Efficient Recognition of Bangla Handwritten Digits Based on Deep Neural Network

Md. Lizur Rahman, Ifrat Jahan, Akash Saha, Md. Nawab Yousuf Ali

Abstract*: *Nw-a-days world has started to move into machine based technologies. Recognition of various features, shapes, images etc., has become extremely excited topics over recent years. Many authors proposed various techniques to recognition of handwritten digits on different languages. This paper presents a new technique based on deep neural network for the purpose of efficiently recognition of handwritten digits for Bangla language. Two datasets are used in this paper including CMATERDB 3.1.1 dataset and ISI (Indian Statistical Institute) dataset. About 24500 samples are used for training purpose and 4800 samples are used for testing purpose and the proposed technique achieves 98.70 percent accuracy. This paper also presents detailed overview on artificial neurons, and deep neural network. In addition, the efficiency of proposed method shown by comparing the results with other existing techniques.

Index Terms: Handwritten Digit Recognition; Perceptron; Sigmoid Neuron; Dropout; Deep Neural Network; Artificial Neuron; Bangla Digit.

I. INTRODUCTION

Now-a-days machines have been designed for performing various activities based on the human behavior. One of the special behaviors of human is visual system. Human brain contains various visual cortices- V1, V2, V3, V4, and V5 [1]. The V1 cortex of human brain is also known as primary visual cortex contains 140 million neurons [1]. There are more than 10 billion connections between those neurons. However, these visual cortices work for different complex image processing of human visual system.

It is easy for humans to detect handwritten digits but for the machines it is difficult [2]. To write a computer program which can detect Bangla handwritten digits is extremely difficult. For example, how we explain digits- '৩' (in English 3) has a loop at the right of the top and an incomplete circle in the top left. It is easy to understand but difficult to explain in programming. We write a neural network based computer program, which can learn to detect handwritten digits. Our program is short in length but it acheives 98.70 percent accuracy.

In general, detection of handwritten digits is one of the perfect prototypes for learning deep neural networks [3]. In various ways deep neural networks can solve the handwritten digits' detection problem. Firstly, it requires a large dataset of handwritten digits. Dataset should be divided into two parts: training data, and test data. Test dataset contains the data which have never been used before for training. Then a system is required, which can learn from training datasets. If the training dataset is very large, the system can learn very efficiently with a good accuracy. Finally, test datasets use for find out the efficiency and accuracy of the system. In this study, we develop a system implementing neural networks for efficient detection of Bangla handwritten digits. We also show an overview about neural networks, and different artificial neurons.

This paper is organized as follows. We start with a discussion of the various previous related works in section II. Then we explain about different artificial neurons and deep neural network in section III. Section IV focuses our research methodology along with dataset information. In section V, we analyze the result of our experiment and compare out result with others related works results. Finally, we show some concluding remarks and future direction in section VI.

II. RELATED WORK

Soltanzadeh & Rahmati [4] proposed a technique for recognition of Arabic handwritten digits. They used support vector machine (SVM) approach for classification of feature of input images. They used data cleaning technique to remove unnecessary data such as extra writings, wrong digits etc. Authors used 4974 images for training purpose and 3939 images for testing purpose. In another similar work by Ashiquzzaman & Tushar [5] proposed a new algorithm for recognition Arabic handwritten numeric images. They used L1 regularization method and activation function on their algorithm to improve their accuracy. They claimed that their proposed method provides 97.4 percent accuracy.

A recent study by Parvin et al. [6] discussed a technique to improve the accuracy of handwritten digit recognition. Authors used binary multi-layer perceptron (MLP) classifier and KNN algorithm to improve the accuracy. Their proposed technique was designed as like as binary tree. Each level of binary tree divided into two sub-classes and this approach continued until every node contains a single class. In this study, authors used Hoda dataset, which consists of 80000 samples. Among them 40000 samples used for training purpose, 20000 samples used for validation purpose, and 20000 samples used for testing purpose. Their proposed method showed 97.12 percent accuracy for MLP and 96.86 percent for KNN.

Rahman et al. [7] proposed a technique to determine the sentiment of human based on social network data by implementing neural networks. Authors also discussed about various techniques of neural networks along with deep learning.

