

# Handwritten Bangla Numerical Digit Recognition Using Fine Regulated Deep Neural Network

**Md. Shahadat Hossain, Md. Anwar Hossain, AFM Zainul Abadin, Md. Manik Ahmed\***

Department of Information and Communication Engineering, Pabna University of Science and Technology, Pabna, **BANGLADESH**

\*Corresponding Contact:

Email: [kironmanik.pust@gmail.com](mailto:kironmanik.pust@gmail.com)

## ABSTRACT

The recognition of handwritten Bangla digit is providing significant progress on optical character recognition (OCR). It is a very critical task due to the similar pattern and alignment of handwriting digits. With the progress of modern research on optical character recognition, it is reducing the complexity of the classification task by several methods, a few problems encounter during recognition and wait to be solved with simpler methods. The modern emerging field of artificial intelligence is the Deep Neural Network, which promises a solid solution to these few handwritten recognition problems. This paper proposed a fine regulated deep neural network (FRDNN) for the handwritten numeric character recognition problem that uses convolutional neural network (CNN) models with regularization parameters which makes the model generalized by preventing the overfitting. This paper applied Traditional Deep Neural Network (TDNN) and Fine regulated deep neural network (FRDNN) models with a similar layer experienced on BanglaLekha-Isolated databases and the classification accuracies for the two models were 96.25% and 96.99%, respectively over 100 epochs. The network performance of the FRDNN model on the BanglaLekha-Isolated digit dataset was more robust and accurate than the TDNN model and depend on experimentation. Our proposed method is obtained a good recognition accuracy compared with other existing available methods.

## Key words:

Bangla Handwritten Recognition, OCR, DCNN, TDNN, FRDNN

7/19/2021

Source of Support: None, No Conflict of Interest: Declared

This article is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Attribution-NonCommercial (CC BY-NC) license lets others remix, tweak, and build upon work non-commercially, and although the new works must also acknowledge & be non-commercial.



## INTRODUCTION

Hand Written Character Recognition (HWCR) has achieved much attention in the modern day due to its applicability in a range of tasks. In our daily life, basic Bangla characters

generally appear in many conventional places such as document identifiers, forms, postal headers, house addresses, encrypted codes with private letters, handwritten flyers, posters, notices, banners, invitation cards, bank checks, tickets, etc. in Bangladesh and a portion of India. HWCR is beneficial in distinct domains number plate identification, OCR, passports, and document analysis, fraud detection, postal system automation, automatic bank cheque processing (Ahmed et al., 2019). Bangla contains similarly shaped numerals that are confused to identify even in printed form and leads Bangla HWCR as a challenging task. Therefore, effort with new techniques for better Bangla HWCR is essential and demands time. To handle the complex shape of the numerical character, researchers are focusing on the structural feature recently. But the extraction of proper structural features from a numeric character is itself a challenging task due to the variation of the characters. If the structural complexity of the Bangla digit is somehow reduced, we will get better recognition precision for those characters during the preprocessing stages (Pramanik & Bag, 2018).

More recently, Deep Convolutional Neural Network (DCNN) have been applied successfully in many pattern classification tasks exposing possible hidden elements of information through their deep design. Among various DCNN models, the convolutional neural network (CNN) has been popular due to the many advancements in the performance of image classification responsibility for traditional NN-based approaches. Our work is based on the basis that classification and recognition of the Bengali handwritten digit not only provides better recognition accuracy but also overcomes the classification complexity as the number of classes decreases. The goal of this investigation is to build up a superior Bangla HWCR system and we consequently examined the deep architecture of DCNN. Unlike other relevant works, it doesn't employ any feature extraction technique, uses regularization parameters only for optimization. To highlight the importance of deep architecture, a traditional three-layer DCNN model was also tested on a similar corpus. It is observed that FRDNN gives higher recognition accuracy than TDNN. The comprehensive investigations reveal that the proposed DCNN based FRDNN model shows the state of the art classification performance passing some existing methods. Our model was compiled by RMSprop and the Keras Fit function was used to train. To improve the recognition of image sizes with a higher degree of information retrieval and accessibility, translation, scaling, and other pattern distortions related to the characteristics of these networks. We experimentally evaluated the effectiveness of our methods in the publicly available Bangla character image database of BanglaLekha-Isolated. Our proposed fine regulated DCNN model obtained better accuracy than the existing models but this model optimized from the traditional three layer DCNN model which is presented in this paper.

