

A Lenet-5 Based Bangla Handwritten Digit Recognition Framework

*Shishir Sarker**, *Songita Sarker*, *Sohanur Rahman*, *Dr Md. Ismail Jabiullah*

*Department of Computer Science and Engineering,
Daffodil International University, Dhaka, Bangladesh*

**Corresponding Author*

E-Mail Id:-shishirsarker50@gmail.com

ABSTRACT

Hand composed Digit recognition in Bangla language is a valuable beginning stage for creating an Optical Character Recognition in the Bengali language. Be that as it may, Absence of huge and honest data collection, recognition of Bangla digit was not build already. In any case, in this outline, a colossal & honest data source known as NumtaDB is utilized for recognition of Bengali digits. The troublesome endeavour is connected to getting the solid presentation and high precision for gigantic, fair, common, natural and particularly extended NumtaDB dataset. So various sorts of pre-processing frameworks are utilized for planning pictures and a significant convolutional neural network is utilized for the request of representation in this paper. The LeNet-5 architecture based convolutional neural network model has indicated superb execution. We have accomplished 97.5% testing exactness which is a decent outcome for huge and fair NumtaDB dataset contrasting with other one-sided datasets. A wide range of pre-processing of pictures is additionally significant before preparing. We utilize some pre-processing strategies for obscure and loud pictures yet these are insufficient for the elite. An examination of the system brings out the EMNIST and MNIST datasets was performed so as to sustain the appraisal.

Keywords: *Handwritten Bangla digit recognition, Convolutional Neural Network, LeNet-5, NumtaDB, Computer vision, unbiased dataset.*

INTRODUCTION

Hand written digit recognition in Bangla language is an old tricky issue in the Computer vision's field. Different sorts of viable use of this framework are, for example, mail, recognition of postal code, optical character recognition, tag acknowledgment, bank checks acknowledgment and so on. [1] The grand number of Bengali digits is finished 10. That is why the acknowledgment assignment is to bunch 10 unmistakable grades. The fundamental endeavor of translated number identification is seeing exceptional composed by hand digits. Writing styles is differing from man to man. However, our commitment is for the all the more testing undertaking. The

troublesome task is connected to getting generous execution and high exactness for gigantic, fair, normal, and significantly expanded NumtaDB dataset. [2] The data source is a blend of five data collection that were assemble from different origin & holding clouding, racket, rotate, understanding, shear, zooming, stature/width move, quality, separation, obstacles, and superimposition. We have handled blur and noisy pictures here. At that point our prepared pictures are characterized by convolutional neural network using its LeNet5 classic network architecture.

This paper is sorted out as pursues: The review works are displayed in Section II.

Our dataset and proposed approach are depicted in Section III. Segment IV shows the preprocessing procedure of images. We have demonstrated the design of the LeNet-5 architecture in Section V. Area VI speaks to our examinations and result investigation lastly end and conclusion is informed in Section VII.

REVIEW WORKS

There are a few investigations dependent on Bangla hand written digit identification utilizing deep learning. Nevertheless, an enormous segment about the assessment works used uneven dataset CMATERDB 3.1.1 [3] since the data source; NumtaDB was not accessible at that point. Beginning late several specialists have displayed a predominant exactness of 99.50% utilizing the auto encoder and noteworthy CNN structure for dataset; CMATERDB 3.1.1. Additionally the pattern of binary was utilized for recognition of Bengali digits [4]. By specific researchers, Le-net was used for identification of Bangla digit [5]. The Convolutional neural framework was exhibited for better oversight learning and precision [6]. Thusly, advancing specialists of Bangla hand digit acknowledgment are utilizing profound CNN plan. Some different classifiers like Neural Network (NN), Support Vector Machine (SVM) and so on for written by hand digit identification. The presentation of Convolutional neural network is superior to different classifiers. Therefore, this network turned into the ongoing pattern transcribed number acknowledgment. Adjusted Convolutional neural framework engineering by including extra layers or further hubs has turned into an approach to break the best in class precision. Other than Bangla language, some colossal research works in English language

physically composed digit acknowledgment; EMNIST [7] & MNIST [8] are the most outstanding dataset for English composed by hand digit and character identification. We inquired about judging EMNIST and MNIST dataset before. The assessment task of this paper is connecting with our past experimentation and that was English physically composed number affirmation & it moves us to come up with Bangla manually written digit acknowledgment also.

