السنة الثالثة والعشرون ملحق 1 العدد 85 / شهر حزيران/ السنة 2025
https://doi.org/10.31272/ijes.v23 i85.1279

المجلة العراقية للعلوم الاقتصادية
Iraqi Journal For Economic Sciences

PISSN : 1812-8742 EISSE ONLIN : 2791-092X Arcif : 0.375

الجمع بين شبكات CNN واكتشاف الزوايا للتعرف على الكاتب العربي

Combining CNNs and Corner Detection for Arabic Writer Identification

ميس زيدان خليف
Mays Zeedan K.
kam4020@uomustansiriyah.edu.iq

ا.م.د مهند تحرير يونس
Muhanad Tahrir Younis
mty@uomustansiriyah.edu.iq

كلية العلوم/ جامعة المستنصرية

المستخلص

تحتل اللغة العربية المرتبة الخامسة في ترتيب اللغات المنطوقة، مما يعني أنّ 420 مليون شخص يتحدثون اللغة العربية. وقد تم التعرف على الأشخاص بيومتريًا باستخدام بصمات الأصابع والوجوه وغيرها من السمات المماثلة. في هذه الورقة، اقترح نموذج تعريف بيومتري للكتابة اليدوية العربية، لأنّ العديد من الحروف العربية لها أشكال متشابهة جدًا ولا يمكن تمييزها إلا من خلال موقع نقطة واحدة أو أكثر، أمّا فوق الحرف أو تحته. يتم تقديم نموذج جديد وفعّال للتعرف على الكتابة اليدوية العربية دون اتصال بالإنترنت. أساسه هو مزيج من عدة طرق مثل كاشف زاوية هاريس، وشي توماسي، والشبكات العصبية التلافيفية (CNNs). يتم استخدام زيادة البيانات أثناء مرحلة تدريب النموذج لتحسين جودة البيانات، دون الحاجة إلى تقسيم الكلمات / الأحرف. الاستفادة من مجموعة كبيرة من المستندات العربية المكتوبة بخط اليد، مثل KHATT و AHAWP تم الوصول إلى معدلات دقة 99٪ و 98٪ على التوالي.

الكلمات الرئيسية: الكتابة اليدوية، اللغة العربية، تحديد الهوية، الشبكات العصبية التلافيفية.

Abstract:

The Arabic language occupies fifth place in the ranking of spoken languages, meaning that approximately 420 million people speak it. People have been biometrically identified using fingerprints, faces, and other similar features. In this paper, a biometric identification model for Arabic handwriting is proposed, as many Arabic letters have very similar shapes and can only be distinguished by the location of one or more dots, either above or below the letter. A novel and efficient offline Arabic handwriting identification model is presented. Its basis is the combination of several methods, such as the Harris corner detector, Shi-Thomasi, and convolutional neural networks (CNNs). Data augmentation is used during the model training phase to improve the data quality, without the need to segment words/characters. Leveraging a large collection of handwritten Arabic documents, such as KHATT and AHAWP, accuracy rates of 99% and 98% were reached, respectively.

Keywords: Handwriting, Arabic language, Identification, CNNs.

1. Introduction

The usage of handwriting serves as evidence that each individual has a unique writing style that sets them apart from others. Handwriting is skill that someone can develop over many years, and it is considered a behavioural characteristic [1]. An issue that is still under research is handwriting identification. Handwriting identification can be very useful, especially in the banking industry to verify the authenticity of receipts and distribute checks, forensics and security sectors can serve as a benchmark for criminal cases involving handwriting. There are two main approaches to handwriting identification: offline and online. In offline recognition, a picture of the written text on a piece of paper is taken using optical scanning, also known as intelligent word recognition or optical character recognition. The data on this form is assumed to be a static handwriting representation. However, online handwriting identification converts text automatically as it is written using specialized digitizer that uses a sensor to identify pen-tip movements, and pen up/pen down switching is referred to as digital ink. The resulting signal is converted into letter codes that can be utilized in computer programs and text processing [2].It is still difficult to identify the writer of a text by handwriting, especially in Arabic, because of the script's cursive style and the subtle variations in letter placement. The problem remains challenging due to the high intra-class variances and sometimes high similarity between two people's writings. Because of this resemblance, the corner detector was used to solve the problem by identifying the corners of handwriting, which are thought to be a reliable feature. Deep learning models were chosen in conjunction with corner detector techniques because they were successful in generating a reliable model that, when combined, could be used to generate a potent model for an identification scenario [3,4]. The rest of this paper is organised as follows: Section 2 discusses related work, Section 3 describes the proposed method, and Section 4 introduces the datasets used. The results and discussions are presented in Section 5, followed by conclusions and future work in Section 6.

