

# Own Handwritten Digit recognition using MLP and CNN in tensorflow

Deepti D.Nikumbh<sup>1</sup>, Rupali S. Kale<sup>2</sup>

<sup>1,2</sup>Mumbai University, Shah and Anchor Kutchhi Engineering College, Chembur

**Abstract-** Object recognition in image is very popular and is widely used in almost all image processing applications. Handwritten digit recognition system is one such application. This paper presents an approach on developing a handwritten digit recognition system using multi-layer neural network and Convolutional Neural Network. These neural network models are trained and tested using the MNIST dataset. Further a real time dataset of authors own handwritten digit where used to test the performance of the system , a comparison of two deep learning models in terms of accuracy i.e successfully classifying digits between 0-9 and computational time taken is presented. The neural network models are developed in python using tensorflow a machine learning library.

**Index terms-** Digit recognition, MLP, CNN

## I. INTRODUCTION

The human visual system is one of the wonders of the world. Human can effortlessly identify and recognize the objects around them. This is because, our visual system is made up of network of billions of neurons which perform complex image processing[1]. Machine learning and deep learning, a branch of AI aims at mimicking these functioning of a neuron and developing systems that can identify things as close to human brain.

Handwritten digit recognition is one of the classical problems in deep learning. It is a challenging task since handwriting differs from person to person. The digits written by different people may not be of same size, thickness, or orientation and pressure applied while writing. Although, this difference in handwriting does not cause any problems to humans, however it is more difficult to teach computers to recognize general handwriting [2].

Handwritten digit recognition finds a lot of application for example online handwriting recognition on tablets and smart phones, recognize

zip codes on mail for postal mail sorting, processing bank check amounts, numeric entries in forms filled up by hand (for example - tax forms).Data entry for business documents. Imagine chemists/drug-stores being able to scan doctors' handwriting without any struggle, Automatic number plate recognition which will help traffic-rule enforcement, Automatic insurance documents being able to extract key information without human intervention, Extracting business card information into a contact list, Book scanning, Make electronic images of printed documents search-able, for example Google Books. A lot of machine learning tools have been developed like scikit-learn, scipy- image etc. and pybrains, Keras, Theano, Tensorflow by Google, TFLearn etc. for Deep Learning. These tools help us to develop various machine learning models , train and test them and deploy these model in state of art applications[3].

### 1.1 Related Works

In this author has developed an algorithm based on deep learning neural networks using activation function and regularization layer which is having more accuracy than existing Arabic neural recognition methods. New developed algorithm gives 97.4% accuracy.[1]

This paper compares three neural network approaches and evaluated with respect to accuracy and performance. These approaches are also compared for their execution time and found that DNN is the most accurate algorithm.[2]

In this paper authors have focused on measuring style compatibility of textured or non-textured 3D furniture models. Style compatibility is measured using triplet CNN and dataset is created containing 420 textured 3D furniture models.[3]

The main focus of the paper is to combine CNN and SVM for recognizing hand-written digits and pattern augmentation for expanding data set. Author showed

that using pre trained convolutional neural network as a feature extractor for handwritten digits is very useful and improves error rate.[4]

Author compared results of widely used machine learning algorithms like SVM, KNN and RFC also compared deep learning algorithms like multilayer CNN using keras and Tensorflow. After all comparison author proves that deep learning techniques gives high amount of accuracy.[5]

Author has studied about the time complexities of KNN and CNN.CNN has been implemented on Keras including Tensorflow and produces accuracy. It has observed that CNN produces high accuracy than KNN.[6]

Author has focused mainly on neural network approaches. Three approaches compared and evaluated for their accuracy and efficiency. Author concluded with more efficient approach as DNN.[7]

Author in this paper compared various machine learning algorithms such as SVM,KNN and RVM with deep learning algorithms like multilayer CNN and found accuracy for it.[8]

## 2 METHODOLOGY

### 2.1 Proposed

The proposed flowchart is as follows:

