Increasing the Accuracy of Writer Identification Based on Bee Colony Optimization Algorithm and Hybrid Deep Learning Method

Hao Libo, Xu Jingqi
Hunan Mechanical and Electrical Vocational and Technical College Changsha 410151, Hunan, China

***Abstract*—It is one of the most important and challenging classification issues to identify the writer's identity from offline handwriting images, which has been the focus of many researchers in recent years. This article presents a novel approach to identifying the author of offline Pertian manuscripts from scanned images based on deep convolutional neural networks. For the first time in the proposed network, the bee colony algorithm has been used in the middle layers of a deep convolutional neural network in order to improve the accuracy of identifying the author and to optimize the parameters, as well as improve the learning performance. In terms of the presented scenario, it was tested independently of the written language in both Persian and English. The proposed method is more accurate than previous studies for the IMA dataset, with an accuracy of 97.60%. Moreover, for the Firemaker dataset, the proposed model has significantly improved over the existing results, with the accuracy of the current model being 99.71%, a value that is 1.78% higher than the results of the previous models.**

***Keywords*—Optimization; bee colony algorithm; deep learning; author identity recognition; handwriting**

I. Introduction

Today, cameras and scanners convert a large number of existing paper documents into digital files. Many applications, such as office automation and digital libraries, require efficient storage, restoration, and management of these image archives. Therefore, it is imperative to obtain effective algorithms to analyze digital images of documents. As a result, author identification through the analysis of documents has become one of the most interesting challenges in the field of image processing and pattern recognition [1], [2], [3]. Identifying the author has become difficult as a result of the type of handwriting, structural differences and different types of writing between people on the one hand and the quality of the obtained images on the other hand [4].

Several recent studies have demonstrated that the emergence of deep learning methods has led to acceptable results for the extraction and classification of image features, along with accuracy, speed, unsupervised training of features, and a reduction of training time and calculation volume [5], [6], [7]. Deep learning is one of the methods for processing images based on different models. So far, few studies have used deep learning techniques for interpreting handwriting versions, and previous approaches have largely relied on conventional machine learning techniques based on feature selection and extraction. A method for offline handwriting recognition has been proposed by Hamdan et al. [8]. A total of five types of features were extracted and analyzed from handwritten texts as part of this study. Principal component analysis and Fisher's linear discriminant analysis were used to reduce the dimensions of the features. Support vector machines and neural networks were used as two classifiers in order to confirm the efficacy of the described method and the extracted features. Chahi et al. [9] presented an operator that shows multiple histograms and emphasizes the extraction of the desired characteristics. An image histogram is constructed by analyzing the distribution of pixels in a small block of pixels. Using the nearest neighbor classification with Hamming distance, they demonstrated that their approach is comparable to or better than contemporary approaches. An online author recognition system based on a recurrent neural network was investigated by Zhang et al. [10]. In order to classify the data, they used the long short-term memory model. An experimental study was conducted on English (135 authors) and Chinese (187 authors) datasets, and the advantages of their method were confirmed compared to other available methods.

According to He and Schomaker [11], handwriting carries explicit and implicit information, i.e., explicit information pertains to the exact lexical content of the words, the number of letters in the word (word length), and the letters themselves. Implicit information, on the other hand, refers to the author's behavior that can be used to identify the author. In order to generate additional information, these researchers suggested that explicit information should be used alongside implicit features. As part of another study, He et al. [12] proposed a deep model called FragNet to extract powerful features from word images. This model consists of two deep architectural paths: the feature pyramid path and the parts path. In the feature pyramid path, the entire word image is used as input, allowing the model to capture global features and contextual information. In the parts path, the input image is divided into smaller segments, which are processed separately to focus on fine-grained details and local features. By combining these two paths, FragNet is able to leverage both global and local features, enhancing its ability to accurately represent and classify word images. Javidi et al. [13] proposed to identify the author offline and independent of the text by using a convolutional neural network and Bayesian network. Based on noisy images, Ni et al. [14] have attempted to recognize the handwriting author offline. Using convolutional neural networks as a method of displaying features has been shown to be efficient. Through the use of convolutional neural networks and traditional machine vision descriptors, they are able to improve applications such as author identification even when images contain noise. Semma et al. [15] propose an approach to identify the writer of a text independently of the text and offline based on the combination of deep and traditional features. Litifu et al. [16]analyzed heterogeneous handwriting data and used deep neural networks to identify the author offline. Using deep learning, Wang et al. [17] have presented an automated method of author identification.

In most studies, English is chosen as the language of analysis of handwritten documents, while few researchers have focused on right-to-left languages, such as Persian, Arabic, Chinese, and Urdu. A gradient-based approach was presented by Helli and Moghaddam [18] for identifying the offline author of Persian texts using general features. Using the gradient descriptor, they proposed a method for extracting three energy-based features and eight angle-based features. In defining a mathematical model with regard to the Arabic language, Abdi and Khemakhem [19] reached a correct percentage of 59% based on the assumption that manuscripts are independent of one another. In the detection process, 92 features are utilized, and the detection criterion is based on a combination of cross-correlation and Fisher's linear criterion. Khosroshahi et al. [20] introduced an offline author recognition system utilizing a deep convolutional neural network model tailored for right-to-left datasets, which was evaluated alongside four other datasets. Sabzekar et al. [21] presented a model capable of identifying authors from handwriting independent of the written language. In a multi-script context, Semma et al. [22] proposed a deep learning-based approach for author identification, which involves identifying key points in the handwriting and extracting small patches around them. This study examined texts in Arabic, English, French, Chinese, and Persian. For effective author identification, it is crucial to extract both abstract features of the author's writing style and subtle details indicative of their writing habits. Current handcrafted features, which capture local shape and gradient information, often face limitations. These limitations arise from reliance on artificial features such as written content information (text) and writing styles (personality). Consequently, a more comprehensive approach that considers both high-level abstractions and fine-grained details is necessary for accurate author identification.

