



An efficient and improved scheme for handwritten digit recognition based on convolutional neural network

Saqib Ali¹ · Zeeshan Shaukat¹ · Muhammad Azeem¹ · Zareen Sakhawat¹ · Tariq Mahmood² · Khalil ur Rehman¹

© Springer Nature Switzerland AG 2019

Abstract

Character recognition from handwritten images has received greater attention in research community of pattern recognition due to vast applications and ambiguity in learning methods. Primarily, two steps including character recognition and feature extraction are required based on some classification algorithm for handwritten digit recognition. Former schemes exhibit lack of high accuracy and low computational speed for handwritten digit recognition process. The aim of the proposed endeavor was to make the path toward digitalization clearer by providing high accuracy and faster computational for recognizing the handwritten digits. The present research employed convolutional neural network as classifier, MNIST as dataset with suitable parameters for training and testing and DL4J framework for hand written digit recognition. The aforementioned system successfully imparts accuracy up to 99.21% which is higher than formerly proposed schemes. In addition, the proposed system reduces computational time significantly for training and testing due to which algorithm becomes efficient.

Keywords Handwritten digit recognition (HDR) · Convolutional neural networks (CNNs) · Feature extraction and classification · MNIST dataset · Deep learning · DL4J

1 Introduction

Recognizing characters from handwritten images, printed text image documents or real images is challenging in the domain of optical character recognition (OCR) [1]. For academic and commercial applications, the trade of handwriting digit recognition (HDR) is of immense concern [2, 3]. HDR is a challenging problem which researchers have been investigating by using machine learning algorithms. HDR is meant for receiving and interpreting handwritten input in the form of pictures or paper documents. However, text extraction from real images is in reality an arduous task due to huge variations in font size and shape, texture and background, etc. Handwriting character recognition is enormously employed in various research areas including processing bank check, automatic number-plate recognition

and postal address checking from envelopes, recognition of ID cards and zip codes [4–7]. Fundamental steps of character recognition (CR) are segmentation, feature extraction and classification [8]. Rapid progress in the field of character recognition is providing an evidence of the advances in learning algorithm and availability of large databases.

Various databases including MNIST, CENPARMI, CEDAR, etc., have promoted advanced research in the field of pattern recognition. Among them, MNIST is considered to have benchmark position for undertaking tasks of pattern recognition. Different classifiers like restrictive Boltzmann machines (RBMs), neural networks (NN) have been tested on MNIST dataset. Recently, recognition of handwritten digits using CNN as classifier is entering into new research zone due to diverse applications in deep learning field [9–11].

✉ Saqib Ali, saqibsaleem788@hotmail.com; Tariq Mahmood, tmsherazi@ue.edu.pk; Khalil ur Rehman, khalil_ur@hotmail.com | ¹Faculty of Information Technology, Beijing University of Technology, Beijing 100124, People's Republic of China. ²Division of Science & Technology, University of Education, Township Campus, Lahore 54000, Pakistan.



To attain higher performances in the domain of character recognition [12–16] and pattern recognition [17, 18], deep learning is fast-advancing field among other machine learning models due to its excellent feature extraction and working as best classifier characteristics. However, the deep neural networking is proved to be time taking network as a consequence of higher amount of hidden (nonlinear) layers and connections. Currently, a convolutional neural network (CNN) is most approvable tool for image recognition because it uses lesser number of hidden layers than DNN [19], relatively few parameters. It is very easy to train the system and used to extract the position-invariant features in a reasonable amount of time for its simple structure, able to map between input dataset to output dataset with temporal subsampling to offer a degree of rotation, distortion and shift invariance [17]. The CNN comprises multiple convolution layers with entirely connected convolution (which equals those in a typical artificial neural network). An image which is taken up as an input by the CNN passes straight through the convolutional sequences, nonlinear function, pooling and entirely connected layers and finally provides us an output [20–23].

Considering one of the most challenging problem in the domain of handwritten digit recognition, hitherto several schemes/algorithms have been proposed [24, 25]. Since handwritten digits can be of various orientations and styles, researchers face many challenges for automated recognition of handwritten digits [26–34]. Ciresan et al. [35] presented convolutional neural network committees for handwritten character classification. Arora [36] employed two architectures: feed-forward neural network (FWNN) and convolutional neural network (CNN) for feature extraction, training and classification of MNIST dataset constituting handwritten images. Outcomes reveal that for the handwritten digit recognition, CNN attains greater accuracy than FWNN. The digit classification accuracy for CNN is 95.63%, whereas with FWNN is 90%. Ghosh et al. [37] carried out a comparative study of deep neural network (DNN), deep belief network (DBN) and CNN on MNIST dataset. According to work, the accuracy of classified digits for CNN is > 98% with some error rates. Anil et al. [38] presented CNN trained with gradient-based learning and backpropagation algorithm for the recognition of Malayalam characters. Their algorithm produced a maximum of 75% accuracy. Their algorithm produced maximum 75% of accuracy. Shobha Rani [39] reported work on recognition of one of the most widely used South Indian script called Kannada. The training of character image samples is carried out by using one of the deep convolution neural networks. Result demonstrates the accuracy of reported CNN model is 92%. Nonetheless, the results of former schemes were not up to the mark in terms of accuracy and computational time for handwritten digit recognition process.