## LITERATURE REVIEW

Recognition of handwritten digit involves important work in documentation correspondence and regular human-PC cooperation (HCI). Convolutional neural networks (CNN) are deep learning calculations that expand the usual false neural network by adding additional requirements to the first layers and expanding the depth of the network. Handwritten detection and recognition are the process of transforming non intelligent information such as images into machine editable content. Ramzan et al. (2018) proposed hand written digit recognition (HWDR) methods that only describe the importance of neural network and its available algorithms with different techniques or modified algorithm (Ramzan et al., 2018). Deep hog is a hybrid model that classify the isolated alphanumeric symbols of the Bangla language. Sharif & Mahboob (2019) proposed a mixed model that classifies all possible classes of the handwritten image and unify existing reference data sets.

The necessity of the individual handcrafted features in this model also has been demonstrated. The results of the experiment show that the proposed deep hog model can surpass the existing sophisticated taxonomy models in Bengali handwritten alpha numerical image classification (Sharif & Mahboob, 2019). Das et al. (2018) introduced a Bengali handwritten recognition scheme for vowels in English characters which is quite fast and it does not need training data to work and based on the structural anatomy of the characters. Some feature acquisition is accomplished by generating curvature scale spaces for characters. Binary images of handwritten letters are treated to identify physical characteristics such as 'bowl', 'lobe', and 'arm' (Das et al., 2018).

Now a days handwriting character recognition is one of the most important issues in the gift of neural networks and machine learning algorithms in pattern recognition applications (Kamran et al., 2018; Hochuli et al., 2018). Shamim et al. (2018) proposed reliable approaches to ensure effective recognition of handwritten digits. The result of this paper shows that the most noteworthy 90.37% accuracy has been gotten for multilayer perceptron (Shamim et al., 2018). Offline handwriting digit recognition using triangle geometry features shows that handwritten recognition of offline numbers is one of the most frequently studied studies today (Arbain et al., 2018). Deep convulsive neural networks a well-known famous problem with computer vision applications is the recognition of numbers from natural images and handwritten numbers. Many machine learning methods have been developed to solve the difficulty of recognizing digits from natural images and handwritten digits. Among them, the most three leading NN strategies are CNN (Rahman et al., 2015), deep neural network (DNN) model, and deep belief network model (DBN) (Thapa & Kumar, 2018). This Paper introduced on CNN, additionally recognized as DCNN that operates on image pixels directly and comparable to the three NN methods, and is evaluated for multiple factors such as accuracy, efficiency, and performance. With the proposed augmentation strategy DCNNs perform better than they do without such augmentation technique. The proposed method indicates the accuracy of amicable recognition and beats other obvious existing methods.

## IMPLEMENTATION

Deep CNNs contain millions of parameters that are learned by training. All parameters as well as gradients and error updates have to be held in memory during training. The main objective of an HWCR system is to recognize numeric BanglaLekha-Isolated (Mohammed et al., 2017) characters, which are in the form of digital images, without any human intervention. This was done by examining the similarity between the extracted features from the given character's images and the database of images.

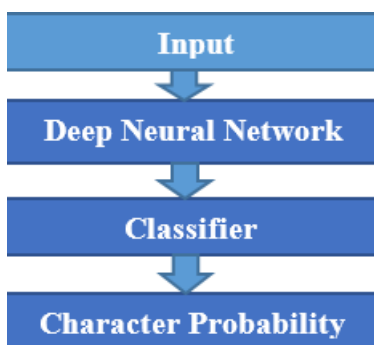


Figure 1: Handwritten bangla character recognition system

The database helps to distinguish features between images; this eliminates the confusion for proper character recognition. Firstly, the HWCR system takes data as input from the user and goes to preprocessing stage, feature extraction, and search similar features of extracted image from image model database and then classifying the characters. Lastly, the deep learning model with regularization a weight initialization is more robust and efficient than compare with the rest three models depending on the time and accuracy factor. To recognize characters from the image, a scheme has been proposed. The basic process for HWCR is shown in Figure 1.