Several approaches have demonstrated significant efficiency in identifying and verifying identities across various datasets in numerous studies. However, challenges persist, particularly with more complex datasets and real-world applications. For instance, variables such as the use of different pens and papers introduce variability that still requires further research. Generally, different environmental conditions have not been adequately considered when preparing databases. For practical applications, it is essential to account for diverse environmental factors. Most studies have focused on English texts, with relatively little attention given to manuscripts in right-to-left languages. This gap underscores the need for more research into handling manuscripts written in various scripts and under different conditions to enhance the robustness and applicability of author identification systems. This research aims to develop an algorithm for offline writer identification under different experimental conditions, including the type of pen and paper used, as well as various uncertainties, based on the use of deep neural networks to process right-to-left handwriting sample images. A variety of conditions have been used to collect the required data from different authors. Then, utilizing the feature addition approach, a two-way hybrid architecture of deep convolutional neural networks was developed in order to extract the most favorable features for the recognition of the identity of the author of Persian texts. In the proposed algorithm, deep network-based features are extracted in one path, and complementary features are extracted in a second path with the aid of an innovative screening system. In the proposed network, the bee colony algorithm is used in the middle layers of the deep convolutional neural network to optimize the parameters, improve learning performance, and increase accuracy. This architecture represents an extension of the 18-layer RosNet neural network model, characterized by the formulation of a two-path architecture model that combines features extracted by the RosNet deep neural network on one path, while global features are combined in a classical manner on the second path. Different classification criteria are used to evaluate and measure the final model. Based on the results obtained by the model on the aforementioned dataset, the proposed method appears to perform well.

In this work, presented a comprehensive analysis of our proposed technique for increasing the accuracy of writer identification based on the Bee Colony Optimization Algorithm and a Hybrid Deep Learning approach. The paper is structured as follows: In Section II, we detail the methodology employed in our study, including the preparation of the database, description of databases used, data pre-processing techniques, the architecture of the proposed hybrid network model, and the incorporation of additional training data. Section III presents the experimental setup and evaluation of our method's performance, followed by a detailed analysis of the results. In Section IV, we discuss the implications of our findings and conclude with insights into future research directions. This structured approach ensures a systematic presentation of our methodology, results, and conclusions, providing readers with a clear roadmap of the paper's content.

II. Proposed Method

A general description of the proposed author recognition system is presented in Fig. 1. In the figure, the first step emphasizes that a database is necessary in order to provide an automatic author identification system. Using information forms, various examples of people's handwriting in the Persian language have been collected in this article. Following the scanning of the samples, the samples are pre-processed. The operations performed at this stage include binarization and noise reduction. As a final step, the image will be transformed into a texture using an efficient algorithm. Afterward, the features that clearly demonstrate the differences between the handwritings are extracted. A new algorithm is used to extract the desired features from the handwritten image, which is considered as a texture in this study. In the following stage, the system is trained, and in the final stage, the system's success rate in identifying the author is evaluated based on the test data. Each step will be explained in more detail in the following paragraphs.

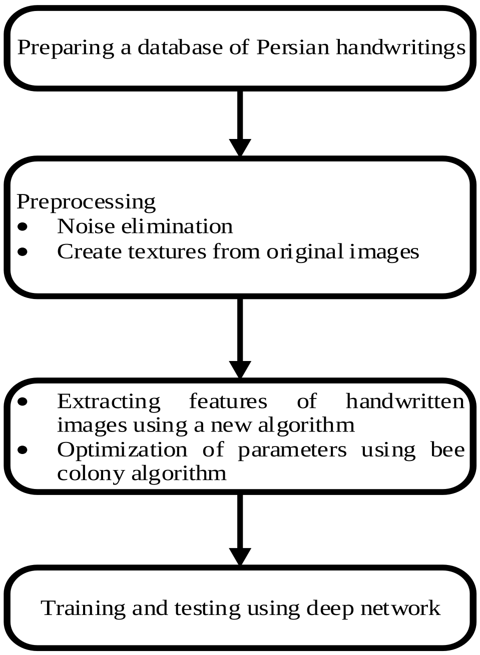


Fig. 1. General steps of the proposed author identification system.

A. Preparation of Database

To evaluate pattern recognition problems, a labeled data set is required. There is no exception to this rule when it comes to identifying the author. There are many databases available for identifying authors from manuscripts in the English language, which is a relatively new field of research. In contrast, there is a much smaller amount of data available for Persian. An example of a handwritten database for the Persian language is reported in study [23]. This database has been collected from their handwriting to identify the mental states and psychological and emotional characteristics of people.

In this study, participants were asked to write a text on an A4 sheet of paper. To investigate the effects of pen type, paper type, and various uncertainties (such as environmental noise), a comprehensive database has been developed. The proposed database consists of handwriting examples composed from 20 participants over a variety of time periods and environments. We asked each participant to write a desired text, preferably one line, and then repeat it nine times at different intervals. For the collection of the writings, two different types of standard paper were used, COPIMAX-Executive and PaperOne. Also, in order to check the type of pen, two different types of Schneider pens with a writing diameter of 0.4 mm and a Xiaomi MJZXBO1WC pen with a writing diameter of 0.5 mm were used. Based on the above parameters, the dataset collected from 20 participants comprises 60 pages and 600 sentences. According to Fig. 2, the sentence samples have a height of 236 pixels and a width of variable pixels. Approximately 2339 x 1656 pixels are the size of sample pages. We scanned each sample at a resolution of 600 dpi and saved it with 256 gray levels in a separate file in the *.tif format for each author.

B. Databases

Since Persian language databases are not available to identify the author, databases of other available languages have been applied to evaluate the accuracy of the proposed technique. A description of the databases that were used in the experiment is provided in this section. Three databases are included in this group, namely QUWI[24], IFN/ENIT [25] and IAM[26]. As part of the QUWI dataset, 1017 Arabic and 975 English authors have provided their handwriting. In the IAM database, there are 1066 samples of English handwriting from 400 different authors. There are also 937 Arabic manuscripts from 411 different authors included in the IFN/ENIT dataset. An example of these datasets can be found in Fig. 2. Additionally, the Firemaker dataset [27] is used, which contains 1000 pages of handwriting from 250 different authors. Approximately 80% of the words are used in the training phase and 20% in the testing phase for each writer.

