

---

## HANDWRITTEN NUMBER ANALYSIS AND PREDICTION OF NUMBERS

Rinku Gopal Jadhav\*<sup>1</sup>, Vivek Vinayak Bidgar\*<sup>2</sup>, Kunal Dnyandev Patil\*<sup>3</sup>,

Ameya Vijaysingh Lodhavat\*<sup>4</sup>, Kanchan V. Dhomse\*<sup>5</sup>

\*<sup>1,2,3,4,5</sup>MET's Bhujbal Knowledge City Savitribai Phule Pune University

Adgaon, Nashik-422003, India.

---

### ABSTRACT

The handwritten number recognition is the capability of computer operations to fit the mortal jotting styles or handwritten integers. It's a hard task for the machine because handwritten integers aren't perfect and can be made with numerous different shapes and sizes. The handwritten number recognition system is a way to break this problem which uses the image of a number and recognizes the number present in the image. The handwritten character and number recognition have been one of the most clamant and absorbing field of pattern recognition and image processing. The main end of this paper is to demonstrate and represent the work which is related to hand written number recognition. The hand written number recognition is veritably clamant task. In this recognition task, the figures aren't directly written or scripted as they differ in shape or size; due to which the point birth and segmentation of hand written numerical script is laborious. We aren't using any pre-collected images of number. But we will prognosticate images live drawn in makeup, so for this we've to collecting images from makeup by our own.

**Keywords:** Machine Learning, Handwritten Digit Recognition, Support Vector Machine, MNIST Dataset.

---

### I. INTRODUCTION

Hand drawn number from an input source (paper or photos) by comparing it with preliminarily trained data. The main ideal is to insure effective and dependable approaches for recognition of handwritten integers. Recognition is relating or distinguishing a thing or a private from the once gests or literacy. Also, Digit Recognition is nothing but feting or relating the integers in any document. Number recognition frame is simply the working of a machine to prepare itself or interpret the integers. Handwritten Number Recognition is the capacity of a computer to interpret the manually written integers from colorful sources like dispatches, bank cheques, papers, filmland, and so forth and in colorful situations for web grounded handwriting recognition on PC tablets, relating number plates of vehicles, handling bank cheques, integers entered in any forms etc. Machine Literacy provides colorful styles through which mortal sweats can be reduced in feting the manually written integers. Deep Literacy is a machine literacy system that trains computers to do what fluently falls into place for people learning through exemplifications. With the application of deep literacy styles, mortal attempts can be lowered in perceiving, learning, feting and in a lot further regions. Using deep literacy, the computer learns to carry out bracket works from filmland or contents from any document. Deep Literacy models can negotiate state-of- art delicacy, beyond the mortal position performance. The number recognition model uses large datasets in order to fete integers from distinctive sources. Handwriting recognition of characters has been around since the 1980s. The task of handwritten number recognition, using a classifier, has extraordinary significance and use similar as online number recognition on PC tablets, fete zip canons on correspondence, recycling bank check quantities, numeric sections in structures filled up by hand (for illustration- duty forms) and so on. There are different challenges faced while trying to break this problem. The handwritten integers aren't always of the same size, consistence, or exposure and position relative to the perimeters. The main ideal was to appear a pattern characterization system to perceive the handwritten integers handed in the MNIST data set of images of handwritten integers (0-9).

### II. LITERATURE REVIEW

This part of the project describes the theoretical background of this project, starting with an explanation of “**Handwritten Digit Recognition**”, followed by overviews of machine learning related work and Predictions of handwritten digits. Various papers are available that contain information on “**handwritten digit recognition using machine learning**”. Some referred papers are mentioned here.