### Data collection and Preparation

There are lots of databases open for handwritten character recognition. Depending on our demand we have chosen the BanglaLekha-Isolated multipurpose comprehensive handwritten isolated character samples dataset for our DNN model (Mohammed et al., 2017). It holds samples of 50 Bangla well known characters, 10 Bangla numerals, and 24 selected compound characters, 2000 handwriting samples were collected, perform the digital operation, and preprocessed individually for each of the 84 characters. The format of every image inside the database was portable network graphics (PNG) format. The collected dataset was organized as follows:

Table 1: Organization of BanglaLekha-Isolated multipurpose comprehensive dataset

Dataset Folder Number	Dataset	Classes
Folder: 1-11	Bangla Vowel Character	11
Folder: 12-50	Bangla Consonant Character	39
Folder: 51-60	Bangla digit Character	10
Folder: 61-84	Bangla Compound Character	24

Depend on our requirement retrieved digit character from folder 51 to 60 that contains 10 classes and over 200 images of individual digit. The character in this dataset contains mixed resolution that may affect our model.



Figure 2: Sample of selected dataset

However, this dataset contains images, taken with varied character making the particular handwritten digit recognition a challenging task. Each image is in PNG format and has a size of  $128 \times 128 \times 3$  but some were different in dimension. Further used grayscale image that remove the computational complexity. Dataset labeling is introduced in this step where corresponding image class is 1 and rest is 0. For training of deep neural networks, characters are written from individual people in individual handwriting styles with different fonts. The two neural network architectures are trained and tested on the Hand Written Character dataset which contains bangla character. A version of that dataset was used for this research. The dataset contains 15600 handwritten digit portable network graphics images from different people. Figure 2 shows a sample of the 9 static hand written digit that have been used for the experiments, taken under different features such as scale, rotation, translation, shape and illumination.

### **Traditional deep neural network**

The traditional deep neural network (TDNN) design is composed of three convolutional layers, two pooling layers and one fully connected layer with ReLU (Rectified Linear Unit). The TDNN structure of the hand written digit recognition model consists of two stages: 1) preprocessing in the first stage, and 2) pattern classification using CNN deep learning algorithm with two different architectures in the second stage. The first convolutional layer has a kernel size of  $3 \times 3$  pixels that contains 32 feature maps with a ReLU activation function. This layer takes images with  $32 \times 32$  pixel values as an input (see Figure 2).

The next hidden layer is a convolutional layer that has a kernel size of  $3 \times 3$  pixels and also contains 32 feature maps with a ReLU activation function. The next layer is a Max-Pooling layer that is configured with a pool size of  $2 \times 2$ . The Max-Pooling layer extracts sub regions of  $2 \times 2$  of the feature map, keeps their maximum value and discards all other values to progressively reduce the spatial size of the representation and reduce the number of parameters during computation, and also control overfitting. Then the next hidden layer is the convolutional layer that has a kernel size of  $3 \times 3$  pixels and also contains 64 feature maps with a ReLU activation function. This layer is conducted by another pooling layer that is the same as the previous layer. Afterward, a layer named Flatten converts the two dimensional matrix data into the vector form, thereby allowing the final output to be prepared by standard fully connected layers to reach the next layers. The first fully connected layer including the ReLU activation function contains 512 neurons.

This is followed by a dropout layer to exclude 20% of neurons to reduce overfitting. The final part of the TDNN structure is the output layer which comprises a Softmax activation function, and contains 10 neurons, one for each handwritten character recognition class. The output is the mapping of the data to the final classes for hand written recognition. Figure 3 represents the traditional CNN method.

### **Proposed fine regulated deep neural network**

Proposed FRDNN model uses the traditional CNN architecture with the fine tuning and optimization techniques applied to obtain the significant result on character recognition. This fine-tuned and optimized model, FRDNN, is shown in Fig. 4. The performance of the traditional CNN was improved by tuning the parameters to include network initialization and regularization. The proposed approach does not include data augmentation. The network architecture of FRDNN model is derived from the TDNN model and consist of some optimization parameters that make the model consistent. The difference between traditional TDNN and FRDNN architectures are illustrated in Figure 3 and Figure 4 respectively.

