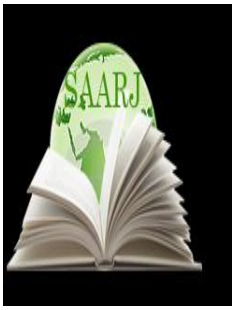
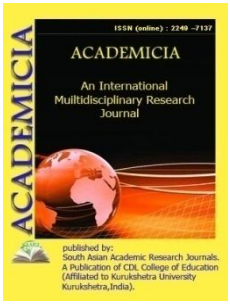


ACADEMICIA
An International Multidisciplinary Research Journal
(Double Blind Refereed & Peer Reviewed Journal)



DOI: 10.5958/2249-7137.2021.02213.8

HANDWRITTEN DIGIT RECOGNITION BASED ON MACHINE LEARNING ALGORITHM

Dr. Kashif Qureshi*

*SOEIT, Sanskriti University,
Mathura, Uttar Pradesh, INDIA
Email id: kqureshi.cse@sanskriti.edu.in

ABSTRACT

OCR is a broad field of study that focuses on pattern recognition. Advanced structures that can identify unique typefaces and have high reliability and accuracy are becoming more popular, and they support a wide variety of report codecs and image formats. Handwriting movement analysis may be utilized as a source of information for handwriting recognition. This will improve the end-to-end process's correctness while also increasing its dependability. One of the most practical problems in pattern recognition is text recognition. Digit recognition is used in a variety of applications, including postal mail sorting, bank processing, and data input. The problem stems from the possibility of creating an algorithm that can identify handwritten numbers and send the data to a scanner, tablet, or other digital device. The article discusses different machine learning-based methods to off-line handwritten digits. The most essential aim of the article is to ensure that effective and reliable methods for recognizing handwritten numbers are developed. For digit recognition, many machine learning methods were employed, including Multilayer Perceptron, Support Vector Machine (SVM), Bayes Net, and Random Forest.

KEYWORDS: *Digital processing, Digit Recognition, Machine Learning, SVM.*

1. INTRODUCTION

Intelligent picture analysis is a fascinating area of Artificial Intelligence study, and it's also critical for a variety of different open studies issues. Handwritten digit recognition is a well-studied subfield that uses mastering models to distinguish pre-segmented handwritten digits. It is one of the most important problems in data mining, machine learning, pattern recognition, and many other artificial intelligence fields[1]. Over the last decade, the most common application of

device mastering strategies has proven effective in conforming decisive structures that compete with human overall performance and achieve far more than manually written classical artificial intelligence systems used in the early days of optical recognition technology. However, not all of the exact models' capabilities have been previously examined[2].

One of the most important jobs in the area of a digit machine is digit identification. In pattern recognition, specialized location sampling methods are utilized to identify such regions. The task of handwritten digit identification is complicated by the large number of different writing styles. To enhance the performance of a handwritten character recognition device, robust function extraction is essential. Because of its usefulness in a variety of areas, handwritten digit recognition has attracted a lot of attention in the world of pattern recognition devices. In the future, a human recognition system may be used to help create a paperless environment by scanning and processing existing paper documents[3].

In recent years, text detection and recognition has become a necessary annoyance. This trend has been fueled by advancements in the fields of computer vision and machine learning, as well as an increase in applications based on textual content identification and recognition. On a global scale, several workshops and seminars, such as the International Conference on Document Analysis and Recognition (ICDAR), are being held, providing a comparable boost to trends in the field of textual content processing from images. Text detection and identification from video subtitles and web pages are also gaining popularity. Text recognition and extraction from real-world situations and images have gotten a lot of attention. There are a variety of optical character recognition methods available as well. Text detection and identification is still a problem that hasn't been addressed completely. Text segmentation and extraction from natural situations are still challenging to do.

In the field of handwritten character identification, pattern recognition and image processing play an important role. The findings in cover a variety of function extraction strategy classifications, including structural function-based methods, statistical function-based strategies, and international transformation approaches[4]. Statistical methods are used to determine how records are chosen. It relies on the statistics of the image's statistical distribution of pixels. The article described an offline handwritten digit recognition device based on SVM[5].

1. Image Acquisition

The acquisition of images is the initial stage, as illustrated in Figure 1. The experiment is carried out on images of the shapes taken using a scanner or a camera on a mobile phone. Following scanning, the photo is subjected to pre-processing to remove historical noise and binarization to produce pixels in 0s and 1s.