Sr.No.	Paper Title	Publication	Abstract
1	Handwritten Digit Recognition Using Machine Learning : A Review	IEEE	In this paper, various methods for handwritten numeral recognition based on MNIST database are compared.
2	Recognition of Handwritten Digit using Convolutional Neural Network in Python with Tensor-Flow and Comparison of Performance for Various Hidden Layer	IEEE	In this paper, the variations of accuracies for handwritten digit were observed for 15 epochs by varying the hidden layers.
3	Handwritten Digit Recognition System Based on Convolutional Neural Network	IEEE	This paper designs a handwritten digit recognition system based on convolutional neural network. The system adopts the method of deep learning and uses the MNIST data set as a training sample.
4	Hand Written Digit Recognition using Machine Learning	IEEE	KNN algorithm to solve the problem of Hand-Written Digit Classification; the dataset used to solve the problem is referred to as the MNIST dataset and Feature Extraction using PCA
5	Usage of Quantum K-Nearest Neighbor Algorithm to improve handwritten digit r	IJTP	Study the efficiency of quantum computing using Grover Algorithm and KNearest Algorithm
6	Strengthening Handwritten Digit Recognition with Two State Q Learning	IEEE	Using two Q state it's become very simple and easy to due to it required less parameter to optimize and easy to function on MNIST Digital dataset, USPS dataset and MATLAB dataset
7	Recognition of handwritten digits with classification of decision tree: A machine learning method	IJECE	This paper tested the standard digital dataset from kaggle for recognition of handwritten digits using a machine learning algorithm decision tree. Accuracy will from 0 -9 digits
8	A comparison of three classification algorithms for the identification of handwritten digits	ICOASE	Comparison of three classification algorithms In other terms Multilayer Perceptron (MLP), Naive Bayes(NB), and K-Star

### III. METHODOLOGY

#### 3.1 MNIST Dataset

Each exploration work needs some estimation, to measure the delicacy and performance of handwritten integers, MNIST dataset is being used for similar reasons. MNIST is the most astronomically employed standard

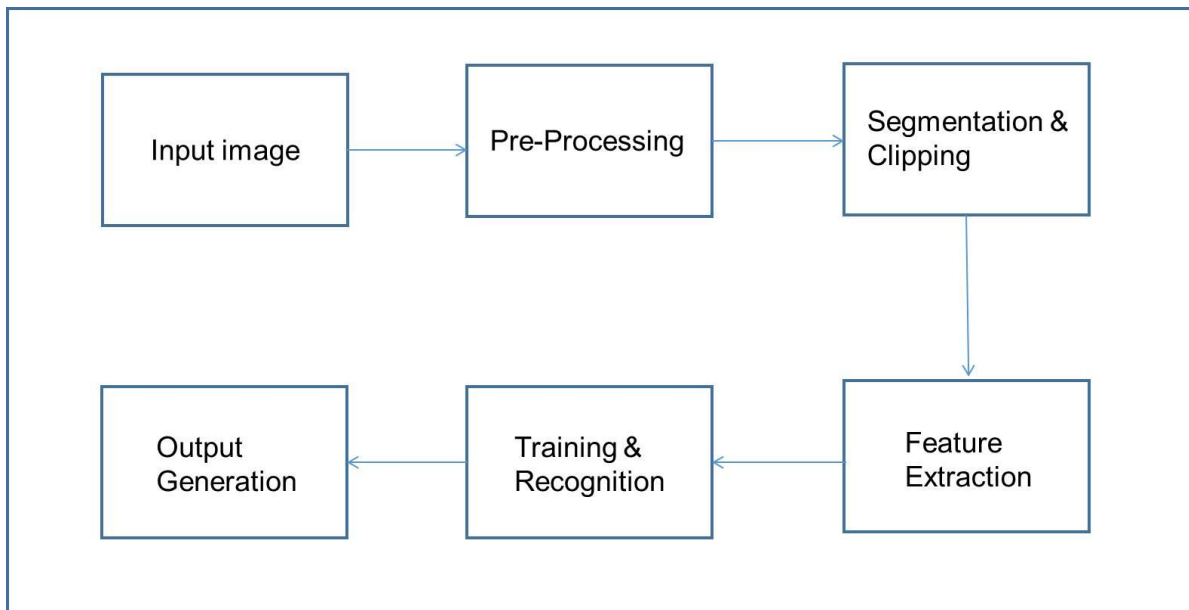
for handwritten number recognition. MNIST is a huge and a standard database of handwritten integers. MNIST dataset has been generally used as a standard for testing bracket algorithms in handwritten number recognition fabrics. The original step to be carried out is to place the dataset, which can be effectively done through the Keras programming interface. Modified National Institute of Norms and Technology dataset, can be defined as the large dataset which is used to train colorful machine literacy models. The black and white images attained from this dataset are regularized in 28x28 pixel bounding box which are anti-aliased and led to the preface of argentine scale situations. The pixels are given as a variety of 784-d pixels and the range extends from 0 to 255 for illustration 0 implies Black and 255 implies White.



