Melegy et al. [16] were the first that apply deep learning for the recognition of complete literal amount words. The proposed system is based on VGG architecture composed of 16 hidden convolution layers and 1 fully-connected layer by using data augmentation.

On the other hand, most researchers have devoted to using handcrafted features. They are divided into three categories. In the first sub-category, the Arabic handwriting was considered as a series of statistical characteristics. Assayony et al [17] proposed a new system of Arabic handwritten literal amount recognition based on a holistic approach using Gabor filters with Bag of Features (BoF). Gabor filter was applied with different scales and orientations to extract local features which will be arranged and fed to BoF frameworks. Hassen et al. [18] proposed a Multi statistical features system for Arabic handwriting literal amount recognition. They used a set of statistical features including Invariant Moments (IV), Histogram of Oriented Gradients (HOG), and Gabor filters. Thereafter, Sequential Minimal Optimization (SMO) classifier was applied.

In the second sub-category, the Arabic handwriting literal amount is considered as a series of structural features. Al-Nuzaili et al. [19] presented an improvement of the Perceptual Feature Extraction Model (PFM) by considering the shapes of loops and dots. In another work [20], the handwriting is considered as a set of distance, angle, vertical and horizontal span features. In the classification stage, three ELM classifiers were combined using the majority vote technique.

The third category considers the handwriting as a mixture of statistical and structural features. In [21], the Arabic handwriting literal amounts were represented using statistical features like Zernike moment invariants (ZMI), local chain code histograms (CCH), zoning, and the density profile histograms (DPH), and some other structural features extracted from the different parts of the image. In the classification stage, SVM was applied based on the extracted features. In [22], the proposed method proceeds by applying Discrete Cosine Transform (DCT) and Histogram of Oriented Gradient (HOG) to extract structural features merged with some other statistical features. An artificial neural network was used in the classification stage.

3 Deep learning for Arabic literal amount recognition

Deep learning has become the core of recent pattern recognition research and it has taken an incredible growth for many computer vision tasks due to its high performance in automatically capturing complex characteristics from the low level to high level. In other words, it aims to model high-level abstractions based on a set of traditional machine learning algorithms using several nonlinear transformations.

From many deep learning techniques, Convolutional Neural Networks CNNs [23] are the most commonly used and popular ones. They proved their efficiency, especially for handwriting recognition and they considered as top solutions in such issues [14, 24–27].

Referring to the aforementioned criteria, we are presenting several CNN architectures for Arabic handwriting literal amount recognition, starting from a simple CNN, until using complex ones like Resnet and VGG. In Spite of the major advantages of deep CNNs that can learn automatically more abstract information by constructing deep architecture, the huge quantity of parameters used could lead to other problems known as over fitting. To deal with such situations and protect the CNN against such cases, we opted to add a dropout layer. Moreover, CNNs need a huge quantity of data to be more efficient, we increase the number of training data by applying data augmentation .

3.1 Convolutional Neural Networks

As we have mentioned above, the Convolutional Neural Network is a type of deep learning technique that proved its efficiency for handwriting character recognition due to its ability to learn the visual patterns from image pixels [28]. Generally, CNN comprises three main layers which are convolutional layers, pooling layer, and fully connected layer (FCL) [29]. CNN performs the nonlinearity by the activation functions and the pooling layers. As an activation function, we have exploited the ReLU function (Rectified Linear Units).

$$f(x) = \max(0, x) \tag{1}$$

FCL is performed as follows:

$$y_j^l = \max(0, \sum_i y_i^{l-1}.w_{i,j}^l + b_j^l) \tag{2}$$

Where: y_i^l represents the j^{th} node in the l^{th} layer, $w_{i,j}^l$ represents the weights between y_i^l and y_i^{l-1}. b_j^l is the bias.

The last step of predicting a distribution p(y_i)is to handle a softmax over the outputs Z_i:

$$y_i = \frac{\exp(Z_i)}{\sum_k \exp(Z_k)} \tag{3}$$

$$Z_i = \sum_j y_i^{l-1}.w_{i,j} + b_i \tag{4}$$

The proposed CNN network scheme based on the dropout layer is illustrated in Fig 1. At first, the network receives the image in the input layer as a sequence of pixels and passed them through convolution layers where the image is convolved with a set of filters. Thereafter, the obtained activation maps are passed to an activation function layer, followed by a MaxPooling layer in order to preserve the pixels of higher values. To protect the network against the over fitting, a dropout layer is added just before the fully connected layers where the classification task is done.