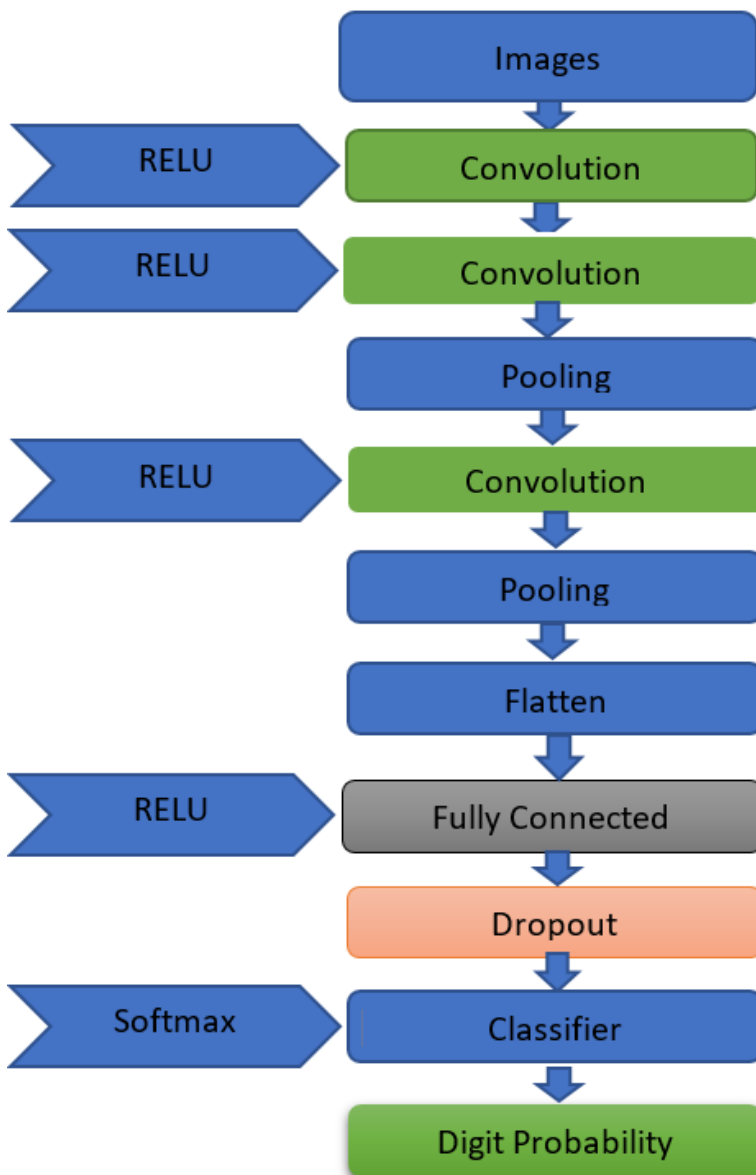


Figure 3. Architecture for traditional deep neural network (TDNN)

**Network initialization:** Initializing the weights in an appropriate way can significantly influence how easily the network learns from the training set. The FRDNN network architecture uses the uniform He initialization (`he_uniform`) for all ReLU layers and the uniform Xavier initialization (`glorot_uniform`) for the output Softmax layer for effective generalization.

**L2 Regularization:** This aims to decrease the complexity of the model while maintaining the same parameter count. L2 regularization does so by penalizing weights with large magnitudes, by minimizing their L2 norm. It uses a hyper parameter  $\lambda = 0.01$  to specify the relative importance of minimizing the norm to minimizing the loss on the training set.

**Weight constraints:** The aims of weight constraints are checking the size or magnitude of the weights and forces weights to be small and can be used instead of weight decay and in conjunction with more aggressive network configurations. This Proposed model uses maxnorm weight constraints for the flattening layer that can improve the generalization. In the FRDNN model, the above optimization parameters make the model consistent and remove overfitting that gives the model better accuracy than traditional deep neural network.