Revised Version Manuscript Received on 30 November, 2018.
Md. Lizur Rahman, East West University, Dhaka, Bangladesh
Ifrat Jahan, East West University, Dhaka, Bangladesh
Akash Saha, East West University, Dhaka, Bangladesh
Md. Nawab Yousuf Ali, East West University, Dhaka, Bangladesh

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

In their study, they developed a system based on their proposed technique to find out the sentiment rate of any text data. They claimed their proposed technique achieved 97 percent accuracy.

Kiani & Korayem [8] proposed a new method for feature learning. For this purpose, they used Boltzmann machine. Authors used deep brief network and spiking neural network for classification. They used 60000 samples for training purpose and 20000 samples for testing purpose and achieved 95 percent accuracy.

III. DEEP NEURAL NETWORK & ARTIFICIAL NEURON

To understand about deep neural networks, we require some knowledge about artificial neurons. Perceptron and sigmoid neuron are two most common artificial neurons.

A. Perceptron

A perceptron contains a group of binary inputs with correspondence weights, an overall bias, and a single binary output [9]. Figure 1 shows the structure of perceptron. Here, x1, x2, and x3 are three binary inputs and w1, w2, and w3 are the corresponding weights for inputs. And b is the overall bias for this perceptron. Eq.1 shows the output calculation process for perceptron, where w = weight, x = input, and b = bias.

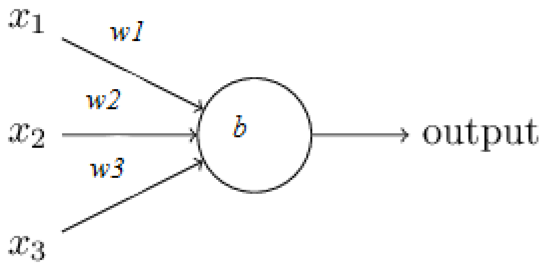

Figure 1: Structure of perceptron [9]

$$Output\ of\ perceptron = \begin{cases} 0 \ ; \ w.x + b \leq 0 \\ 1 \ ; w.x + b > 0 \end{cases} \ \dots\dots\dots (1)$$

Perceptron is not only used in simple neural networks, it is also used for making decision in complex scenarios [9]. Figure 2 shows a multi-layer perceptron (MLP) structure of neural network. This figure contains three layers of perceptrons. Here, first layer of perceptrons performs to produce simple decisions. When the outputs of each perceptrons of first layer use as inputs for second layer perceptrons, networks produce complex decisions. Even through this neural networks can be used for more and more complex decisions using second layers perceptrons outputs as third layers perceptrons inputs.

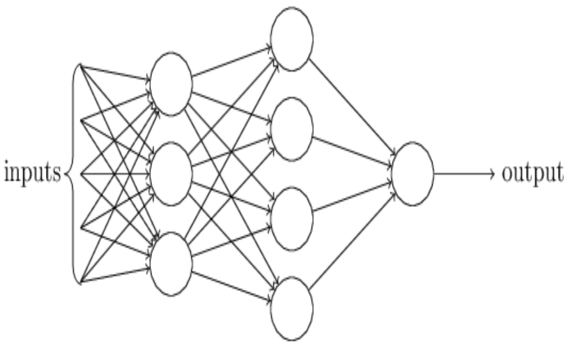

Figure 2: MLP Perceptron in Neural Network

B. Sigmoid Neuron

The change in biases and weights make a system more and better to learn. When a small change in a bias and a weight make a small change in output, it becomes easy for us to make a system as like we want. For example, a system identifies an image as a '৯' (in English 9), which is not correct. The correct answer should be a '১' (in English 1). If we able to make a small change in bias and weight, the network gets more closer to identify that image as a '১'. But when a system contains perceptrons, sometimes a small change in biases and weights cause fully change in the output of a system. Because, a perceptron produces only binary output : 0 or 1.

$$\sigma(z) \equiv \frac{1}{1 + e^{-z}} \dots\dots\dots\dots\dots\dots (2)$$

Sigmoid neuron is the solution of this problem. Sigmoid neuron is an artificial neuron, which ensures small changes in biases and weights cause only a small change in output [10]. So, a network can learn gradually through changing biases and weights. Similar to the perceptrons, a sigmoid neuron also takes several inputs, an overall bias, and produce an output. Each input of sigmoid neuron also contains a weight. But the main difference is instead of taking binary input, sigmoid neuron takes any value (e.g., 0.5611, 0.2374 etc.) between 0 and 1. Sigmoid neurons also produce output of any value between 0 and 1.

