Tech Science Press

Hybrid Trainable System for Writer Identification of Arabic Handwriting

Saleem Ibraheem Saleem[*] and Adnan Mohsin Abdulazeez

Department of Information Technology, Technical Informatics College of Akre, Duhok Polytechnic University, Duhok, 42004, Iraq
[*]Corresponding Author: Saleem Ibraheem Saleem. Email: saleem.saleem@dpu.edu.krd
Received: 30 December 2020; Accepted: 22 February 2021

Abstract: Writer identification (WI) based on handwritten text structures is typically focused on digital characteristics, with letters/strokes representing the information acquired from the current research in the integration of individual writing habits/styles. Previous studies have indicated that a word's attributes contribute to greater recognition than the attributes of a character or stroke. As a result of the complexity of Arabic handwriting, segmenting and separating letters and strokes from a script poses a challenge in addition to WI schemes. In this work, we propose new texture features for WI based on text. The histogram of oriented gradient (HOG) features are modified to extract good features on the basis of the histogram of the orientation for different angles of texts. The fusion of these features with the features of convolutional neural networks (CNNs) results in a good vector of powerful features. Then, we reduce the features by selecting the best ones using a genetic algorithm. The normalization method is used to normalize the features and feed them to an artificial neural network classifier. Experimental results show that the proposed augmenter enhances the results for HOG features and ResNet50, as well as the proposed model, because the amount of data is increased. Such a large data volume helps the system to retrieve extensive information about the nature of writing patterns. The affective result of the proposed model for whole paragraphs, lines, and sub words is obtained using different models and then compared with those of the CNN and ResNet50. The whole paragraphs produce the best results in all models because they contain rich information and the model can utilize numerous features for different words. The HOG and CNN features achieve 94.2% accuracy for whole paragraphs with augmentation, 83.2% of accuracy for lines, and 78% accuracy for sub words. Thus, this work provides a system that can identify writers on the basis of their handwriting and builds a powerful model that can help identify writers on the basis of their sentences, words, and sub words.

Keywords: Writer identification; handwritten Arabic; HOG features; artificial neural network; segmentation; CNN features

1 Introduction

Handwritten Arabic text is the foundation of our writer identification (WI) program. The Arabic language has a wide spectrum of uses, with 347 million individuals collectively speaking it [1]. However, more than 1.2 billion Muslims in the world use Arabic every day to recite the Quran and to pray. Several other languages use Arabic letters in their writing or the same letter types but with small modifications (population of about 700 million) [2]. Examples include Hausa, Kashmiri, Tatar, Ottoman Turkish, Uyghur, Urdu and Cashmiri, Malay, Moriscco, Pashto. Kazakhstan, Kurdish, Punjabian, Sindhi, and Tataran [3]. Arabic is also one of the five mostly commonly spoken languages in the world (including Chinese, French, Spanish, and Hindi/Urdu) [4]. The current research focuses on author recognition (WI) on the basis of handwritten Arabic text. WI is a mental biometric that utilizes a handwritten document in any language. It is regarded as useful because personal handwriting is an early learning and redefined habit. Humans are also valuable biometric tools through the recognition of their signature patterns and the digitalization of theological and historical records, forensics, violent incidents, and terrorist attacks. Technically, the automated recognition of an individual's handwriting may be treated in the same manner as any biometric identification. It involves the retrieval of digital vector representation, the availability of adequate samples from multiple consumers of these vectors, and the presence of a certain measurement/distance between such vectors. Such distance may represent a correlation between the persons from which the samples are collected and correct classification. This research may profit from current WI work utilizing handwriting in other languages, such as English. However, in our attempt to tackle this issue, it's should understand the form and characteristics of Arabic language. The success of existing schemes is influenced by several factors, such as the variations in the national and educational histories of Arabic text authors [5].

Many researchers have conducted writer recognition systems using other languages such as English, persion, and others. Existing approaches rely on a single type of recognition (e.g., online *vs.* offline similarity) or on other text contents and functions [6]. Other studies have utilized articles, chapters, sections, terms, or characters in writer recognition. Certain methods are more involved than others in terms of small sections, such as character or stroke pieces found in text written in Latin, Persian, and other languages. Such methods have been employed in various languages. Researchers should clarify the observable traits of current handwritten text analytics with brief samples. Considering that most of the existing literature is focused on offline text, we structure our analysis according to the collection of text elements that are deemed to be biased and widely used in writer recognition. Our analysis is not strictly focused on Arabic because the growing curiosity in WI lies in Arabic handwritten script. Meanwhile, several preprocessing methods and classification algorithms in many languages are available for use. Although the biometric recognition of a 15-dimensional vector is by no way an inhibitory element, such features are naturally associated. Researchers should demonstrate that all 15 attributes of the application decrease the average accuracy of WI instead of improving it. Therefore, an appropriate approach is either to use a functional selection strategy to achieve the shortest range of functional attributes and thereby guarantee the highest precision or to use a reduction process to achieve smaller met feature vectors, which have linear combinations of 15 dimensions but have considerably low interrelationship attributes. In terms of their discriminatory power, their use a proper system of feature selection to recognize feature rankings and to iteratively fuse the features one by one according to their grades; otherwise, advancement inconsistency may occur [7]. The experimental research in this segment is therefore conducted in stages: (i) the quality of every feature is checked and the single feature test is named; (ii) every feature element in the test is related to weights;

and (iii) numerous characteristics are chosen using an incremental responsive addition method to identify the best set of requirements for ideal identification precision.

The recognition of an Arabic text writer can also be accomplished from a small document irrespective of any logical grouping and building of a sentence. The Arabic text writer identification system is the major one, particularly if the case is small such as the presence of small gaps within a phrase, particularly in cursive text, may result in unclear features may affected as a distinction among two terms. This condition, in effect, leads to inaccurate WI. The main purpose of the current study is to establish and check the efficiency of an automated digital writer recognition scheme for handwritten Arabic text. The program is structured to capture a writer's habit/style, which is learned through training and education. We adopt conventional WI schemes that are used as pattern and biometric recognition functions for different languages. A variety of important structures have been developed over the years with a focus on the features of the world and numerous languages. Nevertheless, existing schemes have faced numerous obstacles, including the restricted reoccurrence of the same terms in one single text. The main contributions of this study can be summarized as follows:

- This study proposes new texture features for WI based on text. Histogram of oriented gradient (HOG) features are modified to extract good features on the basis of the histogram of the orientation for different text angles. These features are fused with convolutional neural network (CNN) features to build a good vector of powerful features. Then, the features are reduced by selecting the best ones using a genetic algorithm. Finally, the normalization method is used to normalize the features and feed them to an artificial neural network (ANN) classifier.

- An efficient and multimodal offline handwriting identification system is proposed in this work. We conduct three types of experiments after preprocessing a given image. First, we utilize the whole image to identify the writer. Our investigation shows that the whole text does not lead to good identification accuracy because we are not focused on the text only as we are dealing with the text and background. The result of this step is an overlap between samples. Second, we extract the sentences by segment and find that the proposed model achieves better results than others. Third, we retrieve good information without the reactive of the background. This stage is described comprehensively herein.

- Our data include 1,000 users, each of whom has 4 samples. We find that this number of samples is not enough to describe a class. Therefore, we randomly select texts from the samples in each class and subsequently generate new samples. We also use rotations in different angles in a process called augmentation. Through this method, we generate around 20 samples for each class.

- We enhance HOG features by using different masks. These features can serve as a good indicator of the angles of edges. We also use a CNN model to generate huge features, select the best ones, and combine them with the HOG features. The ANN is used as a classifier for WI.

The rest of the paper is organized as follows. Section 2 discusses the related work and highlights the research problem. Section 3 introduces the proposed system that can identify writers on the basis of their handwriting. Section 4 presents the experimental results for the preprocessing and segmentation stages for writer recognition based on Arabic handwriting. Finally, Section 5 details the conclusions and future works.