2. Related work : This section explains the work most relevant to our task, writer identification. In addition, will discuss the unique challenges of Arabic handwriting, such as its slanted nature and its textual diversity, and how recent advances in deep learning have addressed these challenges. By reviewing these key areas, we aim to highlight the current state of research in writer identification and place our work in this context.The issue of using Arabic handwriting samples for writer identification was reviewed in [5]. The proposed method is based on extracting short text segments that are identified by two textual descriptors: grey-level running length matrices (GLRL) and histograms of oriented gradients (HOG). The classification process involves calculating the distance between the parts of the sample that need to be compared. A variety of fusion rules are used to combine the similarity scores obtained using GLRL and HOG features. Three popular Arabic handwriting databases are used to evaluate the system: the QUWI database, which includes 1017 writers, the KHATT database, which includes 1000 writers, and the IFN/ENIT database, which includes 411 writers. The highest recognition accuracy rates were reported by Fusion using the "sum" rule, with values of 96.86, 85.40, and 76.27% on the IFN/ENIT, KHATT, and QUWI databases, respectively.Two methods were used: associative temporal classification (CTC) and multidimensional long-short-term memory (MDLSTM), both based on networks in [6]. The advantage of MDLSTM is that it can scan lines of text, in all directions (vertical and horizontal), to accommodate dots, spelling marks, and dashes. The obtained accuracy rate was 75.7%.The issue of Arabic handwriting recognition was introduced in [7]. By proposing a new architecture based on the character model approach that combines CNNs and BLSTM with CTC decoding. For experiments, the Arabic handwriting database KHATT is used. Using a horizontal sliding window that scans the image from right to left, CNNs are applied to a series of 64 × 64 images extracted from the text line image. With 256 filter maps in the final convolution layer, they produce a multi-channel output with dimensions of 1 × 16 × 256. The remaining two dimensions are determined by the amount of pooling that occurs in the CNNs. A new contribution to Arabic handwriting recognition is the Deep CNN-BLSTM combination based on the character model. The obtained accuracy was 79.83%. A method for identifying authors from their handwritten documents was presented in [8]. The main contribution is its recommendation to use co-occurrence features to improve the performance of the writer recognition. From the pre-processed regions of interest (sub-images), a contour-based feature is extracted, the outer contour of which is mainly determined by the ink trace width and shape (IWSL) and modified local binary pattern (MLBP) measurements, to determine the similarities between different handwriting images. Joint probability distributions of MLBP and IWSL are calculated on distinct pixels by treating the contour as a texture and applying these texture descriptions. Chi-square distance and the nearest neighbour rule are used for identification. The proposed system is tested using eight popular handwriting databases (KHATT in Arabic, among others). An accuracy rate of 95% is achieved. A novel and effective system for recognizing handwritten texts in Arabic is demonstrated in [9]. The new method is based on the combination of bidirectional long short-term memory (BLSTM) and CNN networks, which is then followed by a connectionist temporal classification layer. The bidirectional memory network receives features extracted by convolutional layers and uses them for classification. The predicted features are aligned along the most probable path using the connectionist temporal classification layer. The proposed method solves a number of problems associated with this point by being able to identify handwritten texts in Arabic without the need to separate letters. It achieved an accuracy was 80.15%.In Table 1 shows a summary of the above-mentioned works that were selected from the literature. After conducting a comprehensive analysis and comparison of the studies, it is noticeable that the effort spent in automatically identifying the writer using Arabic as a full text is less compared to other languages. This is because Arabic is a difficult language, where letters have a wide range of shapes and positions.