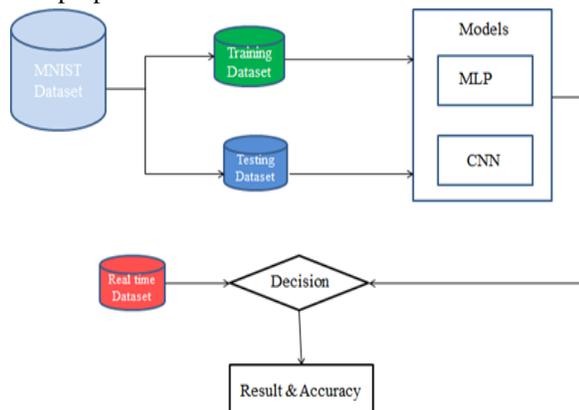


Fig.1. Flowchart for handwritten recognition

MNIST dataset is a classical dataset in deep learning. It contains images of handwritten single digits from 0 to 9.The images are 28\*28 pixels. The following code will automatically download the MNIST dataset and it can be used.

```

from tensorflow.examples.tutorials.mnist import input_data
mnist = input_data.read_data_sets("MNIST_data/", one_hot=True)
    
```

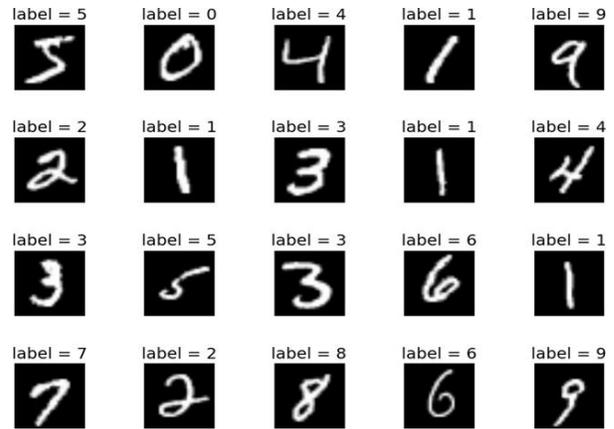


Fig. 2. MNIST PLOT[10]

The MNIST dataset is split into three parts: 55,000 training images (mnist.train), 10,000 testing images (mnist.test), and 5,000 validation image set (mnist.validation)[4].A single digit can be represented by an 2d array of size 28\*28 ,where '0' represents white pixel,'1' represents black pixel and values in between 0 and 1 are grayscale values. This 2d array is flattened in 1d vector of 784 values i.e (784,1) or (1,784).Thus that mnist.train.images is a tensor (an n- dimensional array) with a shape of [55000, 784].

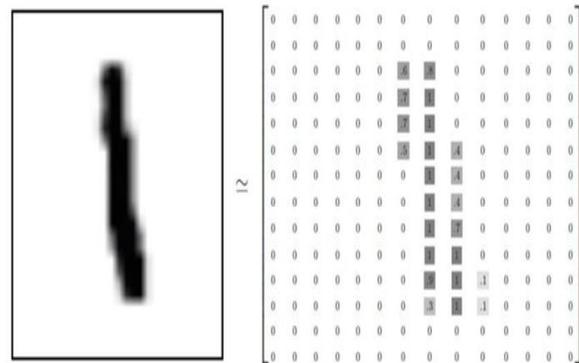


Fig.3. Handwritten image and its 2D array representation [9]

For the labels MNIST uses one hot encoding. Instead of using labels like 'one', 'two', 'three' it uses a single array for each image. The label is represented based upon the index position in label array. The corresponding label will be '1' at index position and '0' elsewhere. For example if the image is of the digit 5 its label array will be [000010000]. Thus, mnist.train.labels is a [55000, 10] array of floats.[4]. The MNIST data is hosted on Yann LeCun's website.[16]

### 2.2 Real Time Dataset

A real time dataset was created where author took handwritten digits samples from different people, kids (in the age group of 7-9) handwritten digit samples were also considered. A small real time dataset of 50 images was created. These raw images are subjected to preprocessing and scaled down to the size of MNIST da- taset images. Labels were created by using one hot encoding technique.

