**Report**

**On**

**Digital Handwritten Recognition**

**A Project Report submitted in partial fulfilment of**

**the requirements for the award of**

**Bachelor in Engineering**

**IN**

**COMPUTER SCIENCE AND ENGINEERING**

**Submitted by**

**Monarch Hasija**

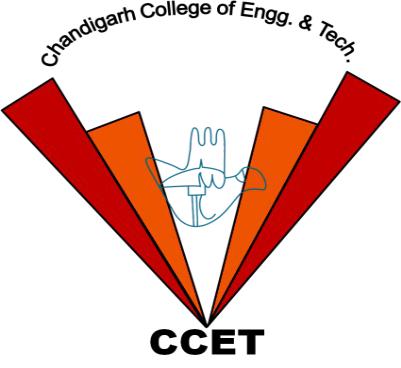
**(Roll no:CO17338)**

**Navneet Lohan**

**(Roll no:CO17340)**

**Ashmeet Kaur**

**(Roll no: LCO17365)**



**CHANDIGARH COLLEGE OF ENGINEERING AND TECHNOLOGY**

**(DEGREE WING)**

Government Institute under Chandigarh (UT) Administration, Affiliated to Panjab University, Chandigarh

Sector-26, Chandigarh. PIN-160019

**Jan, 2020**

**Handwritten Digit Recognition**

**INTRODUCTION:**

There is an ever-increasing amount of image data in the world, and the rate of growth itself is increasing. Infotrends estimates that in 2016, cameras and mobile devices captured more than 1.1 trillion images. According to the same estimate, in 2020 the figure will increase to 1.4 trillion. Many of these images are stored in cloud services or published on the Internet. In 2014, over 1.8 billion images were uploaded daily to the most popular platforms, such as Instagram and Facebook.

Going beyond consumer devices, there are cameras all over the world that capture images for automation purposes. Cars monitor the road, and traffic cameras monitor the same cars. Robots need to understand a visual scene in order to smartly build devices and sort waste. Imaging devices are used by engineers, doctors and space explorers alike.

With time the numbers of fields are increasing in which deep learning can be applied. In deep learning, Convolutional Neural Networking (CNN) is being used for visual imagery analyzing. Object detection, face recognition, robotics, video analysis, segmentation, pattern recognition, natural language processing, spam detection, topic categorization, regression analysis, speech recognition, image classification are some of the examples that can be done using Convolutional Neural Networking.

The accuracies in these fields including handwritten digits recognition using Deep Convolutional Neural Networks (CNNs) have reached human level perfection.

The goal of this article is to observe the influence of hidden layers of a CNN for handwritten digits. We have applied a different type of Convolutional Neural Network algorithm on Modified National Institute of Standards and Technology (MNIST) dataset

**Problem statement**

Digits contained in image files can be located and identified automatically. This is called object detection and is one of the basic problems of computer vision. As we will demonstrate, convolutional neural networks are currently the state-of-the-art solution for object detection. The main task of this thesis is to review and test convolutional object detection methods.

**Technology and Concepts:**

**Machine Learning**

Learning algorithms are widely used in computer vision applications. Before considering image related tasks, we are going to have a brief look at basics of machine learning.

Machine learning has emerged as a useful tool for modelling problems that are otherwise difficult to formulate exactly. Classical computer programs are explicitly programmed by hand to perform a task. With machine learning, some portion of the human contribution is replaced by a learning algorithm. As availability of computational capacity and data has increased, machine learning has become more and more practical over the years, to the point of being almost ubiquitous.

It can be used in two ways:

* Supervised Learning
* Unsupervised Learning

**Neural networks**

Neural networks are a popular type of machine learning model. A special case of a neural network called the convolutional neural network (CNN) is the primary focus of this thesis. Before discussing CNNs, we will discuss how regular neural networks work.

Neural networks were originally called artificial neural networks, because they were developed to mimic the neural function of the human brain.

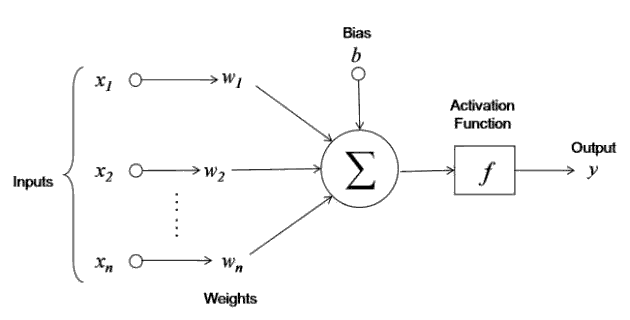


Fig: Neural Network Working

The neuron is trained by carefully selecting the weights to produce a desired output for each input.

**Multi-layer networks**

A neural network is a combination of artificial neurons. The neurons are typically grouped into layers.