This study differs from other existed work because it reveals the effectiveness of CNN in terms of high accuracy and low computational time to classify handwritten digits [36–39]. With extensive literature survey, it comes to know that accuracy of various reported CNN models is not up to the mark as mentioned in Table 3. Therefore, we proposed a CNN-based framework reinforced by DL4J for HDR, which resulted in better performance rates as compared to other CNN-based methods. Java-based DL4J framework helps to obtain more diverse features from each handwritten digit image. The established CNN model can determine and recognize handwritten digits with high accuracy, as it combines the weights of convolution layers during feature extraction with fully connected layers. Afterward, for further verification of accuracy, the system is also checked by changing the number of CNN layers. We also established a variability link between error rates and accuracy for individual handwritten digit (0–9).

Moreover, architectures like CNN are computationally expensive and lead to wastage of resources when used with less complex research problems [40]. Here, we attempt to reduce the overall classification time by reducing the feature space used to train the model to get an optimal model to classify handwritten digits. The feature map reduction has been done by selecting the filter maps of a convolutional layer in the CNN randomly. The experimented results provide conclusive evidence for the usefulness of CNNs with reduced feature space to deal with less complex problems. The results signify that the projected CNN model leads to an improvement in the recognition rate compared with other CNN-based algorithms with greater accuracy. This work will open a new way toward digitalization. Furthermore, this work could be extended to letters reducing humanistic efforts, as the digit recognition performance of our proposed framework (CNN-DL4J) was beyond what can be achieved by a skilled human.

2 Literature work

Numerous researchers marked their contribution in the field of digit recognition. Distinct priorities for individual features relying on the accuracies of different features were implemented on character recognition system by Hanmandlu and Murthy [41]. Hidden Markov model using recurrent neural network (RNN) has been used to determine sequence of character in handwritten script. Graves and Schmidhuber [42] implemented this model for the classification of the handwritten Arabic words with 91% accuracy. Pal and Singh [43] utilized multilayer perceptron (MLP) for recognizing handwritten English characters and achieved accuracy up to 94% and improved computation time for training the dataset.

Neves [44] did comparison with MLP, and his model recognized offline handwritten characters with better accuracy for standard dataset NIST SD19 by implementing support vector machine (SVM). In spite of the fact, MLP was best classifier for nonlinear classes (separable) segmentation; nevertheless, it get effortlessly trapped into local minima. Younis and Alkhateeb established the implementation of deep neural network (DNN) models on MNIST dataset for solving handwritten OCR problem. Models are capable of extracting important features without preprocessing with accuracy 98.46% [45]. Dutt and Dutt [46] demonstrated multilayer CNN using Keras and Theano libraries which attained 98.7% recognition accuracy on MNIST dataset. Ghosh and Maghari [47] did comparative study on three neural network approaches demonstrating that DNN was the best algorithm with 98.08% accuracy. However, every neural network has some error rate due to similarity in digit shape (e.g., 3 and 8 and 6 and 9). After deep analysis of the related literature, it comes to know that that CNN was supposed to be the best classifier than support vector machine (SVM), K-nearest neighbor (KNN) and random forest classifier(RFC) for HDR. Therefore, in this project, the task of HDR is accomplished by using the CNN, incorporating a Deeplearning4j (DL4J) framework, with rectified linear units (ReLU) activations that have never been reported before. The goal is achieved by establishing a model that can recognize and determine the handwritten digits from its image with high accuracy and low computation time. We aim to complete this by using the concepts of convolution neural network. The proposed CNN framework is well equipped with suitable parameters for high accuracy of MNIST digit classification. The time factor is also considered for training the system. Furthermore, high accuracy is counter verified by changing the amount of CNN layers. Employment of additional pooling layers removes discretionary details in images and implants other higher-level characteristics. The MNIST dataset was used to train the network in experiments. MNIST is a handwritten digit dataset, which consists of 60,000 training images and 10,000 images in the test set [48]. The digits are centered in a fixed size (28×28) image. These algorithms are employed to determine the accuracy with which these digits are classified. CNN classification proposed for HDR seems to be superior to other approaches used for handwritten characters'/pattern identification in terms of high accuracy and low computational time. It was noticed that DL4J training and prediction speed was efficient and quite good. By implementing this approach (CNN-based framework reinforced by DL4J for HDR), higher and precise results were obtained. Though the goal is to just create a model

