Farsi handwritten phone number recognition using deep learning

Maryam Akhlaghi¹ · Vahid Ghods²

Received: 12 December 2019 / Accepted: 10 February 2020 / Published online: 14 February 2020
© Springer Nature Switzerland AG 2020

Abstract

An application of artificial intelligence in mobile phones which might be widely accepted by users of these phones is the intelligent system used to automatically detect, search and dial phone numbers using an image taken from a handwritten phone number. In this paper, a reliable method is presented for Farsi handwritten phone number recognition using deep neural networks. In order to recognize a Farsi handwritten digit string, the digit string is first converted to single digits using the proposed segmentation algorithm, and then each segment is classified using a single Farsi handwritten digit recognition algorithm. By classifying each segment, finally, the digit string of the Farsi handwritten phone number image is created. Since there is no database for Farsi handwritten phone numbers, this paper first collects a database of Farsi handwritten phone numbers. Accuracy of the proposed algorithm for Farsi handwritten phone number recognition is 94.6%. After recognizing digits of the phone number, the proposed algorithm is able to search in the phonebook.

Keywords Phone number · Recognition · Farsi · Convolutional neural network · Segmentation

1 Introduction

In recent years, machine learning has become a new interdisciplinary field which tries to make the systems intelligent. As one of the widely applicable areas of artificial intelligence, machine learning adjusts and explores methods and algorithms based on which, computers and systems can learn and train. Idea of machine learning is that a system can sense, remember, learn and detect like a human [1, 2]. One of the widely-applied machine learning methods is deep learning. Idea of deep learning was first presented in the 1950s by definition of a perceptron. Perceptron was the first machine which could recognize and learn. Then, multi-layer perceptron structure with a limited number of hidden layers was defined in 1980s. However, learning ability of perceptron was limited until 2000. Then, the algorithms changed their path towards deep learning and could attract attention of scientists. Deep learning or hierarchical learning is a subfield of machine learning

[3–5]. In deep learning, it is tried to use hierarchical architectures to learn high-level features of data. This method is widely applied as a new method in artificial intelligence. In fact, deep learning is one of the machine learning techniques which is used for supervised or unsupervised feature extraction and converting them for analysis of patterns from multiple layers of linear processing units [6].

Fenrich and Krishnamoorthy [7] proposed a method based on explicit recognition for segmentation problem. In this algorithm, vertical histograms and peaks and valleys of contours are used. Vertical histogram is applied to segment the digit string. Column with minimal value is considered as a candidate for segmenting the vertical line. If the candidate segmentation points from the first part of the algorithm in which vertical histogram is applied do not exist, upper and lower contours are used for segmentation. If a peak of the lower contour and a valley of the upper contour can be connected with a line segment that satisfies slope threshold, a piecewise linear split is made. If they

✉ Vahid Ghods, v.ghods@semnaniau.ac.ir; Maryam Akhlaghi, talatom2012.01@gmail.com | ¹Department of Electronics Engineering, Semnan Branch, Islamic Azad University, Semnan, Iran. ²Young Researchers and Elite Club, Semnan Branch, Islamic Azad University, Semnan, Iran.