**Figure 1:** MNIST Dataset

### 3.2 SYSTEM ARCHITECTURE

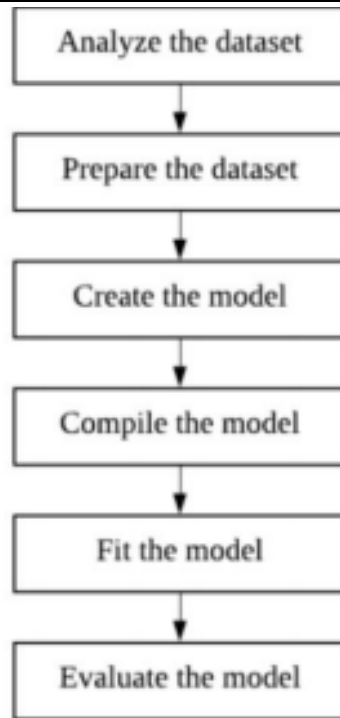
The reason behind this document is to look into the design possibilities of the proposed system, similar as armature design, block illustration, sequence illustration, data inflow illustration and stoner interface design of the system in order to define the way similar as pre-processing, feature extraction (point birth), segmentation, classification and recognition (bracket and recognition of integers).



**Figure 2:** System Architecture

The above Figure illustrates the architecture illustration of the proposed system. The proposed model contains the four stages in order to classify and descry the integers

- A. Pre-processing
- B. Segmentation
- C. Feature Extraction
- D. Classification and Recognition.



**Figure 3:** Handwriting Recognition Processing Steps

### 3.2.1 Pre-processing

The part of the pre-processing step is it performs colorful tasks on the input image. It principally upgrades the image by making it reasonable for segmentation. The abecedarian provocation behind pre-processing is to take off a fascinating illustration from the background. For the utmost part, noise filtering, smoothing and standardization are to be done in this stage. The pre-processing also characterizes a lower depiction of the illustration. Binarization changes over a argentine scale image into a double image. The original approach to the training set images that are to be reused in order to reduce the data, by thresholding them into a double image.

### 3.2.2 Segmentation

Once the pre-processing of the input images is completed sub-images of individual integers are formed from the sequence of images. Pre-processed number images are segmented into a sub-image of individual integers, which are assigned a number to each number. Each individual number is resized into pixels. In this step an edge discovery fashion is being used for segmentation of dataset images.

### 3.2.3 Feature Extraction

Point Birth After the completion of pre-processing stage and segmentation stage, the pre-processed images are represented in the form of a matrix which contains pixels of the images that are of veritably large size. In this way it'll be precious to represent the integers in the images which contain the necessary information. This exertion is also called point birth. In the point birth stage redundancy from the data is removed.

### 3.2.4 Classification Recognition

In the bracket and recognition step the uprooted point vectors are taken as an individual input to each of the following classifiers. In order to showcase the working system model uprooted features are combined and defined using following three classifiers K-Nearest Neighbor , Random Forest Classifier ,Support Vector Machine .

