IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Handwritten Digit Recognition and Text Conversion using MNIST Dataset

[1]K. Deepthi Krishna, [2]Sathvika. R, [3]K. Kavya Sree, [4]Raga Deepika K, [5]Satya Priyanka
[1]Assistant Professor, [2]Student, [3]Student, [4]Student, [5]Student, [1]Computer Science and Engineering,
[1]Vignan's Institute of Engineering for Women, Visakhapatnam, India.

Abstract **-** The aim of this paper is to implement a better and accurate approach to perceive and foresee manually written digits from 0 to 9 and determine the identity of the digit. It is done by extracting data from the famous MNIST dataset, using Histogram of Orientated Gradients (HOG) algorithm and Support Vector Classifier (SVC) to fit this data. The analysis is done considering the parameters, recognition rate, error rate, misclassified image rate and computing time and the results are declared on the basis of accuracy and performance.

Keywords **- Histogram Orientation of Gradients, Support Vector Classifier, Accuracy, Performance**

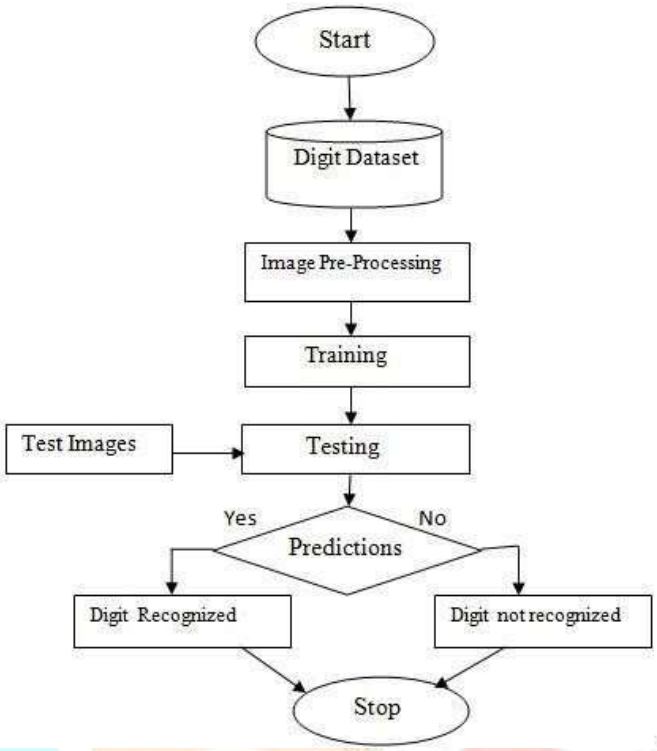
I. INTRODUCTION

Machines are capable of doing high-frequency repetitive tasks with high accuracy without getting bored. When a human does the task, there is proximity that the quality of outcome would vary. They get exhausted after a few hours of work and there is diversity in opinions which impacts the outcome unlike machines. Machine Learning algorithms receive and analyze input data and predict output values within an acceptable range. When new data is fed to these algorithms, they learn and optimize their operations to improve performance. Handwritten Digit Recognition is the process of receiving illegible handwritten input from sources like as paper documents, photographs, touch screens and other such devices and interpret it into legible format. This system is capable of taking handwritten digit input in two ways, through webcam and direct image uploading. Then the raw digit input is made to undergo formatting through preprocessing techniques like blurring, thresholding and contouring. Finally, handwritten digit input is converted into digital format.

II. LITERATURE SURVEY

In 1959, Grimsdale made an attempt in the area of character recognition. Later in 1968 Eden proposed an approach termed as analysis-by-synthesis method to carry on the research work. Eden showed that all handwritten characters have some schematic features. T. Siva Ajay also proposed that the higher rate of accuracy in handwritten digit recognition task can be achieved by the use of convolution neural networks. Shashank Mishra, D.Malathi and K. Senthil Kumar attempted the handwritten recognition using Deep Learning. A. Brakensiek, J. Rottland, A. Kosmala and J. Rigoll proposed a system for offline cursive handwriting recognition which is based on Hidden Markov Models (HMM). Diagonal Feature Extraction has also been proposed for offline character recognition. It is based on ANN model. Here, two approaches using 54 features and 69 features are chosen to build this Neural Network recognition system. Simon Bernard, Laurent Heutte, S´ebastien Adam took up Handwritten Recognition using Neural Networks in 2007 which have seemingly less recognition rates.