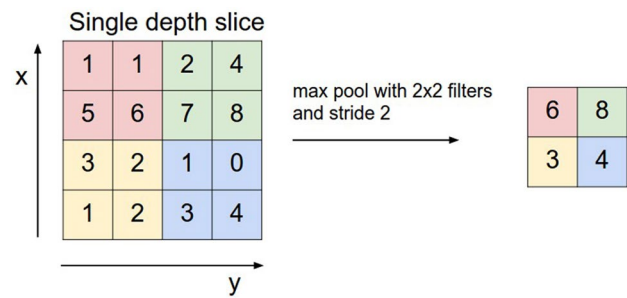


Fig. 1 Data reduction in pooling layer

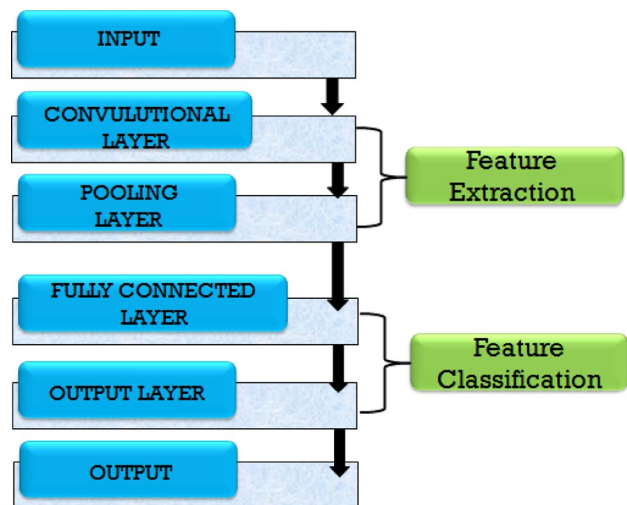


Fig. 2 Overview of CNN framework

which can recognize the digits, it can be extended to letters and then a person's handwriting.

3 CNN framework

A special kind of artificial neural network comprises of input layer, output layer and multiple hidden layers known as CNN. Hidden layer constitutes network of repetitive convolutional and pooling layers thus finally ends at one or more fully connected layer [17]. A detailed architecture of CNN is described in the following sections and shown in Fig. 2.

3.1 Convolutional layer (CNL)

CNL is the first layer in CNN which memorizes the features of input image covering its entire region during scanning through vertical and horizontal sliding filters. It adds a bias for every region followed by evaluation of scalar product of both filter values and image regions. For thresholding

element-wise activation function, such as $\max(0, \mathbf{x})$, sigmoid and tanh, is applied to output of this layer via rectified linear unit.

3.2 Pooling layer (PL)

At second, there comes pooling layer which is also called as max pooling layer or subsampling. In pooling layer (PL), shrinkage in the volume of data takes place for the easier and faster network computation. Max pooling and average pooling are main tools for implementing pooling. This layer obtains maximum value or average value for each region of the input data by applying vertical and horizontal sliding filters through input image and reduces the volume of data as shown in Fig. 1.

3.3 Fully connected layer or dense layer

Lastly, there is fully connected layer after convolution and pooling layer in the standard neural network (separate neuron for each pixel) which is comprised of n numbers of neurons, where n is the predicted class number. For example, there are ten neurons for ten classes (0–9) in digit character classification problem. However, there should be 26 neurons for 26 classes (a–z) for English character classification problem (Fig. 2).

4 Proposed methodology

Key rationale toward optical character recognition (OCR) from handwritten image includes features extraction technique supported by a classification algorithm for recognition of characters based on the features. Previously, several algorithms for feature classifications and extraction have been utilized for the purpose of character recognition. But, with the advent of CNN in deep learning, no separate algorithms are required for this purpose. However, in the area of computer vision, deep learning is one of the outstanding performers for both feature extraction and classification.

However, DNN architecture consists of many nonlinear hidden layers with a enormous number of connections and parameters. Therefore, to train the network with very less amount of samples is a very difficult task. In CNN, only few set of parameters are needed for training of the system. So, CNN is the key solution capable to map correctly datasets for both input and output by varying the trainable parameters and number of hidden layers with high accuracy [49]. Hence, in this work, CNN architecture with DeepLearning4j (DL4J) framework is considered as best fit for the character recognition from the handwritten digit images. For the experiments and verification of system's performance, the normalized standard MNIST dataset is utilized.

4.1 MNIST database used for training and testing

The subset of NIST database is MNIST dataset [48]. Out of 70,000 images of handwritten digits, 60,000 images are used for training and 10,000 images are utilized for testing [48]. Resolution of every image is 28×28 with pixel values in the range of 0–255 (gray scale). As shown in Fig. 3, 0 gray value (in black) is representing background of digit, while digit itself is appeared as 255 gray value (in white).

The MNIST dataset comprised of labeled training and test files. For training and test set files, the pixel values are arranged in row form. Therefore, training set file (images) and test set file (images) consist of 60,000 rows and 784 columns and 10,000 rows and 784 columns, respectively. On contrary, in the training and test label files, the labels' values are 0–9. Hence, 10,000 rows and ten columns for testing files followed by 60,000 rows and ten columns (0–9) for training label file.

