International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 4 Apr 2022, pp: 1053-1060 www.ijaem.net ISSN: 2395-5252

Handwritten Digit Recognition Based on Convolution Neural Network

Shreya Agarwal[1], Harsh Gupta[2], Palleda Soma Chandra[3]
[1,2,3] *Students, SRM Institute of Science and Technology, Delhi NCR*

Submitted: 10-04-2022 Revised: 19-04-2022 Accepted: 22-04-2022

ABSTRACT: Handwritten Character Recognition is a subfield of Image Processing that deals with extracting texts from images or scanned documents.Because of its wide range of applications, such as automatic bank check processing, billing, and automatic postal service, handwritten digit recognition has a large research area.It is difficult for a machine to comprehend handwritten numerals because of the variations in shape and orientation.In this paper, Convolution Neural Network (CNN) was employed to recognize isolated handwritten digits (0 to 9). We have chosen to take the handwritten images for both training and testing from the MNIST dataset. The handwritten recognition system allows users to draw a digit on a webpage, renders it into an image and then recognizes the features of the image and classifies it into one of the ten predefined digits (from 0-9). It thendisplays a performance curve based on the acquired accuracy. The main objective is to recognize the handwritten digits with maximum accuracy.

KEYWORDS: Image Processing, Convolution Neural Network, Handwritten Digit Recognition, Image Rendering.

I. INTRODUCTION

When the eye sees a particular image, the brain quickly segments it and recognises its many aspects. That process is carried out by the brain on its own. This entails not only the examination but also the comparing of these images in order to be aware of their various traits and what it already knows to be capable of identifying these elements. In computer science, there is a field called Image Processing that attempts to do the same thing for machines. Handwritten character or digit recognition is one of the most difficult subfields of pattern recognition. This technique translates images into a machine readable language.More precisely, in this technique, the characters in the input image are detected and the recognized characters from the image are translated into ASCII or other similar machine readable languages.Handwritten character recognition has various useful applications and has been a sought after area of research with the research starting as early as the beginning of computer science due to its intrinsic nature of communication between humans and computers.

Handwriting recognition systems can be used to recognize different types of characters such as alphabets, digits or even symbols. Offline handwriting recognition and online handwriting recognition are the two types of handwriting recognition. Offline handwriting recognition occurs when handwriting is scanned and then understood by a computer. It is referred to as online handwriting recognition when the handwriting is identified while writing on a touch pad with a stylus pen.

II. LITERATURE REVIEW

Many studies on feature extraction and classifier techniques for handwritten digit recognition have been conducted. The majority of them had high recognition accuracy. [1] Mane and Kulkarni (2018), for example, suggested a Customized Convolutional Neural Network (CCNN) that can automatically learn features and forecast numerical categories in a large extended data set, such as Marathi, one of India's most widely spoken regional languages. Furthermore, utilising K-fold cross-validation, the CCNN's performance averaged 94.93 percent accuracy. The suggested CCNN model does not set any limitations on the number of layers, but rather optimises it to meet the issue's requirement.

[2] Sadri, Suen, and Bui reported in 2007 that using context knowledge correctly in segmentation, assessment, and search might significantly enhance the overall performance of a handwritten digit recognition system. As a consequence, employing NN and SVM classifiers, the recognition system was able to achieve 95.28 percent and 96.42 percent recognition accuracy on handwritten numeric strings, respectively. According to Hochuli et al. (2018) [3], the CNN

classifier was able to handle the intricacies of touch numbers better than any other segmentation approach available in the literature. Experiments on two well-known datasets, Touching Pairs 26 Dataset and NIST SD19, show thatthe suggested method is capable of obtaining a recognition accuracy of 97 percent.

In recent years, researchers in the field of pattern recognition have focused more on multi-classifier systems, particularly Bagging and Boosting. [4] Bernard, Adam, and Heutte (2007) investigated the effect of parameter values on the performance of the RF using a traditional feature extraction technique based on a greyscale multi-resolution pyramid. They tested the Forest-RI algorithm, which is known as the Random Forest reference method, on the MNIST handwritten digital database and found that it can recognise handwritten digits with an accuracy of more than 93 percent.

[5] Hanmandlu and Murthy advocated the use of exponential membership functions as fuzzy models to recognise handwritten Hindi and English numerals in 2007. Furthermore, the double layer perceptron was used to improve the defuzzification parameters, and the fuzzy rules were created using the ID-3 approach. Traditional handwritten character recognition syntactic methods were overcome by this strategy, which attained a 95 percent accuracy for Hindi digits and a 98.4 percent accuracy for English digits.