III. FLOWCHART



IV. TECHNIQUES USED

Handwritten character recognition is a field of image processing and pattern recognition. There are two approaches for the pattern recognition i.e., statistical and structural. In statistical approach, the set of characteristic measurements of the input data is generated on the statistical basis and is assigned to one of the n-classes. The structural description of the object is based on the inter- connections and inter-relationships between features of input data. Both approaches are widely used in pattern recognition. Since the handwriting style of different writers varies with another, building a general recognition system that would recognize all characters with good reliability is not possible in every application. Thus, recognition systems are developed specific to applications for achieving reliable performance. In particular, the handwritten digit recognition has been applied to recognize amounts written on cheques for banks and zip codes on envelops for postal services. Handwritten digit recognition system is divided into four stages;

- Data acquisition
- Pre-processing
- Feature extraction
- Classification

(i) Data acquisition: The data used in this project is a set of handwritten digits from 1 to 10. The data has been divided into two categories which form the training set and the test set. Sets of data were collected which are the phone numbers, zip-codes, and address plates for testing purpose.

(ii) Pre-processing: The size of the image used is this project is 28x28 pixels. The images used for testing consists of more than one digit which needs to be separated into individual digits before applying the pre-processing steps. The pre-processing steps include 4 processes;

- Blurring
- Thresholding
- Contouring
- Bounding-Box Formation

A. Blurring: Blurring is the process of removal of sharpness from digits contained in the image input to get rid of irregularities in the raw handwritten digits. For this process, Gaussian Blur technique is used which is a visual effect blurring technique to reduce image noise and reduce detail.

B. Thresholding: Thresholding is the simplest method of segmenting image. Thresholding is used to create binary images from a grayscale image. Simple thresholding methods replace each pixel in an image with a black pixel depending on a constant value. If the image intensity I is less than some fixed constant T or a white pixel if the image intensity is greater than that constant. This results in the dark areas becoming completely black, and the white areas becoming completely white. In our system this ranges between 0 and 255. Anything in the image in the range 0 to 255 is converted to white whereas below 0 or above 255 are converted to black.

C. Contouring: Contours are defined as the line joining all the points along the boundary of an image that are having the same intensity. Contours are present in shape analysis, finding the size of the object of interest, and object detection. So, contouring is applied on individual digits to make it easy for bounding-box formation in the next phase.

D. Bounding-Box Formation: The circumscribed rectangle, or bounding box, is the smallest rectangle that can be drawn around a set of points such that all the points are inside it, or exactly on one of its sides. The four sides of the rectangle are either vertical or horizontal, parallel to the x or y axis. It is called bounding box because it forms a boundary, like a fence, around the shape or set of points. The bounding box is formed taking the top-left corner and bottom-right corner of the digit. The top-right corner is calculated by adding width, and bottom-left corner is calculated by adding the length. Finally, one bounding-box is formed around each digit in the image. The bounded image is further cropped and sent for recognition.

(iii) Feature extraction: Feature Extraction is the process of converting given raw data into set of instance points embedded in a standardized, distinctive and machine understandable space. Feature extraction in digit recognition is the process of segmenting digits from the input of many words and then transforming them into a feature set which is usable by the classifier.

(iv) Classification: The process of classification is performed using Support Vector Classifier (SVC). The objective of an SVC is to fit into the data that is provided, returning a "best fit" hyper- plane that divides, or categorizes, your data. Thereafter, getting the hyper-plane, some features are then fed to the classifier to "predict class".

a) MNIST Dataset

Samples provided from MNIST (Modified National Institute of Standards and Technology) dataset includes handwritten digits total of 70,000 images consisting of 60,000 examples in training set and 10,000 examples in testing set, both with labeled images from 10 digits (0 to 9). This is a small segment that form the wide set from NIST where size was normalized to fit a 20x20 pixel box

and considering the aspect ratio unchanged. Handwritten digits are images in the form of 28*28 grayscale intensities of images representing an image along with the first column to be a label (0 to 9) for every image. The same is opted for the case of the testing set as 10,000 images with a label of 0 to 9.

table:1 Dataset distribution

Digits	Training Set	Testing Set
0	1000	980
1	1000	1135
2	1000	1032
3	1000	1010
4	1000	982
5	1000	892
6	1000	958
7	1000	1028
8	1000	974
9	1000	1009
Total	10000	10000

b) Classifier - SVM (Support Vector Machine)

SVM falls into the category of supervised learning, and with the bonus of classification as well as regression problems. Generally, SVM draws an optimal hyper-plane which classifies into different categories. In two-dimensional space, to start with, we plot the data points of the independent variable corresponding to the dependent variables. Then, the classification process is begun from looking the hyper-plane or any linear or nonlinear plane differentiated the two class at its best.