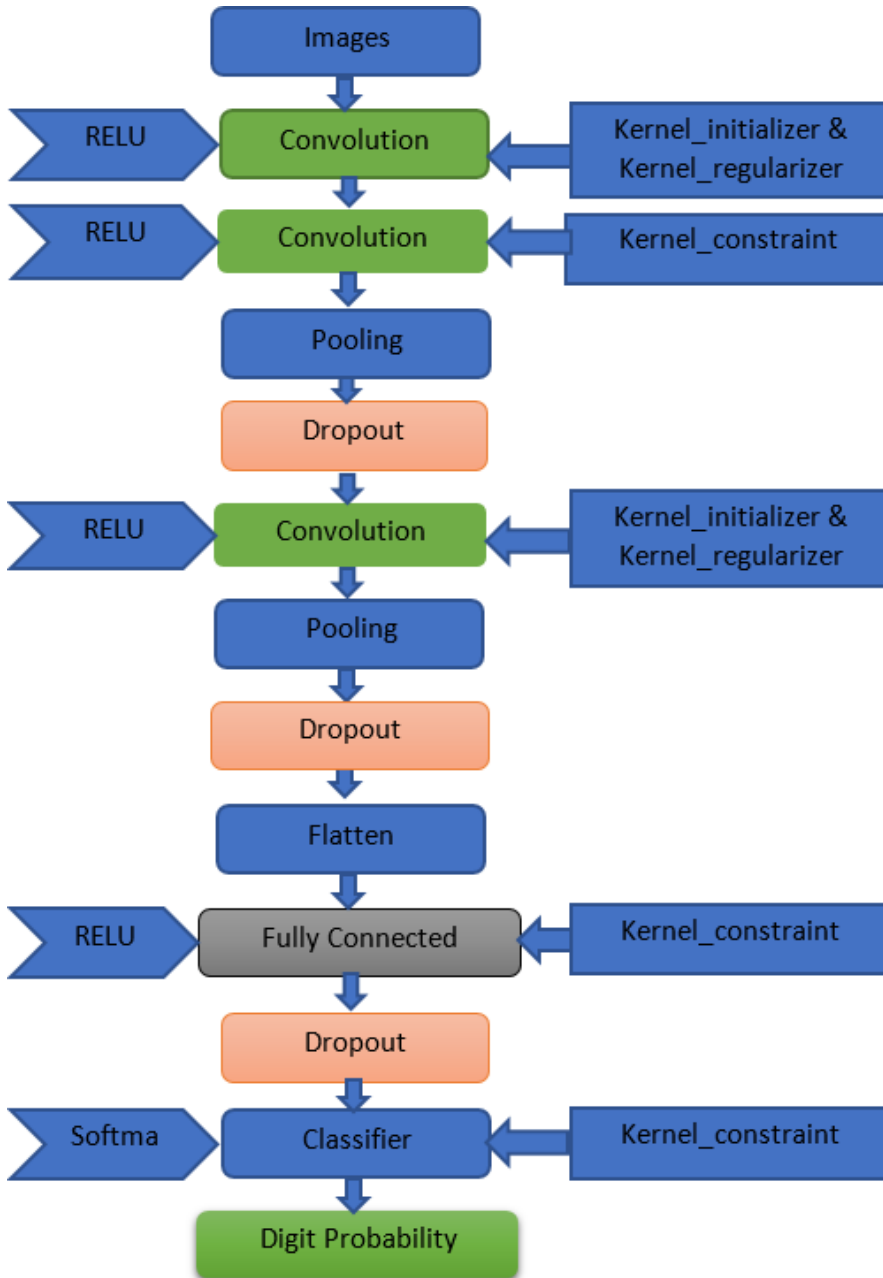


Figure 4: Architecture for proposed fine regulated deep neural network (FRDNN)

## RESULTS AND DISCUSSION

This section first provides a description of the dataset which was used to perform the experiments, and the preprocessing method for preparing the images before they are fed into the traditional deep neural network (TDNN) and proposed FRDNN models. Then, it discusses the measures adopted for evaluating classification accuracy. The convolutional layers before the first pooling layers are the bottleneck of the computational efficiency and memory requirements. In order to make a realtime classification, the images are converted to gray scale. Thus, applying image preprocessing reduces the number of parameters in the first convolutional layer and reduces computational requirements. The labeled image takes for classification extracted from the dataset pool have to be split into training, validation and test sets to train and evaluate the models. Therefore, the following procedure is used to create these files. First, parameters are given for the percentages of each split. In the evaluation process, we divided the datasets 12480 images for training and 3120 images for validation and testing images from 15600 individual images as shown in Table 2. We splitted the dataset manually but we can use split parameter 20% during evaluation which randomly create the test dataset 20 percent of total 15600 images as training set.

Table 2: Dataset for Training and testing the models

Dataset Name	Total Data	Dimension	Training	Testing	Validation
BanglaLekha-Isolated	15600	(28,28,1)	12480	3120	3120

The results obtained when using the traditional and fine regulated CNN architectures. Experiments were performed using the BanglaLekha-Isolated dataset (see Table 2), to learn features from the training data and to test the models using the testing data.