2 Related Work

The feature extraction stage is concerned about the way a text image can be extracted and represented in a series of vectors for use in the classification stage. Up to 163 features may be derived by vectors to obtain good precision in classification [8]. Such features can be held to the minimal point by purifying the collection of unacceptable feature vectors. Feature selection is basically about how reliably broad data can be represented by minimal features of vectors. The inclusion of numerous variables in data analysis results in the culmination of complicated problems. Such massive number of factors also results from efficient computing and broad memory. Therefore, the extraction process blends more than one element with the consequence of satisfaction. The features are broken down according to a set of criteria. The extraction features commonly used in devices for the identification of unique features are defined as follows [9]. Low level: image segmentation, tracking corner, detection of blob, feature extraction, scale-invariant transformation feature; curvature: active contours, parameterized shapes; image motion: image text extraction programs and tools, such as MathWorks, MATLAB, Scilab, and NumPy [10].

For certain words, this is perhaps the most studied method. Authors typically seek to operate for a limited number of chosen words. According to [11] performed WI on the basis of a single term "characteristic" translated into Greek and English. They checked the algorithm for a dependent database of 50 readers with 45 copies of the same title. Unique and morphologically fine colored images of the scanned language have been used in this study. To gain 20-dimensional vectors, scholars have developed horizontal projection templates that are divided into 10 segments and interpreted using two scales of texture features. Classifiers have been used in Bayesian and neural networks. In the English and Greek languages, the proposed approach achieves a 95% precision score. A collection of Latin words was used by [12]. The words included "been," "Cohen," "medical," and "referred." The terms are derived by 1,027 scholars from a transcript of a text. The characteristics derived are level, structural, and concavity characteristics. For sorting, KNN has been used. With this method, the reliability of WI is 83% while the precision is 90.94%. The authors claimed that the degree of recognition of whole written words is higher than that of characters. A WI algorithm of the characteristics of English words was proposed in [12]. A total of 25 different English words written thrice by 1,000 authors were adopted. These words were a sequence of characters. With these definitions, several characteristics have been derived; examples include the GSC, word model recognizer, form curvature (SC), and form description (SCON) types. Relatively long terms strengthen the algorithm to increase efficiency. The algorithm achieves a precision limit of 66% for top 5 authors with the KNN classifier. Thus, letters (G and F) forming terms result in a precision limit of 67% for the Top 5. Words dependent on gradient characteristics also increase precision levels for the Top 10: 82% for authentication and 62% for recognition.

In other dialects, such as Chinese, the number of characteristics is also limited for WI. The study proposed by [13] adopted the well-known principal component analysis dimension reduction system for a Chinese handwriting identification system. A database comprising 40 words that were repeated 10 times by 40 separate readers was adopted. Half of the database was used for device formation; the second half was used for research. The best outcome for a single phrase was an accuracy rate of 86.5% while the combination of 10 words achieved a precision rate of 97.5%. Another study proposed an algorithm for Arabic identification on the basis of the characteristics from images of scanned words [14]. Such attributes involve multiple edge angles, instant invariants, and word dimensions that are also recognized as word structural characteristics (e.g., distance, weight, height, baseline weight to upper edge, and underlying length to the bottom edge).

A database with 27 words copied 20 times by 100 authors was also used in the study. One-fourth of this database was checked during the remainder of the activities. A classifier was used as the K-nearest neighbor. Even the top 10 identifiers achieve a consistency rate of 90% for common terms and 53%–75% for other names.

A few investigators have tried to extract characteristics from pages as the next natural concentration of WI from text. The process may be called a combination of similar line-based WI systems but with numerous branches. In a study by [15] on an offline system of text-independent Arabic identification, an A4 page compiled by 25 authors was used. Each page was divided into four sections with three training blocks and one test block. The texts were preprocessed, and functionality with Gabor filters was extracted. The range parameters are Euclidean, weighted Euclidean, and X2. The latest project that uses a reliable 40 databases by reports 82.5% accuracy in its identifying rate [16]. Arabic article author recognition schemes were suggested in [17] by utilizing various methods of extraction, such as hybrid spectral statistical tests, multichannel (Gabor), and gray-level GLCM similarity matrix. The characteristics were used to determine the best feature subset. Many classifiers, such as the support vector machine (SVM), linear combined distance, and KNN, have been used widely in existing studies. The classification accuracies of these classification methods are 57%, 47%, 69%, and 90%. The program used a database created by 20 authors who were required to compose two A4 versions: one for machine preparation and another for checking. "Text-independent" writing naming method for Latin and Farsi was also proposed by [18]. For extracting features, the authors used Gabor filters. To assess the system, the authors asked 100 writers to write five different pages in Farsi. Thirty other handwritten texts were chosen from the IAM database, in which seven distinct pages were requested by each writer. The study used 60% of the texts to experiment with the proposed system, and the remaining was used to test it. Top1 achieved a 98% precision for the Farsi test and 94.4% precision for the English exam. Another study by [19] proposed an offline Farsi text-independent WI program. Their task was founded on the characteristics of gradients used with Latin characters in the past. The system was tested on 250 manuscript samples, which were collected from 50 authors, each of whom wrote five different sheets. The use of neural networks as a classificatory resulted in 94% accuracy relative to the use of a fuzzy cluster classificatory with 90% accuracy [20].

Existing studies have clearly obtained high WI levels, and they are expected to catch more author habits in any dialect with the accessibility of samples or longer texts from the same writers. However, WI with short texts and/or few examples is by far the most complicated and is generally highly demanded in various areas, such as crime and forensic science. Many researchers have conducted writer recognition systems using other languages such as English, Persian, and others. Existing approaches rely on a single type of recognition (e.g., online *vs.* offline similarity) or on other text content and functions, as explained and mentioned in the preceding chapter (text dependent *vs.* nontext dependent). Other studies have utilized articles, chapters, sections, terms, or characters in writer recognition approach. Certain methods are more involved than others in terms of small sections, such as character or stroke pieces found in text written in Latin, Persian, and other languages. Such methods have been employed in various languages. Researchers should clarify the observable traits of current handwritten text analytics with brief samples. Considering that most of the existing literature is focused on offline text, we structure our analysis according to the collection of text elements that are deemed to be biased and widely used in writer recognition. Our analysis is not strictly focused on Arabic because the growing curiosity in WI lies in Arabic handwritten script.

3 The Proposed Method

Analyzing the related works, we notice that styles and habits in handwriting are embedded in specific components/parts of written texts in entire languages while the arrangement of these parts inside paragraphs reflect significant signs about the identity of the writer. For Arabic writers, word-based WI is one of the most popular techniques in the literature. Similarly, we find that WI was traditionally dependent on word(s), character(s), paragraph(s), line(s), or/and a portion of a character. When handwritten, the Arabic language is a cursive language. Different from the case of other languages, every Arabic word contains one or more subwords, and there is a related stroke at the end of every subword. Several experts believe that the strokes of subwords are outstanding characteristics that are unique to writers and thus suggest their consideration in recognizing Arabic text writers. The diversity of diacritics, which need to be written within subwords, and words are distinct characteristics of Arabic writing; without them, pronunciation and meaning become challenging in several situations. Many people speak Arabic fluently ignore those whose original language is Arabic. Recognized subwords are common in any written text, and several subwords occur as a portion of numerous words or as single words on their own. This information inspires the improvement of our hypothesis that subwords for WI could show important development in precision relative to current methods. The remainder of the study is dedicated to testing the validity of our hypothesis and improving the advanced technique to achieve optimum performance.

This study is mainly focused on automatic WI based on handwritten text. In automating Arabic WI, the initial challenges include the following:

(1) Enhancing images and highlighting text;
(2) Segmenting written text into its words and lines;
(3) Obtaining related measurements to represent digital feature vectors of various words;
(4) Defining the recognition pattern strategy to be accepted. We need to present observed clues to determine the validity of the declared hypothesis when the automatic WI system is advanced. This requirement calls for the employment of standard experimental biometric procedures and assessment measures utilizing suitable databases.