4.2 Experimental design of CNN architecture

The performance of a CNN for a particular application depends on the parameters used in the network. In general, CNN architecture comprised of two main units or parts: (a) feature extractor and (b) feature classifier. In the feature extraction unit, every layer of network collects the

Fig. 3 Example of MNIST dataset



output from the immediate previous layer (as input) and forward the current output to the immediate next layer as inputs; contrarily, classification unit generates the predicted outputs. So, the CNN architecture with convolutional layers is implemented for MNIST digit recognition as shown in Fig. 4. The overall architecture of suggested network is enlisted below.

At first, convolutional layer having a filter map with size 5 takes (28×28) one image as input and provides output feature map with shape $(24 \times 24) 20$. After that, a pooling layer is employed. The pooling layers reduce the resolution of features. It is the technique of moving window across the 2D window space, and the maximum value in the window is the output. This depends on the size of pooling layer taken by the user. The down-sampling operation is performed by using a pool size 2×2 with stride by 2, and it reduces the output size from $(24 \times 24) 20$ to $(12 \times 12) 20$. Afterward, ReLU activation was done. ReLU or rectified linear unit has been used as the activation function. There is a wide range of activation function available when training neural network models. The mainly used activations are sigmoid, tanh, ReLU and leaky ReLU. The ReLU nonlinearity is a popular activation function used in machine learning algorithm because: (a) With ReLU, it is easier to train larger neural networks; (b) it is a simple and efficient function, which helps to solve the problem of vanishing gradients in neural networks. It removes any negative values from the output and makes sure that input and output layer sizes are the same. It replaces all the negative entities in feature maps to zero, and (c) ReLU activation function is added in

each layer so that network learns about nonlinear decision boundaries. Function for nonlinearity used has a filter map with size 5. It takes $(12 \times 12) 20$ image as input and provides output feature map with shape $(8 \times 8) 50$. Along with ReLU layer, a layer with max pooling function is used, which helps to make assumptions about features, thus reducing overfitting and also the training time. The outcome of the ReLU is directed to max pooling layer where it progressively reduces the spatial size of the feature map representations, thereby reducing the number of parameters and computations in the network. Max pooling performs the overfitting on the linearized convolved outcomes with the help of max filter and produces more abstract representation of the convolved outcomes. Down-sampling operation was performed using a pool size 2×2 with stride by 2 to reduce the output size from $(8 \times 8) 50$ to $(4 \times 4) 50$. After that, a fully connected layer is used with 1024 output nodes. Finally, another fully connected layer with ten output nodes is used to get network results for ten digits (0–9).

4.2.1 Mathematical operation

Each input node on convolution operation in convolution layer is meant for extracting features from input images, whereas the input nodes on average or maximum operation in max pooling layer abstract the features from image. The outputs of $S - 1$ th layer are utilized as input for the n th layer, and then, the inputs pass through set of kernels followed by nonlinear function ReLU (f). Suppose, if x_i^{S-1}

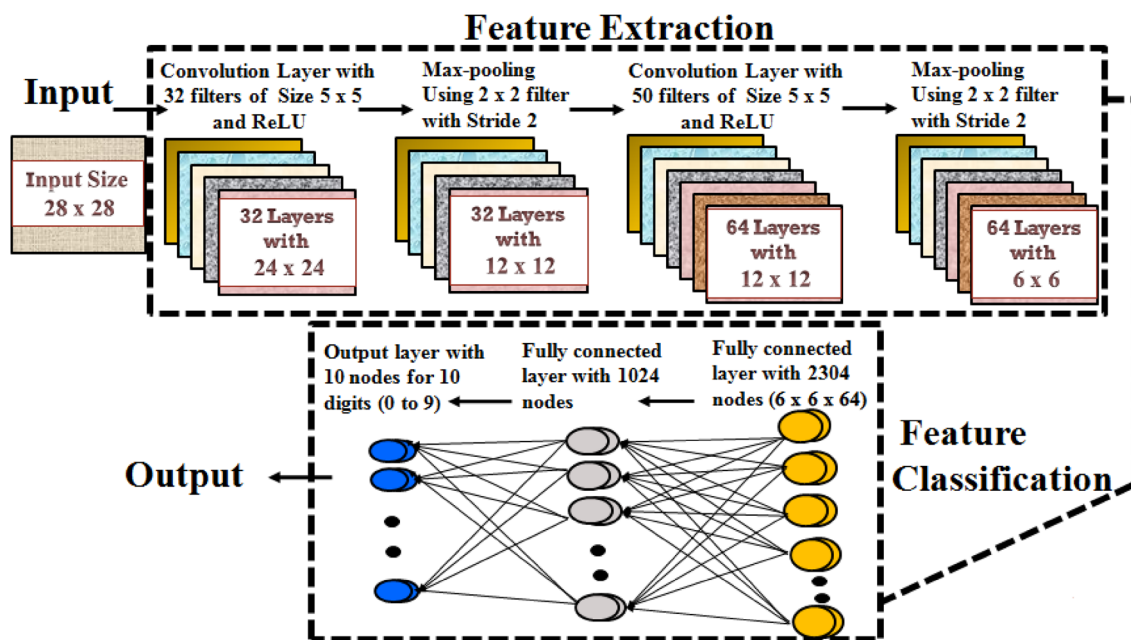


Fig. 4 Proposed design of CNN architecture

inputs from $S - 1$ th layer, k_{ij}^S are the kernels of 1th layer. b_j^S represents the biases of 1th layer. Then, the convolution operation can be stated as follows:

