

**A LENET-5 BASED BANGLA HANDWRITTEN DIGIT RECOGNITION
FRAMEWORK**

BY

SHISHIR SARKER

ID: 161-15-7417

SONGITA SARKER SETU

ID: 161-15-7019

SOHANUR RAHMAN

ID: 161-15-7205

This Report Presented in Partial Fulfillment of the Requirements for the Degree
of Bachelor of Science in Computer Science and Engineering

Supervised By

Dr. Md. Ismail Jabiullah

Professor

Department of CSE

Daffodil International University

Co-Supervised By

Mr. Majidur Rahman

Lecturer

Department of CSE

Daffodil International University



DAFFODIL INTERNATIONAL UNIVERSITY

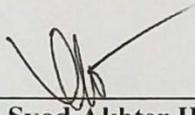
DHAKA, BANGLADESH

DECEMBER 2019

APPROVAL

This Project/internship titled “A LENET-5 BASED BANGLA HANDWRITTEN DIGIT RECOGNITION FRAMEWORK” submitted by Shishir Sarker, ID No: 161-15-7417, Songita Sarker Setu, ID No: 161-15-7019 and Sohanur Rahman, ID No: 161-15-7205 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 05-12-19.

BOARD OF EXAMINERS

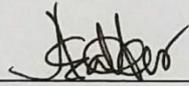


Dr. Syed Akhter Hossain

Professor and Head

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Chairman



Abdus Sattar

Assistant Professor

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner

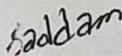


Farah Sharmin

Senior Lecturer

Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Internal Examiner



Dr. Md. Saddam Hossain

Assistant Professor

Department of Computer Science and Engineering
United International University

External Examiner

DECLARATION

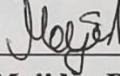
We hereby declare that, this project has been done by us under the supervision of **Dr. Md. Ismail Jabiullah, Professor, Department of CSE, Daffodil International University**. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:



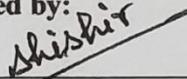
Dr. Md. Ismail Jabiullah
Professor
Department of CSE
Daffodil International University

Co-Supervised by:

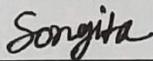


Mr. Majidur Rahman
Lecturer
Department of CSE
Daffodil International University

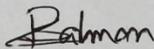
Submitted by:



(Shishir Sarker)
ID: 161-15-7417
Department of CSE
Daffodil International University



(Songita Sarker Setu)
ID: 161-15-7019
Department of CSE
Daffodil International University



(Sohanur Rahman)
ID: 161-15-7205
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/internship successfully.

We really grateful and wish our profound our indebtedness to **Dr. Md. Ismail Jabiullah, Professor, Department of CSE**, Daffodil International University, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of “*Machine Learning*” to carry out this project. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Syed Akhter Hossain, Professor and Head, Department of CSE** for his kind help to finish our project and also to other faculty member and the staff of CSE department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

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 endeavor is connected to getting the solid presentation and high precision for gigantic, fair, common, natural and particularly extended NumtaDB dataset. So various sorts of preprocessing 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 preprocessing of pictures is additionally significant before preparing. We utilize some preprocessing 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.

TABLE OF CONTENTS

CONTENTS	PAGES
Board of examiners	i
Declaration of the Student	ii
Acknowledgement	iii
Abstract	iv

CHAPTER

CHAPTER 1: INTRODUCTION **1-2**

1.1 Introduction	1
1.2 Motivation	1-2
1.3 Objectives	2
1.4 Report Layout	2

CHAPTER 2: BACKGROUND **3-4**

2.1 Review work	3
2.2 Bangla community perspective	3-4

CHAPTER 3: RESEARCH METHODOLOGIES **5-11**

3.1 Working procedure	5
3.1.1 Dataset	5-7
3.1.2 Suggested approach	7
3.1.2.1 Pre-processing of pictures	7
3.1.2.1.1 Resizing and Gray scaling	7
3.1.2.1.2 Interjection	8
3.1.2.1.3 Expelling Blur from Images	8
3.1.2.1.4 Sharpening Images	8
3.1.2.1.5 Taking out clutter from Images	8-9
3.1.2.2 LeNet-5 architecture based convolutional neural network	9
3.1.2.2.1 Architecture of the suggested model	9-10
3.1.2.2.2 The Parameters of training model	11