Fig. 1 shows the key topics in this study. At the preprocessing stage, in addition to the simple image preparations, such as trimming the region outside the text, this study focuses on how to remove noise and unwanted objects, improve images, and rotate images to the correct position while preserving the salient tissue boundaries in the image bounds. The purpose is to find the structural boundaries and the locations of structures accurately. The first stage in our proposed model involves capturing scanned images and using them to pretrain and evaluate the system. The captured text images include different types of noise appearing as text and poor-quality writing, all of which will increase the likelihood of missing the correct text structure. In addition, poor-quality images might present a challenge because some of them could have dark regions while others may be of good quality. In some cases, we may have dark and light regions in the same image. Moreover, text orientation may be changed by scanning written text, or the writer may not be able to control his hand to write on lines. Therefore, in the preprocessing stage, we propose a model that enhances the image, highlights the text, and then increases the contrast of the dark regions. As shown in Fig. 1, the captured image is used as an input for the preprocessing stage to enhance the image quality and make the text clear. In this stage, we enhance the histogram equalization technique to improve the image quality. Then, binarization is performed to obtain the text as a binary matrix. This stage will help us estimate the angle of rotations and use it to

return the image to the correct angle. This stage requires scanning the whole image and using the training data and trainable model to identify the pixels belonging to the text. The outcome of this stage is a binary image. This step is followed by text orientation that can be changed by scanning written text. This stage begins with an investigation into the text orientation (skew) problem. For the post segmentation stage, we extract the lines and words from the binarized image. We then evaluate the existing features to improve the affective one and use it in our proposed system.
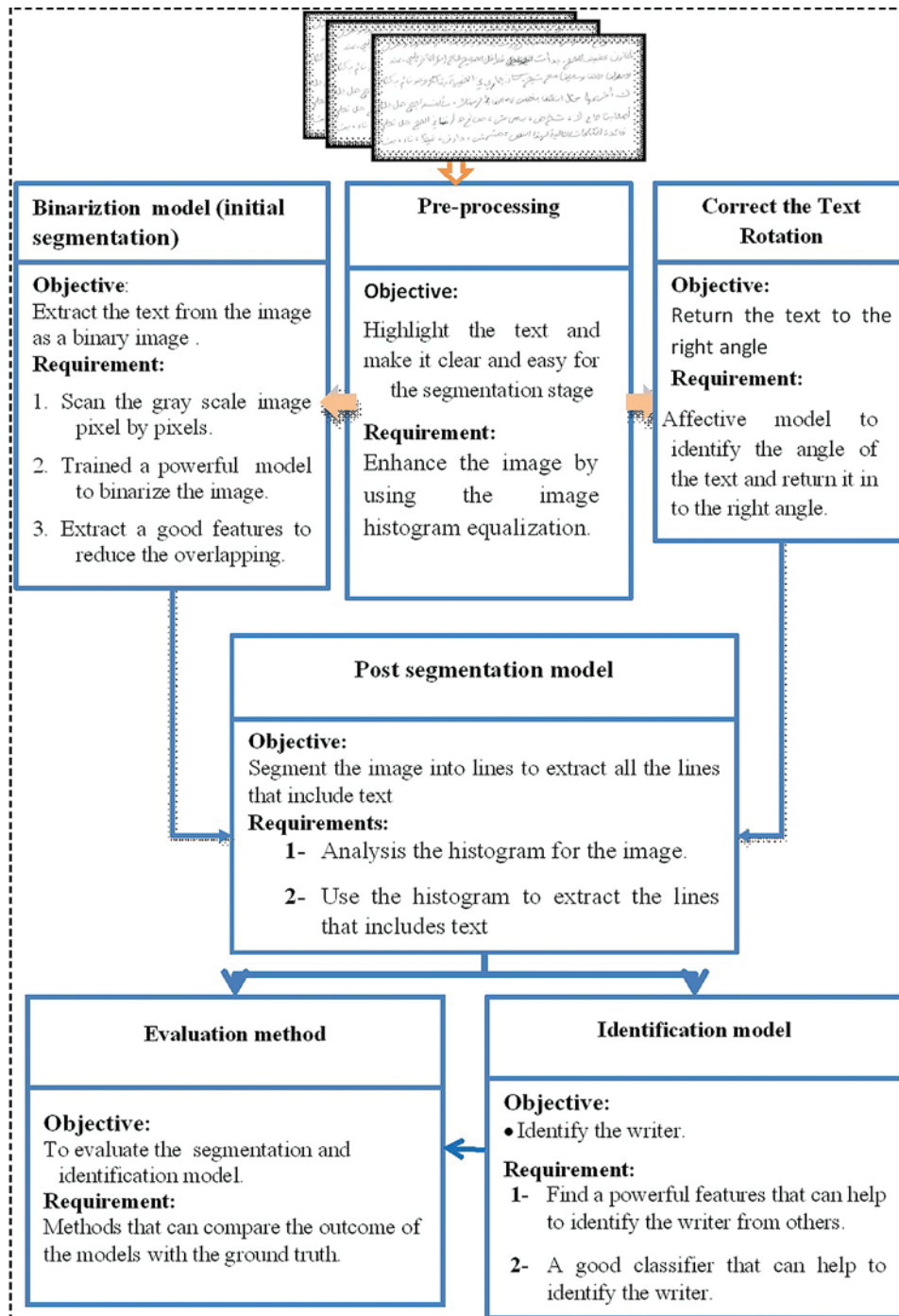

Binariztion model (initial segmentation)

Objective:
Extract the text from the image as a binary image .
Requirement:

1. Scan the gray scale image pixel by pixels.

2. Trained a powerful model to binarize the image.

3. Extract a good features to reduce the overlapping.

Pre-processing

Objective:
Highlight the text and make it clear and easy for the segmentation stage

Requirement:
Enhance the image by using the image histogram equalization.

Correct the Text Rotation

Objective:
Return the text to the right angle
Requirement:

Affective model to identify the angle of the text and return it in to the right angle.

Post segmentation model

Objective:
Segment the image into lines to extract all the lines that include text
Requirements:
1- Analysis the histogram for the image.

2- Use the histogram to extract the lines that includes text

Evaluation method

Objective:
To evaluate the segmentation and identification model.
Requirement:
Methods that can compare the outcome of the models with the ground truth.

Identification model

Objective:
• Identify the writer.

Requirement:
1- Find a powerful features that can help to identify the writer from others.

2- A good classifier that can help to identify the writer.