C. Data Pre-Processing

During the pre-processing stage, binarization, noise removal, and texture extraction are performed on each image. Upon saving the samples, we aligned the pages based on the left margin at 0.1 degrees accuracy so that the lines would be vertical and the page would be smooth. In the next step, the image of each sentence was saved in a separate file with a height of 236 pixels and a variable width. Following the separation of 6000 sentences, sentences size was first changed to 570 pixels with variable width, and the size of the image was later changed to 225 pixels after normalization. Following this, each sentence is divided into 225 x 225-pixel samples using the segmentation method. Fig. 3 shows a division operation on one of the sentences.

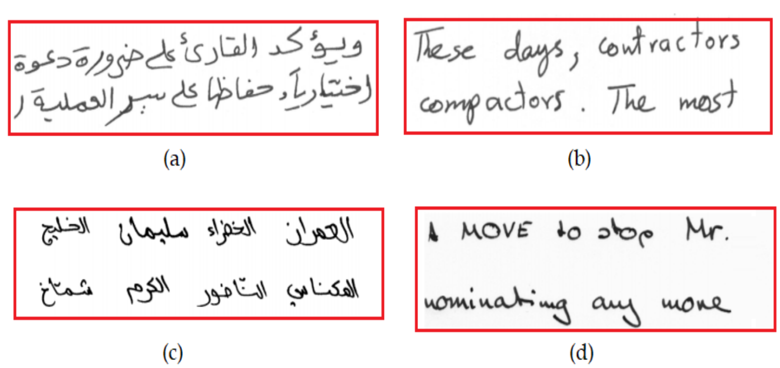


Fig. 2. Handwriting samples from different datasets (a) Arabic QUWI, (b) English QUWI, (c) IFN/ENIT and (d) IAM.

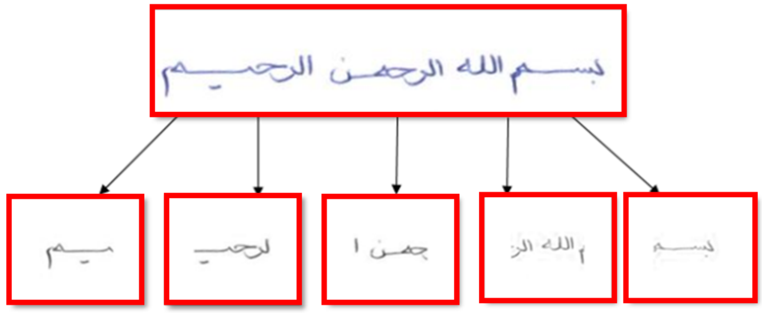
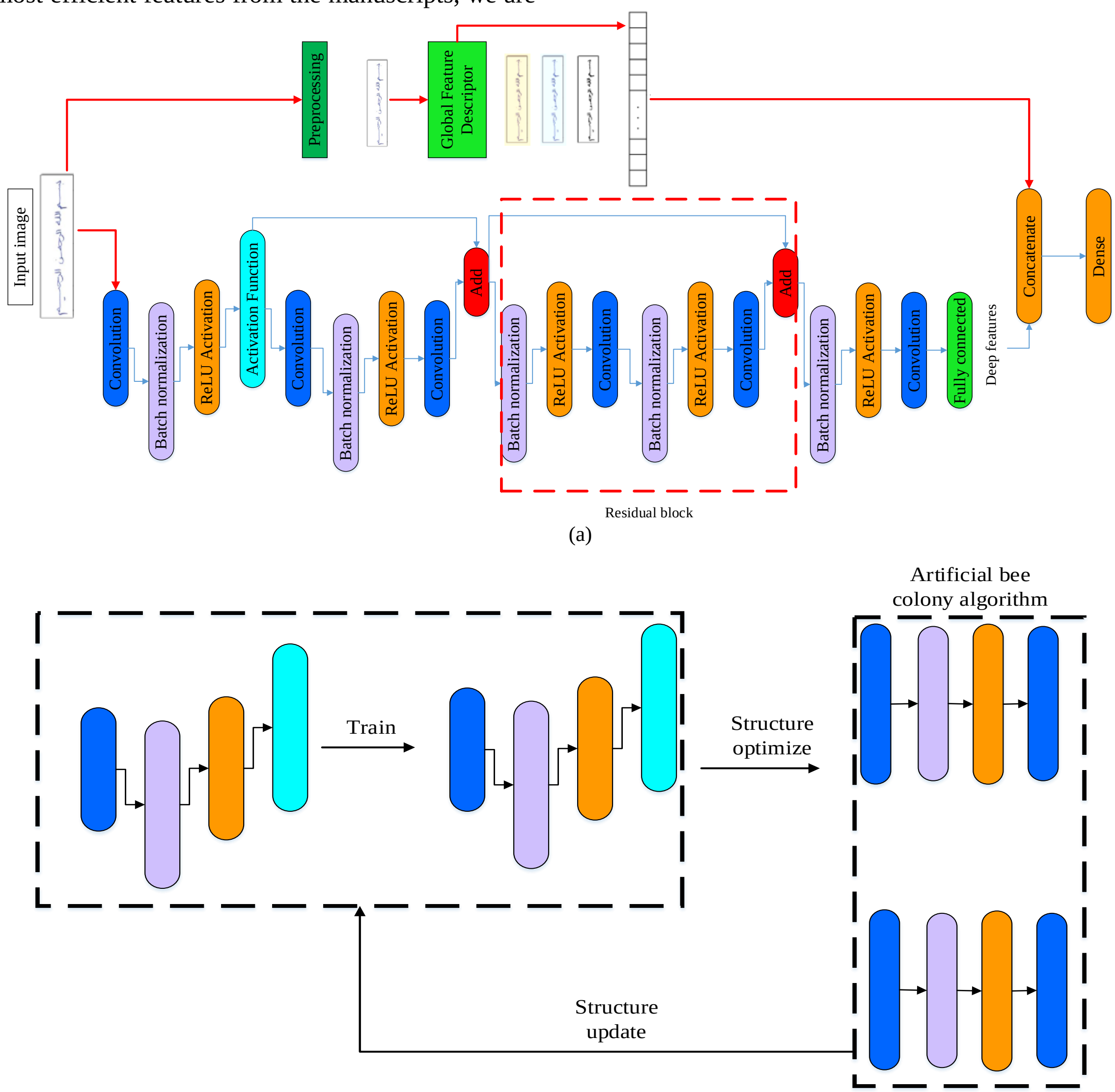