### 2.3 Techniques used for Classification

**Multi-layer Neural Network(MLP):**A MLP classifier is developed consisting of one passive input layer, one active output layer and one active hidden layer. The number of neurons in input layer are 784 , since images in MNIST are of 28\*28 pixel .The number of neurons in output layer at 10 .The number of neuron in hid- den layer at 300 which is decided by trial and error method. Each neuron uses sigmoidal activation function. The MLP is trained using 55,000 training images (mnist.train) and tested using 10,000 testing images (mnist.test). Adam optimizer is used to update network weights iterative based in training data.

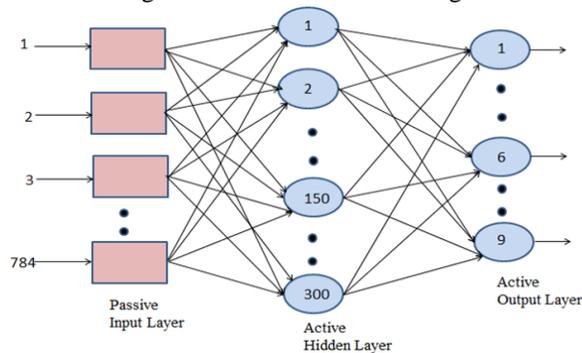


Fig.4. MLP architecture for handwritten digit recognition

### 2.3 Convolutional Neural Network (CNN)

CNN consisting of two convolutional layer and two pooling (subsampling) layer followed by a densely connected neural network was developed. For convolution layer, filter size of [5,5] is considered along with a stride of 1 in all dimension. For pooling layer,max pooling is done with a window size of[2,2] and with stride of 2.Pooling operation is important as it reduces the memory usage and load on computer. Using the above mentioned window size and stride will reduce the input image by 75%. During training and testing drop out is used where units are randomly dropped along with their connections to prevent overfitting.

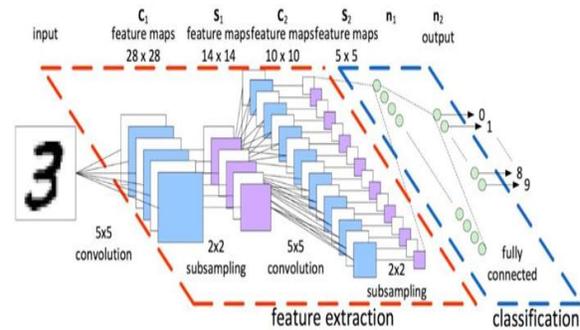


Fig.5. CNN architecture for handwritten digit recognition [17]

## 3 EXPERIMENTAL RESULTS

The MNIST dataset was used for training and testing MLP and CNN further the models where tested with state of art real time dataset. Fig below shows samples of real time dataset.

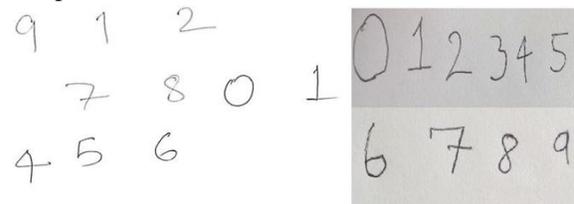


Fig. 6. Samples of some handwritten digits

### 3.1 Experimental Environment used

- Windows 8 operating system as test platform,
- CPU is Intel Core I5, which has dual cores running on 2.4GHZ. RAM 16GB.

### 3.2 Evaluation Factors Considered

The two algorithms are evaluated based on the following Factors:

- Accuracy
- Learning rate
- Execution Time

### 3.3 Experimental Results obtained

In this section we will discuss the results obtained after using both the architectures for standard MNIST dataset and real-time dataset.

For both the architectures the number of epochs kept was fixed i.e 20,000. Af- ter 20,000 epochs with learning rate of 0.001, MLP gave training accuracy as 0.96,testing accuracy as 0.95 and real time dataset accuracy as 0.72.The perfor- mance plot of training, testing and real time dataset is shown below.

