Hand Written Digit Recognition Using Deep Learning

Gaganjot Kaur,
Manav Rachna University
Faridabad Haryana,India

Deepak Singh
Manav Rachna University
Faridabad Haryana, India

Rohit Pandey,
Manav Rachna University
Faridabad Haryana, India

Tapur Tanwar
Manav Rachna University
Faridabad Haryana

Vishwesh Pandey
Manav Rachna Univesity
Faridabad Haryana, India

Abstract-

Machine Learning and deep learning plays very crucial role in AI and computer technology. IT arises due to enhancement in technology in Machine Learning, Deep Learning and Computer Vision calculations. With collaborative use of deep learning techniques with human efforts recognition became much easier. This paper presents handwritten digit recognition from famous MNIST dataset to test whether data is perused accurately or not. Classifiers like KNN,SVM,RFC are compared in this paper on basis of accuracy, time etc. The confusion matrix drawn is used to summarize different algorithm in terms of accuracy**.**

Keywords-KNN, RFC, SVM, CONFUSION MATRIX

I.INTRODUCTION

Manually written digit recognition is the value of a PC framework that perceives the transcribed data sources like numerical, documentation and from various sources like messages, papers, image, texts, reports etc. This has been a subject of research from quite a while. Various research area include territories postal compose translation from envelopes, divider encumbers preparing, signature verification and so on.A great deal of characterization strategies using deep Learning and machine learning is created and utilized for this approach like K-NN, SVM, RFC and so forth. However these techniques in spite of thefact that having the precision of 96.9% are sufficient for this present reality applications.As an illustrationif a letter is written and send to address as name "Aram" and the structure diagnose it wrongly with the name "Tanya" so letter is not gone to "Aram" but it is delivered to "Tanya".Later on there is quite possibility that it may reach to correct address but if the mail could be very urgent, a person has to pay loss due to delay so that means, the accuracy in these types of applications is very important. Also the shortcoming of this approach is that these techniques cannotbe completely relied on as they do not give the exact precision due to lack of complete information regarding task.

II. LITERATURE REVIEW

Machine and deep learningare the latest instrument for processing of an image, object identification, manually written numeric and character acknowledgement. Many AI instruments are created like scikit learn, scipy-picture and so on from this. These apparatuses make it strong and, in this way progressively precise the neutral networks can practically emulate the human cerebrum and is a key fixing in picture preparing are like CNN along back propagation for image processing etc. Quite a research is done on the method of hand recognition using deep learning method, artificial intelligence. Researchers have analysed and also some have established various techniques to find more accurate method of supervised learning to

recognise the different articles be it in form of number, text or an image.

III.MINIST DATASET

The subset of MINST has 70,000 database of handwritten digit recognition. Further it is divided into around or more than50,000 training samples and some testing sample around 10,000.The dataset size is 28*28 values representing images with label.

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000801(2049)	magic number (MSB first)
0004	32 bit integer	60000	number of items
0008	unsigned byte	??	label
0009	unsigned byte	??	label
.....			
xxxx	unsigned byte	??	label

The labels values are 0 to 9.

Fig -1: Training set label

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000803(2051)	magic number
0004	32 bit integer	60000	number of images
0008	32 bit integer	28	number of rows
0012	32 bit integer	28	number of columns
0016	unsigned byte	??	pixel
0017	unsigned byte	??	pixel
.....			
xxxx	unsigned byte	??	pixel

Fig-2: Training set image

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000801(2049)	magic number (MSB first)
0004	32 bit integer	10000	number of items
0008	unsigned byte	??	label
0009	unsigned byte	??	label
.....			
xxxx	unsigned byte	??	label

Fig-3: Test set label

[offset]	[type]	[value]	[description]
0000	32 bit integer	0x00000803(2051)	magic number
0004	32 bit integer	10000	number of images
0008	32 bit integer	28	number of rows
0012	32 bit integer	28	number of columns
0016	unsigned byte	??	pixel
0017	unsigned byte	??	pixel
.....			
xxxx	unsigned byte	??	pixel