$$x_j^S = f(x_i^{S-1} * k_{ij}^S) + b_j^S \tag{1}$$

The input nodes on average or maximum operation in max pooling layer abstract the features from image. Here, 2×2 down-sampling kernel is applied and each output dimension will be the half of the corresponding input dimension for all the inputs. The pooling operation can be expressed as follows:

$$x_j^S = \text{down}(x_i^{S-1}) \tag{2}$$

CNN propagates features of lower-level layers to produce features of higher-level layers contrary to traditional neural networks (NN). With the feature propagation, dimension of feature reduced depending upon the size of convolution and pooling masks. However, for the enhanced accuracy for classification, one can select the feature mapping for extreme suitable features of the input images. Fully connected layer uses the output of last layer of CNN as inputs. Softmax operation is usually utilized to get the classification outputs. Softmax operation for the i th class of input sample(x), weight vector (w) and distinct linear functions (K) can be stated as follows:

$$P(y = i|x) = \frac{\exp(x^T w_i)}{\sum_{k=1}^K \exp(x^T w_k)} \tag{3}$$

5 Experimented results

All experiments are conducted on Intel 2.2 GHZ processor with 4 GB RAM, and Java is used as programming language. In this Deeplearning4j framework (DL4J)-based HDR project, MNIST is employed as a standard database. This database consists of digital handwritten images, and a number characterizes each image. All experiments in this study were performed on MNIST dataset. The research trained the system on 60,000 images from MNIST training dataset, and 5130 from testing dataset are tested.

Table 1 explains the amount of test images of each digit (0–9) used for the recognition. A result signifies that the only 42 images are classified incorrectly. This is in correspondence with higher accuracy of proposed CNN-DL4J framework. Table 2 explains the results with respect to recognition rates and execution times of dataset comprising 0–9 digits. Obtained results show significant improvement in the recognition rates of handwritten digits. In addition, the accuracy and error rates obtained through the proposed system (CNN-DL4J) show a small variation.

Table 1 Number of test images employed and its recognition accuracy (%)

Digit	Test images	Incorrectly classified images	Correctly classified images	Accuracy (%)
0	500	02	498	99.60
1	470	03	467	99.36
2	640	04	636	99.37
3	420	03	417	99.28
4	470	03	467	99.36
5	430	12	618	98.09
6	510	03	507	99.07
7	540	05	535	99.41
8	630	01	429	99.76
9	520	06	514	98.84
Total	5130	42	5088	99.21

Table 2 Recognition rates and execution time (s) of dataset comprising 0–9 digits

Digits	Recognition rate of test database	Error rate	Reject rate	Execution time (s)
0	99.60	00.40	00.00	06
1	99.36	00.64	00.00	05
2	99.37	00.63	00.00	07
3	99.28	00.72	00.00	05
4	99.36	00.64	00.00	04
5	98.09	01.91	00.00	08
6	99.41	00.59	00.00	06
7	99.07	00.93	00.00	05
8	99.76	00.24	00.00	04
9	98.84	01.16	00.00	07

Figure 5a shows error rates and accuracy among different written digits. The minimum error rate was found with digit 8 while maximum with digit 5. However, the result of accuracy is in reciprocal to error rate. The highest accuracy (99.76%) was found with digit 8 and minimum with digit 5 (98.09%).The reason behind this variability is that we employed different writing styles and digit sizes while taking testing dataset. Consequently, proposed schemes attained the average accuracy of 99.21%. To increase the accuracy rate and decrease the error, we recommend a larger size dataset.

In addition to these results, the study hypothesized that increasing the layers of CNN could result in high accuracy rates of handwritten digits. Therefore, hypothesis was tested by employing ten CNN layers. Figure 5b shows a comprehensive comparison between the accuracies achieved in percentage using five CNN and ten CNN layers. The results demonstrated an overall decrease