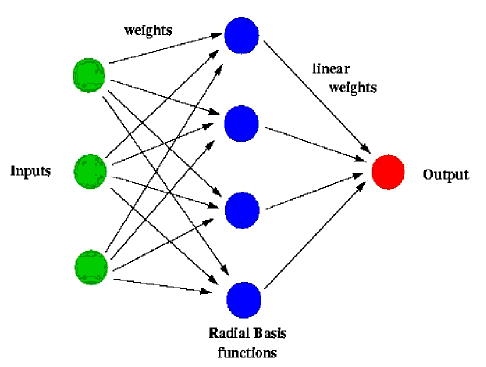
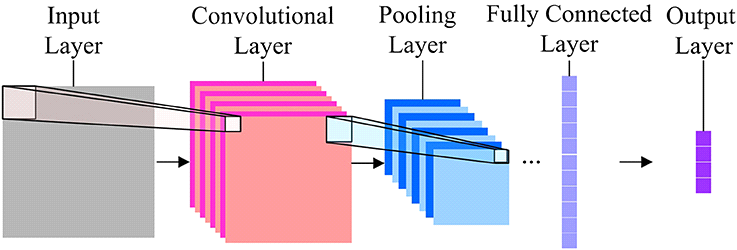


Fig: Multi-Layer Networks

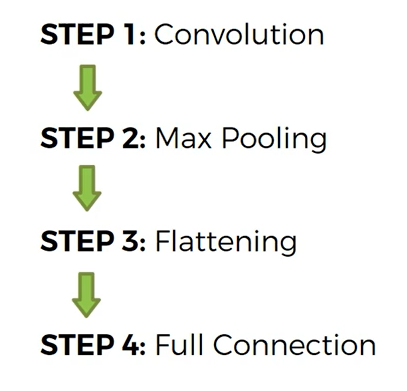
A multi-layer network typically includes three types of layers: an input layer, one or more hidden layers and an output layer. The input layer usually merely passes data along without modifying it. Most of the computation happens in the hidden layers. The output layer converts the hidden layer activations to an output, such as a classification. A multilayer feed-forward network with at least one hidden layer can function as a universal approximator.

**Convolutional neural networks**

 Fig: CNN

The basic idea of the CNN was inspired by a concept in biology called the receptive field. Receptive fields are a feature of the animal visual cortex. They act as detectors that are sensitive to certain types of stimulus, for example, edges. They are found across the visual field and overlap each other.

**Model Structure:**



**Convolution**

Convolution in CNN is performed on an input image using a filter or a kernel. To understand filtering and convolution you will have to scan the screen starting from top left to right and moving down a bit after covering the width of the screen and repeating the same process until you are done scanning the whole screen.

# Pooling

After a convolution layer once you get the feature maps, it is common to add a pooling or a sub-sampling layer in CNN layers. Similar to the Convolutional Layer, the Pooling layer is responsible for reducing the spatial size of the Convolved Feature. This is to decrease the computational power required to process the data through dimensionality reduction. Furthermore, it is useful for extracting dominant features which are rotational and positional invariant, thus maintaining the process of effectively training of the model. Pooling shortens the training time and controls over-fitting.

There are two types of Pooling:

* Max Pooling
* Average Pooling

Max Pooling returns the maximum value from the portion of the image covered by the Kernel.  
Max Pooling also performs as a Noise Suppressant. It discards the noisy activation altogether and also performs de-noising along with dimensionality reduction.

Average Pooling returns the average of all the values from the portion of the image covered by the Kernel. Average Pooling simply performs dimensionality reduction as a noise suppressing mechanism. Hence, we can say that Max Pooling performs a lot better than Average Pooling. So as we have used max pooling in out project.

**Flattening**

Flattening is converting the data into a 1-dimensional array for inputting it to the next layer. We flatten the output of the convolutional layers to create a single long feature vector. And it is connected to the final classification model, which is called a fully-connected layer. In other words, we put all the pixel data in one line and make connections with the final layer.

# Fully Connected Layer (FC Layer)

Adding a Fully-Connected layer is a (usually) cheap way of learning non-linear combinations of the high-level features as represented by the output of the convolutional layer. The Fully-Connected layer is learning a possibly non-linear function in that space.

**Description of the dataset:**

Modified National Institute of Standards and Technology (MNIST) is a large set of computer vision dataset which is extensively used for training and testing different systems. It was created from the two special datasets of National Institute of Standards and Technology (NIST) which holds binary images of handwritten digits. The training set contains handwritten digits from 250 people, among them 50% training dataset was employees from the Census Bureau and the rest of it was from high school students [26]. However, it is often attributed as the first datasets among other datasets to prove the effectiveness of the neural networks.