Algorithm -

1. Identify the correct hyper-plane which segregates the two classes better.

2. Observe for the maximum distance between nearest data point (of either any class) and hyper- plane, the distance is measured as a margin. So, identify the hyper-plane with maximum margin both sides equally. Hyper-plane with higher margin is more robust, whereas low margin has changed for mis-classification.

3. SVM selects the classifier accurately to maximized margin.

4. SVM is robust to the classifier and have a feature to ignore outliers and try to identify the hyper- plane with maximum margin.

c) Feature Descriptor - Histogram of Oriented Gradients (HOG)

The histogram of oriented gradients (HOG) is a feature descriptor used for the purpose of object detection in the field of computer vision and image processing. The technique considers occurrences of gradient orientation in localized regions of an image. This method is identical to that of the edge- orientation histograms scale-invariant feature transform descriptors, and shape contexts, but differs when computed on a dense grid of uniformly spaced cells and uses overlapping local contrast normalization for improving accuracy.

Gradient of an image defines edges of an object. Here, the horizontal f x (x, y) and vertical gradient $f_y(x, y)$ of image $f(x, y)$ are calculated without smoothing. The gradient magnitude and orientation are calculated as;

$$m(x,y) = \sqrt[2]{f_x(x, y)^2 + f_y(x, y)^2} \quad (1)$$

$$\varphi(x,y) = \tan^{-1}\left(\frac{f_x(x,y)}{f_y(x,y)}\right) \quad (2)$$

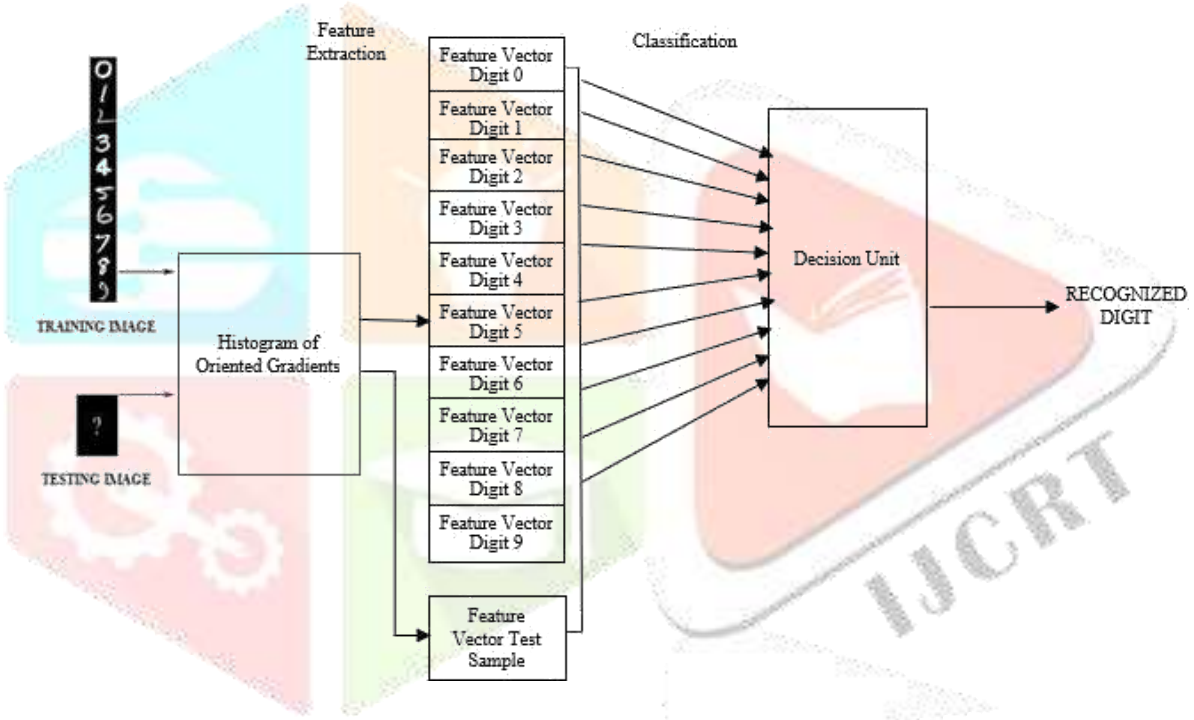


figure.1 Handwritten Digit Recognition using HOG

I. EXPERIMENTAL RESULTS

A. Recognition

As the input image can be given in two forms i.e., either through webcam or direct image upload, the results in each case are as follows.

Webcam

i) Capturing image through Webcam: Image is captured instantly through webcam using save button on the interface and the image converted into grayscale form is displayed on the other half of the screen.

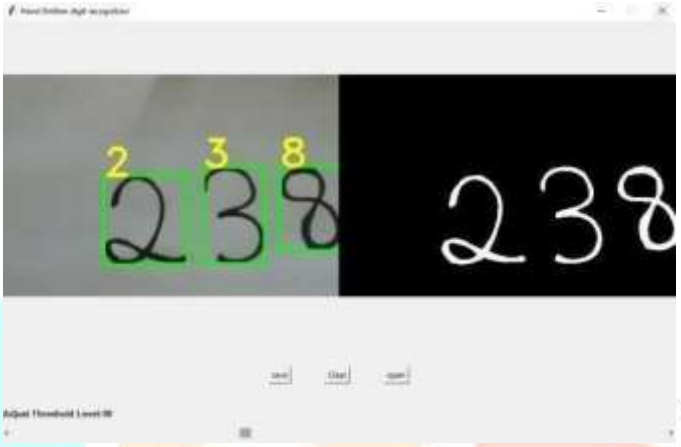


figure.2 Input through Webcam

ii) Recognition of digits on webcam- captured image: The digits on the image captured by webcam are separately made visible using bounding boxes and each recognized digit is displayed against the respective bounding box.

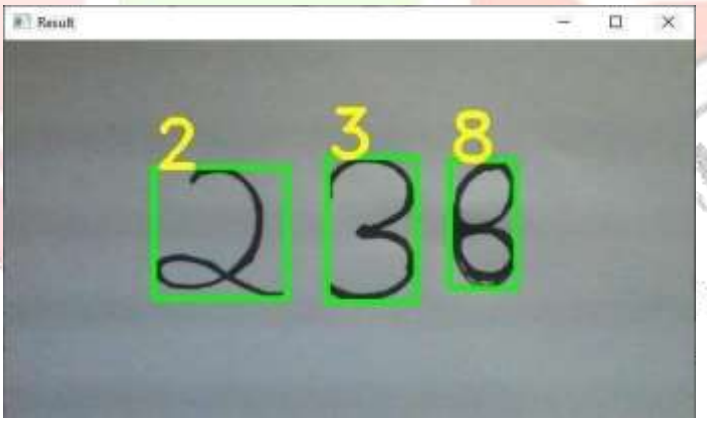


figure.3 Webcam Input Result

Image-Upload

i) Uploading image: Image is uploaded from the local files on the PC using open button on the interface and the image converted into grayscale form is displayed on the other half of the screen.

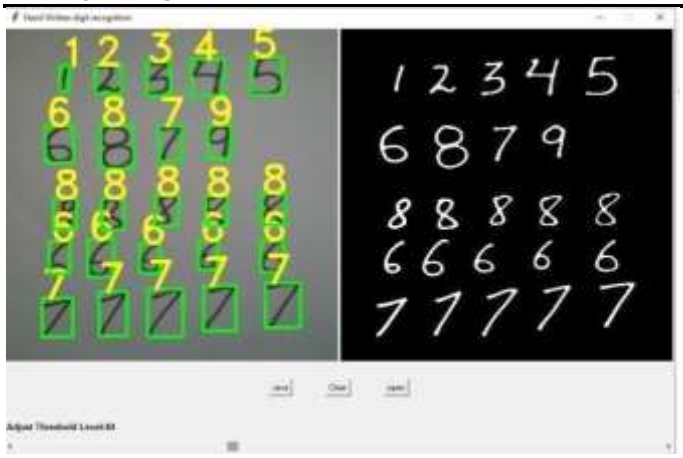


figure.4 Input through Image Upload

ii) Recognition of digits on uploaded image: The digits on the uploaded image are separately made visible using bounding boxes and each recognized digit is displayed against the respective bounding box.

figure.5 Image-Uploaded Input Result

B. Display of recognized digits in text file

The recognized digits are converted into text format and displayed in a separate text file.