CHAPTER 4: RESULT AND OBSERVATIONS	12
4.1 Environment of experiment	12
4.2 Train, Validation, and Test	12
4.3 Outcome with prediction	12
4.4 Result analysis	12
4.5 Run time analysis	12
CHAPTER 5: CONCLUSION	13
REFERENCE	14-15

LIST OF FIGURE

FIGURE NAME	PAGES
Figure-3.1: The images of NumtaDB dataset	5
Figure-3.2: The augmented pictures of NumtaDB dataset	6
Figure-3.3: Preprocessed images	9
Figure-3.4: Proposed LeNet-5 CNN architecture design	10
Figure-4.1: Outcomes with their prediction	12

LIST OF TABLE

TABLE NO.	PAGES
Table-3.1: Shows the Dataset summary	7
Table-3.2: Summery of our suggested LeNet-5 model	10
Table-3.3: The parameters of training model	11

CHAPTER 1

INTRODUCTION

1.1 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 managed as looks for after: The review tasks are appeared in part II. Our data and suggested approach are depicted in portion III. Segment IV represents the method of pre phase of processing of pictures. We demonstrated the arrangement of the LeNet-5 plan in Zone V. Zone VI tends to our appraisals and result assessment at last end and later works are taught in Section VII.

1.2 Motivation

There are a couple of researches reliant on Bangla hand written digit recognition using profound learning. Some gigantic research works on physically composed digit recognition of English language; EMNIST [7] and MNIST [8] are the well-known data record for recognition of English composed by hand digit and character.

Beforehand we got some information about considering EMNIST and MNIST dataset and it inspires us to come up with in Bangla hand written digit acknowledgment.

1.3 Research Objective

Speeding up the process of Bangla digit recognition in document processing is the primary objective of the research. This research is an intention to contribute in Bangla OCR. Viable use of this framework is many. For example, mail, OCR, postal code acknowledgment, tag acknowledgment, bank checks acknowledgment and so on. This project will bring the result of recognizing any sort of hand-written Bangla digit no matter how it is written.

1.4 Report Layout

This segment follows the parts of each progression that we utilized in our report in short.

1. Chapter 1: Introduction
2. Chapter 2: Background
3. Chapter 3: Research Methodologies
4. Chapter 4: Result & Observation
5. Chapter 5: Conclusion

CHAPTER 2

BACKGROUND

2.1 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.

2.2 Bangla community perspective

As Bangla is our mother tongue, every day in every steps we have to gone through Bangla. Still now the official works (govt. and non govt.) are done in Bengali language. There are much research are found for digit recognition in Bangla but not

much available. So, Bangla hand written digit recognition can be a big deal in computer vision in Bangla community perspective.

CHAPTER 3

RESEARCH METHODOLOGIES

3.1 Working procedure

As like all other recognition methodology, digit recognition is only a recognition procedure. A few stages are available for acknowledgment explicitly digit recognition framework. Here the technique is described below.

3.1.1 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.

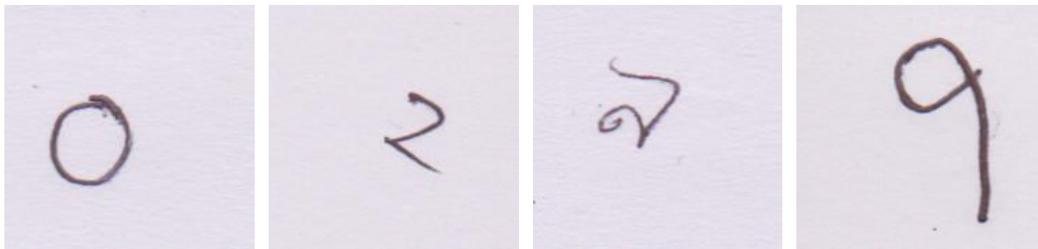


Figure-3.1: The images of the NumtaDB dataset

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].