Fig-4: Test set image

Byte Value	Data Type
0x8	Unsigned Byte
0x9	Signed Byte
0x0B	Short(2Bytes)
0x0C	Int(4Bytes)
0x0D	Float(4Bytes)
0x0D	Double(8Bytes)

Table-1: MNIST Dataset

IV. READINGTHE MNIST DATASET

The capacity that peruses the picture information restores the picture data in the way of representation

the marks. This information is utilized in each program for building forecasts. To test the data is perused accurately, we print for a couple of the names. The yield is as per the following:-

Fig 6: Picture information

3 main algorithms are used for analysis and comparison:-

1.KNN

2.SVM

3.RFC

V.CLASSIFICATION USING RANDOM FOREST CLASSIFIER(RFC)

RFC creates a group of decision tress from randomly selected subset of coaching set the subsequent formula used is below:-

$f^{\wedge}=1/b\ {}^{b}\sum_{b=1}\ f^{\wedge}b\ (x')$

THE MINST DATA AND THE RFC WORKS IS CLASSIFIED IN BELOW STEPS:

1.Fill MNIST information.

2.Separate the information and mark it as testing and training of images and names.

3. To prepare the classifier utilize cross approved for partitioning the training information into training and testing data.

4. RFC algorithms is been utilized by train classifiers. Information is given and then marked it as contribution to train classifier. RFC stand in need of quantity of trees in timberland, numeral of highlights to search for finest split, most extreme profundity of trees and so forth even as the information.

5. The number perceived utilizing RFC is coordinated with the given training labels to acquire the exactness of testing classifiers.

6. For testing data the trained classifier needs to be utilized again.

7. The tested picture information are utilized to see names of numbers, they are contrasted with the provided test marks and to anaylse the precision of calculation.

8. The Confusion Matrix is mark that furnishes level of precision which every numbers has been perceived.

Classification report for classifier RandomForestClassifier(max_depth=, random_state=0):

	precision	recall	f1-score	support
0	0.87	0.99	0.93	88
1	0.90	0.42	0.57	91
2	0.66	0.80	0.72	86
3	0.55	0.82	0.66	91
4	0.91	0.82	0.86	92
5	0.74	0.62	0.67	91
6	0.87	0.99	0.93	91
7	0.70	1.00	0.82	89
8	0.92	0.26	0.41	88
9	0.74	0.82	0.77	92
accuracy			0.75	899
macro avg	0.79	0.75	0.73	899
weighted avg	0.79	0.75	0.73	899

Fig-7: Classification report for RFC

VI. Classification based on K-Nearest Neighbors (KNN) algorithm [2]

KNN is one of the most popular algorithm of machine learning. It is used most used classification algorithm. It is a supervised learning algorithm. It compares training data with a test data and gives a value K which is nearest to the training data point.

For every test data point it calculates Euclidean and Hamming distance between test data and training data. Based on the distance it sorts them in ascending order. Form sorted array it chooses K rows and assigns a class to the test point based on most frequent class of these rows.

KNN stores all available cases and classifies new cases based on some similarities. A new object is classified by a large number of neighbor classes. The new object is assigned to the most common class of its nearest neighbors.

Figure 6 gives clear understanding of KNN approach

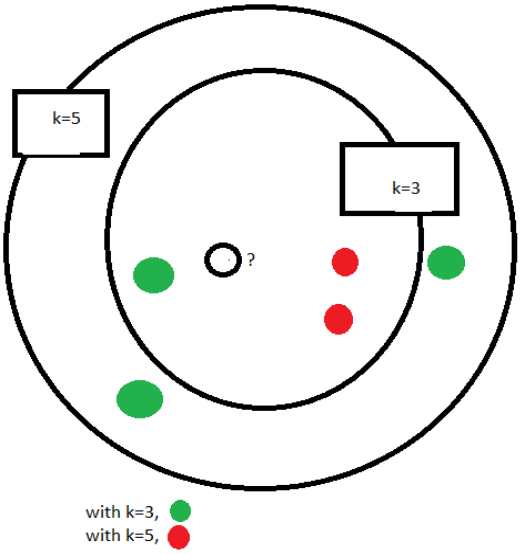

Fig 8: Working of KNN

In the above figure, the empty circle is the test data and identificationof case to be assigned is required to be done.