Figure 1: Main focus of WI of Arabic handwriting system

We study handwritten Arabic text segmentation and the usage of lines and sub words (containing individual diacritics as a single text unit) for WI based on handwritten Arabic text sub words. In this section, we discuss text segmentation for partitioning text into lines and sub words with diacritics. We check the versions of lines and sub word performance depending on WI schemes with varying numbers of sub words utilized. Fig. 2 shows the identical sub word projection profile written by an identical writer (once their diacritics are involved). Figs. 2a and 2c show the horizontal and vertical projections for the sub word in Fig. 2a, respectively. Figs. 2b and 2d show the vertical and horizontal projections for the sub word in Fig. 2b, respectively.
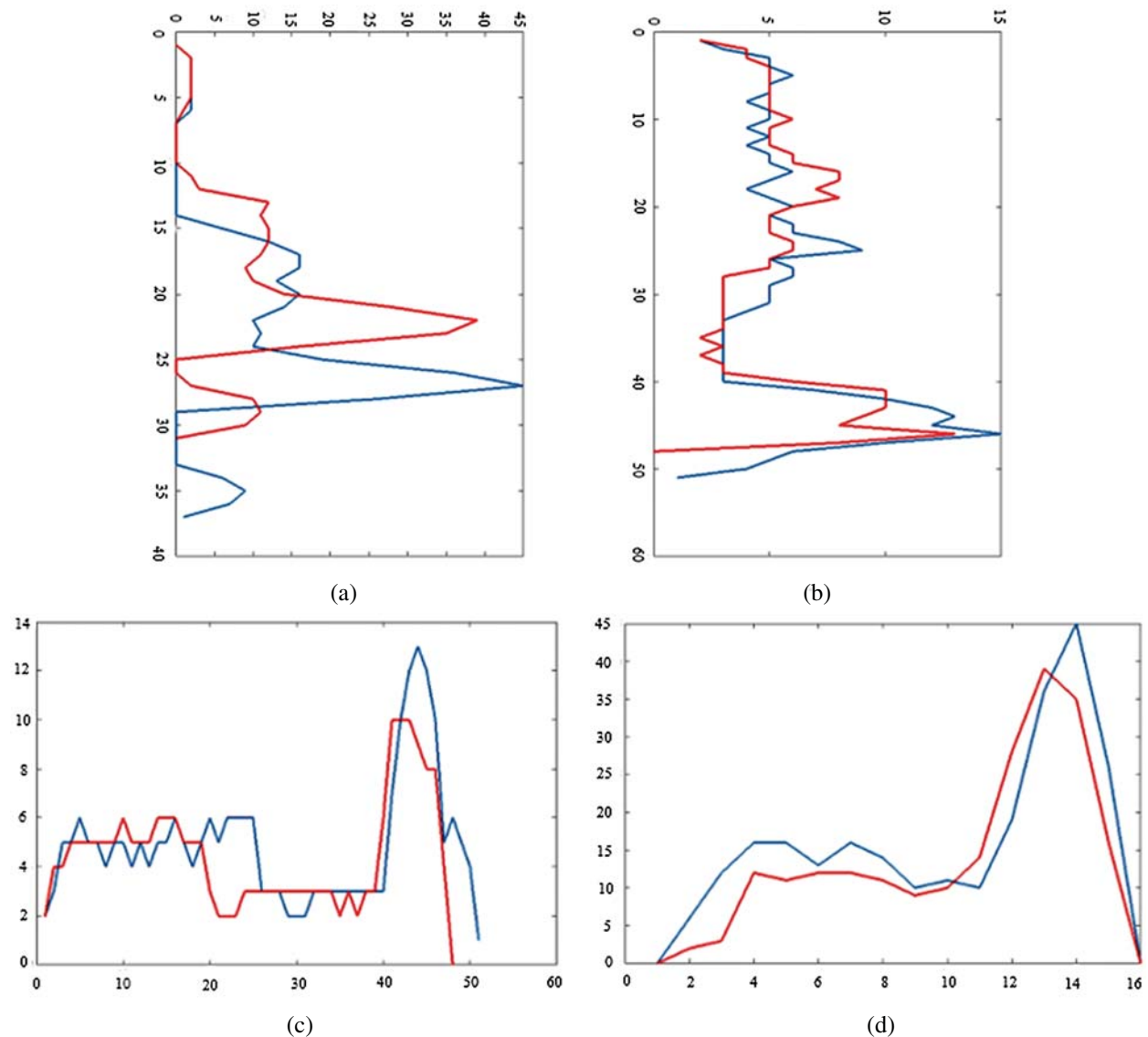

Figure 2: Projections of lines and words for randomly selected text (a) Projections of lines for sample 1 (b) Projections of lines for sample 2 (c) Projections of words for the first line in sample 1 (d) Projections for the second line in sample 2

These differences cause variations in the features derived from several samples of written sub words that would result in false acceptance and/or false rejection in certain situations. Thus, this issue may exert an opposing influence on the accuracy ratio of WI. In our proposed model, we extract two types of features.

- First, we use the CNN features with the SVM classifier for WI.
- Second, we use HOG features as signs for WI. We combine the two features and feed them to the SVM classifier to build a good classifier that can facilitate WI.

Being hierarchical, neural multilayer networks with a strongly supervised learning architecture that is trained using backpropagation algorithms, CNNs consist of trainable classifiers and automated feature extractors. CNNs are used to learn high-dimensional, complex information, and they vary in terms of the subsampling and convolutional layers investigated. The alteration is in their architecture. Several architectures of CNNs are recommended for various challenges in handwriting character/digit recognition and object recognition. They achieve superior pattern recognition task performance. Additionally, CNNs combine three major hierarchical features, namely, spatial subsampling, weight sharing, and local receptive regions, to ensure a degree of invariance to distortion, shift, and scale. As presented in Fig. 3, the net denotes a standard CNN architecture for handwritten character recognition. It contains a group of many layers. First, the input is transformed with a series of filters (C hidden layers) to obtain the feature map values. Second, every convolution layer is followed by a layer of subsampling to reduce the dimensionality (S hidden layers) of the feature map spatial resolution. Convolutional layers substituting subsampling layers establish extractor features to recover the features from raw images. Finally, the layers are followed by two fully connected layers (FCLs) and an output layer. The former layer's output is obtained through every layer as the input.
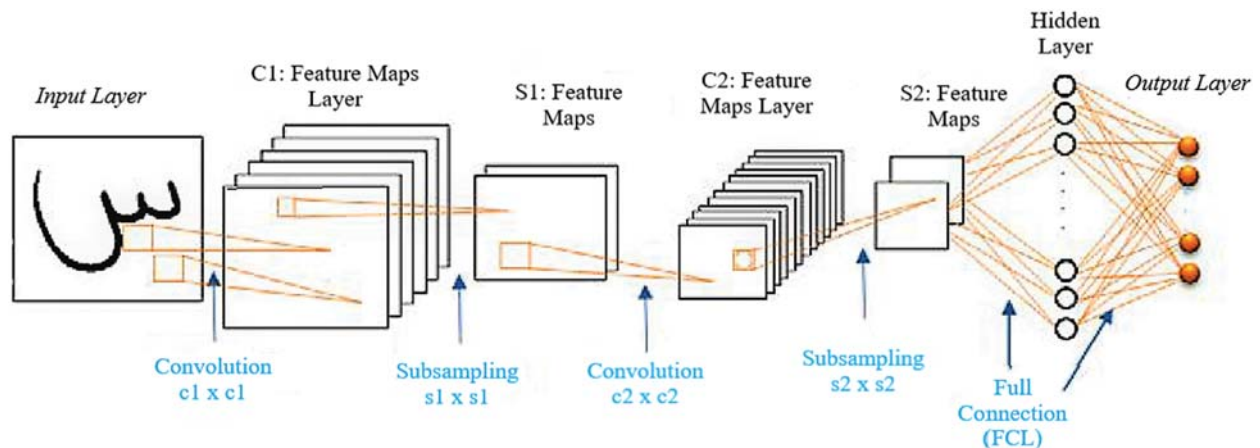
Figure 3: Proposed CNN architecture using many features in the WI of Arabic handwritten system

SVM is an effective discriminative classifier that has been extensively used with progressive outcomes in several pattern recognition/classification activities [21–25]. It is considered as the advanced tool to resolve nonlinear and linear (Fig. 4) challenges in classification [26], and its prediction capacity, global optimum character, parsimony, and flexibility are widely recognized. Structural risk minimization is the foundation of its formulation rather than the minimization of empirical risk that is usually utilized in ANNs [27]. ANN is principally utilized for determining

the best splitting hyperplane (Eq. (1)) or decision surface through the acceptance of a new method depending on the points of sample mapping to a high-dimensional feature space. It is characterized using a nonlinear transformation Φ even when the information is inseparable linearly. The best hyperplane is obtained by fixing a quadratic programming challenge that is dependent on regularization parameters. This transformation is performed using kernel functions, such as sigmoid, polynomial kernel types, and linear and radial basis functions.