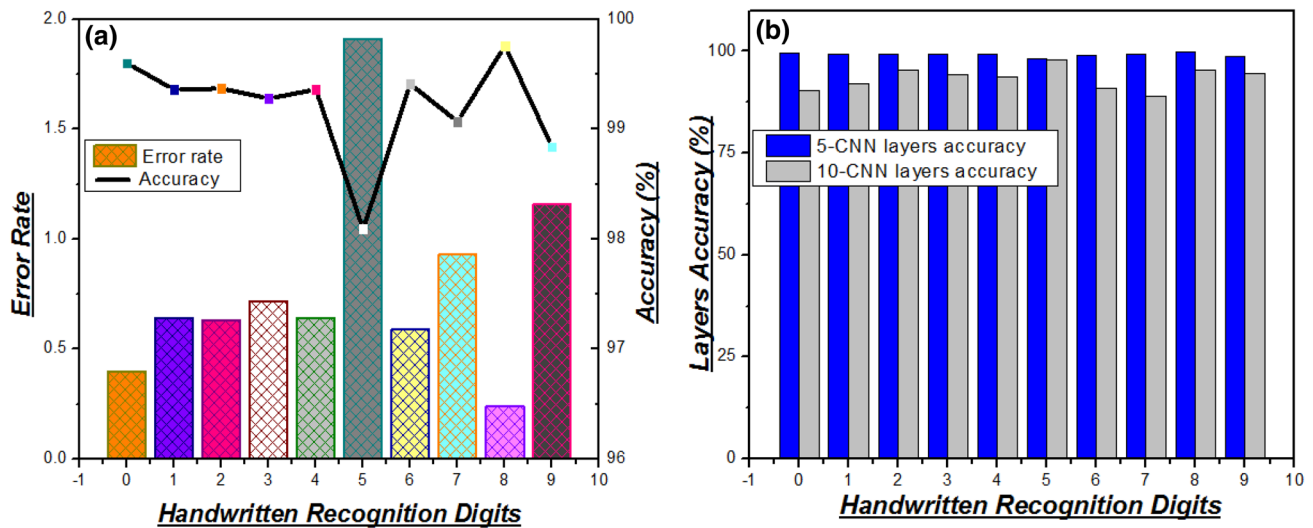


Fig. 5 **a** Plot comparing error rate and accuracy (%) with handwritten recognized digits and **b** comparison of five CNN layers accuracy with ten CNN layers in percentage

in the accuracy by increasing CNN layers. The increase in error rate was significant in digit 0 as displayed in Table 3. These results are concordant with many other studies that favor the use of less number of CNN layers when dealing with small datasets and vice versa. Furthermore, besides the number of CNN layers, accuracy in digits recognition is also dependent on other factors (e.g., size of dataset, shape of images, writing style of digits, etc.).

From all the above results, the research concludes that five CNN layers with DL4J framework are preferred and highly effective model for HDR. Table 3 shows the accuracy comparison between former schemes and currently proposed system by using CNN framework.

6 Conclusion

Character recognition is a key step toward artificial intelligence and computer vision field. To test the classification performance, pattern classification and machine learning communities use the problem of handwritten digit recognition as a model.

In this paper, recognition of handwritten digit using convolutional neural network (CNN), incorporating a Deeplearning4j (DL4J) framework, with rectified linear units (ReLU) activation is implemented. The proposed CNN framework is well equipped with suitable parameters for high accuracy of MNIST digit classification. Time

Table 3 Accuracy comparison of various CNN frameworks

Authors	Method	Accuracy (%)	References
Younis and Alkhateeb	CNN	98.46	[45]
Dutt and Dutt	CNN (Keras+Theano)	98.70	[46]
Ghosh and Maghari	DNN	98.08	[47]
Jana	CNN+CNN in Keras	98.85	[50]
Lee et al.	LeNet-4-CNN	98.9	[51]
Sahuand et al.	SVM	98.6	[52]
Espana-Boquera et al.	Robert edge	99.01	[53]
Katiyar and Mehfuz	Adaptive MLP	98.30	[54]
Siddique et al.	Artificial neural network (ANN)	97.32	[55]
Hossain and Ali	CNN	99.15	[56]
Edwin Antonio Enriquez et al.	CNN	98	[57]
Chakraborty et al.	CNN	99	[40]
Zhao and Liu	CNN and multi-level fusion	≥ 98	[58]
Present work	CNN+DL4J	99.21	–

factor is also considered for training the system. Afterward, for further verification of accuracy, the system is also checked by changing the number of CNN layers. It is worth mentioning here that CNN architecture design consists of two convolutional layers, one with 32 filters with window size 5×5 and second with 64 filters with window size 7×7 . The experimented results demonstrate that the proposed CNN framework for MNIST dataset exhibits high performance in terms of time and accuracy as compared to previously proposed systems. Consequently, handwritten digits are recognized with high accuracy (99.21%) in small computation time.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