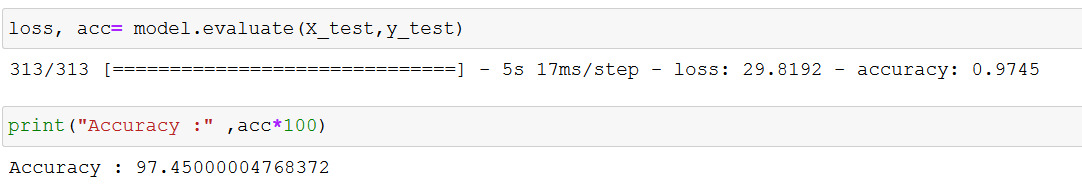
The MNIST database consisting of offline handwritten digits ranging from 0-9. The database was constructed from Special Database 3 (SD-3) and Special Database 1 (SD-1) that contain binary images of handwritten digits. SD-3 was collected among Census Bureau employees, while SD-1 was collected among high-school students. For the results to be independent of both datasets, MNIST dataset was built by mixing NIST SD-1 and SD-3. The total number of digit image samples (70,000), the total number for training (60,000) and testing (10,000), and the subtotal number for each digit are shown in table 1. Each digit is a gray-level fixed-size image with a size of 28 x 28 (or 784 pixels) in total as the features.

|  |  |  |  |
| --- | --- | --- | --- |
| **Table 1.** Dataset Digits | # Training | # Testing | Subtotal |
| 9 | 5949 | 1009 | 6958 |
| 8 | 5851 | 974 | 6825 |
| 7 | 6265 | 1028 | 7293 |
| 6 | 5918 | 958 | 6876 |
| 5 | 5421 | 892 | 6313 |
| 4 | 5842 | 982 | 6824 |
| 3 | 6131 | 1010 | 7141 |
| 2 | 5958 | 1032 | 6990 |
| 1 | 6742 | 1135 | 7877 |
| 0 | 5923 | 980 | 6903 |
| Total | 60,000 | 10,000 | 70,000 |

final classification.

**Performance:**

Applying the CNN model, the accuracy received was around 97%.

****

**Outputs:**

After generating and training the model, we further have added the feature of generating the number (by the user) and getting a suitable result (by model). The graphic user interface is created so as a user can create its own handwritten number ranging from 0-9 and could find an accurate response with suitable accuracy.