Fig. 3. Classification of one of the sentences.

D. Proposed Hybrid Network Model Architecture

Our study presents a method for identifying writers by extracting the features of their texts using an innovative and highly efficient approach. Consequently, since we are seeking to extract the most efficient features from the manuscripts, we are jointly using the RosNet deep neural network model and classical extraction of global features. The architecture of the proposed system is displayed in Fig. 4. The details of the proposed network architecture will be discussed in the following paragraphs.



(b)

Fig. 4. An overview of the proposed hybrid model architecture. The first path extracts global features in a classical manner, while the second path extracts and integrates modern features based on the RosNet deep network model.

This figure indicates that a deep residual network, like the popular ResNet model, has been used to identify the writer. Considering the large number of parameters in deep networks, network training requires a considerable amount of learner effort. Furthermore, the vanishing gradient problem arises as the network deepens, resulting in a large error during the training stage. A modular architecture was developed to address this defect, where residual blocks are arranged one after another along with the connections related to the same mapping. The following calculations are performed in a residual block [28]:

$$x_{l+1} = x_l + F(x_l.W_l) \tag{1}$$

where, x_l and x_{l+1} are, respectively, the input and output of the residual block. The variable of W_l represents the set of weights, and F is the residual function. Based on the above relationship, if the added layer acts as a layer with the same mapping performance, the network performance in the deep model should not have more training error than the shallow model. Recursively, one can write the following relation for each residual block L:

$$x_L = x_l + \sum_{i=1}^{L} F(x_i.W_i) \tag{2}$$

As well, the derivative chain law in the backpropagation network provides the following relationship for the cost function:

$$\frac{\partial \varepsilon}{\partial x_l} = \frac{\partial \varepsilon}{\partial x_L}\frac{\partial x_L}{\partial x_l} = \frac{\partial \varepsilon}{\partial x_l}\left(1 + \frac{\partial}{\partial x_l}\sum_{i=1}^{L} F(x_i.W_i)\right) \tag{3}$$

Based on the above relation, it can be concluded that the gradient $\partial\varepsilon/\partial x_l$ propagates information with or without weight layers. The RosNet network architecture used in this study consists of 18 layers, including four residual blocks of the same structure. As shown in Fig. 3, the batch normalization layer and the ReLU activation function are applied before the convoluted layers. There are also the same number of filters used in each residual block, with the exception of the last convolution layer, which has double the number of filters to maintain its computational complexity. Accordingly, as the depth of the network increases, the number of filters in the residual blocks will be 64, 128, 256, and 512, respectively. As in the basic version of RosNet, the input consists of a set of image fragments; however, in this paper, additional input is also provided to the network in order to represent the features of the image better. In order to accomplish this, global features have been used for the first time in this work. This feature combination is then combined with the features extracted by the deep network in the final residual block, and the final classification is then made based on the combination of these two features. RosNet-18 architecture provides 512 dimensions, which are combined with the proposed global feature vector, which is 189 dimensions, to produce a 701-dimensional feature vector. This number of features constitutes the input of the dense layer, and its output is determined by the number of classes in each dataset. The existing experiments demonstrate that the proposed feature vector effectively represents the global and local information in the image. Following our analysis of the deep residual network base model and the proposed hybrid network method on several databases of handwriting images in the next section, we demonstrate that the joint feature vector produced by this method is more efficient than the feature vector generated by the base model and other methods of similar nature. To obtain the best convergence rate, the proposed network's hyperparameters have been carefully adjusted. After extensive experimentation, the cross-entropy error function and the bee colony optimizer with a learning rate of 0.001 were selected. The network was trained using the conventional error backpropagation method with a batch size of 64. This choice of parameters was driven by a series of trials aimed at optimizing the network's performance. Various combinations of learning rates, optimizers, and error functions were tested to identify the configuration that provided the best convergence rate and accuracy. The sensitivity of these parameters on the results was carefully analyzed. For instance, adjustments in the learning rate directly impacted the speed and stability of convergence, while different optimizers influenced the network's ability to escape local minima and achieve a more global optimum. The selected parameters demonstrated the best performance in our experiments, achieving a balance between training speed and accuracy. However, we acknowledge that further tuning and experimentation with alternative sets of values could potentially enhance the results. Future work could involve a more systematic exploration of the parameter space, perhaps using automated hyperparameter optimization techniques to fine-tune the network further. Slight changes in the learning rate (e.g., ±0.0005) significantly impacted convergence speed and final accuracy, underscoring the importance of precise tuning. Increasing the number of layers beyond 20 showed diminishing returns and a higher risk of overfitting, while fewer layers reduced the model's capacity to learn complex features.

It is necessary to determine the optimal hyperparameters when working with deep networks. Manually determining the optimal parameter will be very challenging if there are a large number of these parameters. To obtain the optimal parameters of the used convolutional network, this article utilizes the bee colony algorithm. The general structure of this algorithm was summarized by He et al. [34]. It is necessary to specify the number of elementary particles in this structure. Randomly assigned particles are used for this purpose. There are actually several different architectures of convolutional neural networks in each of these particles.