- Epshtein B, Oyek E, Wexler Y (2010) Detecting text in natural scenes with strokes width transform. In: IEEE conference on computer vision and pattern recognition, pp 1–8
- Arica N, Vural FTY (2001) An overview of character recognition focused on offline handwriting. *IEEE Trans Syst Man Cybern Part C Appl Rev* 31(2):216–233
- Al-Hmouz R, Challa S (2010) License plate localization based on a probabilistic model. *Mach Vis Appl* 21:319–330. <https://doi.org/10.1007/s00138-008-0164-9>
- Barroso J, Dagless EL, Rafael A, Bulas-Cruz J (1997) Number plate reading using computer vision. In: Proceedings of the IEEE international symposium on industrial electronics. Guimarães, 7–11 July 1997, pp 761–766
- Bušta M, Neumann L, Matas J (2017) Deep TextSpotter: an end-to-end trainable scene text localization and recognition framework. In: IEEE international conference on computer vision (ICCV). Venice, 22–29 Oct 2017, pp 2223–2231
- Neumann L, Matas J (2010) A method for text localization and recognition in real-world images. In: Asian conference on computer vision. Springer, Berlin, pp 770–783
- Neumann L, Matas J (2012) Real-time scene text localization and recognition. In: IEEE conference on computer vision and pattern recognition (CVPR). Providence, RI, 16–21 June 2012, pp 3538–3545
- Yang JB, Shen KQ, Ong CJ, Li XP (2009) Feature selection for MLP neural network: the use of random permutation of probabilistic outputs. *IEEE Trans Neural Netw* 20(12):1911–1922
- Lee S, Son K, Kim H, Park J (2017) Car plate recognition based on CNN using embedded system with GPU. In: 10th international conference on human system interactions (HSI). Ulsan, 17–19 July 2017, pp 239–241
- Raus M, Kreft L (1995) Reading car license plates by the use of artificial neural networks. In: Proceedings of the 38th midwest symposium on circuits and systems, vol 1, pp 538–541
- Toselli AH, Romero V, Pastor M, Vidal E (2010) Multimodal interactive transcription of text images. *Pattern Recogn* 43:1814–1825. <https://doi.org/10.1016/j.patcog.2009.11.019>
- Ciresan DC, Meier U, Gambardella LM, Schmidhuber J (2011) Convolutional neural network committees for handwritten character classification. In: IEEE, Sept 2011, vol 1, pp 1135–1139
- Ciresan D, Meier U, Schmidhuber J (2012) Multi-column deep neural networks for image classification. In: 2012 IEEE conference on computer vision and pattern recognition, 1, 3, 4, pp 3642–3649
- Goodfellow IJ, Bulatov Y, Ibarz J, Arnoud S, Shet V (2013) Multi-digit number recognition from street view imagery using deep convolutional neural networks. [arXiv:1312.6082](https://arxiv.org/abs/1312.6082)
- Simard PY, Steinkraus D, Platt JC (2013) Best practices for convolutional neural networks applied to visual document analysis. In: 2013 12th international conference on document analysis and recognition, vol 2, pp 958–958
- Wang T, Wu DJ, Coates A, Ng AY (2012) End-to-end text recognition with convolutional neural networks. In: International conference on pattern recognition, vol 1, pp 3304–3308
- Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. In: Advances in neural information processing systems, 1, 3, pp 1097–1105
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2014) Going deeper with convolutions, 1, 3. [arXiv:1409.4842](https://arxiv.org/abs/1409.4842)
- LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
- Nielsen M (2015) Neural networks and deep learning. Determination Press, San Francisco
- https://en.wikipedia.org/wiki/Convolutional_neural_network
- <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/>
- O'Shea K, Ryan N et al (2015) An introduction to convolutional neural networks. [arXiv:1511.08458](https://arxiv.org/abs/1511.08458)
- Boukharouba ABA (2017) Novel feature extraction technique for the recognition of handwritten digits. *Appl Comput Inform* 13(1):19–26
- Mohebi E, Bagirov A (2014) A convolutional recursive modified self organizing map for handwritten digits recognition. *Neural Netw* 60(C):104–118
- Xu Q, Pan G (2017) SparseConnect: regularising CNNs on fully connected layers. *Electron Lett* 53(18):1246–1248
- Singhal V, Aggarwal HK, Tariyal S, Majumdar A (2017) Discriminative robust deep dictionary learning for hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 55(9):5274–5283
- Polania LF, Barner KE (2017) Exploiting restricted Boltzmann machines and deep belief networks in compressed sensing. *IEEE Trans Signal Process* 65(17):4538–4550
- Wang H, Chen P, Kwong S (2017) Building correlations between filters in convolutional neural networks. *IEEE Trans Cybern* 47(10):3218–3229
- Ercoli S, Bertini M, Bimbo AD (2017) Compact hash codes for efficient visual descriptors retrieval in large scale databases. *IEEE Trans Multimed* 19(11):2521–2532
- Katayama N, Yamane S (2017) Recognition of rotated images by angle estimation using feature map with CNN. In: IEEE 6th global conference on consumer electronics (GCCE), 21 Dec 2017, pp 1–2
- Dong Z, Zhou Z, Li ZF, Liu C, Jiang YN, Huang P, Liu LF, Liu XY, Kang JF (2017) RRAM based convolutional neural networks for high accuracy pattern recognition and online learning tasks. In: Silicon nanoelectronics workshop (SNW), 4–5 June 2017, pp 145–146
- Teow MYW (2017) Understanding convolutional neural networks using a minimal model for handwritten digit recognition. In: IEEE 2nd international conference on automatic control and intelligent systems (I2CACIS), 21 Oct 2017, pp 167–172
- Tan HH, Lim KH, Harno HG (2017) Stochastic diagonal approximate greatest descent in convolutional neural networks. In: IEEE