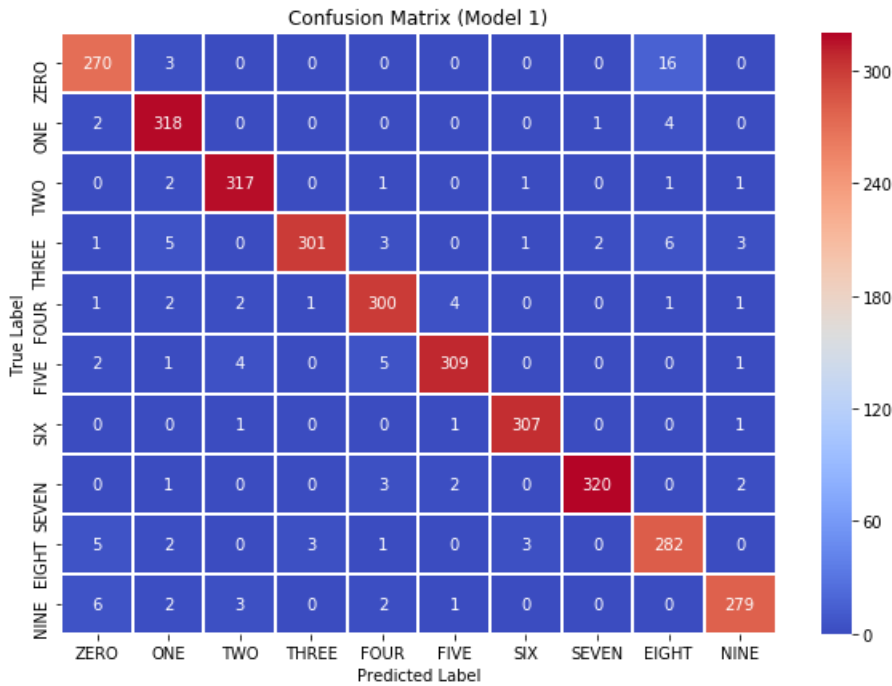
The CNN training process is based on the combination of the back propagation algorithm with the stochastic gradient descent method. The cost function chosen is the categorical Crossentropy loss function, and the RMSprop optimizer is used as the optimization function. Table 3 presents the results when using the TDNN and the proposed FRDNN architectures applied to the handwritten digit recognition dataset. The experiment results show that the FRDNN architecture achieved 96.99% accuracy, which was 0.75% higher than the performance obtained by the CNN based traditional deep neural network architecture (i.e. 96.25%). But the training step of FRDNN model is well generalized than the TDNN model.

Table 3: Classification results

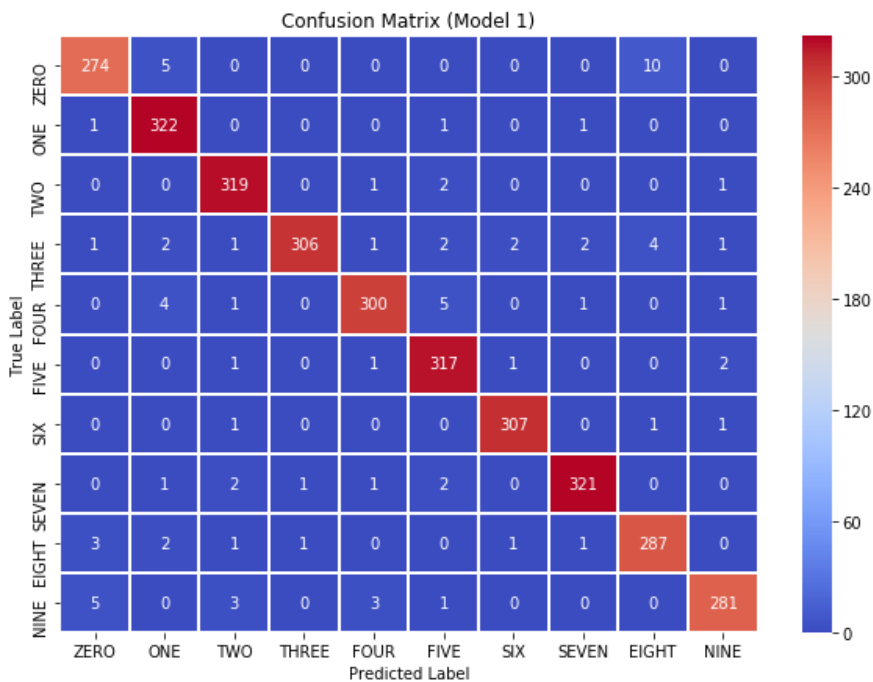
Methods	Number of epochs	Accuracy
Traditional DNN	100	96.25%
Fine Regulated DNN	100	96.99%

The confusion matrices of both TDNN and FRDNN trained with 100 epochs are shown in Fig.5 (a) and Fig.5 (b). The diagonal elements of the confusion matrices represent the number of images which were correctly classified (i.e. predicted label of the number images is equal to the true label), while non diagonal elements are those that are not labeled efficiently by the classifier. The higher the diagonal values of the confusion matrix the better the classification performance.