E. Additional Training Data

In order to train the proposed model, training data samples are required since it is a supervised learning method. Among the challenges of deep learning is the dependence on a large amount of training data. In this work, we have increased the number of training samples in order to overcome this challenge with conventional methods while the proposed model is used in the training process. The selection of data augmentation techniques must be carefully considered, and not all data augmentation techniques can be applied to handwritten data, so this study used a combination of random grayscale techniques, color jitter techniques, and random rotation techniques as data augmentation techniques after attempting many different techniques. As a result of the use of data augmentation techniques, the training dataset is increased by approximately 70%.

III. TESTS AND RESULTS

This section provides the details of the experiments performed to validate the proposed technique, along with a discussion of the results. The proposed author identification method and all experiments were conducted using Python, leveraging various libraries, with NumPy and PyTorch being the most significant. The parameters of the proposed deep network are listed in Table I.

TABLE I. PARAMETERS RELATED TO THE IMPLEMENTATION OF THE PROPOSED DEEP NETWORK

Parameter	Value
Optimizer	Adam
Loss function	Binary cross-entropy
Performance metric	Accuracy
Total Classes	2 (Gate and Non-Gate)
Batch Size	64
Epoch	150

The choice of parameters was made after extensive experimentation to ensure the best convergence rate and performance. The Adam optimizer was selected for its adaptive learning rate properties and efficient handling of sparse gradients. The binary cross-entropy loss function was chosen as the problem involves two classes, making it suitable for the classification task. Accuracy was used as the primary performance metric to evaluate the effectiveness of the model in distinguishing between the two classes (Gate and Non-Gate). The batch size of 64 was found to provide a good balance between computational efficiency and training stability, while the number of epochs (150) ensured sufficient training time for the model to learn the underlying patterns in the data. All results and reviews of the proposed method were conducted using this setup, and the experiments confirmed the robustness and reliability of the approach in identifying authors based on their handwriting. Further experimentation with different parameter sets could potentially enhance the results, as discussed in the sensitivity analysis section. Overall, the experimental results validate the effectiveness of the proposed method, demonstrating its potential for practical application in author identification tasks.

A. Evaluation Criteria

In order to evaluate the performance of the proposed method, 4 evaluation criteria, including recall (see Eq. (4)), precision (see Eq. (5)), accuracy (see Eq. (6)) and F-measure (see Eq. (7)), are used. Also, based on these parameters, the system performance characteristic (ROC) is calculated.

$$Re = \frac{TP}{TP + FP} \tag{4}$$

$$Pr = \frac{TP}{TP + FN} \tag{5}$$

$$\text{ACC} = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

$$\text{Fvalue} = \frac{2\, Pr \times Re}{Pr + Re} \tag{7}$$

where, TN remains the number of correctly identified negative samples, TP stands number of correctly identified positive samples, FP is the number of false positive identifications, and FN is the number of false negative identifications.

B. Evaluation and Comparison of the Presented Method

In this section, the evaluation results of the proposed method are presented based on the combination of the global features extracted by the classical method with the features extracted by the deep neural network model. In this evaluation, the RosNet model, which is the basis for the proposed architecture, and some methods proposed in previous studies are evaluated using the database discussed under the same conditions. The test data set, which includes handwriting images, is input to both the basic deep neural network and the proposed deep neural network, and then the results of the proposed method are presented. For the dataset collected in this study, these results are shown in Table II. As can be seen, the accuracy of the proposed method with two-way hybrid architecture has improved significantly compared to the RosNet basic method for writer identification and has increased from 84% to 99.4%. Compared to the basic method, this result was achieved with a significant and proportional improvement in all components of sensitivity and identification, indicating the efficacy of the proposed methodology. In spite of the improvement in accuracy, the proposed method requires a short amount of training time. Additionally, the two components of accuracy and efficiency are calculated and listed in Table II. Based on these two components, Fig. 5 illustrates the Precision-Recall curve. The best state of this curve is when the value of the area under these curves is equal to the number one, and vice versa. The closer it is to zero, the weaker prediction is obtained [29].

TABLE II. COMPARISON OF THE PROPOSED METHOD AND THE BASIC METHOD FOR THE CURRENT RESEARCH DATASET

Parameter	Method	
	Proposed method	***ResNet***
TP	99%	84%
TN	91%	85%
FN	5%	11%
FP	2%	15%
recall	0.98	0.84
precision	0.93	0.88
F-measure	0.95	0.83
accuracy	98%	87%
Time	0.651 sec/image	0.463 sec/image

The results of the suggested network for the three datasets are presented in Table III. The results obtained using the new hybrid architecture are superior to those obtained using the basic network in all three databases. Additionally, in this table, the results achieved at the beginning of the feature extraction process are also compared with the final results. However, even in cases where the results of the suggested method and the basic method are similar, the proposed method has achieved the desired result much more rapidly. In other words, by combining features extracted by the deep method with those extracted by the classical method, we are not only able to reach the maximum desired result much more quickly, but it is also often more accurate. Due to the small number of writers, the IMA database could not show the influence of the suggested method very well and converged very quickly to the maximum possible accuracy value. The two QUWI and IFN/ENIT databases, which have more data and have a more difficult situation in separating handwriting, show the effect of the presented method well. This conclusion can also be understood based on the graphs displayed in Fig. 6.

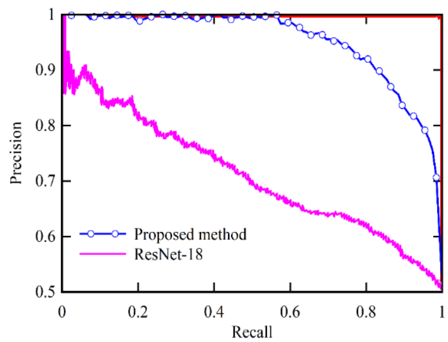


Fig. 5. PR curve of proposed method for the dataset collected in this research.