Linear kernel: $K(x, y) = x \times y$ (1)

Polynomial kernel: $K(x, y) = [(x \times y + 1)]^d$ (2)

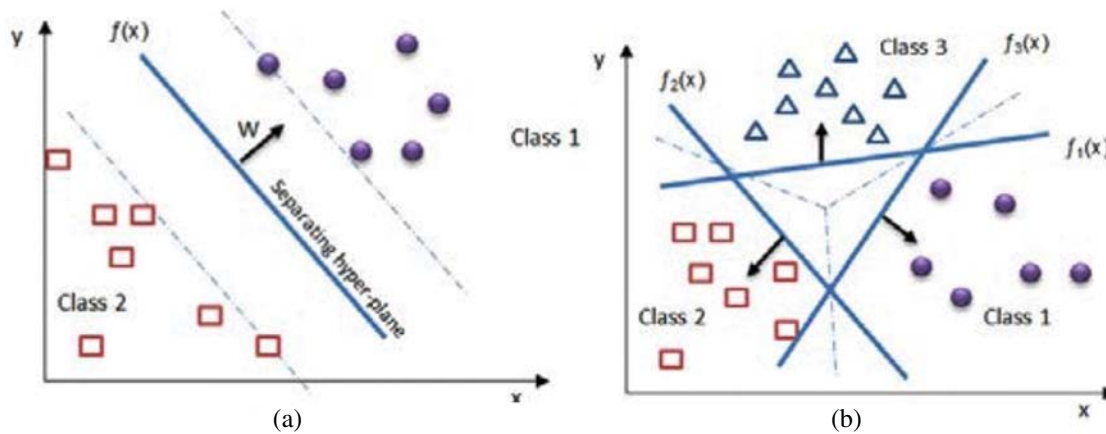Sigmoid kernel: $K(x, y) = \tanh(\beta_0 xy + \beta_1)$ (3)

Radial basis function (RBF) kernel:

$$K(x, y) = \exp(-\gamma \parallel x - y \parallel^2)$$ (4)

Here, d, β_0, β_1, and γ are parameters to be determined empirically.

$$f(x) = W^T \emptyset(x) + b$$ (5)

where $W \in R^n$, $b \in R$, and $\emptyset(x)$ is a feature map.

Figure 4: Basics of SVM approach (a) two classes of hyperplanes, (b) one-*vs.*-all method

The design of our OAHR system depending on SVM and CNN is discussed in this section. The CNN is treated as a deep learning algorithm, in which the dropout method is implemented through training. Our suggested scheme is fitted by replacing the trainable CNN classifier with an SVM classifier. Our aim is to combine the respective capacities of the CNN and SVM so as to obtain a novel and effective recognition system on the basis of two formalisms. We display the CNN architecture-based SVM model in Fig. 5.

- The initial layer accepts the pixels of the raw image as an input.
- The fourth and second network layers are convolution layers alternating with subsampling layers that generate pooling maps as input. Hence, they are capable of obtaining features that are generally unchanged for the local image transformation input. FCL represents the sixth layer containing N neurons. The last layer is replaced through SVM with an RBF kernel for classification. Given the use of an enormous number of parameters and

data, overfitting becomes likely. To protect our network from this challenge and improve it further, we employ the dropout method. This method involves the momentary elimination of a unit from the network. The eliminated unit is only chosen accidentally through training. Dropout is only used at the FCL layer and is specifically applied to feedforward contact perceptron. This selection is dependent on the point because convolutional layers do not have many parameters and overfitting is not a challenge; in this case, the dropout method could not exert much impact. The hidden units' outputs are obtained through the SVM as a feature vector for training. Then, training is performed until decent training is achieved.

- Lastly, the test set classification is carried out through the SVM classifier with obtained features automatically. The adopted CNN structure-based SVM model with dropout is shown in Section 4.
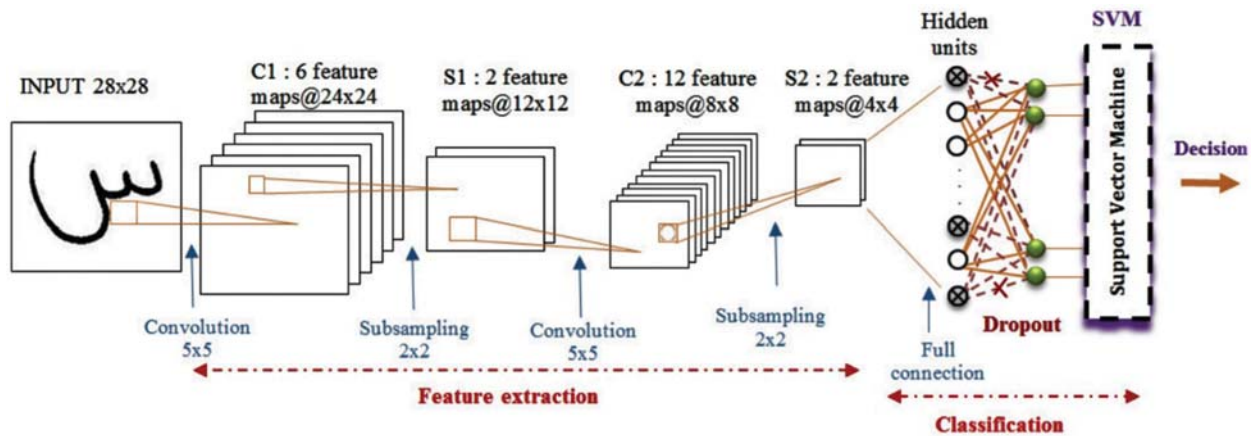

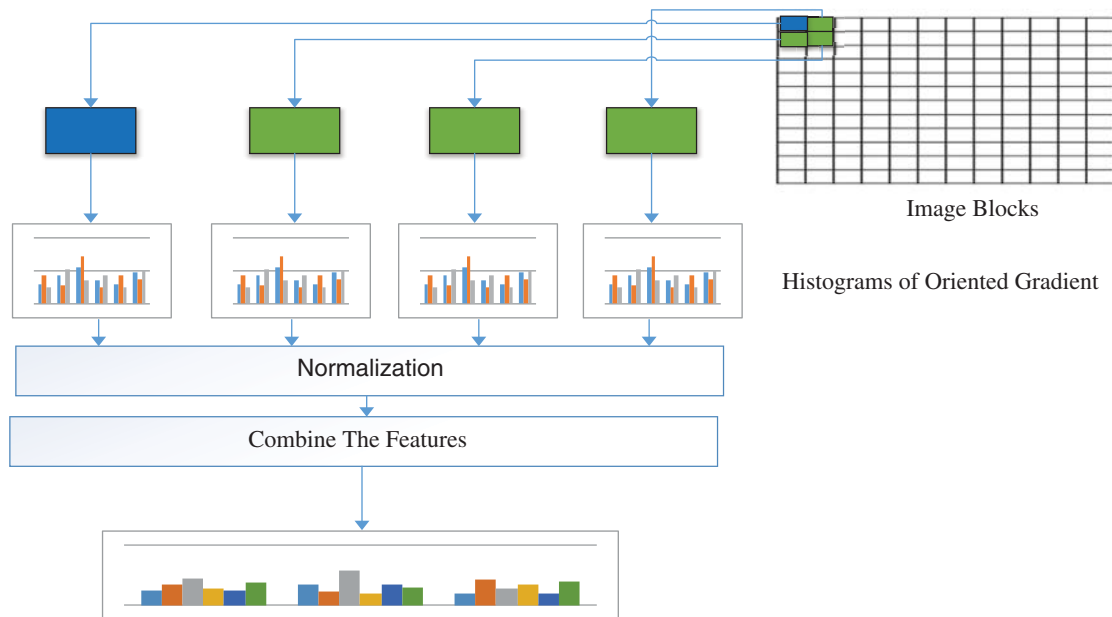**Figure 5:** Proposed CNN-based SVM architecture

Figure 6: HOG process

HOG features are used as the second features in WI by adopting the same SVM classifier. Then, the HOG features are combined with the CNN features to build a good classifier. On the basis of our investigation, we use the feature selection technique to select the best features among the CNN and HOG features for WI. Fig. 6 shows the process of using HOG features.

4 Experiment Results

A wide range of experiments are carried out to evaluate the efficiency of the proposed segmentation method, that is, the HOG and deep learning method for WI. We compare its execution with that of other methods in related works. The datasets utilized in this study are described in terms of the experiments used and the application of data augmentation to increase the number of samples. The execution of the proposed segmentation framework is assessed. Finally, broad tests are performed to discover the optimal trainable model for WI. The model is then compared with existing methods in terms of rank 1 identification rate and identification time.