When k = 3, red circle will be assigned as number of red circles 2 is greater than number of green circles 1.

Whereas when k = 5, green circle will be assigned as number of green circles 3 is greater than number of red circles 2.

For the calculation of Euclidian distance following formula is used:

$d(p, q)=d(q, p) =\sqrt{(q_1 - p_1)^2 -(q_3 - p_3)^2 .....(q_n - p_n)^2}$

$=\sqrt{\sum_{i=1}^{n}(q_i - p_i)^2}$

Classification of MNIST data with KNN:

	precision	recall	f1-score	support
0	1.00	0.99	0.99	88
1	0.99	0.97	0.98	91
2	0.99	0.99	0.99	86
3	0.98	0.87	0.92	91
4	0.99	0.96	0.97	92
5	0.95	0.97	0.96	91
6	0.99	0.99	0.99	91
7	0.96	0.99	0.97	89
8	0.94	1.00	0.97	88
9	0.93	0.98	0.95	92
micro avg	0.97	0.97	0.97	899
macro avg	0.97	0.97	0.97	899
weighted avg	0.97	0.97	0.97	899

Fig 9: Classification report for KNN

From the previous image, it is evident that the accuracy of trained classifier is 97.88% and for handwritten digit prediction, the accuracy comes down to 96.67%. This means that this algorithm lacks the accuracy by 3.33%. The error of 3.33% is large when written documents are considered.

VII. SVM(SUPPORT-VECTOR-MACHINE)

SVM means support vector machine are the supervised learning model associated with the algorithm learning and analyse data. SVM is the very popular machine which offers the solutions to any problem. There are 2 types of SVM Linear and Non-Linear. A SVM model is a representation where in points in space are mapped so that the image gets separated in categories which finally gets divided as wide as possible.

VIII. WORKING OF SVM

Consider a dataset which has two tags and two features x1 and x2, so a classifier is required that can classify the pair(x1,x2) of coordinates. In 2d space it is easy to separate the two classes by a straight line, but there can be multiple lines that can separates the classes. SVM also helps to find out the proper decision that bounded these boundaries

and are said to be as hyper plane. These algorithms helps us to find the nearest points are called support vectors. The distance between the vectors and hyper plane is called margin. The distance between the maximum margin is called optimal hyper plane.