## 3.3 ALGORITHM SURVEY

### 3.3.1 Support Vector Classifier (SVC)

We used Support Vector Classifier algorithm for classification.

We used Support Vector Classifier algorithm for bracket. Support Vector Classifier (SVC) system applies a kernel function to perform bracket and it performs well with a large number of samples. SVC is a

nonparametric clustering algorithm that doesn't make any supposition on the number or shape of the clusters in the data. In our experience it works best for low-dimensional data, so if your data is high dimensional, a preprocessing step, e.g. using top element analysis, is generally needed.

### 3.4 IMPLEMENTATION

Steps In this design we divided perpetration in four corridor as.

- Screen Capture
- Induce dataset and Cargo it
- Fit the model using SVC and calculate delicacy
- Prediction of image drawn in makeup converting our design in GUI.

#### 3.4.1 Screen Capture

We have to collect images of integers (from 0 to 9). To collect images, pyscreenshot package can be used. This package can be downloaded using pip which is a package operation tool that are written in python and used to install python packages.

#### 3.4.2 Induce dataset and Cargo it (Generate dataset and Load it)

We generated our dataset using images that we've collected in 1st step. To induce dataset, we assign 1 to the drawn region and 0 to the background. That means, in our dataset, we will be having only two values i.e., 0 and 1. Pixel value ranges from 0 to 255. Generally, 0 represent black and 255 represent white. we will assigning 0 to pixel value from 0 to 100 and 1 to pixel value from 100 to 255. Now our pixel value isn't from 0 to 255, it's only 0 and 1. In this way, we will generating dataset (csv train). Final step is to open the dataset, shuffle it i.e., change the position of each row of data and display it.

#### 3.4.3 Fit the model using SVC and calculate delicacy

We have to train our model and calculate the delicacy. For this, we separate dependent (Y) and independent variable (X). Our independent variable will be the pixel value (i.e. 0 and 1). Each number is represented by huge number of 0 and 1. And our dependent variable will be our number (i.e. from 0 to 9). We used scikit-learn to train our model.

#### 3.4.4 Prediction of image drawn in makeup converting our design in GUI

We have to prognosticate number drawn in makeup. Now, we don't give 20 we draw number in makeup and at the same time, pass that number to model, our model has to prognosticate that number.

## IV. EXPERIMENTAL ANALYSIS & RESULTS

This is our Final GUI for "Handwritten Digit Recognition using Machine Learning Algorithm".