Sigmoid neuron is defined in eq. (2). Where σ is the sigmoid function and z = w· x+b. When the value of z = w· x+b is very large, then e(-z) ≈0 and so σ(z)≈1. But, when the value of z = w· x+b is very negative, then e(-z) ≈∞ and so σ(z)≈0. Hence, when the value of z = w· x+b is very large and very negative, sigmoid function behaves like perceptron. Figure 3 shows the graph of sigmoid neuron.

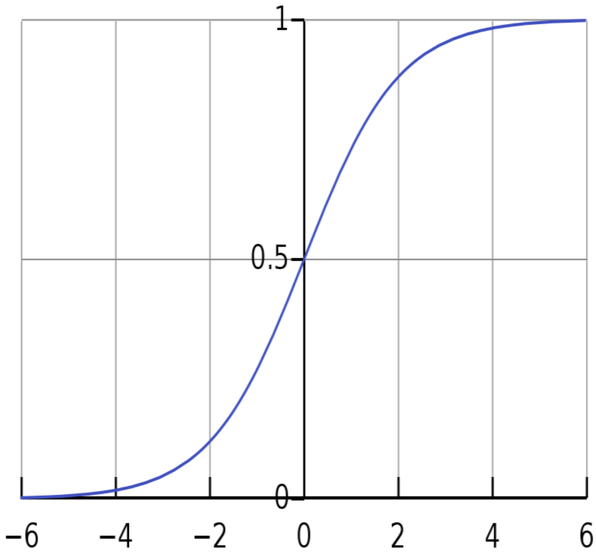

Figure 3: Graphical Representation of Sigmoid Neuron

We have developed a system implementing neural network, where we have used the partial derivatives of the sigmoid neurons to increase the smoothness of sigmoid neuron.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

C. Architecture of Neural Network

Mainly a neural network contains three types of layer including: input layer, hidden layer, and output layer [11]. Figure 4 shows the structure for multi-layer perceptron (MLP) of a neural network. The left most layer of a neural network called input layer and neurons of this layer called input neurons. The right most layer of a network known as output layer and neurons of this layer called as output neurons. Rest of the layer(s) of a neural network is known as hidden layer. A neural network contains only one input layer and one output layer [11]. But hidden layer might be one or multiple based on the complexity of the network [11]. The connection between one layer neurons to other layer neurons is fully connected [11].

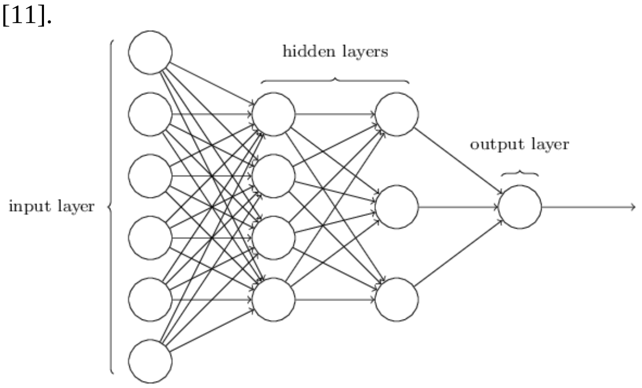

Figure 4: Structure of Neural Network

A neural network contains very simple structured input layer and output layer. Suppose we want to find out whether a handwritten image is a '৩' or not. If the size of handwritten image is 32*32 pixels, then the input layer contains 32*32 = 1024 input neurons. Output layer produces a output value in between 0 to 1. If the output value is greater or equal to 0.5, then the input image will consider as '৩'. Otherwise input image is not '৩'.