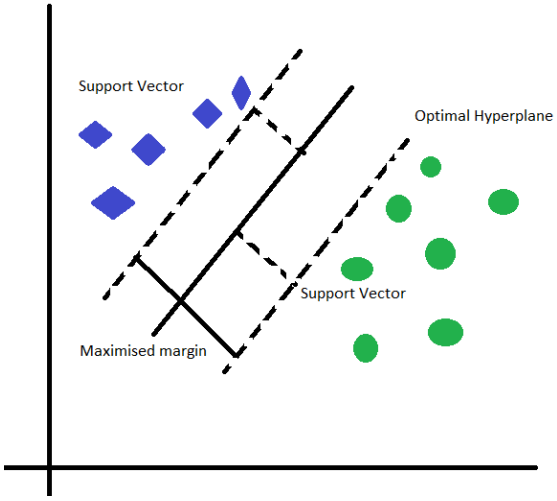

Fig10(a): Working of SVM

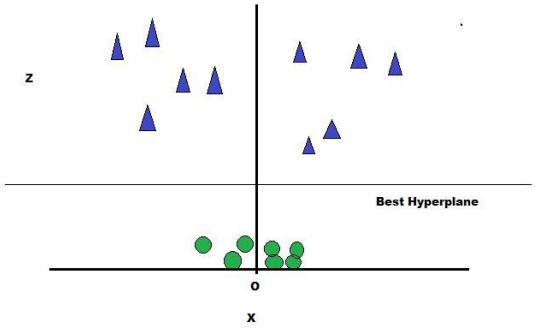

Fig 10(b): Working of SVM

IX. CONFUSION MATRIX

Every output shows a confusion matrix. It explains a particular table that permits a visual image of the performer of an associate degree by giving the accuracy such as every input and output categories. In this, the starting category is shown because the columns represent an information and the rows represent the instances in actual categories. Here the matrix C is specified as C[X, Y] which represented the quantity which familiar to be in cluster X however same expected to be in cluster Y, therefore in binary classification, the counting of all negative is C_[0,0], false negatives is C_[1,0],true positives is C_[1,1] and false positives is C_[0,1] With this process, the confusion matrix within the pictures is higher than it shows the actual accuracy of this digit which is found as a victim that algorithm rules as severally

Digit	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	**973**	1	1	0	0	1	2	1	0	0
1	0	**1133**	2	0	0	0	0	0	0	0
2	10	9	**996**	2	0	0	0	13	2	0
3	0	2	4	**976**	1	13	1	7	3	3
4	1	6	0	0	**950**	0	4	2	0	19
5	6	1	0	11	2	**859**	5	1	3	4
6	5	3	0	0	3	3	**944**	0	0	0
7	0	21	5	0	1	0	0	**991**	0	10
8	8	2	4	16	8	11	3	4	**914**	4
9	4	5	2	8	9	2	1	8	2	**96**

Fig 11: Confusion matrix representing the accuracy of each digit

The CNN for handwritten Digit Recognition has three different phases.

PHASE 1:
This section is toinput all the information. Then MNIST information is declared as 784-D array of pixels, thus first it has a tendency to convert it into grayscale pictures exploitation which are of 28 X 28 pixels.

PHASE 2:
This section, led an outline models to simply make a convolution neutral network. In this, the tendency to use the consecutive category from Keras source to make a network from this network, 3 different layers are selected “CONV=>RELU=>POOL”.

a) FIRST CONVOLUTION LAYER:

The primary layer tend to take twenty convolution channels that comes as a window of size 5x5 over all the
Photos of 28x28 network size and analysis is done at with a most sensible cost of the pixel.

b) RELU FUNCTION:

The convolution could be the techniques utilized in reverse proliferation. Hence, the RELU works the enactment whenever a convolution layer decline the odds of evaporating the inclination and furthermore stays away from poverty because of this it loses the important information and even got the information on a ton and furthermore ton inside the pixels.

c) POOLING FUNCTION:

This layer gets all the information from the RELU function to perform and also provides the major steps to the 3D TENSOR.In brief it pools all the layers whichis obtained from previous layer and that forms a replacement image matrix which is of smaller size. These pictures square measure once more input into the

second set of layer i.e. "CON=>RELU=>POOL" and method will carry on until we have a tendency to identify the digit which is of small set.

PHASE 3:

This layer has to attach every previous layer to succeeding layers. It consists of five hundred neurons and finally, a SoftMax category can be used which gives us a list of possibilities for every ten class labels. The category label is biggest likelihood and is selected because the last output is shown from network.

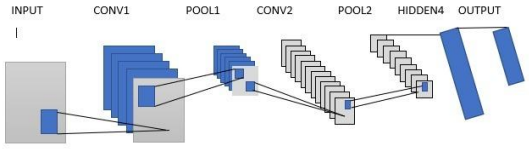

Fig 12: CNN Layers for Handwritten Digit Recognition

Fig 13: CNN Prediction Results on MNIST

The output produced is applied to form a confusion matrix for the model. Throughout the output there is a tendency for unit area and is able to add a larger variety of layers. However adding a lot of layers would possibly have an effect on the system. Since different layers are used, so it's referred as Deep Learning system.

X. RESULT AND ANALYSIS

The correct form for industrial application are digits, character and various others. Due to this it lead to a great interest of speed recognition. The accuracy and comparison of different techniques which are used is shown by the image given below. Accordingly it is analysed that the CNN have three hidden layers it provides the quantity of an accuracy of 98.72% although this accuracy is not exactly similar to others which are achieved. But mistreatment of Google's Tensor Flows the accuracy of 99.70% is achieved.