4.1 Experiment Setup

Extensive experiments are conducted to evaluate the WI of Arabic handwriting using the hybrid trainable model. Specifically, 2D maximum embedding difference is utilized in feature extraction. The experiments are performed in MATLAB using a 64-bit Windows 10 Proficient machine equipped with an Intel Pentium 2.20 GHz processor and 8 GB RAM. The sum of 10 distinctive tests is used to evaluate WI of Arabic handwriting from different datasets.

4.2 Data Preparation

An unlimited handwritten Arabic text database published by 1,000 individual authors contains a dataset that applies to a list of KHATT's (KFUPM Handwritten Arabic Text). This database was developed by Professor Fink from Germany, Dr. Märgner, and a research group headed by Professor Sabri Mahmoud from Saudi Arabia. The database comprises 2,000 images with identical texts and 2,000 images of different texts with the text line images removed. The images are followed by a manually verified reality and a Latin ground-truth representation. For other handwriting recognition-related studies, including those on writer detection and text recognition, the same database may be used. For further details about the selected database, reviewers may refer to [28,29]. For academic purposes, version 1.0 of the KHATT platform is open to academics. Description of the database: (1) forms written by a thousand different authors; (2) scanned in various resolutions (200, 300, and 600 DPI); (3) authors from various cultures, races, age classes, expertise levels, and rates of literacy with accepted writings with clear writing styles; (4) 2,000 specific images of paragraphs (text origins of different subjects, such as art, education, fitness, nature, and technology) and their segmented line images; (5) 2000 paragraphs, each of which includes all Arabic letters and types, and their line images, including related materials.

Free paragraphs written by writers about their selected topics: (i) manually checked ground realities are given for paragraph and line images; (ii) the database is divided into three rare collections of testing (15%), education (70%), and validation (15%); and (iii) encourage research in areas such as WI, line segmentation, and noise removal and binarization techniques in addition to manual text recognition. Fig. 7 shows four samples from one writer, that is, KHATT_Wr0001_Para1, where the first part of the label refers to the name of the data set (KHATT), the second part refers to the number of the writer, and the third part refers to the paragraph number. In Fig. 7, we introduce four images that include texts from the same writer. The first and second images involve similar text paragraphs, and the remaining images show different texts. Fig. 7 also

shows paragraphs from a second set of writers. The texts in the first and second images are the same, but they are written by different writers. As explained previously, the database includes 4,000 samples from 1,000 writers, each of whom produced 4 samples. In this study, forms are gathered from 46 various sources covering 11 various topics. Tab. 1 presents the sources' topics, the number of gathered passages in each topic, and the number of pages from which the passages are extricated. The forms were filled out by individuals from 18 nations. The distribution of the number of forms by nation is presented in Tab. 2. Individuals from Syria, Bahrain, Algeria, Sudan, Australia, Libya, Canada, Lebanon, and Oman also filled out the forms.

Figure 7: Sample of four images of text from the same writer. The first and second images show similar text paragraphs, and the remaining images use different text. The sample on the right shows paragraphs from second writers

Table 1: Source data's topics, paragraphs, and source pages

Category	Reference	Paragraph	Pages
Art	3	399	182
Economy	4	78	59
Education	1	46	4
Health	3	105	72
History	6	174	107
Literature	5	635	319
Management	5	86	128
Nature	7	136	70
Social	5	130	720
Technology	4	189	103
World	3	22	14
	46	**2000**	**1788**

Table 2: Writers' countries of origin

Country	Number of writers (%)	Country	Number of writers (%)
Saudi Arabia	675 = 67.5	USA	16 = 1.6
Morocco	90 = 9	Egypt	13 = 1.3
Jordan	81 = 8.1	Tunisia	13 = 1.3
Yemen	45 = 4.5	Kuwait	11 = 1.1
Palestine	29 = 2.9	Others	27 = 2.7

On the basis of the data, we build and test our proposed model. As mentioned in the previous section, the deep learning method is used to identify writers, and its features are fused with HOG features. Deep learning is influential, but it commonly needs to be trained on gigantic sums of samples to fully assesment. Such requirement can be considered as a main restriction. Deep learning methods prepared on small databases appear lower execution within the angles of flexibility and generalization from approval and testing samples. Subsequently, these methods endure from issues caused by overfitting. A few strategies have been proposed to diminish the overfitting issue. Data augmentation, which increments the sum and differing qualities of information by "augmenting" them, is an successful procedure for decreasing overfitting in models and making strides the differing qualities of datasets and the generalization execution. Within the application of image classification, different enhancements have been made; cases incorporate flipping the picture horizontally or vertically and deciphering the picture by many pixels. Data augmentation is the best and most common technique for image classification. It falsely extends the training database by utilizing distinctive procedures. In this section, we propose an augmentation method that can help to enlarge the training data. Fig. 8 shows the proposed augmenter.

As shown in Fig. 8, we use two methods for generating more samples from the training samples. The first method is based on images generated by using the line text of images and reconstructing new images. In this method, we randomly select the lines and obtain new images that include different orders of text lines. Then, rotation augmentation is performed by rotating the image to right or left on an axis between 1° and 359°. The safety of rotation augmentation

is heavily determined by the rotation degree parameter. Slight rotations, such as that between 1 and 20 or between -1 and -20 could be useful in digit recognition tasks, such as MNIST. However, as the rotation degree increases, the label of the data is no longer preserved after the transformation [30–39].
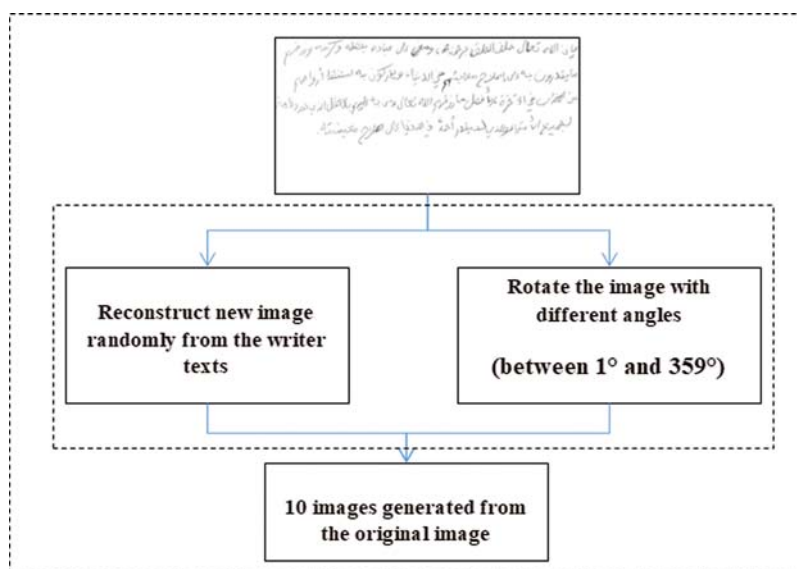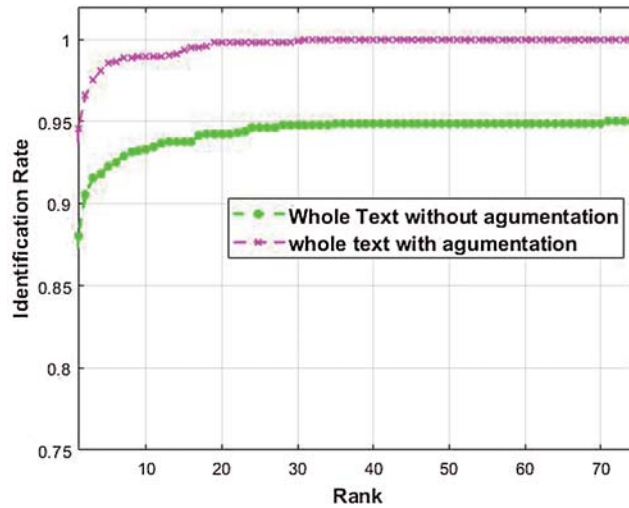

Figure 8: Proposed augmenter

4.3 Writer Identification Stage Results

The writer identification scheme implements multiple assessment from a big dataset with handwriting images of distinguished writings, and comeback a conceivable grade of candidates. Meanwhile, in the case of writer identification, the method includes a one-to-one comparison with clarification as to whether or not the two tests were composed by the same individual. One of the methods that have been used to enhance existing models and provide enough data for identification is the augmentation method. The cumulative match characteristic (CMC) curve is a metric used to measure the performance of identification and recognition algorithms on the basis of the precision for each rank. It is a popular measurement in biometric systems in which identification performance is measured according to the relative ordering of match scores corresponding to each sample (in closed-set identification). As shown in Fig. 9, the CMC is applied to the output of two experiments. The first experiment involves the whole text and the original data. The second experiment involves the use of the augmentation method to increase the training dataset. The method obviously improves the results of the proposed model.