cannot be connected with a line segment which satisfies the slope threshold, a straight line is used to create the segmentation cut [8]. Success of this method using 450 images of postal-code is achieved 94.9%. Although solution of this algorithm is explicit, it has inspired many other scientists. Yu and Yan [9] have presented a method using a morphological structural technique for segmentation. In the pre-processing step, smoothing and linearization is performed and structural points of the image contours used for defining segment cuts are detected. Success rate for 3287 images extracted from NIST database was between 85 and 95% [9]. Suwa and Naoi [10] presented an algorithm based on segmentation-recognition which uses image skeleton as input. Edges are extracted from the skeleton and a connected graph is made. Authors have reported a success rate of 88.75% for 2000 images adopted from NIST SD19 database [10]. In [11], binary PSO is used to present an algorithm for reducing number of features required for recognition of Farsi handwritten digits. Results of implementing this algorithm on Hoda database has shown a recognition accuracy of 99.4% without feature reduction (400 features) and recognition accuracy of 99.28% with feature reduction (64 features). In [12], an approach based on apparent features has been presented which processes data using histogram of oriented graphs. HOG is an efficient feature descriptor for handwritten digits; since it is a gradient-based descriptor, it is robust against illumination variations. In general, main idea of this structure is to use features of HOG features with SVM. Implementing this structure on MNIST database has given a recognition accuracy of 97.25%. In [13], histogram has been used as a set of features for handwritten digits. Then, the features are considered as inputs of the SVM network. Simulation results indicate a recognition accuracy of 95.6% for MNIST database. In [14], a new structure based on CNN has been used to recognize handwritten digit string. This network is a combination of feature extraction layer, dimension transfer layer and bottom to top output layer. Simulation results indicate a significant improvement compared to other methods implemented on ORAND-CAR-A and ORAND-CAR-B databases and the recognition accuracies obtained for these databases is 92.2% and 94.02%, respectively.

Author of [15] investigates presented methods in histogram modification based image watermarking. It shows that using techniques like a wisely selection of the adjacent bins, constant points of cover images and secret keys make them be a good nominee for image watermarking to resist against signal processing attacks.

In [16], the approach is presented that incorporates the global relative spatial information to the inverted index of the BoVW model. They computed histograms based on the magnitude of orthogonal vectors between PIVW

relative to the geometric center of an image. Authors of [17] examined the changes in results for handwritten digits with changes in the hidden layers. The accuracy curves were generated for the different factors using CNN MNIST digit dataset. The layers were taken at random in a periodic series so that each case performs in a different way during the test.

In [18], they have presented a complete literature review on different techniques for CBIR and image representation. They argue image features representation using low-level visual features. Due to diversity in non-homogeneous image properties, they cannot be represented by using single feature representation. To increase the efficiency of CBIR and image representation, one way is to use low-level features in fusion. On the other hand, researchers used deep learning methods to extract features from images and this approach can be useful in improving the performance of different areas such as the face recognition [19] and human pose estimation [20]. In these methods, they used CNNs for proper feature extraction and decision making. Meanwhile, [21, 22] investigated the role of asymmetric left and asymmetric right face images accurate age estimation and used the data augmentation for makeup-invariant face recognition.

The subject of this paper is to recognize and search for handwritten phone numbers in a phonebook. For example, the phone number images are captured from a camera or other inputs. The proposed algorithms recognize digits from the images using deep neural networks.

The rest of this paper is organized as follows. Section 2 reviews components of the proposed algorithm including database collection and Farsi handwritten phone number recognition. Section 3 presents results of the proposed algorithm implementation in MATLAB. Finally, the paper is concluded in Sect. 4.

2 Proposed method

In this section, components of the proposed algorithms for Farsi handwritten phone number recognition are studied.

2.1 Collecting Farsi handwritten phone number data

Since purpose of this paper is to recognize and search for Farsi handwritten phone numbers in a phonebook and there is no database for Farsi handwritten phone numbers, a set of such data is collected first. Examples of such handwritten phone numbers are given in Fig. 1.

Phone numbers are collected from 45 individuals. Each individual has written the digits with fonts of 7 different pen sizes such that 40 individuals have written 6–7 digits

Fig. 1 Examples of various Farsi handwritten phone numbers

and 5 individuals have written 12–14 digits with each font size.

2.2 Proposed Farsi handwritten digit string segmentation algorithm

First step is to implement the digit string segmentation algorithm. Since final purpose of this paper is to recognize Farsi handwritten phone numbers which is a digit string, these phone numbers should be segmented as a single digit. Block diagram of this segmentation algorithm is shown in Fig. 2.