	RFC	**KNN**	**SVM**	**CNN**
Trainer classifier accuracy	99.71%	97.88%	99.91%	99.98%
Accuracy on test images	96.89%	96.67%	97.91%	98.72%

Table 2: Percent Accuracy of Each Classification Technique

Model	**Test Error Rate**
RFC	3.11%
KNN	3.33%
SVM	2.09%
CNN	1.28%

Table 3: Classifier Error Rate Comparison

By now the image should be acknowledgement by the speed and the recognition system. Given below are the times of coaching and testing. The timings from CPU of coaching and testing are noted below. Using of GPU will increase the coaching and testing time.

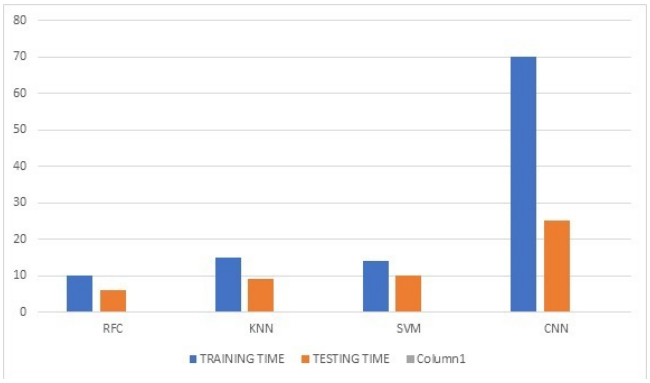

Fig 14: Classifier Training & Testing Comparison

Model	**Training Time**	**Testing Time**
RFC	10 min	6 min
KNN	15 min	9 min
SVM	14 min	10 min
CNN	70 min	20 min

Table 4: Training & Testing Time Comparison

XI. CONCLUSION

An overview of Deep learning has been done through this paper, a number of the most used Machine Learning algorithms i.e. RFC, KNN and SVM are trained and tested on the identical data to draw a comparison and this is why there is a need of deep learning method which is used in many critical applications like Handwritten Digit Recognition.

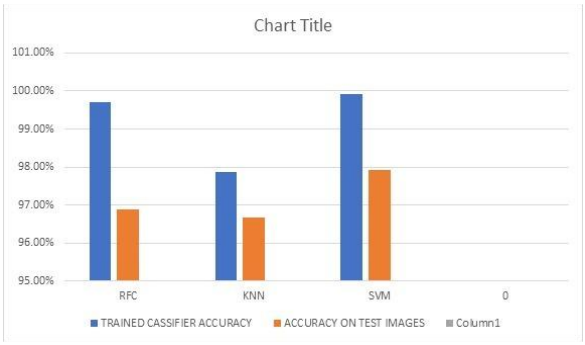

Fig15: Accuracy Comparison

Throughout the paper, it's shown that by use of Deep Learning technique a very high amount of accuracy can be achieved. Using the Convolution Neutral Networkit's given that accuracy is of 98.72%. Sopresenting the implementation of CNN using Tensor Flow it gives us a good result of accuracy which is about 99.70%. Every tool has its own complexity and accuracy. So now the process and the code complexity is normal.

XII. REFERENCE

[1] Fatahi, M., 2014. MNIST Handwritten Digits.
[2] http://yann.lecun.com/exdb/mnist/
[3] http://towardsdatascience.com/the-best-machine-learning-algorithm-for-handwritten-digit-recognition-2c6089ad8f09
[4] http://scikit-learn.org/stable/modules/generated/sklearn.ensemle.RandomForestClassifier.
[5] http://scikit-learn.org/stable/modules/generated/sklearn.neighbours.KNeighborsClassifier.html
[6] http://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html
[7] Kumar, R., Goyal, M.K., Ahmed, P. and Kumar, A., 2012, December. Unconstrained handwritten numeral recognition using majority voting classifier. In Parallel Distribution and Grid Computing (PDGC),2012 2nd IEEE International Conference on (pp. 284-289). IEEE.
[8] http://en.wikipedia.org/wiki/Support_vector_machine
[9] http://www.javapoint.com/machine-learning-support-vector-machine-algorithm