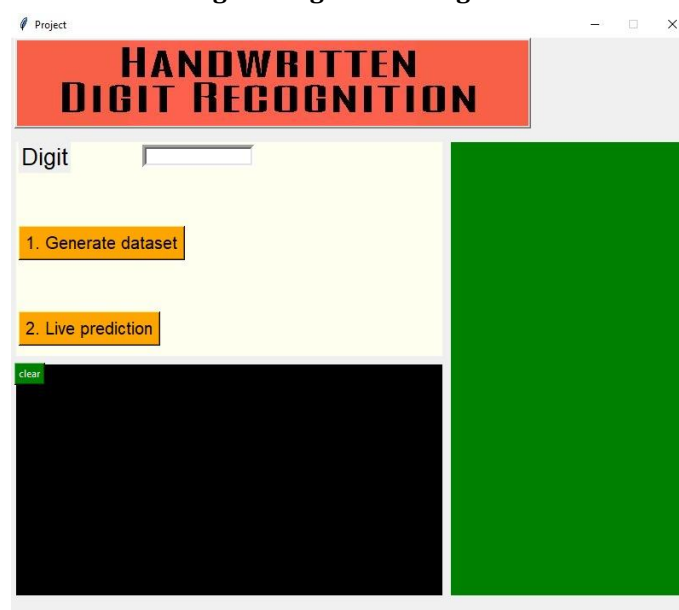


Figure 4: GUI Screen

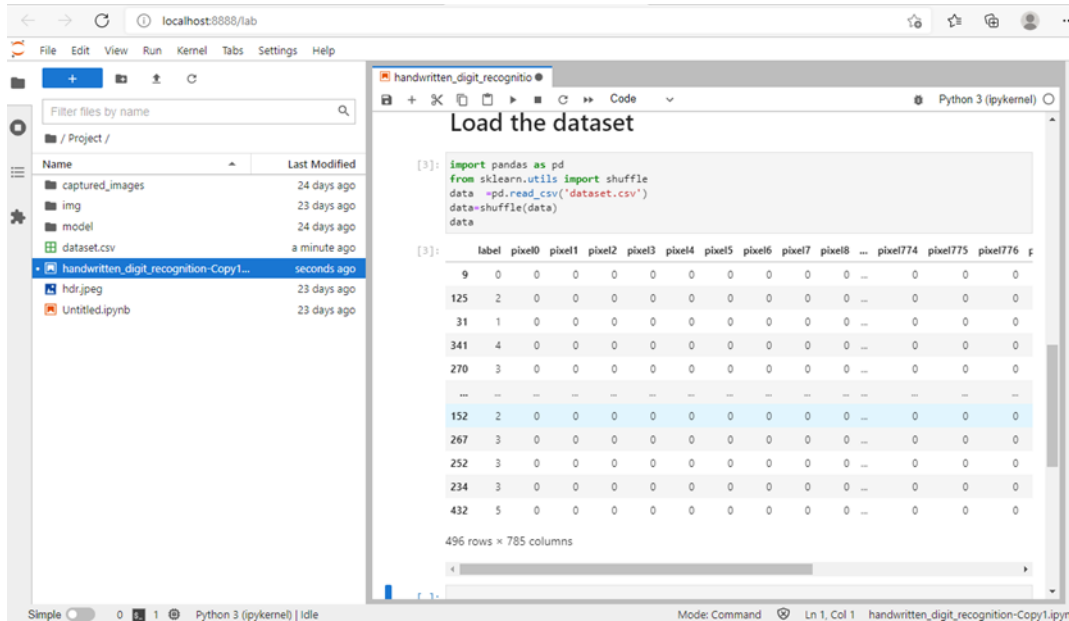


Figure 5: LOAD DATASET

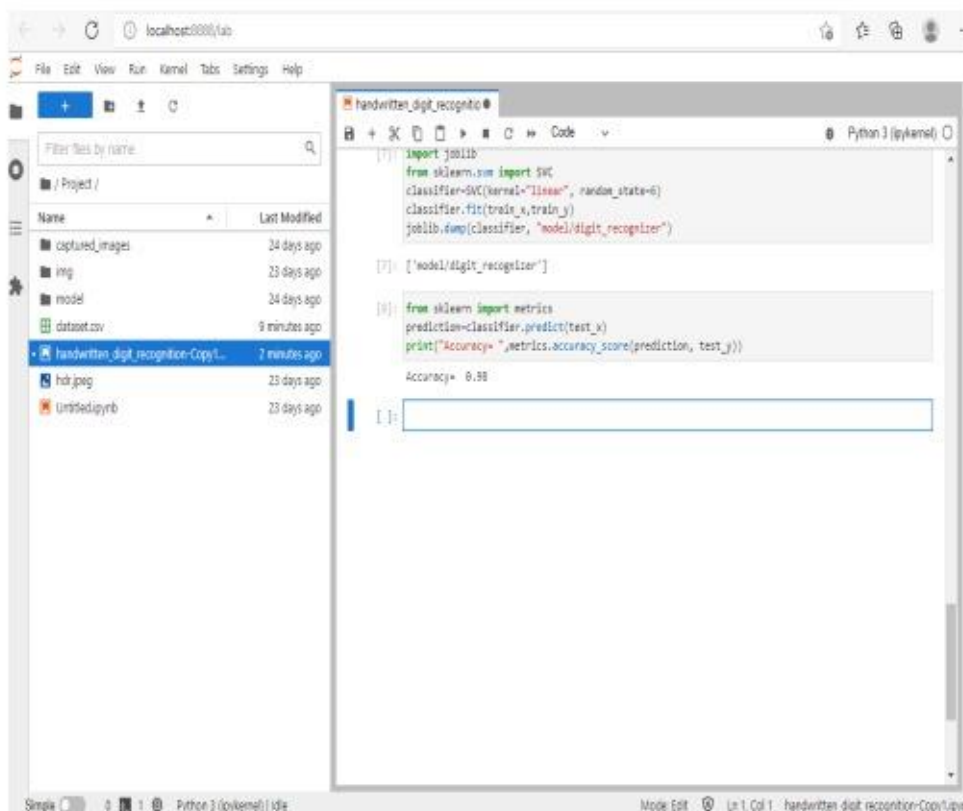


Figure 6: SVC Model & Calculate Accuracy



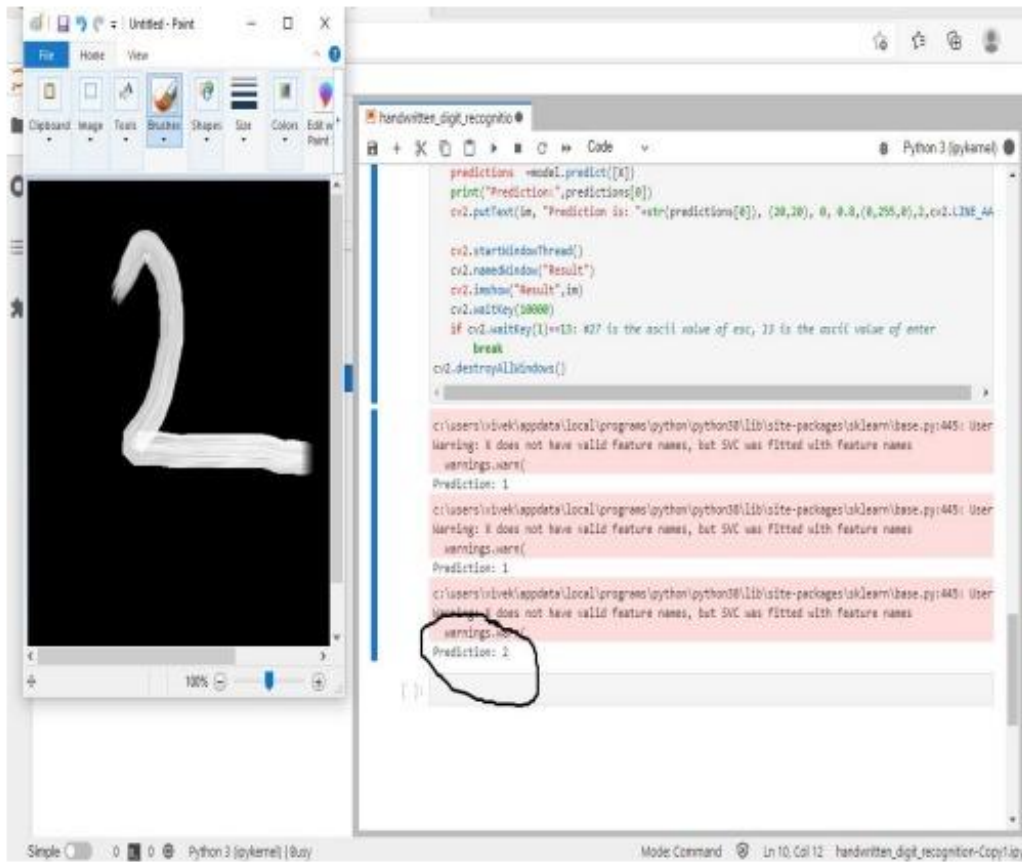


Figure 7: Prediction of image



Figure 8: Prediction of Digit



Figure 9: Result

## V. CONCLUSION

In this work, we examined the concepts of machine learning and how they might be used in the prediction of Hand-written number recognition using the efficient SVC supervised machine learning algorithm. All the experiments were implemented or executed in Jupiter Notebook, a web application which runs on local machine server the datasets used were training and testing MNIST datasets.