In Fig. 2, the input image including Farsi handwritten digit string is introduced to the block which converts the image into grayscale. Then, the converted image is

Fig. 2 Block diagram of the proposed segmentation algorithm

introduced to the binary block to be converted to binary. Because we should define the threshold for each image to convert directly the input image to binary, and this is not constant. Because it requires binary images to continue the preprocessing process such as morphological operation and connected component detection. Meanwhile, all the digit information is preserved in the binary image. In the proposed algorithm, median and morphology filter are applied to the binary image comprising a Farsi handwritten digit string. Median filter reduces the impulse noise of the images. After applying median filter to the image, the image is filtered through morphology operator. In morphology, opening and closing operators are used. In fact, these operators are used to open points of the image which are connected incorrectly and connect points of the image which are separated incorrectly. The operators detect correct connection using filters considering connection, disconnection, distance and thickness of the image (digit) and apply it to the image. Finally, output image of the previous block is applied to the framework detection block as input. In this block, range of each digit is specified as in Fig. 3 considering value of each pixel in digit string image. In this segmentation method, the input image is binary and value of pixels in white areas of each digit is represented with 1 and black background represents 0 such that distance between digits is specified with zero. Thus, segmentation can be done considering value of each pixel and its neighborhood values, if a sudden change occurs in pixel value from zero to one or vice versa. This indicates range of each digit is the pixels in which a sudden change has occurred. Therefore, all digits of each input image including several digits are segmented using this method. Since in this segmentation method, range of each digit is determined considering pixel values and each segment including one digit is selected, the inserted block sticks to the digit from top and bottom as well as left and right.

2.3 Proposed recognition system for single digit

First step in implementation of the proposed Farsi handwritten phone number recognition system is to implement the recognition system which can recognize single Farsi handwritten digits with high accuracy. The implemented recognition system is trained using Hoda

Fig. 3 Determining range of digits in a digit string

database [14]. Block diagram of this recognition method is shown in Fig. 4.

According to Fig. 4, it can be seen that this recognition system is a single Farsi handwritten digit recognition system; the classifier used in the proposed recognition algorithm is a CNN algorithm with three convolution layers and two pooling layers. The pooling layer used in this classifier is the max-pooling layer. Size of convolution layers from first to third is 32, 64 and 128. Size of each pooling layer is 2×2 . The proposed structure is comprised of 7 separate blocks. In this recognition system, digits image is first introduced to the convolution operator block, then convolution operator is applied to all instances of the images and the outputs obtained from this layer enter the next block or the pooling layer. In this block, pooling layer is applied to the samples transferred from the previous block and outputs are created considering type of this operator. For the second time, these samples enter the next convolution layer and after applying convolution, they enter the pooling layer and integration is applied to them. For the last time, outputs transferred from the previous stage enter the convolution layer and after applying the convolution layer to all these samples, they are applied as input to fully-connected layer. The handwritten digit recognition system is trained using this layer through updating a set of weights. Since classification has 10 classes, last layer

of the CNN classifier is a fully-connected layer with 10 neurons.

2.4 Creating a new database using Farsi handwritten phone numbers

In Sects. 2.2 and 2.3, a digit string segmentation system and a single Farsi handwritten digit recognition system are presented. By integrating these two algorithms, a single Farsi handwritten digit recognition system from collected Farsi handwritten phone numbers is developed. Structure of the proposed handwritten digit string recognition is shown in Fig. 5.

As can be seen in Fig. 5, this recognition structure is comprised of 4 main blocks. First, images of phone numbers including digit string are applied to segmentation block which is based on recognition. In fact, this segmentation block is the segmentation block described in Fig. 2. After segmenting digit strings, segmented digits are applied to CNN so that each segment is recognized. CNN classifier is the system trained for recognizing single Farsi handwritten digit using Hoda database and it is shown in Fig. 4. After recognizing segment of each image including digit string, each recognized digit is stored in the file regarding its class which is a digit between 0 and 9 for creating a new database. This is done for all images collected from Farsi handwritten phone numbers and finally, a new database is created for Farsi handwritten digits from 0 to

Fig. 4 **a** Block diagram and **b** schematic of the proposed single Farsi handwritten digit recognition system based on CNN classifier

Fig. 5 Block diagram of the proposed Farsi handwritten digit string recognition system

9. Hereafter, this new database is called handwritten digit database (PAD), which is used along with Hoda database for other sections of digit and phone number recognition.