We evaluate the KHATT dataset containing 4,000 Arabic handwritten documents from 1,000 authors. For each writer, we take one image for testing. The testing stage includes 1,000 samples. We conduct four experiments that include the whole text for WI without augmentation. The same experiments are performed on the dataset by using the proposed augmenter to increase the training samples. The model is then used in WI by using the segmented line. The same model is applied to the sub words. We provide a list of abbreviations used to generate the tables. These abbreviations are as follows: N. S: number of samples, CI: correct identification, MCN:

misclassified samples, and AC: accuracy. Tab. 3 and Fig. 10 show the results of the proposed system by using the whole text with and without the augmenter. Fig. 10 reveals that the augmenter enhances the results for the HOG features and ResNet50, as well as the proposed model, because of the increase in the amount of data. The augmenter also helps the system to secure further information about the nature of writing patterns.

Figure 9: CMC curves showing the identification accuracy of whole text images with and without augmentation

Table 3: Accuracy of proposed model for augmented and non-augmented data

	HOG and ANN model				ResNet50				HOG and CNN features			
	N. S	CI	MCS	AC	N. S	CI	MCS	AC	N. S	CI	MCS	AC
Whole paragraph without augmentation	1000	579	421	57.9	1000	792	208	79.2	1000	854	146	85.4
Whole paragraph with augmentation	1000	723	277	72.3	1000	803	197	80.3	1000	942	58	94.2
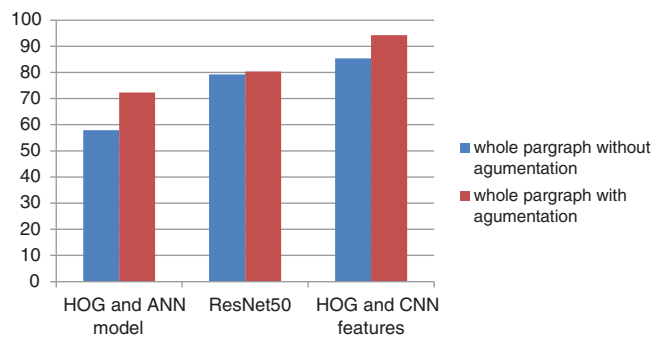
Figure 10: Accuracy of proposed model for augmented and non-augmented data

Tab. 4 shows the affective results of the proposed model for whole paragraphs, lines, and sub words by using different models. We compare these results with those of the CNN and ResNet50. The whole paragraph produces superior results across all models because it provides rich information and the model can have more features for different words.

Table 4: Affective results of models for whole paragraphs, lines, and sub words

	HOG and ANN model				ResNet50				HOG and CNN features			
	N. S	CI	MCS	AC	N. S	CI	MCS	AC	N. S	CI	MCS	AC
Whole paragraph with augmentation	1000	723	277	72.3	1000	803	197	80.3	1000	942	58	94.2
Lines	1000	621	379	62.1	1000	711	289	71.1	1000	832	168	83.2
Sub words	1000	594	406	59.4	1000	621	379	62.1	1000	780	220	78

We test the proposed model on different lines and find that these lines can provide good identification results. The result for each line are obviously close to others. In our future work, we will conduct further investigations to identify affective features that can facilitate WI based on text lines and sub words with high accuracy as shown in Tab. 5.

Table 5: Affective results of models for whole lines

	HOG and ANN model				ResNet50				HOG and CNN features			
	N. S	CI	MCS	AC	N. S	CI	MCS	AC	N. S	CI	MCS	AC
Line1	1000	615	385	61.5	1000	693	307	69.3	1000	822	178	82.2
Line2	1000	632	368	63.2	1000	724	276	72.4	1000	820	180	82
Line 3	1000	625	375	62.5	1000	699	301	69.9	1000	797	203	79.7

5 Conclusion

This study proposes new texture features for WI based on text. The HOG features are modified to extract good features by extracting the histogram of orientation for different angles of texts. These features are fused with the CNN features to build a good vector of powerful features. Then, the features are reduced by selecting the best ones using a genetic algorithm. The normalization method is used to normalize the features and feed them to the ANN classifier. This study presents all experimental results to accomplish the proposed objectives. Series of experiments are conducted using a database comprising 2,000 images with identical texts and 2,000 images with different texts and with their text line images removed. The images are followed by a manually verified reality and a Latin ground-truth representation. For other handwriting recognition-related studies, such as those on writer detection and text recognition, the same database may be used. The proposed technique is employed to evaluate WI of Arabic handwriting by using the hybrid trainable model. The experimental results are analyzed, interpreted, compared, validated, and discussed in depth. Different analyses and discussions of the results are also performed. The strengths and weaknesses of the proposed technique are highlighted accordingly. One of the limitations of this study is the preprocessing stage, which should help us to highlight text and distinguish between texts and background. However, this type of method exerts a negative effect on the dataset, in which text in some images may be missed and may even disappear. This point results in the segmentation method being unaffected for these types of images. In addition, our

proposed trainable method needs samples for the training stage. The expert using the method should provide system samples from different types of text and select the samples from different regions with inherent challenges. For our future work, we will use frequency domain filters to enhance images and highlight texts. By using this method, we can realize our two objectives, namely, to remove noise and highlight text and to recover the regions with dark areas. We will also use the frequency domain to identify and remove unwanted objects.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] K. Fergani and A. Bennia, "New segmentation method for analytical recognition of arabic handwriting using a neural-markovian method," *International Journal of Engineering and Technologies*, vol. 14, pp. 14–30, 2018.

[2] S. K. Jemni, Y. Kessentini and S. Kanoun, "Out of vocabulary word detection and recovery in arabic handwritten text recognition," *Pattern Recognition*, vol. 93, pp. 507–520, 2019.

[3] S. N. Srihari and G. Ball, "An assessment of arabic handwriting recognition technology," in *Guide to OCR for Arabic Scripts*. London, UK: Springer, pp. 3–34, 2012.

[4] A. Cheikhrouhou, Y. Kessentini and S. Kanoun, "Hybrid HMM/BLSTM system for multi-script keyword spotting in printed and handwritten documents with identification stage," *Neural Computing and Applications*, vol. 32, no. 13, pp. 9201–9215, 2020.

[5] M. Elleuch, R. Maalej and M. Kherallah, "A new design based-svm of the cnn classifier architecture with dropout for offline arabic handwritten recognition," *Procedia Computer Science*, vol. 80, pp. 1712–1723, 2016.

[6] S. Husham, A. Mustapha, S. Mostafa, M. Al-Obaidi, M. Mohammed *et al.,* "Comparative analysis between active contour and otsu thresholding segmentation algorithms in segmenting brain tumor magnetic resonance imaging," *Journal of Information Technology Management*, vol. 12, pp. 48–61, 2020.

[7] Y. Qawasmeh, S. Awwad, A. F. Otoom, F. Hanandeh and E. Abdallah, "Local patterns for offline arabic handwritten recognition," *International Journal of Advanced Intelligence Paradigms*, vol. 16, no. 2, pp. 203–215, 2020.

[8] H. Ghadhban, M. Othman, N. Samsudin, M. Ismail and M. Hammoodi, "Survey of offline arabic handwriting word recognition," in *Int. Conf. on Soft Computing and Data Mining*, Langkawi, Malaysia, pp. 358–372, 2020.