[6] Cecotti in 2016 developed a novel active machine learning technique for handwritten numeral classification. By querying experts to provide labels for individual occurrences, dynamic learning approaches tackle the problem that enormous databases are not always immediately available. Cecotti tested the approach on four databases corresponding to different scripts (Latin, Bangla, Devanagari, and Oriya) and found that it was accurate to 98.54 percent on the MNIST training database.

[7] To identify Gurmukhi's handwritten characters, Mahto, Bahtia, and Sharma (2015) used a combination of horizontal and vertical projection feature extraction. To classify handwritten characters, the experiment used linear SVM and k-NN (k = 1, k = 3, k = 5, k = 7) with a maximum accuracy of 98.06 percent.

[8] Roy et al. published a research in 2017 that looked at how to use a novel deep learning technique to identify compound characters in handwritten Bangla. The researchers employed the RMSProp technique to improve the training process and achieve faster convergence using layered training on deep convolutional neural networks (DCNN). When compared to the traditional shallow learning model, the suggested DCNN demonstrated a considerable improvement in recognition rates, reaching 90.33 percent.

[9] Karimi et al. (2015) has suggested a method for recognising Persian handwritten digits that includes three primary sections: pre-processing, feature extraction, and classification. In the feature extraction step, 115 features abstracted from Persian handwritten digits are used to create a set of relevant and complementary features. The ensemble classifier methods like as Boosting and Bagging are used to distinguish the classes of samples during the classification phase. Furthermore, the results of this experiment were evaluated using the Tarbiat Modares University (TMU) digital database, with the greatest recognition accuracy of 98.06 percent for Persian handwritten numbers.

Table 1 Comparative description of the classification techniques RR for the handwritten characters from the different databases. [10]

WORK REFERENCE	TECHNIQUES	DATABASE	RECOGNITION RATE
Mane & Kulkarni, 2018	CCNN	Self created	94.93%
Sadri et al., 2007	NN, SVM	NIST SD19	96.42%
Hochuli et al., 2018	CNN	NIST SD19	97%
Bernard et al., 2007	RF	MNIST	93%
Hanmandlu & Murthy, 2007	ID-3	CEDAR	90.33%
Cecotti, 2016	K-NN	MNIST	98.54%
Mahto et al., 2015	SVM, KNN	Self created	98.06%

Roy et al., 2017	DCNN	CEDAR	90.33%
Karimi et al., 2015	Bagging, Boosting	TMU	98.06%

III. HDR MODELS

Some of the most popular and commonly used algorithms used for handwritten digit recognition are KNN (K nearest Neighbours), SVM (Support Vector Machine) andRF (Random Forest).

1. K Nearest Neighbours (KNN)

The K-nearest neighbour algorithm is used for both regression as well as classification. It's a versatilealgorithm that may also be used to fill in missing values and resample datasets.It belongs to the supervised learning domain and finds intense application in pattern recognition, data mining and intrusion detection.It's also known as a lazy learner algorithm since it doesn't learn from the training set right away; instead, it saves the dataset and performs an action on it when it comes time to classify it.By computing the distance between the test data and all of the training points, KNN tries to predict the proper class for the test data. It then selects the K number of points that are the most similar to the test data. The KNN algorithm analyses the likelihood of test data belonging to each of the 'K' training data classes, and the class with the highest probability is chosen.

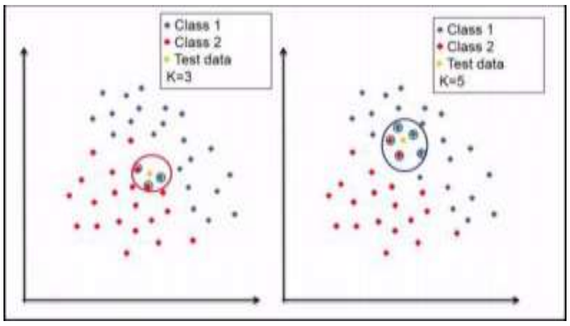

Fig 1 The working of the KNN Classifier [11]

2. Support Vector Machine (SVM)

The Support Vector Machine or SVM algorithm developed by Vapnik and Cortes in 1998, is a powerful discriminant classifier that has yielded positive results when used to a variety of pattern recognition and classification issues. It is also considered the most sophisticated tool for tackling linear and nonlinear classification problems due to its simplicity, flexibility, prediction capability, and global optimality. It employs structural risk minimization rather than the empirical risk minimization that has been used in artificial neural networks in the past.For classification, the SVM tries to find the optimum hyperplane by separating the points of two classes to the maximum extent possible, resulting in data points that are accurately classified. The SVM method only looks for the separating hyperplane with the highest margin when dealing with linear separable jobs. Additionally, SVM features a kernel method that can improve accuracy when dealing with non-linear data.