- international conference on signal and image processing applications (ICSIPA), 12–14 Sept 2017, pp 451–454
35. Ciresan DC, Meier U, Gambardella LM, Schmidhuber J (2011) Convolutional neural network committees for handwritten character classification. In: 2011 international conference on document analysis and recognition (ICDAR). IEEE, pp 1135–1139
 36. Arora S, Bhatia MPS (2018) Handwriting recognition using deep learning. In: Keras, international conference on advances in computing, communication control and networking (ICACCCN2018), vol 18, pp 142–145
 37. Ghosh MMA, Maghari AY (2017) A comparative study on handwriting digit recognition using neural networks. In: IEEE, 2017
 38. Anil R, Manjusha K, Kumar SS, Soman KP (2015) Convolutional neural networks for the recognition of Malayalam characters. In: Proceedings of the third international conference on frontiers of intelligent computing: theory and applications (FIC-TA) 2014. Springer, Cham, pp 493–500
 39. Shobha Rani N, Chandan N, Sajan Jain A, Kiran HR (2018) Deformed character recognition using convolutional neural networks. *Int J Eng Technol* 7(3):1599–1604
 40. Chakraborty S, Paul S, Sarkar R, Nasipuri M (2019) Feature map reduction in CNN for handwritten digit recognition. Springer, Singapore, pp 143–148
 41. Hanmandlu M, Murthy OVR (2007) Fuzzy model based recognition of handwritten numerals. *Pattern Recogn* 40(6):1840–1854
 42. Graves A, Schmidhuber J (2009) Offline handwriting recognition with multidimensional recurrent neural networks. In: Advances in neural information processing systems NIPS'22, vol 22. MIT Press, Vancouver, pp 545–552
 43. Pal A, Singh D (2010) Handwritten english character recognition using neural network. *Int J Comput Sci Commun* 1(2):141–144
 44. Neves RFP, Filho ANGL, Mello CAB, Zanchettin C (2011) ASVM based off-line handwritten digit recognizer. In: International conference on systems, man and cybernetics, IEEE Xplore, Brazil, 9–12 Oct 2011, pp 510–515
 45. Younis KS, Alkhateeb AA (2017) A new implementation of deep neural networks for optical character recognition and face recognition. In: Proceedings of the new trends in information technology, Jordan, Apr 2017, pp 157–162
 46. Dutt A, Dutt A (2017) Handwritten digit recognition using deep learning. *Int J AdvRes Comput Eng Technol* 6(7):990–997
 47. Ghosh MMA, Maghari AY (2017) A comparative study on handwriting digit recognition using neural networks. In: International conference on promising electronic technologies, pp 77–81
 48. <http://yann.lecun.com/exdb/mnist/>
 49. Jain N, Rahul K, Khamaru I, Jha AK, Ghosh A (2017) Hand written digit recognition using convolutional neural network (CNN). *Int J Innov Adv Comput Sci* 6:260–266
 50. Jana R, Bhattacharyya S (2019) Character recognition from handwritten image using convolutional neural networks In: Recent trends in signal and image processing, advances in intelligent systems and computing, vol 922, pp 23–30
 51. Lee S-W (1996) Off-line recognition of totally unconstrained handwritten numerals using multilayer cluster neural network. *IEEE Trans Pattern Anal Mach Intell* 18(6):648–652
 52. Sahuand MK, Dewangan NK (2017) A survey on handwritten character recognition. *Int Adv Res J Sci Eng Technol*. <https://doi.org/10.17148/IARJSET.2017.4120>
 53. Espana-Boquera S, Castro-Bleda MJ, Gorbe-Moya J, Zamora-Martinez F (2011) Improving offline handwritten text recognition with hybrid HMM/ANN models. *IEEE Trans Pattern Anal Mach Intell* 33(4):767–779
 54. Katiyar G, Mehruz S (2016) A hybrid recognition system for off-line handwritten characters. SpringerPlus 5:357
 55. Siddique MAB, Khan MMR, Arif RB, Ashrafi Z (2018) Study and observation of the variations of accuracies for handwritten digits recognition with various hidden layers and epochs using neural network algorithm. In: 4th international conference on electrical engineering and information and communication technology, pp 118–123
 56. Hossain MA, Ali MM (2019) Recognition of handwritten digit using convolutional neural network (CNN). *Glob J Comput Sci Technol D Neural Artif Intell* 19:27–33
 57. Enriquez EA, Gordillo N, Bergasa LM, Romera E, Huelamo CG (2019) Convolutional neural network vs traditional methods for offline recognition of handwritten digits, vol 855. Springer, Basel, pp 87–99
 58. Zhao H, Liu H (2019) Multiple classifiers fusion and CNN feature extraction for handwritten digits recognition. *Granul Comput*. <https://doi.org/10.1007/s41066-019-00158-6>

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