2. Pre Processing

On color, grey-degree, or binary record pictures with textual content and/or images, pre-processing techniques are used. Size normalization is one of the stages in pre-processing, when interpolation is performed for the same old sized image. Binarization is the process of converting a grayscale image to a binary image via the use of thresholding. The pixels in a binary picture have both a 0 and a 1 value. White pixels make up the backdrop, while black pixels make up the foreground. The erosion and dilation techniques are used to smooth the edges of things. Dilation is used to monitor opening clean away erosion. To remove tiny objects from the foreground,

filtering is used. The final filter to dilate is the closing filter, which is accompanied by an erosion mechanism. It fills in tiny gaps in the foreground and moves small sections of background into the forefront. Because they remove the margins of characters, edge detection using morphological gradient operators is used in area detection.

3. Segmentation

To distinguish text from backdrop and retrieve bounded textual information from an image, the segmentation technique is employed. Integrated methods that concentrate on phrase matching/recognition often combine or replace sophisticated segmentation with recognition, while stepwise approaches utilize segmentation to extract characters that may then be fed to recognition algorithms. Character segmentation may be handled using horizontal histogram profiles (for line segmentation), vertical histogram profiles (for word segmentation), and associated problem analysis. The picture is broken down into its constituent parts. Handwritten sentences are scanned to find valid segmentation points between characters using a basic heuristic segmentation method. The segmentation is based on the discovery of minima or arcs between letters, which is not uncommon in handwritten cursive writing. Lines, words, and characters are separated from the pre-processed image.

To divide the textual content lines, line segmentation is used. Word Segmentation calculates the distance between words, whereas Character Recognition calculates the space between characters. For character recognition, a variety of neural network methods, including SVM (Support vector machine), are employed. Essentially, the output of both sets of rules produces nearly equal accuracy, and they may be used to educate and compare accuracy.

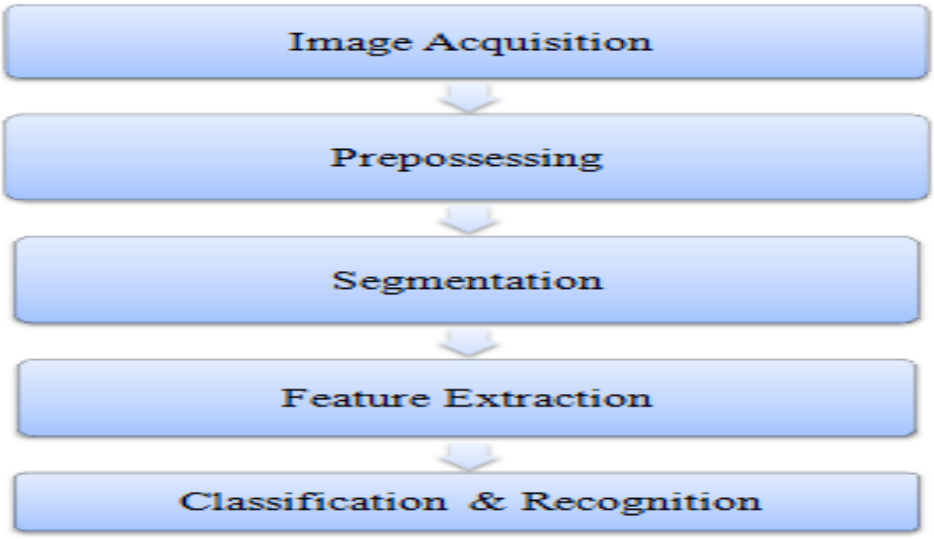


Figure 1: Illustrates the Processes involved for Digit recognition

4. Feature Extraction

Character Geometry-based Feature Extraction extracts the kind of line that creates a particular character. The picture is divided into windows of equal length, and a feature is applied to each window's starters—the pixel at the intersection of two windows has many neighbors. Gradient Feature Extraction. The Sobel operator is used to extract variations in depth of tiny neighbor

pixels. Each pixel's gradient vector is obtained, and the gradient picture is deconstructed using chain code. Crossings and Distances is a common statistical function that counts the number of times a contour is crossed by a line segment in a certain direction. The character's body is divided into a number of regions in different directions, and then characteristics from each are retrieved. Projections in which the characters are represented by projecting pixel gray values onto traces in a variety of ways. This depiction converts a two-dimensional photo into a one-dimensional signal that may be used to represent the character image in Figure 2.