D. Dropout

Over fitting is a major problem in neural network. Over fitting is the close and similar set of data, which has been used before in the system. Overfitting occurs when a system starts to memorize the training data instead of learning [12]. We need to detect when overfitting occurs. Dropout is the technique which solves the overfitting problem [13]. Dropout not only solves the overfitting problem; it also increases the efficiency of the system. Dropout is a technique, which prevents the complex co-adoptions of training data. 'Dropout' means dropping out some units from neural network. This technique basically drops some unnecessary neurons from any layer of a network. Figure 5 shows a neural network after applying dropout.

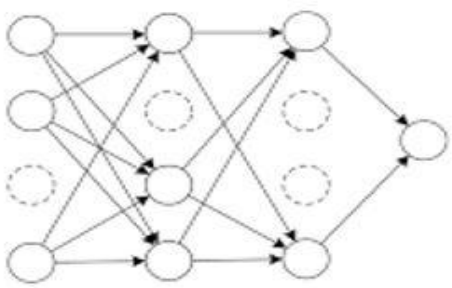

Figure 5: After Appling Dropout in Neural Network

IV. METHODOLOGY & DATASET

Various persons write a digit in different styles, shapes, and sizes. Therefore, recognition of handwritten digits is not an easy task. Moreover, Bangla digits are more complex than English digits. In this study, we have used two well-known Bangla handwritten numeric datasets. CMATERBD 3.1.1 dataset, consists of 6000 images [14]. Another one is ISI dataset which prepared by Indian Statistical Institute contains 23299 images [15]. In our experiment, we divide CMATERBD dataset images into two parts: 4500 images for training purpose and 1500 images for testing purpose. For the same purpose we also divide ISI dataset, where 20000 images for training and 3299 images for testing purpose. Figure 6 shows some handwritten digit's images.

We have used MLP based deep neural network approach to recognition of Bangla handwritten digits. Firstly, we have taken a grayscale image of handwritten digit to find out which digit it is. If the image size is more than 32*32= 1024 pixel, we will resize it to 32*32-pixel size. In our experiment, we use ReLU (Rectified Linear Unit) function to activate the neurons in both input layer and hidden layer. We also use softmax function in our output layer. We use dropout technique to prevention of overfitting problem. The input layer of our system consists 1024 neurons. The first output layer consists 512 neurons, the immediate layer consists 128 neurons, and the outer layer consists 10 neurons. Each of the outer layer's neuron represents a Bangla digit. Figure 7 represents our overall methodology.

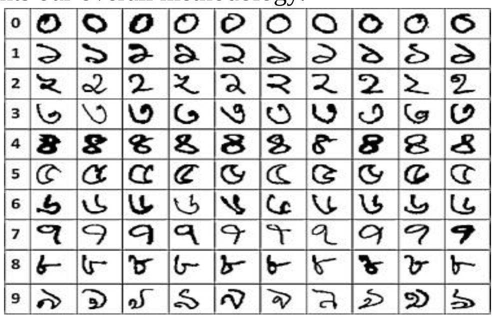

Figure 6: Bangla Handwritten Digits

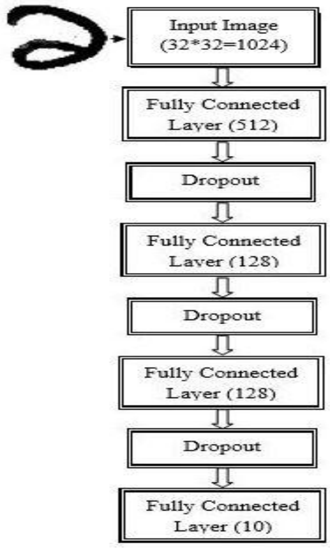

Figure 7: Description of our Method

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

V.RESULT ANALYSIS

In this section, we analyze our result to show the efficiency of our proposed method. Our proposed method achieves 98.70 percent of the accuracy, which is quite good. Since, Bangla handwritten digits' dataset is not so large, we use limited amount of sample for training and testing session. We compare our experimental result with various existing works and method outperforms well. Table 1 shows the accuracy comparison of various existing techniques.