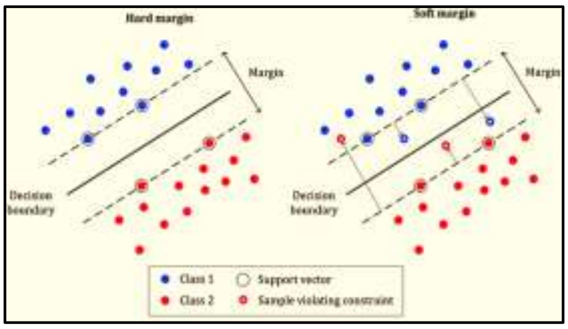

Fig 2 The working of the SVM Classifier

3. Random Forest (RF)

The Random Forest or RF Random Forest (RF) refers to a set of classifiers that use the L-tree classifier h (x, k), k = 1,... L, where k is an independent random vector from the same distribution and x is the input. Random forest can be defined as a set of approaches that includes numerous algorithms based on this notion. Breiman created the random forest concept in 2001, based on the bagging principle. Bagging is utilised in the Forest-RI method using the notion of random feature selection. Several scholars have presented versions of the Forest-RI algorithm in recent years. Breiman (2001), for example, created Forest-RC, a programme for generating RF in which each node's segmentation is based on a linear mixture of features rather than a single function. This means that the processing is limited to a small amount of data, and the Forest-RI approach is difficult to use. Robnik attempted to improve the original Forest-combination RI's process in 2004 by introducing a weighted voting approach.

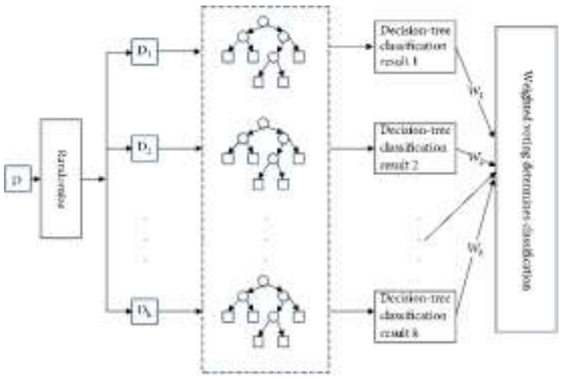
D
Randomise
D_1
D_2
D_k
Decision-tree classification result 1
Decision-tree classification result 2
Decision-tree classification result k
W_1
W_2
W_k
Weighted voting determines classification

Fig 3 Structure of Weighted Random Forest [12]

IV. MNIST DATASET

The MNIST database (Modified National Institute of Standards and Technology database) is the conventional database used for handwritten digit recognition and thus is applied here. The dataset consists of 60,000 training and 10,000 testing images.It was developed using binary images of handwritten numbers from two NIST datasets. The training set includes handwritten numbers from 250 persons, with 50 percent of the training dataset coming from Census Bureau employees and the balance from high school students. It is frequently credited as one of the first datasets among others to demonstrate the efficiency of neural networks.

The numbers are all grayscale and are all the same size, with the intensity in the centre of the image at 28×28 pixels. Because all of the photos are 28×28 pixels wide, they create an array that can be flattened into a 28*28=784 dimensional vector.Each pixel's intensity is described by a binary value in each component of the vector. Before application, any faulty pictures, such as missing values, were checked. The dataset has no missing values.