DATASET & PROPOSED APPROACH

Dataset

In our suggested procedure, we offer significance to the dataset named NumtaDB that contains 85,000+ Bangla manually written numeral pictures. Since NumtaDB is an honest, fair, enormous, natural, and checked on the dataset. The dataset; NumtaDB can confirm suggested methodology execution impeccably and we trust that our suggested methodology will get around a similar precision for genuine manually written digit identification. The images of the dataset; NumtaDB are authentic pictures with no pre-maintenance. The dataset; NumtaDB is an accumulated dataset from five grand origin. As indicated by this data source, the origins are named from 'a' to 'e'. The preparation and trial sets have different subgroups depending upon the wellspring of the data (setting up a, test-a, etc.). The completely datasets have been distributed into planning and preliminary defines with the objective that handwriting from a comparable subject/sponsor is missing in both. Each picture of this data source; NumtaDB is concerning to 180×180 pixels.

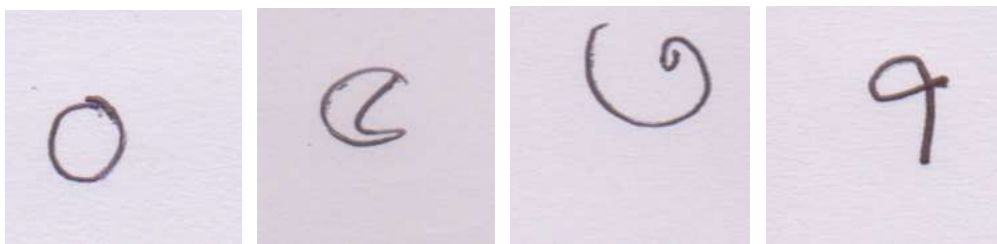


Fig.1:-The images of the NumtaDB dataset

In the opinion of NumtaDB, a pair of expanded datasets (expanded from trial pictures of dataset 'A' and 'C') are joined

to the try out set which comprises of the accompanying enlargements" [9].

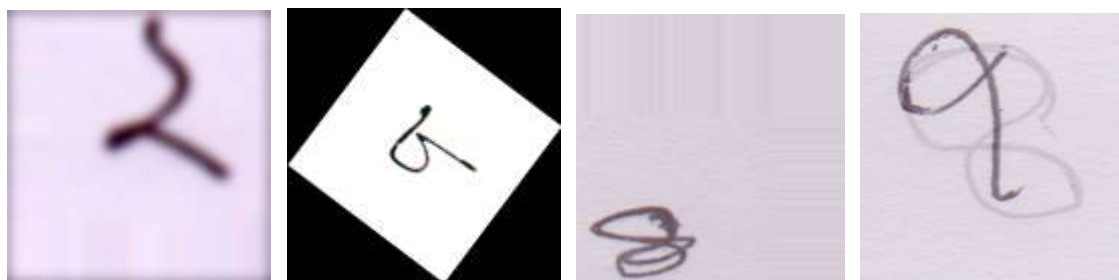


Fig.2:-The augmented pictures of the NumtaDB dataset

- Spatial modification: Spinning, Relocation, Trim, High point/Span Shift, Channel reposition, Sprint.
- Brilliance, Opposition, Logged, Hue reposition, Crash.
- Obstructions.
- Overlying

In this way, dataset; NumtaDB has made hand written digit identification for Bangla, all the more testing by increased pictures.

EMNIST [7], the comprehensive version of dataset; MNIST where we utilized EMNIST digits and EMNIST balanced data for numeral and letter identification. EMNIST digits dataset includes 280,000 models with 10 classes and 30,000 models for respectively class.

MNIST dataset [8] is among the most mainstream adjusted dataset in English language composed by hand and has a readiness lot of 60,000 models and a trial set of 10,000 models. For respectively

class there are 6,000 getting ready models and 1000 trial models.

Proposed Approach

We suggest clustering NumtaDB dataset composed by hand, following two critical advances. These methods are:

- Pre phase of processing of pictures.
- LeNet-5 architecture based CNN model.

The subtleties of our duo significant advances are portrayed in part IV and V.

PREPROCESSING PROCEDURE OF IMAGES

Resizing and Gray scaling

The initial size of NumtaDB dataset pictures contains 180×180 pixels which are unreasonably enormous for pre phase of processing productively. Therefore we shrink the size of pictures to 32×32 pixels. Additionally transform all RGB (Red, Green and Blue) pictures to GRAY scale pictures. The concealing channel is changed over to 1 channel from 3 channels.

Interjection

Images can lose a lot of significant data due to resizing. Between territory addition is favored strategy for picture devastation. This strategy is resampling utilizing pixel zone connection. We use between zone addition in the wake of resizing pictures.

Expelling Blur from Images

We utilize Gaussian blur to include obscure from the start and afterward subtract the obscured picture from the first picture. At that point we include a weighted bit of the cover to get de-obscured picture [10].

$$g_{mask}(x, y) = f(x, y) - f'(x, y) \quad (1)$$

$$g(x, y) = f(x, y) + k * g_{mask}(x, y) \quad (2)$$

Here the blurred image is represented by $f'(x, y)$ and weight is represented by k for generality.