TABLE 1. Accuracy Comparison of Various Method

Work	Accuracy
Khan et al. [15]	94%
Kiani & Korayem [8]	95%
Parvin et al. [6] (KNN)	96.86%
Parvin et al. [6] (MLP)	97.12%
Ashiquzzaman & Tushar [5]	97.40%
Das et al. [16]	97.70%
Proposed Method	98.70%

VI. CONCLUSION

This research aims to efficiently recognition of Bangla handwritten digits. In this study, firstly we have developed a system which can efficiently learn to recognize handwritten digits. We have used two Bangla numeric dataset: CMATERBD 3.1.1 and ISI. Totally, those two datasets consist of 29300 samples. Among them, 24500 samples are used for training and 4800 samples are used for testing. We have found that, our proposed system achieves 98.70 percent of accuracy. Accuracy comparison with other existing techniques shows the efficiency of our system. If we use more samples for training purpose, our system will become more accurate. Since, Bangla handwritten digits' datasets are not very large enough, we use a limited amount of samples in our system. In future we will extend our work by adding Bangla word recognition feature.

REFERENCES

1. Blakemore, C., & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. The Journal of physiology, 203(1), 237-260.
2. Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In Advances in neural information processing systems (pp. 3104-3112).
3. Belongie, S., Malik, J., & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. CALIFORNIA UNIV SAN DIEGO LA JOLLA DEPT OF COMPUTER SCIENCE AND ENGINEERING.
4. Soltanzadeh, H., & Rahmati, M. (2004). Recognition of Persian handwritten digits using image profiles of multiple orientations. Pattern Recognition Letters, 25(14), 1569-1576.
5. Ashiquzzaman, A., & Tushar, A. K. (2017). Handwritten Arabic numeral recognition using deep learning neural networks. arXiv preprint arXiv:1702.04663.
6. Parvin, H., Alizadeh, H., Moshki, M., Minaei-Bidgoli, B., & Mozayani, N. (2008, November). Divide & conquer classification and optimization by genetic algorithm. In Convergence and Hybrid Information Technology, 2008. ICCIT'08. Third International Conference on (Vol. 2, pp. 858-863). IEEE.
7. Rahman, L., Sarowar, G., & Kamal, S. (2018). Teenagers Sentiment Analysis from Social Network Data. In Social Networks Science: Design, Implementation, Security, and Challenges (pp. 3-23). Springer, Cham.
8. Kiani, K., & Korayem, E. M. (2015, November). Classification of Persian handwritten digits using spiking neural networks. In Knowledge-Based Engineering and Innovation (KBEI), 2015 2nd International Conference on (pp. 1113-1116). IEEE.
9. Gardner, M. W., & Dorling, S. R. (1998). Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmospheric environment, 32(14-15), 2627-2636.
10. Glorot, X., Bordes, A., & Bengio, Y. (2011, June). Deep sparse rectifier neural networks. In Proceedings of the fourteenth international conference on artificial intelligence and statistics (pp. 315-323).
11. Hagan, M. T., Demuth, H. B., Beale, M. H., & De Jesús, O. (1996). Neural network design (Vol. 20). Boston: Pws Pub..
12. D. M. Hawkins, "The problem of overfitting," Journal of chemical information and computer sciences, vol. 44, no. 1, pp. 1–12, 2004.
13. N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting." Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929–1958, 2014.
14. Shopon, M., Mohammed, N., & Abedin, M. A. (2017, February). Image augmentation by blocky artifact in Deep Convolutional Neural Network for handwritten digit recognition. In Imaging, Vision & Pattern Recognition (icIVPR), 2017 IEEE International Conference on (pp. 1-6). IEEE.
15. Bhattacharya, U., and Chaudhuri, B. B., Handwritten numeral databases of indian scripts and multistage recognition of mixed numerals, In IEEE transactions on pattern analysis and machine intelligence, 31(3) on pp.444-457. IEEE,2009.
16. Das, N., Reddy, J. M., Sarkar, R., Basu, S., Kundu, M., Nasipuri, M., & Basu, D. K. (2012). A statistical–topological feature combination for recognition of handwritten numerals. Applied Soft Computing, 12(8), 2486-2495.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication