

# Digit Classification using Convolutional Neural Network

Shourya Bhatnagar<sup>1</sup>, Kartikay Sharma<sup>1</sup>, Deeksha Sharma<sup>1</sup>, Danish Vij<sup>1</sup>  
<sup>1</sup> Student,  
Department of Computer Science & Engineering,  
Dr. Akhilesh Das Gupta Institute of Information Technology and Management,  
Shastri Park, Delhi, India

Mrs. Neha Sharma<sup>2</sup>  
<sup>2</sup> Assistant Professor,  
Department of Computer Science & Engineering,  
Dr. Akhilesh Das Gupta Institute of Information Technology and Management,  
Shastri Park, Delhi, India

**Abstract:-** Neural networks are a group of algorithms that are identical to the human brain and are programmed to identify different patterns. A neural network learns from numerous levels of representation and reacts appropriately to different levels of abstraction where different patterns are learned by each layer. Nowadays CNNs have been commonly used in pattern recognition, sentence identification, speech recognition, face recognition, categorization of text, etc. In this work we use CNN to identify handwritten digits using different numbers of hidden layers and epochs to achieve highly accurate results. This research was carried out using the database of the Modified National Standards and Technology Institute (MNIST)

**Keywords:-** CNN, MNIST, Image Processing, Digit Recognition

## 1. INTRODUCTION

Traditional methods of machine learning (such as multilayer perception machines, vector support machines, etc.) mostly use shallow structures to manage a small number of samples and computing units. The quality and generalization ability of complex classification problems are obviously inadequate when the target items have rich definitions. The convolution neural network (CNN) developed in recent years has been widely used in the field of image processing because it is good to deal with problems of object classification and recognition and has brought great improvement in the accuracy of many tasks of machine learning. It has become a dominant and standard model for deep learning.[1]

Through taking images as input and then classifying them into certain classes, CNN is primarily used in object recognition. Here we use it for handwritten digit identification. We're going to have a set of handwritten digits from 0-9 with their labels. Using 60,000 inputs, the model is

## 2. RELATED WORK

Different studies have investigated various CNN models over the period.

ARDIS: A handwritten Swedish digital dataset. This paper introduces a new handwritten image-based historical digital dataset called Arkiv Digital Sweden (ARDIS). The photographs in the ARDIS dataset were drawn from 15,000 documents of the Swedish church, written in the nineteenth and twentieth centuries by various priests with different styles of handwriting. The database consists of three single-digit datasets and a single-digit sequence dataset. The digit string dataset contains 10,000 samples of red-green-blue color space, with 7,600 single-digit images in different color spaces in the other datasets. A systematic study of machine learning techniques was carried out on a number of digital datasets. In addition, the link between ARDIS and current electronic databases is being studied by the Modified National Institute of Standards and Technology (MNIST) and the US Postal Service (USPS). Experimental results show that machine learning algorithms, including deep learning approaches, have low accuracy of recognition when educated on existing datasets and evaluated on datasets. Accordingly, convolutionary neural networks trained on MNIST and USPS and evaluated on ARDIS provide the highest accuracies, respectively 58:80 percent and 35:44 percent. This dataset is publicly available to the research community to further advance handwriting virtual recognition algorithms.[3]

Across different fields, CNN has achieved expert-level performance, medical research is no exception. Gulshan et al, Esteva et al, and Ehteshami Bejnordi et al. demonstrated the capacity of deep learning for diabetic retinopathy screening, identification of skin lesions, and detection of lymph node metastasis. There has been an increase in interest among radiology researchers in the potential of CNN, and several studies have already been published in areas such as lesion detection, identification, segmentation, image recovery, and natural language processing. [4]

### 3. METHODOLOGY

The approach used here is the simulation of CNN. CNN object classification model takes, processes and classifies an input image, in our case digits, under a certain category.

#### 3.1 Dataset

MNIST Dataset: It is a 60,000 28x28-pixel grayscale dataset with handwritten single-digit images ranging from 0 to 9. The task is to allocate a given image of a handwritten digit to one of 10 groups representing values between 0 and 9.

It is a database that is commonly used. Top-performing models are deep learning CNNs that achieve a classification accuracy of over 99%, with an error rate of between 0.4% and 0.2% on the hold-out test dataset.

#### 3.2 Image Processing

In this step, by dividing each region into different contours, we first extract each drawn digit from the surface.

Then we're working on every contour. We first change the image from rgb 3 channel image to grayscale 1 channel image before operating on the picture. A threshold of the given frequency is then applied to the image in order to remove all the different types of noise. After removing the noise from the selected contour, we begin to extract the digit form from the image. Through making a mask, this is achieved.