Sharpening Images

There are numerous channels for honing pictures. We utilize the Laplacian channel & our filter is 3×3 .

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Taking out Clutter from Images

We took out pepper and salt clamor from datasets pictures. And utilize middle channel for taking out salt and pepper clamor. In the wake of pre phase of maintenance, the pictures become understandable, keen and salt and pepper free.

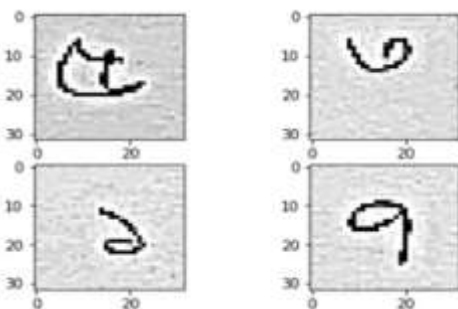


Fig.3:-Preprocessed images

LeNet-5 architecture based convolutional neural network

CNN has been giving exceptional execution in the stage of hand written digit recognition throughout the previous couple of years. CNNs are the establishment of current best in class profound learning-based PC vision. Different types of architecture we use LeNet-5 architecture of CNN for our work. LeNet-5 was utilized on huge scale to naturally characterize written by hand digits. LeNet-5 is a significant design is that before it was imagined, character acknowledgment had been done for the most part by utilizing highlight building by hand, trailed by an AI model to figure out how to arrange hand-built highlights. LeNet-5 made hand designing highlights repetitive on the grounds that the system takes in the best interior portrayal from crude pictures consequently.

Architecture of the model

Our suggested architecture contains of 3 convolutional layers and 3 completely associated layer. The initial layer has 32 channels and every channel size is 5×5 . The second layer has 64 and size of each channel is 3×3 . And the third layer has 128 channels with 3×3 channel size. ReLu (Rectified Linear unit) [11] is utilized as an enactment work for all surfaces. Max-pooling layers & Batch normalization are utilized after each layer. 2×2 is the pool scale of Max-pooling surface. Batch normalization is utilized for accelerate research [12]. Among the three completely associated layers, the first has 120 channels and the middle one has 84 channels along with the final one has 10 channels for the 10 digits/numbers. The last actuation function is a soft-max work for arrangement. Here we utilize Adam [13] streamlining agent for refreshing loads. Figure 4 represents the structure of suggested LeNet-5 construction based CNN engineering. From contribution to yield each design is stamped appropriately.

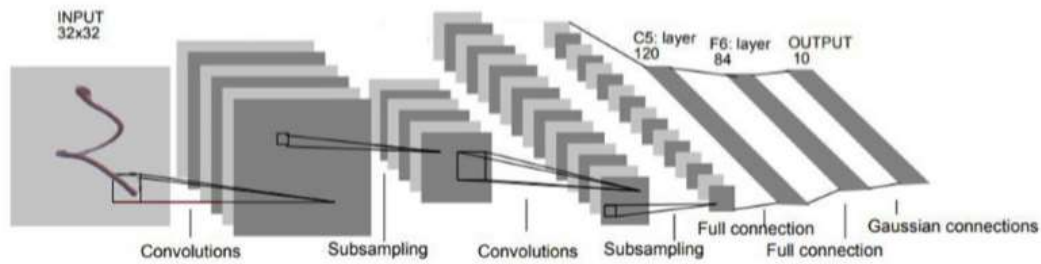


Fig.4:-Proposed LeNet-5 CNN architecture design

Table 1 represents the summary of our suggested LeNet-5 CNN architecture.

Table 1:-Summary of our proposed lenet-5 model

Layer (type)	Output Shape	Param #
conv2d_4	(None, 28, 28, 32)	832
max_pooling2d_4	(None, 14, 14, 32)	0
batch_normalization_4	(None, 14, 14, 32)	128
conv2d_5	None, 12, 12, 64	18496
max_pooling2d_5	None, 6, 6, 64	0
batch_normalization_5	None, 6, 6, 64	256
conv2d_6	None, 4, 4, 128	73856
max_pooling2d_6	None, 2, 2, 128	0
batch_normalization_6	None, 2, 2, 128	512
flatten_2	None, 512	0
dense_4	None, 120	61560
dense_5	None, 84	10164
dense_6	None, 10	850

The parameters of training model

The course of action model of our

arrangement is needy upon explicit parameters.

Table 2:-The parameters of training model

Parameter's Name	Value
Learning rate	10 ⁻³
Batch Size	64
Epoch	30
Shuffle	True

EXAMINATIONS AND RESULT INVESTIGATION

Environment of Experiment

Our environment of experiment is setup with Intel core-i5 processor, AMD Radeon R5 M255 GPU and 4GB of RAM.