[9] A. Chahi, Y. Ruichek and R. Touahni, "Cross multi-scale locally encoded gradient patterns for off-line text-independent writer identification," *Engineering Applications of Artificial Intelligence*, vol. 89, pp. 103459, 2020.

[10] Z. Kaoudja, B. Khaldi and M. L. Kherfi, "Arabic artistic script style identification using texture descriptors," in *1st Int. Conf. on Communications, Control Systems and Signal Processing*, El Oued, Algeria: IEEE, pp. 113–118, 2020.

[11] E. Zois and V. Anastassopoulos, "Morphological waveform coding for writer identification," *Pattern Recognition*, vol. 33, no. 3, pp. 385–398, 2000.

[12] B. Zhang and S. Srihari, "Analysis of handwriting individuality using word features," *Document Analysis and Recognition*, vol. 2, pp. 1142, 2003.

[13] L. Zuo, Y. Wang and T. Tan, "Personal handwriting identification based on pca," in *Second Int. Conf. on Image and Graphics*, Hefei, China: International Society for Optics and Photonics, pp. 766–771, 2002.

[14] S. Al-Ma'adeed, E. Mohammed, D. Al Kassis and F. Al-Muslih, "Writer identification using edge-based directional probability distribution features for arabic words," in *Int. Conf. on Computer Systems and Applications IEEE/ACS*, Doha, Qatar, pp. 582–590, 2008.

[15] F. Shahabi and M. Rahmati, "Comparison of gabor-based features for writer identification of farsi/arabic handwriting," in *Tenth Int. Workshop on Frontiers in Handwriting Recognition*, La Baule, France, pp. 104466, 2006.

[16] F. Shahabi and M. Rahmati, "A new method for writer identification and verification based on farsi/arabic handwritten texts," in *Ninth Int. Conf. on Document Analysis and Recognition*, Parana, Brazil, IEEE, pp. 829–833, 2007.

[17] A. Al-Dmour and R. Zitar, "Arabic writer identification based on hybrid spectral—Statistical measures," *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 19, no. 4, pp. 307–332, 2007.

[18] B. Helli and M. E. Moghadam, "Persian writer identification using extended gabor filter," in *Int. Conf. Image Analysis and Recognition*, Póvoa de Varzim, Portugal, Springer, pp. 579–586, 2008.

[19] S. Ram and M. Moghaddam, "A persian writer identification method based on gradient features and neural networks," in *2nd International Congress on Image and Signal Processing*, Tianjin, China, IEEE, pp. 1–4, 2009.

[20] S. Ram and M. Moghaddam, "Text-independent persian writer identification using fuzzy clustering approach," in *Int. Conf. on Information Management and Engineering*, Kuala Lumpur, Malaysia, IEEE, pp. 728–731, 2009.

[21] M. Mohammed, S. Gunasekaran, S. Mostafa, A. Mustafa and M. Ghani, "Implementing an agent-based multi-natural language anti-spam model," in *Int. Sym. on Agent, Multi-Agent Systems and Robotics*, Putrajaya, Malaysia, IEEE, pp. 1–5, 2018.

[22] S. Mostafa, A. Mustapha, A. Hazeem, S. Khaleefah and M. Mohammed, "An agent-based inference engine for efficient and reliable automated car failure diagnosis assistance," *IEEE Access*, vol. 6, pp. 8322–8331, 2018.

[23] M. Mohammed, M. Ghani, S. Mostafa and D. Ibrahim, "Using scatter search algorithm in implementing examination timetabling problem," *Journal of Engineering and Applied Sciences*, vol. 12, pp. 4792–4800, 2017.

[24] S. Mostafa, M. Ahmad, A. Mustapha and M. Mohammed, "Formulating layered adjustable autonomy for unmanned aerial vehicles," *International Journal of Intelligent Computing and Cybernetics*, vol. 10, no. 4, pp. 430–450, 2017.

[25] S. Mostafa, S. Gunasekaran, A. Mustapha, M. Mohammed and W. Abduallah, "Modelling an adjustable autonomous multi-agent internet of things system for elderly smart home," in *Int. Conf. on Applied Human Factors and Ergonomics*, Washington, USA, Springer, pp. 301–311, 2019.

[26] K. Abdulkareem, M. Mohammed, S. Gunasekaran, M. Al-Mhiqani, A. Mutlag *et al.,* "A review of fog computing and machine learning: Concepts, applications, challenges, and open issues," *IEEE Access*, vol. 7, pp. 153123–153140, 2019.

[27] A. Mutlag, M. Ghani, N. Arunkumar, M. Mohammed and O. Mohd, "Enabling technologies for fog computing in healthcare iot systems," *Future Generation Computer Systems*, vol. 90, pp. 62–78, 2019.

[28] S. A. Mahmoud, I. Ahmad, M. Alshayeb, W. Al-Khatib, M. Parvez *et al.,* "Khatt: Arabic offline handwritten text database," in *Int. Conf. on Frontiers in Handwriting Recognition*, Bari, Italy, IEEE, pp. 449–454, 2012.

[29] S. Mahmoud, I. Ahmad, W. Al-Khatib, M. Alshayeb, M. Parvez *et al.,* "Khatt: An open arabic offline handwritten text database," *Pattern Recognition*, vol. 47, no. 3, pp. 1096–1112, 2014.

[30] M. A. Mohammed, S. S. Gunasekaran, S. A. Mostafa, A. Mustafa and M. K. Abd Ghani, "Implementing an agentbased multi-natural language anti-spam model," in *Int. Sym. on Agent, Multi-Agent Systems and Robotics*, Putrajaya, IEEE, pp. 1–5, 2018.

[31] I. J. Hussein, M. A. Burhanuddin, M. A. Mohammed, M. Elhoseny, B. Garcia-Zapirain *et al.,* "Fully automatic segmentation of gynaecological abnormality using a new viola-jones model," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 3161–3182, 2021.

[32] S. Albahli, A. Alsaqabi, F. Aldhubayi, H. T. Rauf, M. Arif *et al.,* "Predicting the type of crime: Intelligence gathering and crime analysis," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 2317–2341, 2021.

[33] A. A. Mutlag, M. K. Ghani, M. A. Mohammed, M. Maashi, S. A. Mostafa *et al.,* "MAFC: Multi-agent fog computing model for healthcare critical tasks management," *Sensors*, vol. 20, no. 7, pp. 1853, 2020.

[34] M. K. Ghani, M. A. Mohamed, S. A. Mostafa, A. Mustapha, H. Aman *et al.,* "The design of flexible telemedicine framework for healthcare big data," *International Journal of Engineering & Technology*, vol. 7, no. 3.20, pp. 461–468, 2018.

[35] D. M. Abdulqader, A. M. Abdulazeez and D. Q. Zeebaree, "Machine learning supervised algorithms of gene selection: A review," *Technology Reports of Kansai University*, vol. 62, no. 3, pp. 233–244, 2020.

[36] D. A. Hasan and A. M. Abdulazeez, "A modified convolutional neural networks model for medical image segmentation," *Machine Learning*, vol. 62, no. 3, pp. 233–244, 2020.

[37] D. A. Zebari, D. Q. Zeebaree, A. M. Abdulazeez, H. Haron and H. N. A. Hamed, "Improved threshold based and trainable fully automated segmentation for breast cancer boundary and pectoral muscle in mammogram images," *IEEE Access*, vol. 8, pp. 203097–203116, 2020.

[38] D. Q. Zeebaree, H. Haron and A. M. Abdulazeez, "Gene selection and classification of microarray data using convolutional neural network," in *Int. Conf. on Advanced Science and Engineering*, Duhok, Iraq, IEEE, pp. 145–150, 2018.

[39] D. Q. Zeebaree, H. Haron, A. M. Abdulazeez and D. A. Zebari, "Machine learning and region growing for breast cancer segmentation," in *Int. Conf. on Advanced Science and Engineering*, Duhok, Iraq, IEEE, pp. 88–93, 2019.