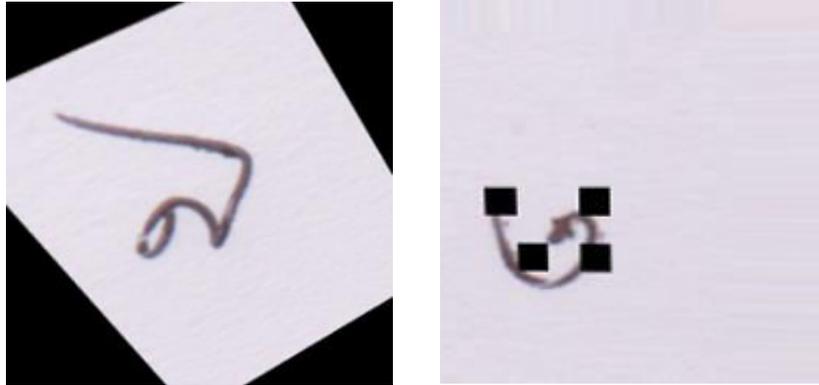


Figure-3.2: The augmented pictures of the NumtaDB dataset

- 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.

Dataset 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.

Dataset attributes are given below [10];

Table-3.1: Shows the Dataset summary

Code Name	Train-Test Split	Total Digit (Training)	Total Digits (Testing)	Total Digits (combined)
a	85%-15%	19702	3489	23191
b	85%-15%	359	69	428
c	85%-15%	24298	4381	28679
d	85%-15%	10908	1948	12856
e	85%-15%	16777	2970	19747

3.1.2 Suggested 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.

3.1.2.1 Preprocessing of pictures

As our dataset is unbiased and it contains huge amount of images of digits different in size, we have to process these before our work. Here these steps are followed;

3.1.2.1.1 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.

3.1.2.1.2 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.

3.1.2.1.3 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 [11].

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

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

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

3.1.2.1.4 Sharpening Images

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

$$[-1 \ -1 \ -1]$$

$$[-1 \ 9 \ -1]$$

$$[-1 \ -1 \ -1]$$

3.1.2.1.5 Taking out Clutter from Images

We took out pepper and salt clamor from dataset's 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.

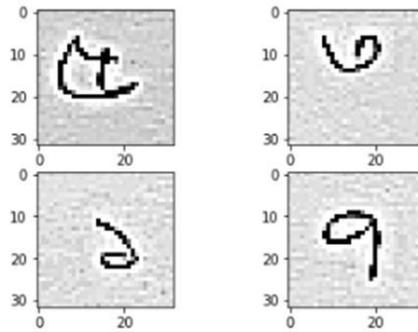


Figure-3.3: Preprocessed images

3.1.2.2 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 architecture of CNN for our work. LeNet-5 was utilized on huge scale to naturally characterize written by hand digits. LeNet 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 made hand designing highlights repetitive on the grounds that the system takes in the best interior portrayal from crude pictures consequently.

3.1.2.2.1 Architecture of the suggested 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) [12] 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 [13]. 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 fuction is a softmax work for arrangement. Here we utilize Adam optimizer

[14] streamlining agent for refreshing loads. Fig. represents the structure of suggested LeNet-5 construction based CNN engineering. From contribution to yield each design is stamped appropriately.