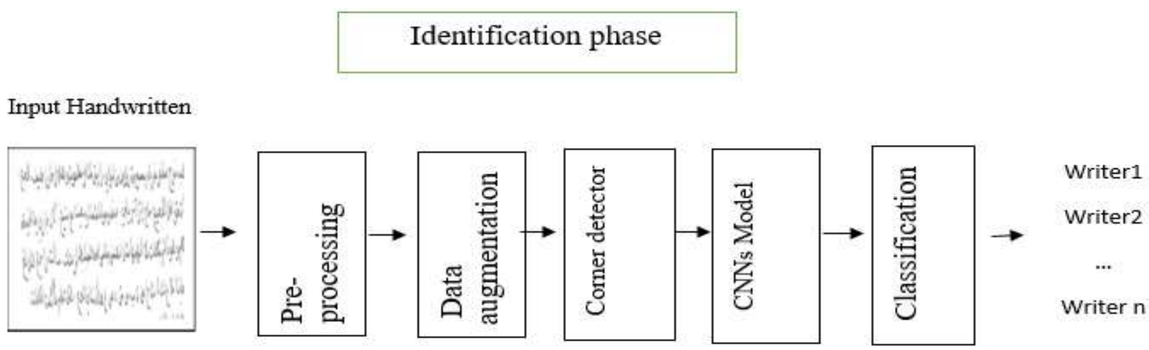
Table 1: Synopsis of notable research on author identification.

Study	Dataset	Identification rate	Method
Hannad Y,et al(2019)	IFN/ENIT, KHATT, QUWI	96.86%,85.40%,76.27%	GLRL+HOG
R.Ahmed, et al(2020)	KHATT	75.7%	MDLSTM+CTC
Z.Noubigh,et al(2021)	KHATT	79.83%	CNN+BLSTM
Tayeb (2022)	KHATT	95%	Chi-square distance and the Nearest Neighbour
Hicham et al(2023)	KHATT	80.15%	CNN+BLSTM+CTC

المصدر: اعداد الباحثين.

3. Proposed Model : This paper presents a composite model for writer identification. More than one technique, such as the Harris corner detector, Shi-Tomasi and convolutional neural networks (CNNs), which used to identify Arabic handwriting. By taking advantage of each strategy, the performance of the model can be improved by combining them. Figure 1 illustrates the model architecture.

Figure 1:The model architecture.



المصدر: اعداد الباحثين.

3.1 System Inputs: In this system, two handwritten datasets are used. In the case of KHATT data, 1000 users were taken, each with four different handwritten images, while in the case of AHAWP data, 82 users were taken, each with a number of images. In the case of words only, the handwriting image is then entered into the system.

3.2 Preprocessing : The two datasets go through a series of preprocessing steps. These steps include resizing the images to make them all the same size, which helps reduce the amount of white space in the images and focuses more on the writing. After that, the handwritten portion of the image is enlarged by a factor of 2 in both dimensions using upsampling to make it more accurate and clear.

3.3 Data Augmentation : Large amounts of data are needed for deep learning networks. The process of collecting enough training data is expensive and time-consuming as well. Applying the data augmentation to training data is a standard procedure to address this problem. Applying a data augmentation technique to artificially increase the size of the training dataset by creating updated copies of the images in the dataset. The learning domain of our model will be expanded, enabling it to adjust and predict new data more effectively. Data augmentation was applied to the images using multiple variations. For example, flipping the images using (0, 1, -

1), and then randomly rotating the images between (90 clockwise, 90 counterclockwise, and 180).