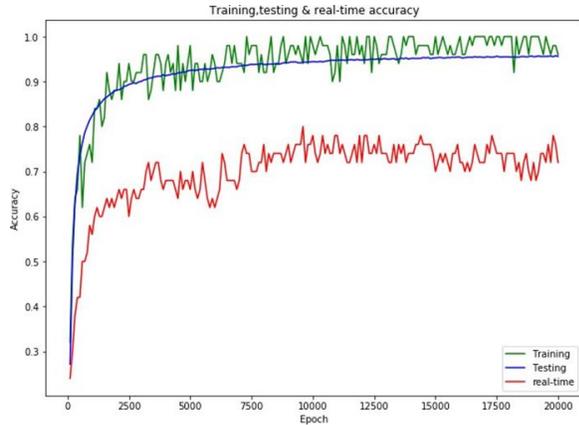


Fig. 7. Performance of training and testing dataset  
 From the plot we can see that model is underfit till approximately 11000 ephocs since testing accuracy is sometimes larger than training accuracy after that testing accuracy is less than training accuracy and both the accuracies don't change much Since the training and testing plot has started forming plateau, it will require some more iterations to converge. Real time dataset accuracy is in the range of 0.6 to 0.7 much lesser than training and testing. This is because images in real time dataset are captured from different cameras thus quality of images are different further some digits are written will less pressure ,some digits are not centered.when such images are subjected to pre-processing steps certain information is lost.

Learning rate plays a very important role in deciding the convergence i.e to stop when maximum accuracy is reached. Figure below shows the training, testing and real-time plots with learning rate of 0.01.Comparing the fig and fig we can see that the noise in the graph has increased.

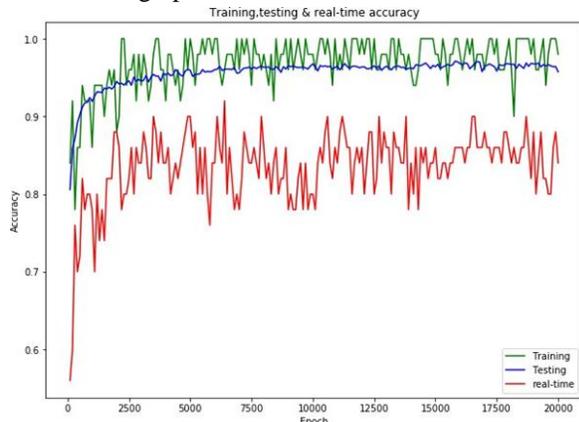


Fig. 8. performance plot with learning rate of 0.01  
 After 5,000 ephocs CNN gave training accuracy as 0.98, testing accuracy as 0.98 and real time accuracy

as 0.84.which is better than performance of MLP. The performance plot of training, testing and real time dataset is shown below.

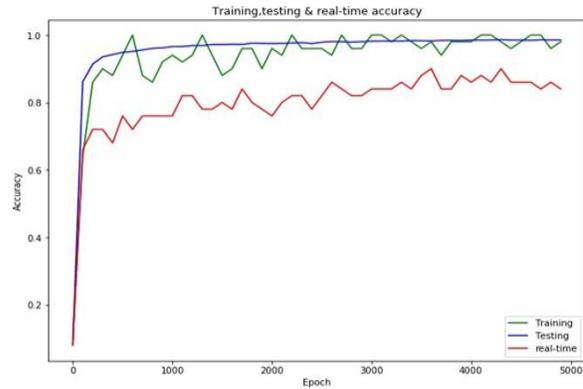


Fig. 9. Performance plot of CNN  
 From the plot we can see that model is initially underfit till approximately 4000 ephocs since testing accuracy is sometimes larger than training accuracy after that testing accuracy has started becoming less than training accuracy. Initially there are huge fluctuations or noise in graph, as the epochs increases noise is reducing the graph is trying to form a line (plateau) Since the testing plot has started forming plateau, training graph will require some more iterations to converge. Real time dataset accuracy is in the range of 0.8 to 0.9 lesser than training and testing and will take few more iteration to converge. After 5000 iterations, the performance in terms of accuracy and execution time is shown in table below.