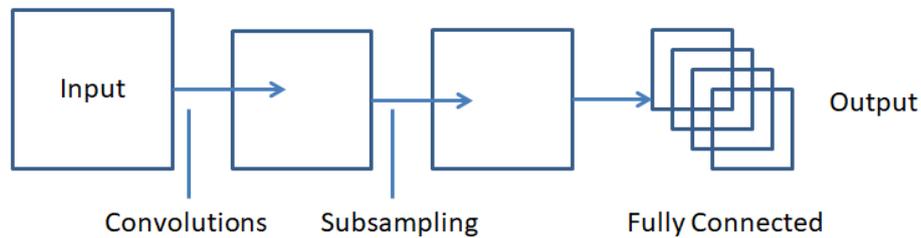


Figure-3.4: Proposed LeNet-5 CNN architecture design

Table 2 represents the summary of our suggested LeNet-5 CNN architecture.

Table-3.2: Summary of our proposed LeNet-5 model

Layer (type)	Output Shape	Param #
Conv2d_4	(None, 28, 28, 32)	832
Maxpooling2d_4	(None, 14, 14, 32)	0
Batch_normalization_4	(None, 14, 14, 32)	128
Conv2d_5	None, 12, 12, 64	18496
Maxpooling2d_5	None, 6, 6, 64	0
Batch_normalization_5	None, 6, 6, 64	256
Conv2d_6	None, 4, 4, 128	73856
Maxpooling2d_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

3.1.2.2.2 The parameters of training model

The course of action model of our arrangement is needy upon explicit parameters.

Table-3.3: The parameters of training model

Parameter	Point
Rate of learning	10^{-3}
Size of Batch	64
Epoch	30
Shuffle	True

CHAPTER 4

RESULT AND OBSERVATIONS

4.1 Environment of experiment

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

4.2 Train, Validation, and Test

Surrounded by 85000+ pictures, train chop portion of dataset; NumtaDB is 85% and the test chop portion is 15% [15]. 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.

4.3 Outcome with prediction

Some images of output with their prediction are given below;

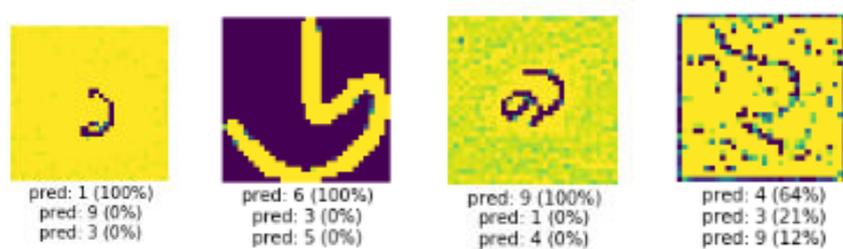


Figure-4.1: Outcomes with their prediction

4.4 Result analysis

After 30 epochs we get outcome, Test accuracy 97.5%.

4.5 Run time analysis

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

CHAPTER 5

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, vol. 31.5, pp. 531-549, 1998.
- [2] S. Alam, T. Reasat, R.M. Doha, and A.I Humayun, "NumtaDBAssembled 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, pp. 64-68.
- [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, pp. 1-4. 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 pp.444-457. 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, vol. 1, no. 1, pp. 16, 2016
- [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 handwrittendigits," 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] "Evaluation of Bengali handwritten digit recognition challenge", 2018.[Online]
- [11] R. Gonzalez, and R.E. Woods, "Digital Image Processing", Third Edition, pp. 162-163.
- [12] 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
- [13] S. Ioffe, and C. Szegedy, "Batch Normalization: Accelerating deep network training by reducing internal covariance shift", arXiv:1502.03167[cs.LG], 2 March, 2015.
- [14] D. Kingma, and J. Ba, "Adam: A method of stochastic optimization", arXiv:1412.6980[cs.LG], 2015.

[15] S. Alam, T. Reasat, R.M. Doha, and A.I Humayun, “NumtaDBAssembled Bengali handwritten digits”, arXiv:1806.02452 [cs.CV], 6 June, 2018.