(a)



(b)

Figure 5: Confusion matrix (a) Traditional deep neural network, (b) proposed fine regulated deep neural network on the bangla lekha isolated dataset.

The overall classification performance of the FRDNN network is high, and the confusion matrices show that nearly all of the images which were misclassified by TDNN, were correctly classified by FRDNN model. Importantly, Fig. 6 and Fig. 7 depict the model accuracy and training loss when adopting TDNN and FRDNN, respectively. These figures show that training and testing performance were close together during different epochs, which indicates that the CNN models were not overfitting the data.

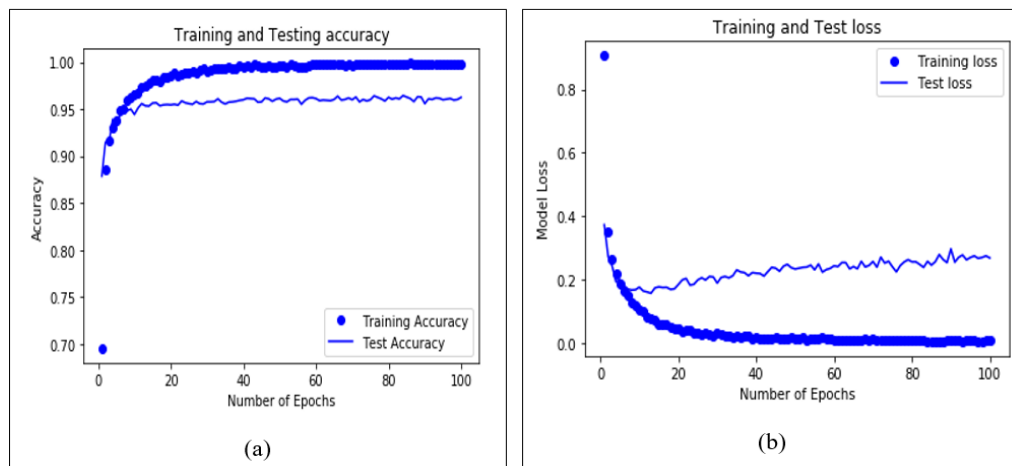


Figure 6: (a) Accuracy for training and testing data, and (b) loss during training process of TDNN model

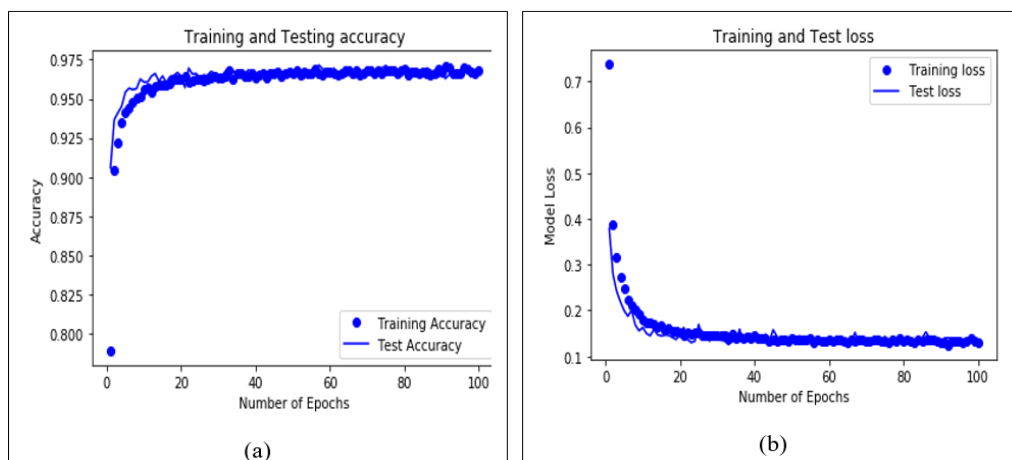


Figure 7: (a) Accuracy for training and testing data, and (b) loss during training process of FRDNN model

The improvement in CNN performance was attributed to the data adjustment, network initialization and regularization parameters which allowed the FRDNN to achieve much higher classification performance than the traditional deep neural network without overtraining. Data adjustment creates new transformed images from the base dataset, and the new images are added into the dataset to improve the model's character computational complexity. Its regularization also makes the model consistent, making the image decentralized for the kernels to recognize them from various positions.