2.5 Proposed Algorithm for Searching Farsi Handwritten Phone Number in Phonebook

In this section, an algorithm is proposed which can be used by the system to receive image of Farsi handwritten phone number as input and search for the equivalent phone number in the phonebook. Structure of the proposed algorithm is shown in Fig. 6. First, CNN classifier is trained through integrating PAD and Hoda Farsi handwritten digit databases. After adjusting parameters of the CNN classifier, the proposed algorithm can be used to search for phone numbers.

Considering the block diagram shown in Fig. 6, this algorithm has 6 main blocks. First, input images considered as test images are applied to the segmentation block. This segmentation algorithm which is based on recognition is the algorithm shown in Fig. 2 and all of its components have been described in previous sections. Segmentation block segments the image including a Farsi handwritten phone number and applies it to the CNN classifier block. In the CNN block shown in Fig. 4, the segmented images are recognized as single digits. Then, the digits are introduced to the phone number recognition block. In this block, recognized digits of each digit string are inserted in an array resulting in the phone number equivalent to the input image. Then, the recognized phone number is

Fig. 6 Proposed algorithm to search for Farsi handwritten phone numbers in phonebook

introduced to this block to be compared with phone numbers of the phonebook; after digit by digit comparison of the recognized phone number with all phone numbers of the phonebook, label of the phone number which has maximum number of similar digits with the recognized phone number is returned as the selected class. By repeating this procedure for all images, accuracy of the proposed algorithm for searching for phone numbers is obtained and reported to the accuracy block.

3 Implementation of the proposed algorithms

In this section, the proposed algorithms are implemented using MATLAB 2017b and their results are investigated.

3.1 Implementing the proposed segmentation algorithm

The algorithm proposed for segmentation of Farsi handwritten digit string which its block diagram is shown in Fig. 2 has multiple subsequent steps: first, an image of the set of Farsi handwritten phone numbers described in Sect. 2.1 is applied to the proposed algorithm as an input which is a string of digits. Figure 7a shows an example of collected Farsi handwritten phone numbers. Since image

Fig. 7 Applying steps of the segmentation algorithm on an input image. **a** input image, **b** grayscale image, **c** binary image, **d** filtered image, **e** applying open morphological operation, **f** applying closed morphological operations and **g** segmented image

color does not offer any additional information about segmentation and recognition of digits, in the next step, if the input image is a color image, it is converted to a grayscale image. The grayscale image is shown in Fig. 7b. In the next step, in order to convert the grayscale image into binary, a threshold value is determined considering values of all pixels of the grayscale image. Considering this threshold, values of the image pixels are represented with 0 or 1, i.e. with black and white. The binary image is given in Fig. 7c. In the next step, median filter is applied to the binary image and the image shown in Fig. 7d is obtained.

In the next step, output image of the previous block or the filtered image is expanded using open morphological operations and the expanded image is shown in Fig. 7e. This operation is used to eliminate discontinuities between digit string and the digits. In the next steps, the expanded image is eroded using closed morphological operations. In fact, purpose of these operations is to eliminate the possible connections among digits in a string of digits. Output

image of this block is shown in Fig. 7f. Finally, the binary image is created which is applied as input to the last block to determine range of each digit. In this step, considering zero and non-zero values of pixels, range of each digit is determined. Segmented image is shown in Fig. 7g. As can be seen in Fig. 7g, range of each digit is determined by red rectangular blocks or the input handwritten digit string is segmented as single handwritten digits. It should be noted that in this step, the digit string is segmented but the algorithm cannot recognize value of each segment. Thus, it is required to integrate this algorithm with a Farsi handwritten digit recognition algorithm.