To evaluate the proposed model, the four data sets described in section B have been used. The results of the identification of the proposed hybrid model and the base model of ResNet-18 with the TTA technique to identify the writers using each of the four data sets are given in Table III. According to Table III, the proposed improved network-based model performs better than the basic model for identifying the writer using each of the four datasets.

Several state-of-the-art methods are compared with the proposed deep network model in the following section. A comparison of the results is provided in Tables IV and V. According to the proposed conjugate method, the accuracy of the IMA dataset is 97.60%, which is a higher level of accuracy than that observed in previous studies. In this table, Khan et al. [30] reported an accuracy of 97.20% for 650 authors, which is close to our results. Moreover, according to Table V, for the Firemaker dataset, our proposed model has improved significantly in comparison to the existing results, with an accuracy of 99.71%, which is about 1.78% better than that of Khan et al. [31].

TABLE III. EVALUATION RESULTS OF THE PROPOSED MODEL AND RESNET-18 MODEL FOR FOUR COMPREHENSIVE DATA SETS

Method	**Dataset**			**Accuracy (%)**
	Dataset	***Language***	***Writer number***	
Proposed hybrid network	Present study	Persian	60	98.2
	IMA	English	400	97.6
	QUWI	Arabic	1017	98.5
	IFN/ENIT	Arabic	411	99.7
ResNet-18	Present study	Persian	60	83.0
	IMA	English	400	96.3
	QUWI	Arabic	1017	89.8
	IFN/ENIT	Arabic	411	96.2

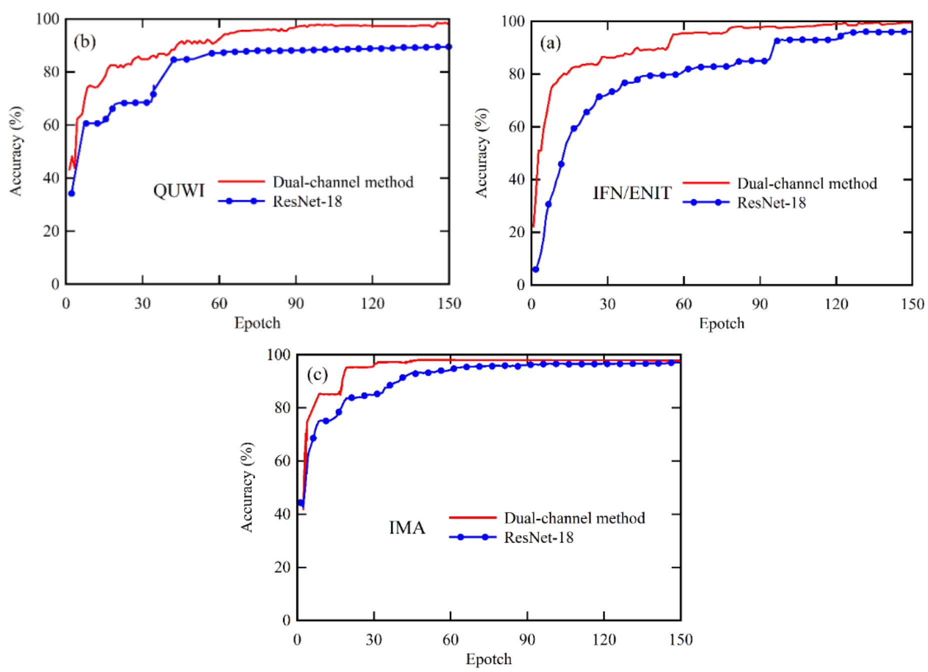


Fig. 6. The evaluation results of the writer identifying with the presented architecture model and RosNet architecture model on three handwriting databases (a) IFN/ENIT, (b) QUWI and (c) IMA.

TABLE IV. WRITER IDENTIFICATION RESULTS RELATED TO THE IMA DATASET

Method	Year	Sample No.	Accuracy
Ref [32].	2016	657	96.92%
Ref. [30]	2017	650	97.20%
Ref. [33]	2017	650	89.90%
Ref. [34]	2018	650	88.57%
Ref. [35]	2019	650	90.12%
Ref. [36]	2019	657	91.17%
Ref. [37]	2020	657	94.06%
Ref. [12]	2020	657	96.30%
Our simple method		657	95.25%
Our conjugate approach		657	97.50%

TABLE V. WRITER IDENTIFICATION RESULTS RELATED TO THE FIREMAKER DATASET

Method	Year	Sample No.	Accuracy
Ref. [38]	2012	250	86.00%
Ref. [39]	2013	250	91.80%
Ref. [40]	2014	250	92.40%
Ref. [41]	2015	250	89.80%
Ref. [30]	2017	250	89.47%
Ref. [35]	2019	250	92.38%
Ref. [31]	2019	250	97.98%
Ref. [37]	2020	250	97.60%
Ref. [12]	2020	250	97.60%
Our simple method		250	90.32%
Our conjugate approach		250	99.71%

C. Evaluation and Comparison of the Presented Method with low Training Data

Generally, the number of training samples is considered to be a challenge in deep learning methods, so methods that can achieve desirable results with fewer training samples are always considered. A method of strengthening training data has been used to compensate for the lack of training samples in this study. Although the generated samples are based on the number of real base samples, the evaluation of the newly developed method shows that it is able to obtain favorable results even with very few base samples. In this experiment, we considered the number of basic training samples equal to the correct multiplier of 10% of the entire dataset and performed the evaluation process with the remaining images. A summary of the results of this experiment, which represents an intermediate result of two implementations of the suggested technique, can be found in Table VI. Also, Fig. 7 demonstrates that the suggested method achieved 68.97% accuracy with 60% of the training sample. However, if only 5% of the training sample is used, this accuracy is reduced to 82%. In this way, the influence of training sample numbers is clearly visible in this image.