## CONCLUSION

Moreover, the proposed FRDNN is enhanced by the presence of dropout regularization and L2 Regularization to eliminate the problem of data overfitting. With these modifications, the FRDNN results demonstrate that it is an effective method of increasing the performance of CNN for hand written numeric character recognition. The proposed FRDNN was compared to a traditional DNN. The results revealed that the proposed approach which includes data adjustment, along with a fine tuned convolutional neural network using network initialization (ReLU and Softmax), and L2 regularization achieved the highest classification recognition accuracy 96.99% that .75% improvement over the traditional deep CNN model which was not tuned using the above mentioned parameters.

## REFERENCES

- Ahmed, M., Akhand, M. A. H., & Rahman, M. M. Hafizur. (2019). Recognizing Bangla Handwritten Numeral Utilizing Deep Long Short-Term Memory. *Int. J. Image, Graph. Signal Process*, 11(1), 23–32.
- Arbain, N. A., Azmi, M. S., Muda A. K., Muda, N. A. & Radzid, A. R. (2018). Offline handwritten digit recognition using triangle geometry properties. *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.*, 10(January), 87–97.
- Biswas, M., Islam, R., Shom, G. K., Shopon, M., Mohammed, N., Momen, S., Abedin, A. (2017). BanglaLekha-Isolated: A multi-purpose comprehensive dataset of Handwritten Bangla Isolated characters. *Data in Brief*, 12, 103-107. <https://doi.org/10.1016/j.dib.2017.03.035>
- Das, P., Dasgupta, T. & Bhattacharya, S. (2018). A Bengali handwritten vowels recognition scheme based on the detection of structural anatomy of the characters. *Adv. Intell. Syst. Comput.*, 518(January), 245–252.
- Hochuli, A. G., Oliveira, L. S., Britto, A. S., & Sabourin, R. (2018). Handwritten digit segmentation: Is it still necessary? *Pattern Recognition*, 78, 1–11.
- Kamran, S. A., Humayun, A. I., Alam, S., Doha, R. M., Mandal, M. K., Reasat, T., & Rahman, F. (2018). AI Learns to Recognize Bengali Handwritten Digits: Bengali.AI Computer Vision Challenge 2018. <https://arxiv.org/abs/1810.04452v1>
- Pramanik, R., & Bag, S. (2018). Shape decomposition-based handwritten compound character recognition for Bangla OCR. *Journal of Visual Communication and Image Representation*, Volume 50, 123-134.
- Rahman M. M., Akhand, M. A. H., Islam, S., Shill, P. C. (2015). Bangla Handwritten Character Recognition using Convolutional Neural Network. *I. J. Image, Graphics and Signal Processing*, 8, 42-49.
- Ramzan, M., Khan, H. U., Awan, S. M., Akhtar, W., Ilyas, M., Mahmood, A., & Zamir, A. (2018). A survey on using neural network based algorithms for hand written digit recognition. *Int. J. Adv. Comput. Sci. Appl.*, 9(9), 519–528. <https://doi.org/10.14569/IJACSA.2018.090965>
- Shamim, S. M., Miah, M. B. A., Sarker, A., Rana, M., & Jobair, A. (2018). AI. Handwritten digit recognition using machine learning algorithms. *Indones. J. Sci. Technol.*, 3(1), 29–39.
- Sharif, S. M. A., & Mahboob, M. (2019). Deep Hog: a Hybrid Model To Classify Bangla Isolated Alpha-Numerical Symbols. *Neural Netw. World*, 29(3), 111–133.
- Thapa, R. & Kumar, D. (2018). Recognizing Digits from Natural Images and handwritten Digits using Deep Convolutional Neural Networks. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT)*, 4(1), 158-165.

--0--

**How to Cite:**

Hossain, M. S., Hossain, M. A., Abadin, A. Z., & Ahmed, M. M. (2021). Handwritten Bangla Numerical Digit Recognition Using Fine Regulated Deep Neural Network. *Engineering International*, 9(2), 73-84. <https://doi.org/10.18034/ei.v9i2.551>