3.2 Implementing Farsi handwritten digit recognition system

The proposed Farsi handwritten digit recognition algorithm which is shown in Fig. 4 has multiple blocks. Subsequent blocks of proposed structure are trained and tested considering Hoda Farsi handwritten digit database. Implementing each step of this algorithm is described in the following.

3.2.1 Applying images of Hoda database

Hoda database is one of the largest Farsi handwritten digit databases [23]. In this section, this database is used to train and test the proposed handwritten digit recognition algorithm. Number of images existing in this database is 78,002 images with 64*54 pixels. In the first step, data is first divided to two training and test data. Among total data, 54,600 images are selected as training data and 23,402 images are selected as test data. As mentioned in Sect. 2.3 and shown in Fig. 4, the classifier used in the proposed recognition algorithm is the CNN algorithm with three convolution layers and two pooling layers. After adjusting parameters of the CNN classifier, the algorithm is trained using the training data. After being trained, the obtained recognition model is stored. In the final step, the trained Farsi handwritten digit recognition system is evaluated using the test data. It should be mentioned that number of epochs is considered to be 10. Obtained recognition accuracy in the test step for the proposed structure is 99.34%.

3.2.2 Creation a new database for Farsi handwritten digits (PAD database)

As mentioned in Sect. 2.2, the proposed segmentation algorithm is able to segment the images of Farsi handwritten phone numbers to single digits. But this system is not able to recognize each segment and cannot detect each segment is associated to which digit. Thus, it is

required to integrate it with a single digit recognition system. On the other hand, the proposed Farsi handwritten digit recognition algorithm which is trained in Sect. 3.2 using Hoda database is able to recognize single digit. By combining the algorithms proposed in Sect. 2.2 and 2.3, a Farsi handwritten phone number recognition algorithm can be developed. Structure of the proposed algorithm is shown in Fig. 5.

In fact, the proposed algorithm segments collected images of Farsi handwritten phone numbers (digit string) first and then each segment is recognized using the recognition system and the class of each segment is determined, i.e. it is specified each segment is associated to which digit. By recognizing each segment, the whole image which includes Farsi handwritten phone number is recognized. By repeating this algorithm on all Farsi handwritten phone number images, a new database is developed for Farsi handwritten digit which is called PAD. Figure 8 shows examples of PAD database.

Characteristics of the developed database for Farsi handwritten digits called PAD are as follows:

- Total number of images is 10,125
- Size of images is 64*54
- Suffix of images is.bmp

Number of data of each class of this database is given in Table 1.

Fig. 8 Examples of images of PAD

Table 1 Number of data of each class of PAD

Digit	0	1	2	3	4	5	6	7	8	9
Number	1015	995	1010	1011	1019	1017	1016	1012	1013	1017

3.3 Urdu handwritten dataset

Urdu Handwritten dataset [24] was collected from the handwritten data of over 900 people in different age groups. After data pre-processing, they extracted data from 900 people and poor text quality data is discarded. Written samples were scanned and stored on a computer. Participants ranged in age from 22 to 60 years. The writers are native and non-native speakers of Urdu, but this does not have a direct impact on the writing style. This dataset contains samples of 10 digits and 40 characters. We used the digit samples of this dataset in this research.

3.4 Training and test of the proposed system for Hoda, PAD and Urdu databases

In order to evaluate the proposed Farsi handwritten digit recognition algorithm trained and tested using Hoda database in Sect. 3.2, it is trained and tested using PAD database and the database obtained through integration of Hoda, PAD and Urdu.

3.4.1 Using PAD database

In order to evaluate PAD database, the proposed Farsi handwritten digit recognition system which was trained and tested using Hoda database is trained and tested using PAD database. Among total number of data of PAD database which is 10,125, 7000 images are considered as training data and 3125 images are considered as test data. Recognition accuracy of this algorithm using PAD database is obtained as 98.4%.