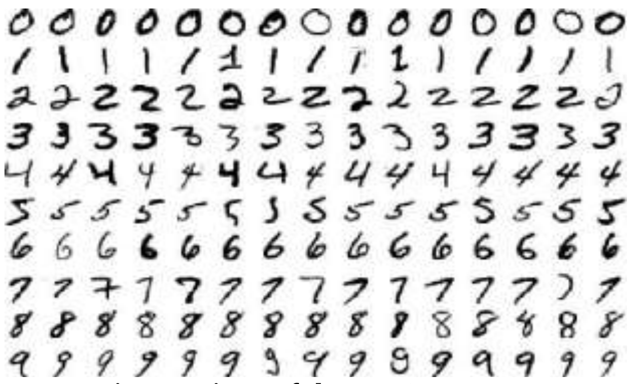

Fig 4 A snippet of the MNIST Dataset

V. CONVOLUTION NEURAL NETWORK

We have used a 7-layered convolution neural network with one input layer followed by five hidden layers and one output layer to recognise handwritten digits. The input layer is made up of 28 by 28 pixel pictures, implying that there are 784 neurons in the network. The input pixels are grayscale, with a white pixel having a value of 0 and a black pixel having a value of 1. There are five hidden layers in this CNN model. The convolution layer 1 is the first hidden layer, and it is responsible for extracting features from input data.By convolving a filter with the previous layer, this layer applies convolution to small localised areas. It also includes multiple feature maps with learnable kernels and rectified linear units (ReLU). The locality of the filters is determined by the kernel size. To improve model performance, ReLU is used as an activation function at the end of each convolution layer as well as a fully connected layer. The pooling layer 1 is the next hidden layer. It reduces the number of parameters and computational complexity of the model by reducing the output information from the convolution layer. Pooling is classified into four types: maximum pooling, minimum pooling, average pooling, and L2 pooling.

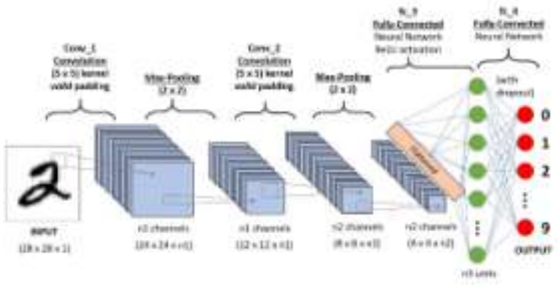
Conv_1 Convolution (5 x 5) kernel valid padding
Max-Pooling (2 x 2)
Conv_2 Convolution (5 x 5) kernel valid padding
Max-Pooling (2 x 2)
fc_3 Fully-Connected Neural Network ReLU activation
fc_4 Fully-Connected Neural Network
(with dropout)
0
1
2
9
INPUT (28 x 28 x 1)
n1 channels (24 x 24 x n1)
n1 channels (12 x 12 x n1)
n2 channels (8 x 8 x n2)
n2 channels (4 x 4 x n2)
n3 units
OUTPUT

Fig 5 7-Layered Structure of Convolution Neural Network

Convolution layer 2 and pooling layer 2 perform the same function as convolution layer 1 and pooling layer 1 and operate in the same manner, with the exception that their feature maps and kernel size differ. After the pooling layer, a Flatten layer is used to convert the 2D feature map matrix to a 1D feature vector, allowing the output to be handled by the fully connected layers. A fully connected layer, also known as the dense layer, is similar to the hidden layer of an ANN, but it is fully connected and connects every neuron from one layer to the next. Dropout regularization method is used at fully connected layer 1 in order to reduce overfitting. The network's output layer is made up of ten neurons that determine the digits 0 through 9.

VI. WORKING PRINCIPLE

The process of handwritten digit recognition using CNN is carried out in six phases. [13] The six phases are as follows -

1. Image Acquisition

Input images for handwritten digits can be obtained by directly writing in the computer or through other means, such as scanners and photographs. We have obtained input images by drawing them digitally.

2. Pre-Processing

The primary goal of the pre-processing steps is to normalise strokes and remove variations that would otherwise complicate recognition and lower the recognition rate. Pre-processing consists of five common steps: size normalisation, interpolating missing points, slant correction, and point resampling.

3. Segmentation

Segmentation is the separation of an image's individual characters. The images with more than a single digit are divided into sub-images and then digits are extracted from those images.

4. Feature Extraction

The primary goal of the feature extraction phase is to extract the most relevant pattern for classification. To extract the features of individual digits, feature extraction techniques such as Principle Component Analysis (PCA), Linear Discriminant Analysis (LDA), Chain Code (CC) etc. may be used. These characteristics are used to train the system.

5. Classification

When an input image is presented to the HCR system, its features are extracted and fed into a trained classifier like CNN. Classifiers compare the input feature to the stored pattern to determine the best matching class.