All values greater or less than the range of input values that the mask function provides are converted to black and if a pixel value is between the given range its value is modified to white. Once the mask is created, the original image is added, and the image is sent to the CNN template with the mask applied to it.

#### 3.3 CNN Modelling

CNN object classification model takes, processes and classifies an input image, in our case digits, under a certain category. Deep learning allows for the training and testing of CNN models. -input image will pass through a series of filter (Kernels), pooling, fully connected layers (FC) convolution layers and use Softmax to classify an object with probabilistic values from 0 to 1.

### 4. CLASSIFICATION

#### 4.1 Proposed architecture

Here we used a convolutionary neural network (CNN) that was implemented using the Keras library. The template consists of 2-layer sets (2 convolution layers and 1 maxpool layer) with a dropout of 0.25 after each set to prevent overfitting. The result was flattened, and later two thick layers were added.

#### 4.2 Convolutional layer

A CNN's main building block is the coevolutionary surface. The parameters of the layer consist of a set of learnable filters with a small receptive field but extending through the full h of the volume of the input. The filter is translated over the width and height of the input volume during the forward transfer, measuring the dot product

between the filter entries and the source, and generating a 2-dimensional activation map of that filter. As a result, the network learns filters that activate when some specific type of feature is detected at some spatial location in the input.

#### 4.3 Pooling layer

Pooling is a non-linear method of down-sampling. The purpose is to gradually reduce the representation's spatial size

#### 4.4 Dense Layer/Fully connected Layer

Within one layer, dense or completely connected layers connect each neuron to each neuron in another layer. It is the same as the conventional neural perceptron multi-layer (MLP) network. To identify the objects, the flattened matrix passes through a fully connected surface.

#### 4.5 Dropout

Dropout is a regularization technique for reducing overfitting in neural networks by preventing complex coadaptation on training data. It is a very efficient way of performing model averaging with neural networks. The term "dropout" refers to dropping out units (both hidden and visible) in a neural network.

#### 4.6 RMSProp

It is used as an optimizer, RMSProp (Root Mean Squared Propagation). Optimization is the search process for parameters that minimize our functions or maximize them. Geoffrey Hinton developed RMSProp. RMSProp aims to overcome the dramatically declining learning levels of Adagrad by using a moving squared gradient average. To normalize the slope, it uses the strength of the recent gradient descents. The learning rate is automatically adjusted in RMSProp and a different learning rate is selected for each parameter. RMSProp divides the speed of learning by the average square gradient exponential decay.

#### 4.7 Rectified linear unit

The rectified linear unit (ReLU) was used as an activation function. A standard computer chip circuit can be seen as a digital network of activation functions that can be "ON" (1) or "OFF" (0), depending on input.

#### 4.8 Data generator

The data generation method for both training and testing data set is used to achieve greater accuracy. A data generator function is useful to ensure that random data is chosen, and the data differs. The category "ImageDataGenerator" produces tensor image data batches with an increase in real-time data.

#### 4.9 Softmax

The effect is a 10-potential range from 0 to 9. The softmax function has been used to get the efficiency. This takes a vector of K real numbers as input and normalizes this into a distribution of probability consisting of K probabilities proportional to the input numbers exponentials. That is, certain vector components may be negative or greater than one before applying softmax and may not add up to 1. But each component will be in the interval (0,1) after applying softmax and the components will add up to 1, so they can be interpreted as probabilities.

## 5. RESULTS

In result we tend to aim to match the potency of the three classifiers. The results are shown in the table below

Table 1. Comparison of different machine learning models

Algorithm Name	Trained Data Accuracy (%)	Testing Data Accuracy (%)
SVM	97.9	97.6
KNN	97.1	96.8
CNN	99.4	98.33

## 6. FUTURE SCOPE AND CONCLUSION

The efficiency of the three classifiers i/e SVM, KNN and CNN. Our best model proved to be CNN with the accuracy the highest training and testing accuracy of 99.4% and 98.33% We have shown a version of CNN that can identify handwritten digits. Further convolution and secret layers can make the results further reliable. Handwritten digit

recognition is an excellent experimental problem for learning about neural networks and it provides an excellent way to improve more sophisticated deep learning techniques. It can be expanded in the future to character recognition, handwriting in real time, reading computerized bank check numbers, signature authentication, interpretation of postal addresses, etc.

### REFERENCES

- [1] <https://link.springer.com/article/10.1186/s13640-019-0417-8#CR1>
- [2] <https://link.springer.com/article/10.1007/s13244-018-0639-9>
- [3] <https://link.springer.com/article/10.1007/s00521-019-04163-3>
- [4] <https://link.springer.com/article/10.1007/s13244-018-0639-9>