3.4.2 Using combination of Hoda and PAD databases

In this step, handwritten digit recognition system which was trained and tested using PAD is evaluated using combination of PAD and Hoda. By integrating these two databases, 88,127 images are created in 10 classes from 0 to 9. In the developed database, there are 8800 instances in each class. Among total data, 61,600 images are used as training data and 26,527 data is used as test data. This algorithm with 10 epochs obtains an accuracy of 99.37%.

3.4.3 Using combination of Hoda, PAD and Urdu databases

In this step, handwritten digit recognition system which was trained and tested using PAD is evaluated using combination

of Hoda, PAD and Urdu. By integrating these three databases, 96,147 images are created in 10 classes from 0 to 9. In the developed database, there are about 9600 instances in each class. Among total data, 67,200 images are used as training data and 28,947 data is used as test data. This algorithm with 10 epochs obtains an accuracy of 98.81%.

3.4.4 Comparing results of the proposed system using Hoda, PAD and Urdu databases and their combination

In this section, results of implementing the proposed Farsi and Urdu handwritten digit recognition algorithm obtained using three different databases are compared. Recognition accuracies obtained for each database are shown in Table 2. Total and average run-time for training and test sets are shown in this table. The average test time shows that the proposed method can be employed in a real time purpose.

3.5 Phone number searching in the phonebook

In this step, the proposed algorithm for searching for Farsi handwritten phone number in the phonebook is implemented. Structure of this algorithm is shown in Fig. 6. Classifier of this algorithm has 3 convolution layers and 2 pooling layers and the output layer is a fully connected layer with 10 neurons. Purpose of this block is to recognize each segment of the image. In the proposed algorithm, the recognized phone number is compared with the phonebook digit by digit; after digit by digit comparison of the recognized phone number with all phone numbers of the phonebook, label of the phonebook which has maximum number of similar digits with the recognized phone number is returned as the selected class. In fact, number of similar digits between the recognized phone number and phone number of the selected class results accuracy of each input image. By repeating this procedure for all input images (test images) and applying arithmetic mean on all search accuracies (input images), total search accuracy is obtained. In order to implement the proposed algorithm for Farsi handwritten phone number recognition, PAD and

Hoda databases including 88,127 images of Farsi handwritten digits from 0 to 9 are integrated to train the algorithm and 45 images including Farsi handwritten phone numbers are used to evaluate the algorithm. (Images used for evaluation of the algorithm are out of the dataset and the network has not seen them previously). Results of implementing this algorithm indicate a recognition accuracy of 94.6% for phone number searching in the phonebook. Figure 9 shows some errors in implementation of this algorithm.

Fig. 9 Segmentation and recognition errors

Table 2 Recognition accuracy of the proposed Farsi handwritten digit recognition algorithm for different databases

Database	Total training time (s)	Number of training data	Average training time (ms)	Total test time (s)	Number of test data	Average test time (ms)	Recognition accuracy (%)
Hoda	5754.1	54600	105.38	132.19	23402	5.6	99.34
PAD	771.50	7000	110.21	18.87	3125	6	98.4
Urdu	549.81	5600	98.18	12.92	2420	5.3	96.57
Hoda + PAD	7077.7	61600	114.9	140.44	26527	5.3	99.37
Hoda + PAD + Urdu	7900.3	67200	117.56	156.44	28947	5.4	98.81

Considering Fig. 9, the results can be analyzed as follows:

In Fig. 9a, the ninth digit is 1 which is incorrectly classified as class 9 due to similarity.

In Fig. 9b, 0 which is the seventh digit is incorrectly classified in class 5 due to similarity.

In Fig. 9c, fourth digit which is 3 is incorrectly recognized as 2 due to similarity.

In Fig. 9d, the segmentation algorithm has considered the fifth and sixth digits as one segment due to their similarity; hence, the recognition system has classified this segment as class 3, incorrectly.