6. Post-Processing

The process of correcting misclassified results using linguistic knowledge is referred to as post-processing. It is the processing of shape recognition output. The addition of language information can improve the accuracy of pure shape recognition.

Some shape recognizers return a single string of digits for handwriting input, while others return a number of alternatives for each digit, often with a measure of confidence for each alternative.

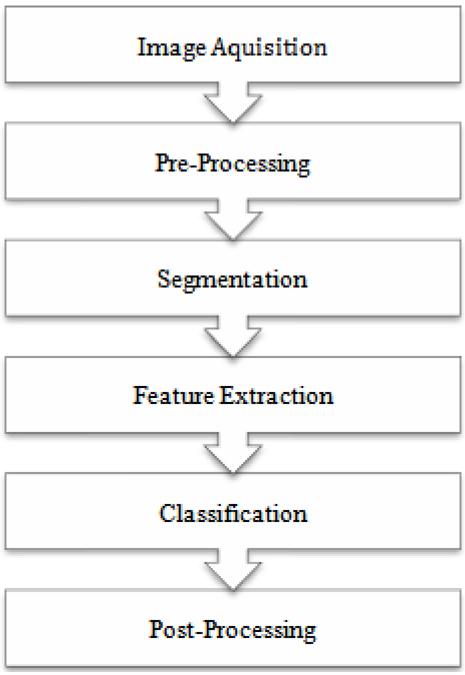

Fig 6 Block Diagram for HDR

VII. RESULT AND DISCUSSION

We applied Convolution Neural Network on the MNIST dataset to examine the variation in accuracies for handwritten digits. The accuracies obtained in the result was obtained by using TensorFlow in Python. The handwritten digit recognition system works as a web application where the user can draw any digit in the drawing board that is provided and click the 'Predict' button. It then displays a bar graph that shows the recognized digit along with the accuracy that is achieved. Fig 7 shows the user interface of the application. It shows that a number '4' was drawn on the drawing board and the resultant graph classified the digit drawn as '4' with 99% accuracy. We have also managed to reach the maximum accuracy of 100% in many cases. Fig 8 is one such example, where it is the resultant graph of the digit '2' drawn on the drawing board.

Although there are many algorithms that have been used for handwritten digit recognition, we have found our model that is based on convolution neural network, to give optimum results.

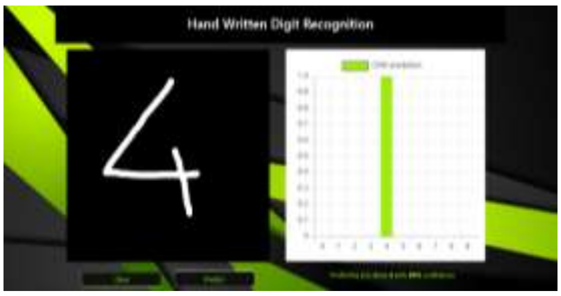

Fig 7 UI for the Handwritten Digit Recognition System

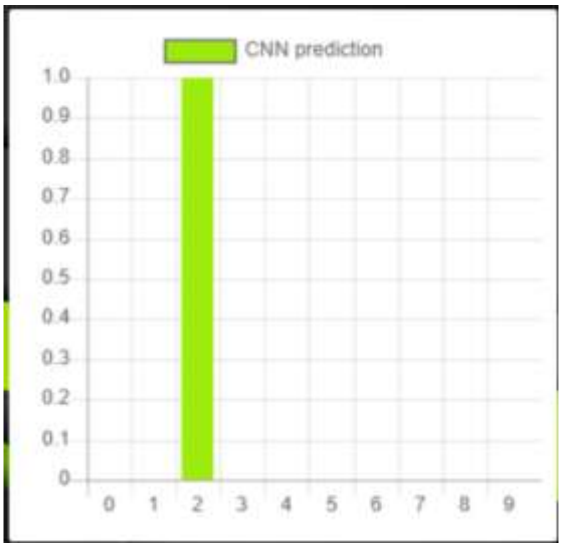

Fig 8 Image Classification Graph fordigit '2'

VIII. CONCLUSION

This paper furnishes a brief idea about the working of the Handwritten Digit Recognition System using Convolution Neural Network. It also delivers a detailed explanation of other algorithms that have been and can be used for handwritten digit recognition. This paper also discusses the advances made in the field of handwritten digit recognition.