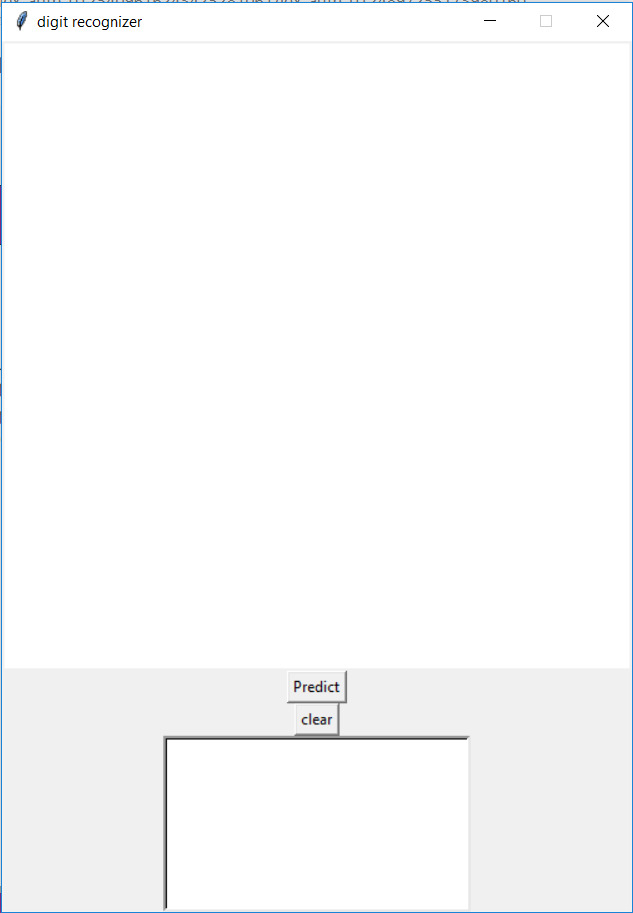
****

Fig: And empty window

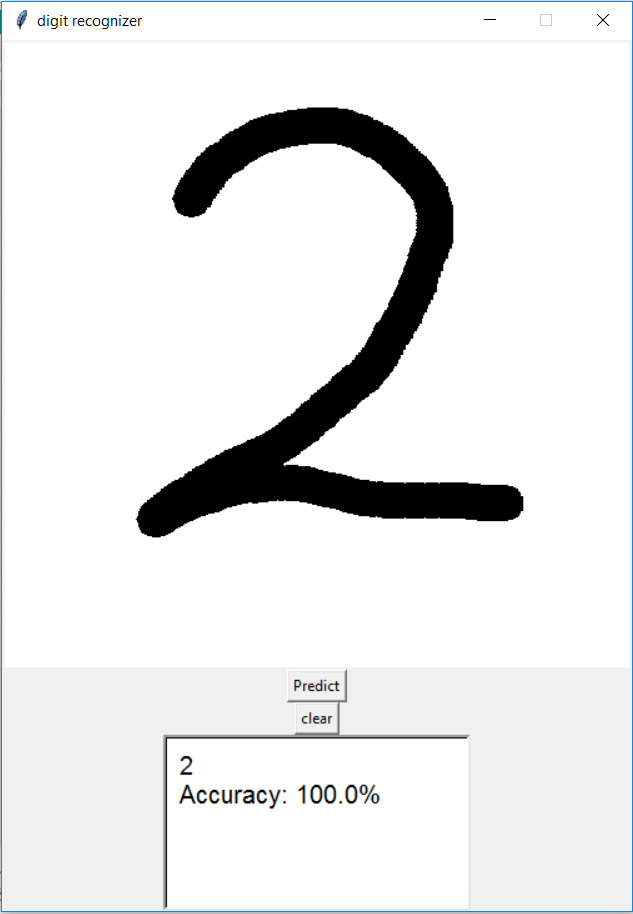
****

Fig: A number Created by user

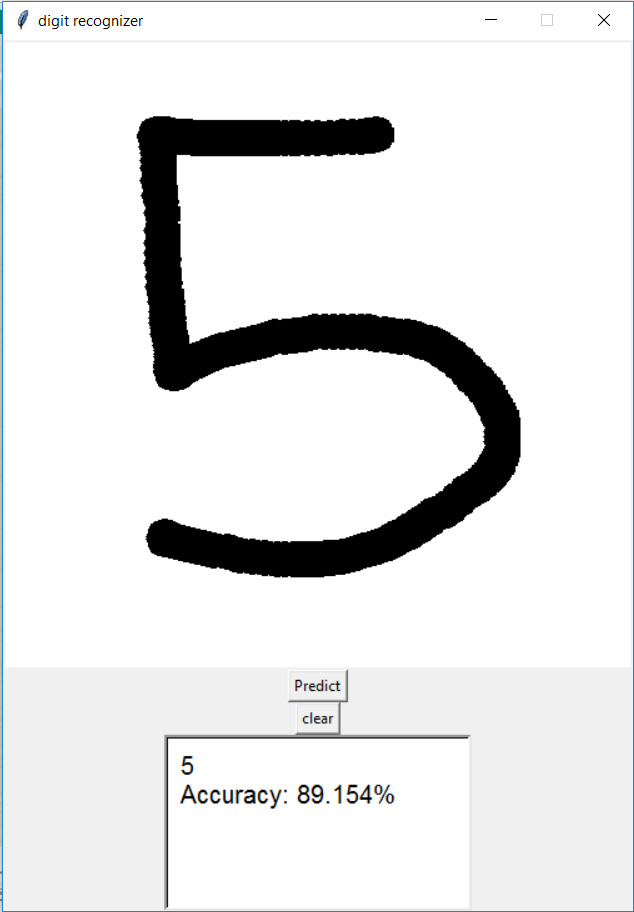
****

Fig: A number Created by user

**Conclusions:**

In this report we have discovered the only way to achieve highest accuracy of predicting handwritten numeral. Thus, this study settled on training a model to recognize the digit and providing a reliable gui where user can test the model. This study built handwritten recognizers evaluated their performances on MNIST (Mixed National Institute of Standards and Technology) dataset and then improved the training speed and the recognition performance. This study discusses in detail all advances in the area of handwritten character recognition. The most accurate solution provided in this area directly or indirectly depends upon the quality as well as the nature of the material to be read.

**Bibliography**

* <https://www.medium.com>
* [https://www.analyticsvidhya.com](https://www.analyticsvidhya.com/)
* <http://yann.lecun.com/exdb/mnist/>
* [https://towardsdatascience.com](https://towardsdatascience.com/wtf-is-image-classification-8e78a8235acb)
* <https://www.geeksforgeeks.org/image-classifier-using-cnn/>
* Handwritten Digit Recognition using Machine Learning Algorithms

By: S M Shamim, Mohammad Badrul Alam Miah, Angona Sarker, Masud Rana & Abdullah Al Jobair

* Recognition of Handwritten Digit using Convolutional Neural Network in Python with Tensorflow and Comparison of Performance for Various Hidden Layers

By: Fathma Siddique , Shadman Sakib , Md. Abu Bakr Siddique

* Digits - A Dataset for Handwritten Digit Recognition

By: Alexander K. Seewald