3.2 Visual Geometry Group (VGG)

Starting from the common idea that says going deeper through a network will give better accuracies, Simonyan et al. [30] introduced the Visual Geometry Group (VGG) architecture which includes mainly two convolutional layers with ReLU as an activation function followed by max-pooling layer. The final layer of VGG architecture is the softmax for classification purposes. One of the advantages of VGG architecture is its simplicity, unlike previous networks by minimizing the kernel-size filters into 3 × 3 which allows learning more complex features.

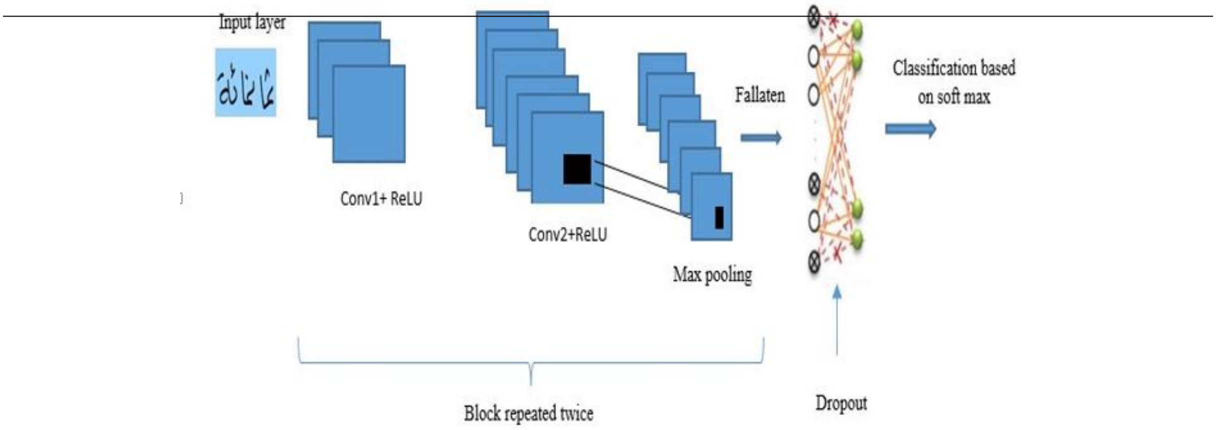

Figure 1: General scheme of our proposed network based on the dropout layer.

3.3 Residual Network (ResNet)

Going deeper into a network by increasing the number of layers, does not mean that the network is more efficient, sometimes going deeper could harm the network efficiency during the training data because of vanishing gradient problem. Recently, Residual Network ResNet network architecture, which was proposed by [31], has becomes very popular tools in the community of deep learning due to solving such a vanishing problem. The main idea behind the residual network is skipping one or more layers by a residual mapping. In another term, adding a parameter to the output from the previous layer to the layer ahead

4 Database presentation

For many Arabic countries, checks with the handwritten format are stilling the fundamental tool for financial transactions where about one hundred billion checks are treated over the world and the majority of them are treated manually basing on human agents. Automatic check reading has become an active area of research. AHDB benchmark database [32] is a publicly available database that contains the 63 different classes representing the Arabic handwritten literal amounts normally used on checks. Each class contains 105 samples written by different writers. As it is using for many researchers, a cross-Validation with three folds (two folds for training and the remainder for testing) is used for this study where each fold contains 2205 samples. A sample of each class is shown in (Table 1):

Table 1: Arabic words used to express amounts on checks extracted from AHDB database

N°	Class	N°	Class	N°	Class	N°	Class	N°	Class
1	أحد	14	خمسة	27	تسعة	40	ستة	53	ثلاثمئة
2	احدى	15	خمسمئة	28	تسعمئة	41	ستمئة	54	ثلاثمائة
3	واحد	16	خمسمائة	29	تسعمائة	42	ستمائة	55	عشرون
4	ثمان	17	أربع	30	تسعون	43	ستون	56	عشرين
5	ثمانية	18	أربعة	31	تسعين	44	ستين	57	اثنان
6	ثمانمئة	19	أربعمئة	32	لا	45	عشر	58	اثنين
7	ثمانمائة	20	أربعمائة	33	سبع	46	عشرة	59	مئتين
8	ثمانون	21	أربعون	34	سبعة	47	ثلاثون	60	مائتين
9	ثمانين	22	أربعين	35	سبعمئة	48	ثلاثين	61	ألفان
10	اثنى	23	مئة	36	سبعمائة	49	ألف	62	ألفين
11	خمسون	24	مائة	37	سبعون	50	آلاف	63	غير
12	خمسين	25	مليون	38	سبعين	51	ثلاث		
13	خمس	26	تسع	39	ست	52	ثلاثة		

5 Experimental Results

For our case, we are dealing with 63 different classes as previously described. The limited lexicon that we used argues the use of the holistic approach where the images were fed directly to the network without any segmentation. Since the CNNs require a huge amount of data to be efficient which is not available in the AHDB database, for all experiments done in this section, we have used data augmentation technique for training images with criteria that are related and interpreted by the Arabic language orientation, zoom, and writing width).

Moreover, the batch size was selected to be 32 batches. All the experiments done in this section were obtained using Python with Keras and Tensorflow installed on a computer with a Core i7 "7th generation" processor, 16 GB of RAM and AMDA Radeon graphical card. As a metric of evaluation, we have used the accuracy metric, which is the quotient of the total number of correctly classified words corrected over the total number of words.

5.1 Results of the proposed Architecture

First, we tested the effect of CNNs on the recognition of the literal amounts based on a simple CNN composed of four layers. After every two layers, a max-pooling layer is added. The aim of using a simple CNN is to study the effect of adding the dropout layer to a network for several epochs where the experiments can be easily done. To achieve such objective, several experiments have been carried out without using a dropout layer, and by adding the dropout layer in both extraction and classification parts, and only adding the dropout layer before the classification layer. The obtained results are illustrated in the following figures (Fig.2, Fig.3, and Fig. 4):

Figure 2: Results without using dropout.

Figure 3: Results by adding the dropout layer in both stages of feature extraction and classification.

Figure 4: Results by adding the dropout layer only before the classification layer.

Based on the above figures, it is clearly shown the positive effect of adding a dropout layer on the test data regardless of its position, as shown in Figure 3 and Figure 4. Moreover, it is obvious that using the dropout layer on just before the fully connected layer gave in some epoch's very high performance compared with those that have been obtained by using dropout out in both feature extraction and classification parts. Whilst, the average recognition rate obtained by adding the dropout layer on both parts is better than which is achieved by adding it just once before the classification stage. It can be caused by the negative influence of the irrelevant characteristics without removing them by using dropout just on the classification.

5.2 VGG results

As we have mentioned in the previous sections, the Visual Geometry Group (VGG) network is one of the simplest and efficient networks. For this study, we used a VGG network architecture based on 13 layers to investigate the impact of adding more layers in the recognition rate. During the first fifth epochs, the recognition rates obtained by the VGG16 network are very low compared to those obtained by the previously proposed architectures. However, the architecture faithfully success in increasing the recognition rate of training samples to be stabilized on 100%. Despite that VGG-16 contains more layers than the previously proposed ones; this hasn't affected positively the recognition rates especially for the test samples (Fig. 5). This can be interpreted by the nature of Arabic writing and the existing features that could be captured just by a few layers.

Figure 5: VGG16 results.

5.3 Resnet results

In order to investigate the effect of more complex architectures by going deeper, we opted to apply Resnet architecture using always a dropout layer, with a depth of 17 convolutional layers with randomly initialized weight. According to the obtained results (Fig. 6), it is clearly shown the positive effect of increasing the number of layers to go more depth basing on Resnet network. This allows weights to be updated correctly through the backpropagation of the gradient error. Resnet tries to find an optimized number of layers to eliminate the vanishing gradient problem. Unlike the previous architectures, the recognition rate is starting with high performance in training and test samples then decrease on some epochs of the test images, and finally stabilizes from the 14th epoch for both stages. The recognition rate is about 100% to 98% for training and test samples respectively, after a number of 100 epochs.

Figure 6: Resnet results.)

5.4 Results discussion

In order to investigate the impact of deep learning algorithms on the recognition of Arabic handwriting, in this paper have tested three architectures with different parameters. Table 2 summarizes all experiments done for this study.

Table 2: Proposed architectures results.

Architecture	Accuracy %
Proposed architecture with dropout	95.7
VGG16	97.14
Resnet	98.57

It is clearly shows the high performance of deep architectures for the recognition with even simple or complex ones. First, we have tested the performance of the proposed architecture with several positions of dropout layer where the best recognition rate was 95.71% by adding a dropout layer just before the classification stage. In spite of the low number of layers used in our proposed architectures, it gives very good results which are very close to them obtained with VGG and Resnet architectures where we are going deep on the network by increasing the number of layers by adding always the dropout layer in the same position. VGG and Resnet architectures are given results close to each other. However, the nature of Resnet eliminating the vanishing problem by trying to find the optimized number of layers allows it to outperform the other architecture with the best average recognition rate of 98.57%.

5.5 Comparison with state of the arte systems

In this work, we studied the effectiveness of several CNN architectures in such a context, where the best recognition was 98.57%, and it was attained when using Resnet architecture. The comparisons of the obtained results with other recent and relevant works made on the AHDB database are summarized in Table 3:

Table 3: Proposed architectures results.

Authors	Accuracy ‘%
Menasria et al [21]	89.13
Assayony and Mahmoud [17]	86.44
Hassan et al [18]	95
Al-Nuzaili et al [19]	92.13
El-Melegy et al [16]	97.8
Amani Ali et al [15]	96.8
Our proposed system	98.57

Based on the above table, it is clear that the implemented CNN architectures have proven their efficiency against handcraft based methods, in the context of Arabic handwriting literal amount images from AHDB database.

6 Conclusion and Perspectives

Arabic handwriting recognition is still a challenging task due to several inherent characteristics of Arabic script and it is still in the level of experiments. Arabic handwritten literal amount recognition is a typical application of the handwriting recognition domain since the checks are considered as the fundamental financial transactions tool for several Arab countries. In this paper, several Convolutional Neural Networks CNNs were implemented like simple CNN, VGG-16 and ResNet, with the application of regularization methods including dropout and data augmentation techniques. The obtained results proved the efficiency of the CNNs architectures that outperform the existing methods based on handcrafted features. As a perspective, we intend to evaluate an enhanced version using transfer learning of the implemented architectures for recognizing Arabic words from other vocabularies and databases.

References

[1] Liana M Lorigo and Venugopal Govindaraju. "Offline Arabic handwriting recognition: a survey." In: *IEEE transactions on pattern analysis and machine intelligence* 28.5 (2006), pp. 712–724.

[2] Aicha Korichi et al. "Off-line Arabic handwriting recognition system based on ML-LPQ and classifiers combination." In: *2018 International Conference on Signal, Image, Vision and their Applications (SIVA)*. IEEE. 2018, pp. 1–6.

[3] Aicha Korichi et al. "Arabic handwriting recognition: Between handcrafted methods and deep learning techniques." In: *2020 21st International Arab Conference on Information Technology (ACIT)*. IEEE. 2020, pp. 1–6.

[4] Ramzi Zouari, Houcine Boubaker, and Monji Kherallah. "A time delay neural network for online arabic handwriting recognition." In: *International Conference on Intelligent Systems Design and Applications*. Springer. 2016, pp. 1005–1014.

[5] Gary F Simons and Charles D Fennig. *Ethnologue: languages of Asia*. sil International Dallas, 2017.

[6] Mohammad S Khorsheed. "Off-line Arabic character recognition–a review." In: *Pattern analysis & applications* 5.1 (2002), pp. 31–45.

[7] Irfan Ahmad and Sabri A Mahmoud. "Arabic bank check analysis and zone extraction." In: *International Conference Image Analysis and Recognition*. Springer. 2012, pp. 141–148.

[8] Irfan Ahmad and Sabri A Mahmoud. "Arabic bank check processing: state of the art." In: *Journal of Computer Science and Technology* 28.2 (2013), pp. 285–299.

[9] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." In: *nature* 521.7553 (2015), pp. 436–444.

[10] Adeline Granet et al. "Transfer Learning for Handwriting Recognition on Historical Documents." In: *ICPRAM*. 2018, pp. 432–439.

[11] Najwa Altwaijry and Isra Al-Turaiki. "Arabic handwriting recognition system using convolutional neural network." In: *Neural Computing and Applications* 33.7 (2021), pp. 2249–2261.