In Fig. 9e, seventh digit which is 4 is segmented as two segments incorrectly to discontinuity and recognition is not performed correctly; recognized segments are classified in classes 7 and 5, incorrectly.

3.6 Comparing handwritten digit recognition systems with the proposed method

In this section, results obtained from the proposed method are compared with other digit recognition algorithms. Table 3 shows results of the proposed Farsi handwritten digit recognition system along with digit recognition systems presented by different researchers.

In a general comparison, it can be seen the proposed algorithm trained and tested with Hoda and combination

of PAD and Hoda has a higher recognition accuracy compared to the other algorithms.

4 Conclusion

Human brain is able to recognize the written digits simultaneously using image features and the knowledge obtained during his life. Thus, systems with higher efficiency are developed for recognizing handwritten digits or pattern recognition using human learning models.

In this paper, an algorithm is presented for recognizing Farsi handwritten phone number. PAD database introduced in this paper is developed by segmenting Farsi handwritten phone numbers using the proposed segmentation algorithm based on Farsi handwritten digits. The proposed recognition system is trained and tested using a combination of Hoda and PAD databases. After segmentation the digit string and recognizing them by the CNN classifier, the recognized digits for each digit string are arranged in the array, and thus the entire phone number is obtained. Also, in this paper, a Farsi phone number search algorithm is also proposed for handwritten phone book which corresponds to the number of digits of the recognized number and their similarities. In fact, the number of similar digits between the recognized phone number and the selected class phone number results in the search accuracy for each input image. Results of implementing

Table 3 Comparing recognition accuracy of various algorithms for handwritten digit recognition

References	Database	Number of data	Recognition method	Recognition accuracy (%)
[11]	Hoda	80000	Neural network SVM	99.28
[12]	MNIST	–	Neural network SVM	97.25
[14]	ORAND-CAR-A and ORAND-CAR-B	–	Neural network CNN	92.2 and 94.02
[25]	–	–	Neural network LVQ	88.76
[26]	–	46000	Neural network CNN	95.7
[13]	MNIST	70000	Neural network SVM	95.6
[27]	MNIST	70000	Neural network CNN	99.1
The proposed method	Hoda	78002	Neural network CNN	99.34
The proposed method	PAD	10125	Neural network CNN	98.4
The proposed method	Urdu	8020	Neural network CNN	96.57
The proposed method	Combination of Hoda and PAD	88127	Neural network CNN	99.37

the proposed method indicated high accuracy of these algorithms.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants performed by any of the authors.

Informed consent Informed consent was obtained from all individual participants included in the study.

The dataset can be requested by the author email.