Table 1. Accuracy and Execution performance

Architecture	Training Accuracy	Testing Accuracy	Real time Accuracy	Execution Time(sec)
MLP	0.98	0.92	0.62	16.58
CNN	0.98	0.98	0.84	3 hrs 3 min

#### 4 CONCLUSION

In this paper we have created own handwritten digit recognizer using two approaches CNN and MLP. MNIST dataset was used for training and testing and above that a real time dataset was created and tested. Both the architectures are evaluated based upon its accuracy and execution time and results are shown in table 1.For MLP it is observed that it is simple to implement and takes less time to train and test. Whereas its performance on real time dataset is not satisfactory. For CNN, it is observed that its architecture is complex hence difficult to implement and takes more time to train and test (preferably GPU

enabled systems) and its performance on real time dataset is high.

#### REFERENCES

- [1] “Handwritten Arabic Numeral Recognition using Deep Learning Neural Networks”
- [2] “A Comparative Study on Handwriting Digit Recognition Using Neural Networks”, 2017 International Conference on Promising Electronic Technologies, Mahmoud M. Abu Ghosh Faculty of Information Technology Islamic University of Gaza Palestine , Ashraf Y. Maghari Faculty of Information Technology Islamic University of Gaza Palestine
- [3] “Deep Model Style: Cross-class Style Compatibility for 3D Furniture within a Scene”, 2017 IEEE International Conference on Big Data (BIGDATA), Tse-Yu Pan, Yi-Zhu Dai, Wan-Lun Tsai, and Min-Chun Hu Dept. of Computer Science and Information Engineering National Cheng Kung University Tainan, Taiwan.
- [4] “Pattern Augmentation for Handwritten Digit Classification based on Combination of Pre-trained CNN and SVM”, Yoshihiro Shima School of Science and Engineering Meisei University Hino-city, Tokyo, Japan Yumi Nakashima, Michio Yasuda School of Information Science Meisei University Hino-city, Tokyo, Japan
- [5] “Handwritten Digit Recognition Using Deep Learning”, Anuj Dutt, Aashi Dutt, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 6, Issue 7, July 2017, ISSN: 2278 – 1323.
- [6] “Analogizing Time Complexity of KNN and CNN in Recognizing Handwritten Digits”, Tanya Makkar Department of Computer Science, Amity School of Engg. and Technology, Amity University Uttar Pradesh, Noida, India, Yogesh Kumar Department of ECE, Amity School of Engg. and Technology, Amity University Uttar Pradesh, Noida, Ashwani Kr Dubey Department of ECE, Amity School of Engg. and Technology, Amity University Uttar Pradesh, Noida, Álvaro Rocha Department of Informatics Engineering, University of Coimbra Portugal, Ayush Goyal Electronics & Communication
- [7] A Comparative Study on Handwriting Digit Recognition Using Neural Networks
- [8] Handwritten Digit Recognition Using Deep Learning
- [9] [https://www.tensorflow.org/versions/r1.0/get\\_started/mnist/beginners](https://www.tensorflow.org/versions/r1.0/get_started/mnist/beginners)
- [10] <https://corochann.com/mnist-dataset-introduction-1138.html>
- [11] Author, F.: Article title. Journal 2(5), 99–110 (2016).
- [12] Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016).
- [13] Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, Location (1999).
- [14] Author, F.: Contribution title. In: 9th International Proceedings on Proceedings, pp. 1–2. Publisher, Location (2010).
- [15] LNCS Homepage, <http://www.springer.com/lncs>, last accessed 2016/11/21.
- [16] <http://yann.lecun.com/exdb/mnist/>
- [17] <https://www.udemy.com/join/login-popup/?next=/complete-guide-to-tensorflow-for-deep-learning-with-python/learn/#overview>