TABLE VI. COMPARISON OF THE ACCURACY OF THE RESULTS FOR DIFFERENT NUMBERS OF TRAINING SAMPLES OF THE DATABASE COLLECTED IN THIS RESEARCH (EPOCH=130)

Training samples	10%	20%	30%	40%	60%	80%	100%
Accuracy (%)	87.54	91.04	94.97	96.12	97.24	97.56	98.60

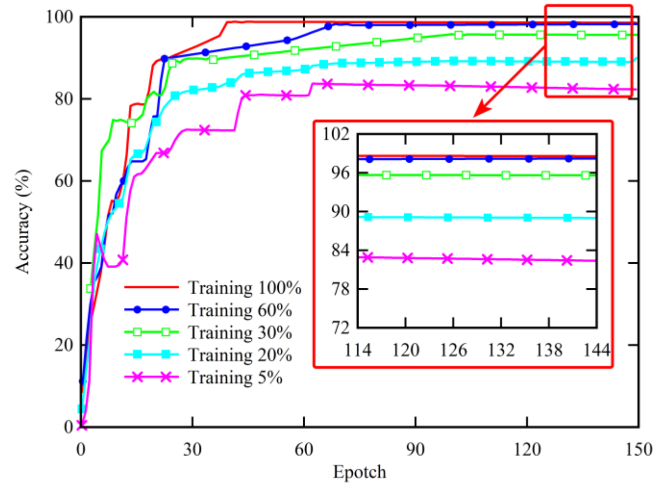


Fig. 7. The influence of training samples number with the suggested method on database collected in this research on the accuracy of the model.

IV. CONCLUSION

This paper introduces a novel approach to offline author handwriting recognition by leveraging the bee colony optimization algorithm and integrating deep and classical image features. The proposed architecture extends the 18-layer RosNet model, incorporating a two-path architecture where deep neural network features from RosNet are combined in one path, and global features are fused in a classical manner in the second path. Notably, the bee colony optimization algorithm is employed to select optimal features and determine the appropriate number of layers for enhanced accuracy in author identification. Through comprehensive evaluation using various classification criteria, the method consistently demonstrates promising performance across different datasets.

In conclusion, the proposed method presents a novel approach to offline author handwriting recognition, leveraging the bee colony optimization algorithm and integrating deep and classical image features. The rationale behind selecting this method stems from its ability to capitalize on the strengths of deep learning techniques for feature extraction, the flexibility offered by the two-path architecture, and the innovative use of the bee colony optimization algorithm for feature selection.

Furthermore, it is important to acknowledge the limitations of existing methods that may hinder their effectiveness in addressing the problem at hand. These limitations include the reliance on handcrafted features, which may not capture the complexity of handwriting styles adequately, and scalability issues that limit their applicability to large datasets or real-world scenarios.

By addressing these limitations and leveraging the strengths of our proposed method, we aim to advance the field of offline author handwriting recognition and contribute to the development of more accurate and robust recognition systems.

REFERENCES

[1] V. Karthikeyan, "Modified layer deep convolution neural network for text-independent speaker recognition," Journal of Experimental & Theoretical Artificial Intelligence, pp. 1–13, 2022.

[2] M. Bibi, A. Hamid, M. Moetesum, and I. Siddiqi, “Document forgery detection using source printer identification: A comparative study of text‐dependent versus text‐independent analysis,” Expert Syst, vol. 39, no. 8, p. e13020, 2022.

[3] A. Srivastava, S. Chanda, and U. Pal, “Exploiting Multi-Scale Fusion, Spatial Attention and Patch Interaction Techniques for Text-Independent Writer Identification,” in Asian Conference on Pattern Recognition, Springer, 2021, pp. 203–217.

[4] T. Bahram, “A texture-based approach for offline writer identification,” Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 8, pp. 5204–5222, 2022.

[5] S. A. El-Moneim et al., “Text-dependent and text-independent speaker recognition of reverberant speech based on CNN,” Int J Speech Technol, vol. 24, no. 4, pp. 993–1006, 2021.

[6] S. Farsiani, H. Izadkhah, and S. Lotfi, “An optimum end-to-end text-independent speaker identification system using convolutional neural network,” Computers and Electrical Engineering, vol. 100, p. 107882, 2022.

[7] M. F. Mridha, A. Q. Ohi, J. Shin, M. M. Kabir, M. M. Monowar, and M. A. Hamid, “A thresholded Gabor-CNN based writer identification system for Indic scripts,” IEEE Access, vol. 9, pp. 132329–132341, 2021.

[8] Y. B. Hamdan and A. Sathesh, “Construction of statistical SVM based recognition model for handwritten character recognition,” Journal of Information Technology and Digital World, vol. 3, no. 2, pp. 92–107, 2021.

[9] A. Chahi, Y. Ruichek, and R. Touahni, “Block wise local binary count for off-line text-independent writer identification,” Expert Syst Appl, vol. 93, pp. 1–14, 2018.

[10] X.-Y. Zhang, G.-S. Xie, C.-L. Liu, and Y. Bengio, “End-to-end online writer identification with recurrent neural network,” IEEE Trans Hum Mach Syst, vol. 47, no. 2, pp. 285–292, 2016.

[11] S. He and L. Schomaker, “Deep adaptive learning for writer identification based on single handwritten word images,” Pattern Recognit, vol. 88, pp. 64–74, 2019.

[12] S. He and L. Schomaker, “Fragnet: Writer identification using deep fragment networks,” IEEE Transactions on Information Forensics and Security, vol. 15, pp. 3013–3022, 2020.

[13] M. Javidi and M. Jampour, “A deep learning framework for text-independent writer identification,” Eng Appl Artif Intell, vol. 95, p. 103912, 2020.

[14] K. Ni, P. Callier, and B. Hatch, “Writer identification in noisy handwritten documents,” in 2017 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2017, pp. 1177–1186.

[15] A. Semma, Y. Hannad, I. Siddiqi, C. Djeddi, and M. E. Y. El Kettani, “Writer identification using deep learning with fast keypoints and harris corner detector,” Expert Syst Appl, vol. 184, p. 115473, 2021.

[16] A. Litifu, Y. Yan, J. Xiao, and H. Jiang, “Writer identification using redundant writing patterns and dual-factor analysis of variance,” Applied Intelligence, pp. 1–16, 2021.

[17] Z. Wang, A. Maier, and V. Christlein, “Towards end-to-end deep learning-based writer identification,” INFORMATIK 2020, 2021.

[18] B. Helli and M. E. Moghaddam, “A text-independent Persian writer identification based on feature relation graph (FRG),” Pattern Recognit, vol. 43, no. 6, pp. 2199–2209, 2010.

[19] M. N. Abdi and M. Khemakhem, “A model-based approach to offline text-independent Arabic writer identification and verification,” Pattern Recognit, vol. 48, no. 5, pp. 1890–1903, 2015.

[20] S. N. M. Khosroshahi, S. N. Razavi, A. B. Sangar, and K. Majidzadeh, “Deep neural networks-based offline writer identification using heterogeneous handwriting data: an evaluation via a novel standard dataset,” J Ambient Intell Humaniz Comput, pp. 1–20, 2022.

[21] M. Sabzekar, R. Khazaei, V. Babaiyan, and Y. Akbari, “Script independent offline writer identification from handwriting samples based on texture using wavelet transform in Persian-English languages,” Journal of Modeling in Engineering, vol. 18, no. 63, pp. 1–13, 2021.

[22] A. Semma, Y. Hannad, I. Siddiqi, S. Lazrak, and M. E. Y. El Kettani, “Feature learning and encoding for multi-script writer identification,” International Journal on Document Analysis and Recognition (IJDAR), vol. 25, no. 2, pp. 79–93, 2022.

[23] M. Ziaratban, K. Faez, and F. Bagheri, “FHT: An unconstraint Farsi handwritten text database,” in 2009 10th International Conference on Document Analysis and Recognition, IEEE, 2009, pp. 281–285.

[24] S. Al Maadeed, W. Ayouby, A. Hassaine, and J. M. Aljaam, “QUWI: an Arabic and English handwriting dataset for offline writer identification,” in 2012 International Conference on Frontiers in Handwriting Recognition, IEEE, 2012, pp. 746–751.

[25] M. Pechwitz, S. S. Maddouri, V. Märgner, N. Ellouze, and H. Amiri, “IFN/ENIT-database of handwritten Arabic words,” in Proc. of CIFED, Citeseer, 2002, pp. 127–136.

[26] U.-V. Marti and H. Bunke, “The IAM-database: an English sentence database for offline handwriting recognition,” International Journal on Document Analysis and Recognition, vol. 5, pp. 39–46, 2002.

[27] L. Schomaker, L. Vuurpijl, and L. Schomaker, “Forensic writer identification: A benchmark data set and a comparison of two systems,” 2000.

[28] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[29] D. Tien Bui, B. Pradhan, I. Revhaug, and C. Trung Tran, “A comparative assessment between the application of fuzzy unordered rules induction algorithm and J48 decision tree models in spatial prediction of shallow landslides at Lang Son City, Vietnam,” Remote sensing applications in environmental research, pp. 87–111, 2014.

[30] F. A. Khan, M. A. Tahir, F. Khelifi, A. Bouridane, and R. Almotaeryi, “Robust off-line text independent writer identification using bagged discrete cosine transform features,” Expert Syst Appl, vol. 71, pp. 404–415, 2017.

[31] F. A. Khan, F. Khelifi, M. A. Tahir, and A. Bouridane, “Dissimilarity Gaussian mixture models for efficient offline handwritten text-independent identification using SIFT and RootSIFT descriptors,” IEEE Transactions on Information Forensics and Security, vol. 14, no. 2, pp. 289–303, 2018.

[32] L. Xing and Y. Qiao, “Deepwriter: A multi-stream deep CNN for text-independent writer identification,” in 2016 15th international conference on frontiers in handwriting recognition (ICFHR), IEEE, 2016, pp. 584–589.

[33] S. He and L. Schomaker, “Writer identification using curvature-free features,” Pattern Recognit, vol. 63, pp. 451–464, 2017.

[34] P. Pandey and K. R. Seeja, “Forensic writer identification with projection profile representation of graphemes,” in Proceedings of First International Conference on Smart System, Innovations and Computing: SSIC 2017, Jaipur, India, Springer, 2018, pp. 129–136.

[35] H. T. Nguyen, C. T. Nguyen, T. Ino, B. Indurkhya, and M. Nakagawa, “Text-independent writer identification using convolutional neural network,” Pattern Recognit Lett, vol. 121, pp. 104–112, 2019.

[36] A. Chahi, Y. Ruichek, and R. Touahni, “An effective and conceptually simple feature representation for off-line text-independent writer identification,” Expert Syst Appl, vol. 123, pp. 357–376, 2019.

[37] A. Chahi, Y. Ruichek, and R. Touahni, “Cross multi-scale locally encoded gradient patterns for off-line text-independent writer identification,” Eng Appl Artif Intell, vol. 89, p. 103459, 2020.

[38] A. A. Brink, J. Smit, M. L. Bulacu, and L. R. B. Schomaker, “Writer identification using directional ink-trace width measurements,” Pattern Recognit, vol. 45, no. 1, pp. 162–171, 2012.

[39] G. Ghiasi and R. Safabakhsh, “Offline text-independent writer identification using codebook and efficient code extraction methods,” Image Vis Comput, vol. 31, no. 5, pp. 379–391, 2013.

[40] X. Wu, Y. Tang, and W. Bu, “Offline text-independent writer identification based on scale invariant feature transform,” IEEE Transactions on Information Forensics and Security, vol. 9, no. 3, pp. 526–536, 2014.

[41] S. He, M. Wiering, and L. Schomaker, “Junction detection in handwritten documents and its application to writer identification,” Pattern Recognit, vol. 48, no. 12, pp. 4036–4048, 2015.