Train, Validation, and Test

Surrounded by 85000+ pictures, train chop portion of dataset; NumtaDB is 85% and the test chop portion is 15% [14]. On this investigation, we chop the preparation information into preparing and approval remaining the chop proportion about 80%-

20%. Here we utilize the approval information to assess the representation execution. Finally, our conclusive outcome is estimated by test data. The test data of NumtaDB are from five one of kind origins (a to e) and furthermore duo expanded datasets were conveyed from test set A and C. These are for our Evolution.

Outcome with Prediction

Some images of output with their prediction are given below;

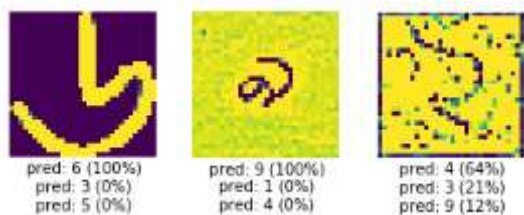


Fig.5:-Outcomes with their prediction

Result Analysis

After 30 epochs we get 97.5% Test accuracy.

Run Time Analysis

The handling time of preparing is additionally estimated. The all-out preparing time is 55.96 min.

CONCLUSION

In this report, we exhibited LeNet-5 CNN based recognition framework of Bangla digit for a standard and testing dataset. We accomplished 97.5% test perfection that is a decent outcome for huge & fair NumtaDB dataset contrasting with extra one-sided datasets. A wide range of preprocessing of pictures is additionally significant before preparing. We utilize some preprocessing procedure for obscure and loud pictures yet these are insufficient for elite. As a later work, we think the propelled information increase system and further developed CNN model can beat the issues of recognizing expanded pictures.

REFERENCES

1. B. B. Chaudhuri & U. Pal. *A complete printed Bangla OCR system*. Pattern recognition.1988.31.5:531-549p.
2. S. Alam, T. Reasat, R.M. Doha and A.I Humayun. *NumtaDB Assenbled Bengali handwritten digits*. arXiv:1806.02452 [cs.CV], 6 June, 2018.
3. M. Shopon, N. Mohammed, and M. A. Abedin, "Bangla handwritten digit recognition using autoencoder and deep convolutional neural network," 2016 International Workshop on Computational Intelligence (IWCI), Dhaka, 2016, 64-68p.
4. T. Hassan and A.H. Khan. *Handwritten Bangla numeral recognition using Local Binary Pattern*. In Electrical Engineering and Information Communication Technology (ICEEICT), 2015 International Conference on.1-4p. IEEE, 2015.
5. U. Bhattacharya and B. B Chaudhuri. *Handwritten numeral databases of Indian scripts and multistage recognition of mixed numerals*. In IEEE transactions on pattern analysis and machine intelligence, 31(3) on .444-457p. IEEE, 2009.
6. M.I.H.R.I. M.A.H. Akhand, and M. Ahmed. *Convolutional neural network training with artificial pattern for Bangla handwritten numeral recognition*. ICIEV.2016.1(1):16p.
7. S. Afshar, G. Cohen and J. Tapson, A. Schaik. *EMNIST: An extension of MNIST to handwritten letters*. arXiv:1702.05373[cs.CV], 2017.
8. Y. Lecun and C. Cortes. *The MNIST database of handwritten digits*. 1998. [Online]. Available: <http://yann.lecun.com/exdb/mnist/> G. Cohen, S. Afshar, and J. Tapson, A. Schaik, "EMNIST: An extension of MNIST to handwritten letters", arXiv:1702.05373[cs.CV], 2017.
9. Bengali Ai. *NumtaDB: Bengali Handwritten Digits*. 2018. [Online]. Available: <https://www.kaggle.com/BengaliAI/numta>. [Accessed: 608-2018].
10. R. Gonzalez and R.E. Woods. *Digital Image Processing*. Third Edition, 162-163p.
11. V.Nair and Hinton G. E. *Rectified linear units improve restricted boltzmann machines*. Proceedings of the 27th International Conference on Machine Learning (ICML-10). 2010
12. S. Ioffe and C. Szegedy. *Batch Normalization: Accelerating deep network training by reducing internal*

- covariance shift.*
arXiv:1502.03167[cs.LG], 2 March, 2015.
13. D. Kingma and J. Ba, *Adam: A method of stochastic optimization.* arXiv:1412.6980[cs.LG], 2015.
14. S. Alam, T. Reasat, R.M. Doha and A.I Humayun. *NumtaDB Assembled Bengali handwritten digits.* arXiv:1806.02452 [cs.CV]. 6 June, 2018.