3.4 Feature extraction using corner detector : By utilising its capacity to identify distinguishing characteristics, like corners, the corner detector can be utilised efficiently for handwriting recognition. It can be used in any language, but Arabic is the most important. Arabic handwriting is distinguished by its cursive style, in which letters can have a wide range of shapes based on where they fall in a word, and characters are frequently joined. It is essential to recognise corners and other distinguishing characteristics to differentiate between various characters and handwriting styles, which include the following:

1: Pre-processing the Handwritten Text

a-Grayscale Conversion: To streamline the processing, convert the handwritten text image input to grayscale.

b- Binarisation: transform into a binary image, in which the background is white and the text is usually black.

2: Applying the Corner Detector

a- Gradient Calculation: To get the intensity variations that correspond to corners, compute the image gradients in the x and y directions.

b- Corner Detection: Utilize the corner detection algorithm to pinpoint the Arabic script's corners—important locations where noticeable intensity shifts occur.

3: Feature Extraction

a- Extract Corners: The corners that have been detected are utilized as features to symbolize the distinctive qualities of the Arabic script.

b- Additional Features: To generate a strong feature set for every character or word, combine the corner features with other pertinent features like stroke direction, curvature, and aspect ratio.

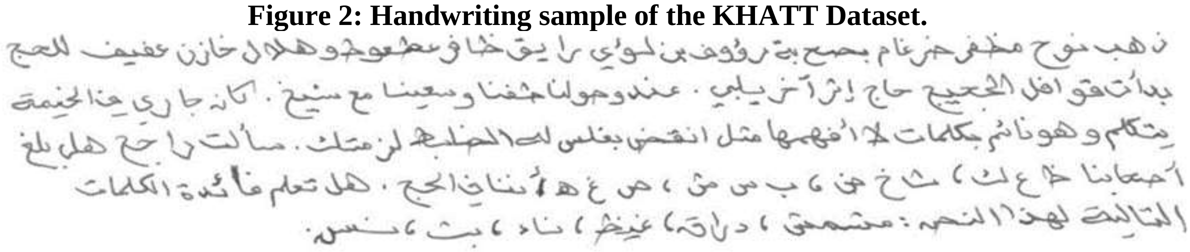
3.5 CNN Model : After obtaining the pixelated images (which contain corners), the feature extraction phase begins. The CNN extracts high-level features from the input image, using convolutional layers followed by pooling layers. These features are passed through fully connected layers, with a Softmax classifier providing the final predictions.

3.6 System Output: The system outputs the author identification results, with 1000 classes for the KHATT dataset and 82 classes for the AHAWP dataset. Each class represents a unique author, identified based on the extracted features.

4. Datasets : Within this section, the datasets used in the proposed model will be described in two cases: full text in the KHATT dataset and words only in the AHAWP dataset. Also, illustrative images will be included for each type.

4.1 KHATT Dataset: The King Fahd University of Petroleum and Minerals (KFUPM) has copyright ownership and copyright protection for its handwritten Arabic text database (also known as KHATT خط). Research teams from TUD Ortmund, Germany, TU Braunschweig, Germany, and KFUPM, Saudi Arabia, developed KHATT. It is made up of 1000 writers' handwritten Arabic documents. Different resolutions of the images (200, 300, and 600 dpi) were scanned. The writers were chosen from different nationalities, genders, ages, hand (left<right), and educational attainment. These naturally occurring writings are unique because they were created with no limitations on writing techniques. Researchers from all over the world can use the database for free to study a variety of handwritten-related issues, including text recognition, writer identification and verification, form analysis, preprocessing, and segmentation. In Figure 2, a handwritten sample is shown; the database has been obtained by numerous international research groups and investigators for their research purposes. The database is divided as follows: each person has four pages. The first, second, third and fourth pages contain paragraphs written by the author [10,11].

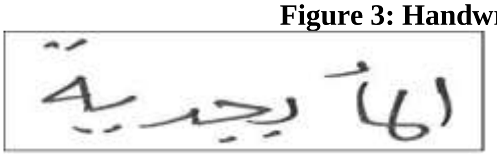
Figure 2: Handwriting sample of the KHATT Dataset.



المصدر: مجموعة بيانات خط.

4.2 AHAWP Dataset :A handwritten Arabic alphabet words and paragraphs dataset, the dataset includes three paragraphs, ten distinct words that span all alphabets, and 65 distinct Arabic alphabets (with variations on begin, end, middle, and regular alphabets). 82 distinct users' anonymous contributions made up the dataset. Every user is instructed to write every letter, word, ten times. Every alphabet, word, and paragraph has a user ID that uniquely and anonymously identifies its author. 53199 alphabet images, 8144-word images, and 241 paragraph images make up the entire dataset. The writers were proficient in both writing and speaking Arabic. It was not recommended that they use any particular brand or shade of writing instrument. This produced a dataset with multiple variants of writing instruments in terms of type and colour. There are several uses for this dataset. It applies to optical handwriting recognition of words and letters. Handwritten Arabic text can also be utilized for writer identification or verification. Using this dataset, it is also possible to assess how different the writing styles of isolated alphabets are from those of the same alphabet written by the same user as a word or paragraph. In this model, we used word images only [12]. Figure 3 shows a handwritten sample.

Figure 3: Handwriting sample of AHAWP Dataset.



المصدر: مجموعة بياناتAHAWP

5. Results

This section presents the results obtained by using the model. By using a series of convolutional layers, with a different number of layer filters on each layer: used sixteen filters, and used thirty-two filters in the feature acquisition stage (hidden layers). This difference can be attributed to several factors, the most important of which is that during testing, several issues, including misfits and situations that run the training for a long time indefinitely, continue to represent a problem to some extent. With the help of the epoch, we have monitored the progress of our model during the training phase and see if it has started to overfit or is still improving. Ten epochs of explicit training were used, which prevented the model from controlling the overfit and helped it produce better results. Remarkably, 70% of the dataset was drawn for training and 30% for the test. The best parameter tuning is given in Table 2.

Table 2 shows the best parameters.

No. of filters	Number of Patch size	Epoch	Data split
Input layer 16 filters Hidden layer 32 filters	128	10	70% for training and 30% for testing

المصدر: اعداد الباحثين.

Evaluate the effectiveness of the suggested model. The model was trained using different cases, each of which was created to tackle a particular issue and support its significance in enhancing the outcomes. The KHATT dataset results are displayed in Table 3, which also assesses the accuracy of training, testing, and validation. In Case 1, where a CNN is used with only raw data, the training, testing, and validation accuracies are 0.36, 0.28, and 0.25, respectively. This suggests that without further preprocessing or feature extraction, the model finds it difficult to extract significant patterns from the dataset. A slight improvement is seen in Case 2, which uses CNN for data augmentation. Training, testing, and validation accuracies are 0.58, 0.41, and 0.36, respectively. These findings nevertheless point to inadequate learning, most likely as a result of weak feature extraction.in Case 3, the inclusion of Harris corner detection leads to a significant performance boost, with training and testing accuracies improving to 0.79 and 0.50, respectively. Despite this improvement, the validation accuracy is 0.44, possibly due to overfitting. Lastly, Case4, which combines data augmentation with Harris corner detection, achieves exceptional results, with training, testing, and validation accuracies reaching 0.9966, 0.9930, and 0.9827, respectively. This highlights the importance of leveraging both data augmentation and robust feature extraction for effectively capturing patterns in the KHATT dataset. In summary, the table demonstrates that while CNNs alone struggle to perform well, the addition of Harris corner detection and data augmentation significantly enhances performance. Case 4's nearly perfect accuracy emphasises the critical role of feature extraction and diverse training data in achieving optimal results.

Table (3) Four cases passed through the KHATT dataset with Harris corner.

	Train Accuracy	Test Accuracy	Val. Accuracy
CNNs + Only data	0.36	0.28	0.25
CNNs + Augmentation	0.58	0.41	0.36
Harris corner + CNNs	0.79	0.50	0.44
Harris corner + CNNs + Augmentation	0.9966	0.9930	0.9827

المصدر: اعداد الباحثين بعد تنفيذ البرنامج.

The AHAWP dataset is new and has not been extensively studied, especially in terms of words. The initial findings for writer identification using this novel dataset are shown in Table 4. The table provides insight into the efficacy of various methods by assessing training, testing, and validation accuracies across four experimental configurations. The model performs moderately in Case 1, where a CNN is used to analyse only raw data, with training, testing, and validation accuracies of 0.72, 0.21, and 0.17, respectively. These findings imply that while CNNs are capable of extracting certain significant features from the dataset, their performance is constrained in the absence of further improvements. Case 2, which applies data augmentation with CNN, shows a drastic decline in performance, with training, testing, and validation accuracies dropping to 0.81, 0.49, and 0.40, respectively. This surprising reduction might indicate that the augmentation process introduces noise or artefacts unsuitable for this dataset. The addition of the Harris corner detector in Case 3 improves performance compared to. The training, testing, and validation accuracies are 0.90, 0.73, and 0.70, respectively. This suggests that while Harris corner detection enhances feature extraction, its effectiveness is limited without data augmentation. Case 4 achieves exceptional results by combining Harris corner detection and data augmentation, with training, testing, and validation accuracies of 0.986, 0.97, and 0.92, respectively. This demonstrates that the synergy between robust feature extraction and diverse training data significantly improves the model's ability to identify writers from the AHAWP dataset.

Table (4) Four cases passed through the AHAWP (words) dataset with Harris corner.

	Train Accuracy	**Test Accuracy**	**Val. Accuracy**
CNNs + Only data	**0.72**	**0.21**	**0.17**
CNNs + Augmentation	**0.81**	**0.49**	**0.40**
Harris corner + CNNs	**0.90**	**0.73**	**0.70**
Harris corner + CNNs + Augmentation	**0.986**	**0.97**	**0.92**

المصدر: اعداد الباحثين بعد تنفيذ البرنامج.

This section will display every outcome that was achieved following the application of Shi-Tomasi. As seen in the preceding tables, the first two cases—the CNN case with data only and the second case, data augmentation with CNN—are shared by both types of data used, thus they don't need to be mentioned again. We will only include Shi-Tomasi with CNN and data only, as well as Shi-Tomasi with data augmentation and CNN, in Table 5.

Table 5: Cases passed on the KHATT dataset with Shi-tomasi.

	Train Accuracy	**Test Accuracy**	**Val. Accuracy**
Shi-Tomasi + CNNs	0.68	0.37	0.30
Shi-Tomasi + CNNs + Augmentation	0.993	0.991	0.980

المصدر: اعداد الباحثين بعد تنفيذ البرنامج.

Table 6. Cases passed on the AHAWP (words) dataset with Shi-tomasi.

	Train Accuracy	Test Accuracy	Val. Accuracy
Shi-Tomasi + CNNs	**0.79**	**0.61**	**0.52**
Shi-Tomasi + CNNs + Augmentation	**0.98**	**0.963**	**0.90**

المصدر: اعداد الباحثين بعد تنفيذ البرنامج.

As mentioned earlier, the AHAWP dataset is new and has not been widely studied, especially when it comes to words, so these are the first results in the case of writer identification. The related work on the Khatt dataset in writer identification is shown in Table 7, including the improvement rate of how much each of the previous models improved.

Table 7. Comparing with techniques on the Arabic KHATT Dataset.

Model	Accuracy	Improvement rate
GLRL+HOG	85.40%	13.6%
MDLSTM+CTC	75.7%	23.3%
CNN+BLSTM	79.83%	19.17%
Chi-square distance and the Nearest neighbor	95%	4.%
CNN+BLSTM+CTC	80.15%	18.85%
(Our-model) CNNs+ corner detector	99%	

المصدر: اعداد الباحثين.