[12] Rania Maalej and Monji Kherallah. "Convolutional neural network and BLSTM for offline Arabic handwriting recognition." In: *2018 International Arab conference on information technology (ACIT)*. IEEE. 2018, pp. 1–6.

[13] Mohamed Elleuch, Rania Maalej, and Monji Kherallah. "A new design based-SVM of the CNN classifier architecture with dropout for offline Arabic handwritten recognition." In: *Procedia Computer Science* 80 (2016), pp. 1712–1723.

[14] Mohamed Elleuch, Najiba Tagougui, and Monji Kherallah. "A novel architecture of CNN based on SVM classifier for recognising Arabic handwritten script." In: *International Journal of Intelligent Systems Technologies and Applications* 15.4 (2016), pp. 323–340.

[15] Amani Ali Ahmed Ali and Suresha Mallaiah. "Intelligent handwritten recognition using hybrid CNN architectures based-SVM classifier with dropout." In: *Journal of King Saud University-Computer and Information Sciences* (2021).

[16] Moumen El-Melegy et al. "Recognition of Arabic Handwritten Literal Amounts Using Deep Convolutional Neural Networks." In: *Iberian Conference on Pattern Recognition and Image Analysis*. Springer. 2019, pp. 169–176.

[17] Mohammed O Assayony and Sabri A Mahmoud. "Recognition of Arabic handwritten words using Gabor-based bag-of-features framework." In: *International Journal of Computing and Digital Systems* 7.01 (2018), pp. 35–42.

[18] Hanadi Hassen and Somaya Al-Maadeed. "Arabic handwriting recognition using sequential minimal optimization." In: *2017 1st International Workshop on Arabic Script Analysis and Recognition (ASAR)*. IEEE. 2017, pp. 79–84.

[19] Qais Al-Nuzaili et al. "Enhanced structural perceptual feature extraction model for Arabic literal amount recognition." In: *International Journal of Intelligent Systems Technologies and Applications* 15.3 (2016), pp. 240–254.

[20] Qais Ali Al-Nuzaili et al. "Pixel distribution-based features for offline Arabic handwritten word recognition." In: *International Journal of Computational Vision and Robotics* 7.1-2 (2017), pp. 99–122.

[21] Azzeddine Menasria et al. "Multiclassifiers system for handwritten Arabic literal amounts recognition based on enhanced feature extraction model." In: *Journal of Electronic Imaging* 27.3 (2018), p. 033024.

[22] Alia Karim Abdul Hassan and Mustafa S Kadhm. "Handwriting word recognition based on neural networks." In: *International Journal of Applied Engineering Research* 10.22 (2015), pp. 43120–43124.

[23] K Fukushima. "A hierarchical neural network capable of visual pattern recognition." In: *Neural Network* 1 (1989).

[24] Yanming Guo et al. "Deep learning for visual understanding: A review." In: *Neurocomputing* 187 (2016), pp. 27–48.

[25] Rashad Ahmed, Wasfi G Al-Khatib, and Sabri Mahmoud. "A survey on handwritten documents word spotting." In: *International Journal of Multimedia Information Retrieval* 6.1 (2017), pp. 31–47.

[26] Luiz G Hafemann, Robert Sabourin, and Luiz S Oliveira. "Learning features for offline handwritten signature verification using deep convolutional neural networks." In: *Pattern Recognition* 70 (2017), pp. 163–176.

[27] Lianwen Jin et al. "Online handwritten Chinese character recognition: from a bayesian approach to deep learning." In: *Advances in Chinese Document and Text Processing*. World Scientific, 2017, pp. 79–126.

[28] Yann LeCun, Yoshua Bengio, et al. "Convolutional networks for images, speech, and time series." In: *The handbook of brain theory and neural networks* 3361.10 (1995), p. 1995.

[29] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks." In: *Advances in neural information processing systems*. 2012, pp. 1097–1105.

[30] Karen Simonyan and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." In: *arXiv preprint arXiv:1409.1556* (2014).

[31] Kaiming He et al. "Deep residual learning for image recognition." In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 770–778.

[32] Somaya Al-Ma'adeed, Dave Elliman, and Colin A Higgins. "A data base for Arabic handwritten text recognition research." In: *Proceedings eighth international workshop on frontiers in handwriting recognition*. IEEE. 2002, pp. 485–489.