## VI. REFERENCES

- [1] Rohan Sethi , Ila Kaushik, "Hand Written Digit Recognition using Machine Learning ," 9th IEEE International Conference on Communication Systems and Network Technologies.
- [2] Wenfei Liu, Jingcheng Wei, Qingmin Meng, "Comparisons on KNN, SVM, BP and the CNN for Handwritten Digit Recognition," 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA).
- [3] Anchit Shrivastava, Isha Jaggi ,Sheifali Gupta, Deepali Gupta, "Handwritten Digit Recognition Using Machine Learning : A Review ," 2nd International Conference on Power Energy, Environment and Intelligent Control (PEEIC) G. L. Bajaj Inst. of Technology and Management Greater Noida, U. P., India, Oct 18-19, 2019.
- [4] Jinze Li, Gongbo Sun, Leiye Yi, Qian Cao, Fusen Liang, Yu Sun has implemented this paper designs a handwritten digit recognition system based on convolutional neural network. The system adopts the method of deep learning and uses the MNIST data set as a training sample.
- [5] Hamayun A. Khan , "MCS HOG Features and SVM Based Handwritten Digit Recognition System" , Scientific Research , 2017.
- [6] Yuxiang Wang , Ruijin Wang , Dongfen Li , Daniel Adu-Gyamfi , Kaibin Tian and Yixin Zhu, "Improved Handwritten Digit Recognition using Quantum K-Nearest Neighbor Algorithm" Springer , 2019.
- [7] Assegie, T. A., Nair, P. S., "Handwritten digits recognition with decision tree classification: a machine learning approach," International Journal of Electrical and Computer Engineering (IJECE), Indonesia, 446-4451, 2019.
- [8] Liu, C., Fujisawa, H., "Handwritten digit recognition: Benchmarking of state-of-the-art techniques," Pattern Recognition, 2003.
- [9] Fialoke, S., "Predicting Digits from their handwritten images, (16.10.2020).
- [10] Gunn, S.R., "Support vector machines for classification and regression," 1998.
- [11] Khan, H. (2017) MCS HOG Features and SVM Based Handwritten Digit Recognition System. Journal of Intelligent Learning Systems and Applications, 9, 21-33.
- [12] Wu S., Wei W., Zhang L. (2018) Comparison of Machine Learning Algorithms for Handwritten Digit Recognition. In: Li K, Li W., Chen Z., Liu Y. (eds) Computational Intelligence and Intelligent Systems. ISICA 2017. Communications in Computer and Information Science, vol 874. Springer, Singapore



- 
- [13] Tokas, R., & Bhadu, A. A comparative analysis of feature extraction techniques for handwritten character recognition. *International Journal of Advanced Technology & Engineering Research*, 2(4), 215-219, 2012.
- [14] Neves, R. F., Alberto Filho, N. G., Mello, C. A., & Zanchettin, C. A SVM based off-line handwritten digit recognizer. In *System, IEEE International Conference on Man, and Cybernetics (SMC)* (pp. 510-515), 2011
- [15] Seewald, A. K. Digits-a dataset for handwritten digit recognition. *Austrian Research Institut for Artificial Intelligence Technical Report, Vienna (Austria)*.
- [16] Das, N., Pramanik, S., Basu, S., Saha, P. K., Sarkar, R., Kundu, M., & Nasipuri, M. (2014). Recognition of handwritten Bangla basic characters and digits using convex hull based feature set. *arXiv preprint arXiv:1410.0478*.