6. Conclusions : The scientific approach to writer identification is handwriting analysis. Handwriting analysis is scientifically known as graphology or graph analysis. Since handwriting is a method of analysing frozen graphic structures created in the brain and written in a printed or slanted style on paper, it is sometimes referred to as "mind writing" or "brain writing". Individual differences exist in handwriting. Handwriting analysis can be used to identify criminals. In this work, we have presented a comprehensive analysis of the latest developments in writer identification methods, focusing on preprocessing and explaining the efficiency of using boosting due to its relationship with deep learning to improve the accuracy of the model. Our system consists of a convolutional neural network and combined with the Harris corner detector/Shi-Tomasi, the KHATT dataset, which is one of the challenging datasets because, contains text documents and manuscripts in the Arabic language. Since each language is different from other languages and has its own characteristics that impose its own challenges, the standard approach does not apply to all of them. For example, features that work for English do not work for Arabic text because Arabic is a connected text. We have seen that while Latin scripts have been widely used for author identification, the Arabic script family (Arabic, Urdu, Persian, etc.) and Chinese scripts have not received as much attention and still perform well below expectations. This is because these scripts are difficult and complex. Although the results reported in this paper are promising, there is still room for further improvement. Creating an unrestricted dataset with a large number of samples from the class is essential.

REFERENCE

1) Najem, W., & Muhanad, T. Arabic writer identification for children using optimized adversarial-attention and dynamic hybrid classification. 2024 Revista científica de sistemas e informática, 4(2), 1.

2) A. Krizhevsky, I. Sutskever, and G. E. Hinton. ImageNet classification with deep convolutional neural networks. 2017 Communications ACM, vol. 60, no. 6, pp. 84–90, May. doi: 10.1145/3065386.

3) Gaata, Methaq Talib, Younis, Muhanad Tahrir. Hasoon, Jamal N., Mostafa, Salama A., Hessenberg factorization and firework algorithms for optimized data hiding in digital images. 2022 Journal of Intelligent Systems, 31(1), 440-453.

4) H. K. Albahadily, I. A. Jabbar, A. A. Altaay, and X.. Ren, “Issuing Digital Signatures for Integrity and Authentication of Digital Documents”, 2023 Al-Mustansiriyah J. Sci., vol. 34, no. 3, pp. 50–55, Sep., doi: 10.23851/mjs.v34i3.1278.

5) Hannad Y, Siddiqi I, Djeddi C ,El-Kettani MEY. Improving Arabic writer identification using score-level fusion of textural descriptors, 2019.

6) Ahmad, R., Naz, S., Afzal, M. Z., Rashid, S. F., Liwicki, M., & Dengel, A. A deep learning based Arabic script recognition system: benchmark on KHAT. 2020 Int. Arab J. Inf. Technol., 17(3), 299-305.

7) Noubigh, Z., Mezghani, A., & Kherallah, M. Contribution on Arabic handwriting recognition using deep neural network. 2021 In Hybrid Intelligent Systems: 19th International Conference on Hybrid Intelligent Systems (HIS 2019) held in Bhopal, India, December 10-12, 2019 19 (pp. 123-133). Springer International Publishing.

8) Tayeb Bahram, “A texture-based approach for offline writer identification. 2022 Journal of King Saud University Computer and Information Sciences (34) 5204–5222.

9) Hicham Lamtougui, Hicham El Moubtahij et al. An Efficient Hybrid Model for Arabic Text Recognition. 2023 CMC, vol.74, no.2.

10) S. A. Mahmoud, I. Ahmad, M. Alshayeb, W. G. Al-Khatib, M. T. Parvez, G. A. Fink, V. Margner, and H. EL Abed., KHATT: Arabic Offline Handwritten Text Database, 2012. In Proceedings of the 13th International Conference on Frontiers in Handwriting Recognition (ICFHR 2012), Bari, Italy, 2012, pp. 447-452, IEEE Computer Society.

11) Sabri A. Mahmoud, Irfan Ahmad, Mohammed Alshayeb, and Wasfi G. Al-Khatib. A Database for Offline Arabic Handwritten Text Recognition,2011 M. Kamel and A. Campilho (Eds.): ICIAR 2011, Image Analysis and Recognition Part II, LNCS 6754, pp. 397--406. Springer, Heidelberg (2011). DOI: 10.1007/978-3-642-21596-4_40.

12) Khan, M.A. Arabic handwritten alphabets, words and paragraphs per user (AHAWP) dataset. 2022 Data Brief, 41, 107947.