The primary goal of this research is to find a representation of isolated handwritten digits that allows effective recognition. The most important problem in any recognition process is addressing feature extraction and correct classification approaches, which we attempted to do in our project. The algorithm we used tries to address both factors and does so well in terms of accuracy and time complexity. This work is being done as a first attempt. In comparison to other studies, this study focused on determining which image pre-processing and feature extraction techniques based on OCR can improve classification model accuracy by more than 99 percent. Based on our initial study, we have concluded that CNN gives more optimum results as compared to other algorithms such as SVM, KNN or RF and we managed to reach the maximum accuracy rate of upto 100%.

REFERENCES

[1]. Mane, D., & Kulkarni, U. (2018). Visualizing and Understanding Customized Convolutional Neural Network for Recognition of Handwritten Marathi Numerals. Procedia Computer Science,132, 1123-1137. doi:10.1016/j.procs.2018.05.027

[2]. Sadri, Javad & Suen, Ching & Bui, T.D.. (2007). A genetic framework using contextual knowledge for segmentation and recognition of handwritten numeral strings. Pattern Recognition. 40. 898-919. 10.1016/j.patcog.2006.08.002.

[3]. Hochuli, A., Oliveira, L., Jr, A. B., & Sabourin, R. (2018). Handwritten digit segmentation: Isit still necessary? Pattern Recognition,78, 1-11. doi:10.1016/j.patcog.2018.01.004

[4]. Bernard, S., Adam, S., & Heutte, L. (2007). Using Random Forests for Handwritten Digit Recognition. Ninth International Conference on Document Analysis and Recognition (ICDAR 2007) Vol 2,1043-1047. doi:10.1109/icdar.2007.4377074

[5]. Hanmandlu, M., & Murthy, O. R. (2007). Fuzzy model based recognition of handwritten numerals. Pattern Recognition,40(6), 1840-1854. doi:10.1016/j.patcog.2006.08.014

[6]. Cecotti, H. (2016). Active graph based semi-supervised learning using image matching: Application to handwritten digit recognition. Pattern Recognition Letters,73, 76-82. doi:10.1016/j.patrec.2016.01.016

[7]. Sfbwt Mahto, Manoj & Bhatia, Karamjit & Sharma, R. (2020). Segmentation of Offline Handwritten Gurmukhi Words Using Projection Features. 10.1109/SMART46866.2019.9117480.

[8]. Roy, S., Das, N., Kundu, M., & Nasipuri, M. (2017). Handwritten isolated Bangla compound character recognition: A new benchmark using a novel deep learning

approach. Pattern Recognition Letters,90, 15-21. doi:10.1016/j.patrec.2017.03.004

[9]. Karimi, H., Esfahanimehr, A., Mosleh, M., Ghadam, F. M., Salehpour, S., & Medhati, O. (2015). Persian Handwritten Digit Recognition Using Ensemble Classifiers. Procedia Computer Science,73, 416-425. doi:10.1016/j.procs.2015.12.018

[10]. Zhao, K. (2018) Handwritten Digit Recognition and Classification Using Maching Learning. M.Sc. in Computing (Data Analytics), Technological University Dublin.

[11]. Gupta, Akanksha & Narwaria, Ravindra Pratap & Solanki, Madhav. (2021). Review on Deep Learning Handwritten Digit Recognition using Convolutional Neural Network. International Journal of Recent Technology and Engineering. 9. 245-247. 10.35940/ijrte.E5287.019521.

[12]. Gao, Xiang & Wen, Junhao & Zhang, Cheng. (2019). An Improved Random Forest Algorithm for Predicting Employee Turnover. Mathematical Problems in Engineering. 2019. 1-12. 10.1155/2019/4140707.

[13]. Purohit, Ayush & Chauhan, Shardul. (2016). A Literature Survey on Handwritten Character Recognition. International Journal of Computer Science and Information Technology. 7. 1-5.

[14]. Siddique, Fathma & Sakib, Shadman & Siddique, Abu Bakr. (2019). Recognition of Handwritten Digit using Convolutional Neural Network in Python with Tensorflow and Comparison of Performance for Various Hidden Layers.

[15]. G. J. S. Padmashali and D. Kumari, “Handwritten Digit Recognition Using Deep Learning”, IJRESM, vol. 4, no. 7, pp. 182–185, Jul. 2021.