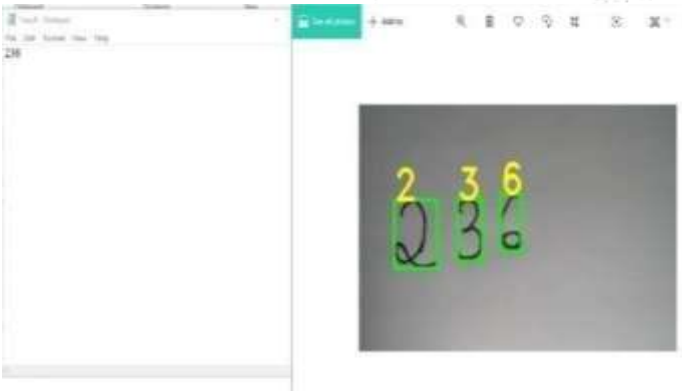


figure.6 Text-Converted Output

II. PERFORMANCE

The presented results show that the Handwritten Digit Recognition system using Histogram Orientation of Gradients (HOG) algorithm achieves an accuracy of 90%. Conversely other methods like neural networks give lesser efficient results but with lesser time complexity.

III. CONCLUSION

Support Vector Classifier (SVC) with the help of Histogram of Oriented Gradients (HOG) algorithm has shown significant performance on the MNIST dataset. The main objective of the project has been achieved by overcoming manual time and effort to recognize, search and extract required information from handwritten documents. The proposed work can further be enhanced in the terms of performance.

REFERENCES

[1] B. Scholkopf, C. Burges,V. Vapnik, Deriving support data for a given task, First International Conf. Knowledge Discovery & Data Mining, AAAI Press, MenloPark, CA, 1995 - B. Scholkopf, P. Simard, A. Smola, V. Vapnik, Prior knowledge in support vector kernels, Advances in Neural Information Processing Systems, vol. 10, MITPress, 1998.

[2] P. Simard, Y. Le Cun,J. S. Denker, Efficient pattern recognition using an advanced transformation distance, in: Advances In Neural Information Processing Systems, vol. 5, Morgan Kaufmann, 1993.

[3] Xiao-Xiao Niu, Ching Y. Suen: A novel hybrid CNN-SVM classifier to recognize handwritten digits. Pattern Recognition (PR) 45(4):1318-1325 (2012)

[4] X. Wang, M. Yang, S. Zhu, and Y. Lin. "Regionlets for generic object detection". In ICCV, 2013.

[5] A. Yuan, G. Bai, L. Jiao, Y. Liu, "Offline Handwritten English Character Recognition Based on Convolutional Neural Networks", 10th IEEE International Workshop on Document Analysis Systems, (2012).

[6] W. M. Pan, T. D. Bui and C. Y. Suen, "Isolated Handwritten Farsi Numerals Recognition With Sparse and Over-Complete Representations", 10th International Conference on Document Analysis and Recognition, (2009).

[7] U. Babu, A. Chintha and Y. Venkateswarlu, "Handwritten Digit Recognition With Structural, Statistical Features and K-Nearest Neighbor Classifier", International Journal Information Engineering and Electronic Business, vol. 6, no. 1, (2014).

[8] B. E. Qacimy, M. Kerroum and A. Hammouch, "Feature Extraction based on DCT for Handwritten Digit Recognition", International Journal of Computer Science Issues, vol.11, no. 2, (2014).

[9] S. Al-Omari, P. Sumari, S. Al-Taweel and A. Husain, "Digital Recognition using Neural Network", Journal of Computer Science, vol. 5, no. 6, (2009).

[10] R. Ebrahimzadeh and M. Jampour, "Efficient Handwritten Digit Recognition based on Histogram of Oriented Gradients and SVM", International Journal of Computer Applications, vol. 104, no. 9, (2014).

[11] A. Lawgali, "Handwritten Digit Recognition based on DWT and DCT", International Journal of Database Theory and Application, vol.8, no.5 (2015).

[12] N. Dalal and B. Triggs. "Histograms of Oriented Gradients for Human Detection",CVPR, (2005).

[13] T. Kobayashi, A. Hidaka and T. Kurita, "Selection of Histograms of Oriented Gradients Features for Pedestrian Detection", (2008).