References

- Cireşan DC et al (2010) Deep, big, simple neural nets for handwritten digit recognition. *Neural Comput* 22(12):3207–3220
- LeCun Y et al (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
- Schmidhuber J (2015) Deep learning in neural networks: an overview. *Neural Netw* 61:85–117
- Arel I, Rose DC, Karnowski TP (2010) Deep machine learning—a new frontier in artificial intelligence research. *IEEE Comput Intell Mag* 5(4):13–18
- Zeiler MD, Fergus R (2013) Stochastic pooling for regularization of deep convolutional neural networks. [arXiv:1301.3557](https://arxiv.org/abs/1301.3557)
- Radford A, Metz L, Chintala S. (2015) Unsupervised representation learning with deep convolutional generative adversarial networks. [arXiv:1511.06434](https://arxiv.org/abs/1511.06434)
- Ribas FC et al (2013) Handwritten digit segmentation: a comparative study. *Int J Doc Anal Recognit (IJDAR)* 16(2):127–137
- Fenrich R, Krishnamoorthy S (1990) Segmenting diverse quality handwritten digit strings in near real-time. In: *Proceedings of the 4th advanced technology conference*
- Yu D, Yan H (2001) Separation of touching handwritten multi-numeral strings based on morphological structural features. *Pattern Recognit* 34(3):587–599
- Suwa M, Naoi S (2004) Segmentation of handwritten numerals by graph representation. In: *IWFHR-9 2004. Ninth international workshop on frontiers in handwriting recognition, 2004. IEEE*
- SadighiNav M, Soleymanilvari A, Khosravi H (2014) Reducing characteristics using binary particles collection for recognition of handwritten Farsi digits. *Sci J Comput Intell Electr Eng* 1:57–68
- Ebrahimzadeh R, Jampour M (2014) Efficient handwritten digit recognition based on histogram of oriented gradients and SVM. *Int J Comput Appl* 104(9):10–13
- Tuba E, Tuba M, Simian D (2016) Handwritten digit recognition by support vector machine optimized by bat algorithm. In: *WSCG 2016 - 24th conference on computer graphics, visualization and computer vision*. pp 369–376
- Zhan H, Lyu S, Lu Y (2018) Handwritten digit string recognition using convolutional neural network. In: *2018 24th International conference on pattern recognition (ICPR)*. IEEE
- Araghi TK (2019) Digital image watermarking and performance analysis of histogram modification based methods. In: Arai K, Kapoor S, Bhatia R (eds) *Intelligent computing*. SAI 2018. *Advances in intelligent systems and computing*, vol 858. Springer, Cham
- Ali N, Zafar B, Iqbal MK et al (2019) Modeling global geometric spatial information for rotation invariant classification of satellite images. *PLoS ONE* 14(7), 0219833
- Siddique F, Sakib S, Bakr A, Siddique M (2019) Recognition of handwritten digit using convolutional neural network in python with tensorflow and comparison of performance for various hidden layers. [arXiv:1909.08490](https://arxiv.org/abs/1909.08490)
- Latif A, Rasheed A, Sajid U et al (2019) Content-based image retrieval and feature extraction: a comprehensive review. *Math Probl Eng* 2019, 9658350
- Ratyal N, Taj IA, Sajid M, Mahmood A, Razzaq S, Dar SH et al. (2019) Deeply learned pose invariant image analysis with applications in 3D face recognition. *Math Probl Eng* 2019:1–21. <https://doi.org/10.1155/2019/3547416>
- Alikarami H, Yaghmaee F, Fadaeieslam MJ (2017) Sparse representation and convolutional neural networks for 3D human pose estimation. In: *2017 3rd Iranian conference on intelligent systems and signal processing (ICSPIS)*. IEEE
- Sajid M, Iqbal Ratyal N, Ali N, Zafar B, Dar SH, Mahmood MT, Joo YB (2019) The impact of asymmetric left and asymmetric right face images on accurate age estimation. *Math Probl Eng* 2019:1–10. <https://doi.org/10.1155/2019/8041413>
- Sajid M, Ali N, Dar SH, Iqbal Ratyal N, Butt AR, Zafar B, Shafique T, Aziz Baig MJ, Riaz I, Baig S (2018) Data augmentation-assisted makeup-invariant face recognition. *Math Probl Eng* 2018:1–10. <https://doi.org/10.1155/2018/2850632>
- Khosravi H, Kabir E (2007) Introducing a very large dataset of handwritten Farsi digits and a study on their varieties. *Pattern Recognit Lett* 28(10):1133–1141
- Ali H, Ullah A, Iqbal T, Khattak S (2020) Pioneer dataset and automatic recognition of Urdu handwritten characters using a deep autoencoder and convolutional neural network. *SN Appl Sci* 2(2):152
- Khatri SK, Dutta S, Johri P (2015) Recognizing images of handwritten digits using learning vector quantization artificial neural network. In: *2015 4th international conference on reliability, infocom technologies and optimization (ICRITO) (Trends and Future Directions)*. IEEE
- Alwzway HA et al (2016) Handwritten digit recognition using convolutional neural networks. *Int J Innov Res Comput Commun Eng* 4(2):1101–1106
- Nimisha J, Kumar R et al (2017) Hand written digit recognition using convolutional neural network (CNN). *Int J Innov Adv Comput Sci IJACS* 6:260–266

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.