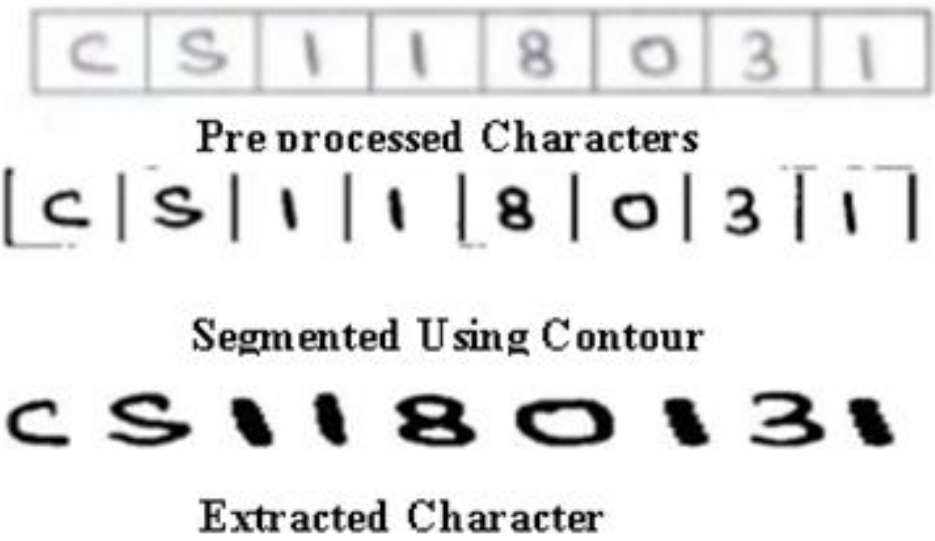


Figure 2: Illustrates the extraction of character processes

Border Transition Technique (BTT): In the border transition method, each character is assumed to be oriented vertically. Each character is divided into four identical quadrants. In each department, the scanning and computation of zero-to-one transitions in vertical and horizontal instructions takes place.

Graph Matching Technique: In a graph matching method, a structural feature of the man or woman is used. Altering the typeface or rotation is a helpful approach. These three characteristics are specified. First, a basic one-pixel endpoint is connected, which contains location statistics. Then a department factor with feature information linked to the branch is assigned to more than three pixels.

5. Classification

Non-text areas next to text regions may be included as positives after text detection and localization degree output. Using class algorithms, the classification step validates text areas and removes non-text parts. This level is also known as verification. Both supervised and unsupervised classification methods exist. The characteristics of text, such as color, size, texture, and so on, are elaborated by supervised algorithms. Unsupervised algorithms don't have any previous knowledge of text characteristics. Type algorithms that are supervised need to be educated before they can be used in class. These algorithms go through a procedure that enables them to extract functions from the text to categorize and utilize them in the category section.

Edge proximity restrictions, as well as area, top, and width limitations on the block obtained at the detection level, are utilized[6].

Unsupervised categorisation algorithms are no longer processed. In contrast to supervised categories, they only extract text functions during the class segment, and they reuse capabilities extracted in prior types for future types. This is similar to adaptive mastery. Wavelet transform, which provides successive approximation through low-bypass clean out and edge and feature information from an excessive-bypass filter. It splits an image into 1616 sized windows and extracts 36 characteristics from each windowed picture for unsupervised categorization of textual and non-textual information. Features like variation of stroke width, difference between contrast of text and historical past, and aspect ratio of bounding box are utilized in the method to create connected components that may be classified utilizing a k-means based entirely classifier. These are global functions taken from a picture that has been split into sub-areas[7].

An effective information modelling method capable of capturing and representing complicated input/output connections is an artificial neural network approach to be used. The desire to create an artificial machine that could do "intelligent" activities comparable to those performed by the human brain motivated the creation of neural community technology. Neural networks are similar to the human brain in that the neural network's information is stored in synaptic weights, which are the intensities of interneuron connections. Because the relative roles of the fields may vary in form, the characteristics of an image are extracted separately to educate the community. This allows it to split the picture into distinct blocks and monitor the neural community on each block.

5.1.Support Vector Machine (SVM)

Support vector machines (SVM) are supervised learning models with related algorithms that analyze data for classification and regression assessment in machine learning (ML). An SVM method creates a version that assigns fresh instances to one of two categories, making it a non-probabilistic binary linear classifier, given a collection of training examples tagged as belonging to at least one or the opposite of two classes (although techniques consisting of Platt scaling exist to apply SVM in a probabilistic type setting). An SVM version is a representation of the instances as points in space, mapped in such a way that the examples of the different classes are separated by as large a distance as feasible. New instances are mapped into the same space and assigned to a category depending on whatever feature of the space they fall into.

SVM (Support Vector Machine) is a kind of supervised machine learning method that aims to categorize data components by maximizing the margin between training in a high-dimensional space[8]. SVM is a representation of instances as components in space, mapped with the help of a reasonable gap as large as feasible between the examples of different courses. Following that, additional instances are mapped into the same space and assigned to a category depending on where they lie on the map. The technique is developed through a "training" phase in which training statistics are used to increase an effective set of rules for discriminating between companies previously described with the help of the operator (e.g. patients vs. controls), and a "testing" section in which the algorithm is used to blind-predict which organization a new belief belongs to. It also generates adequate search area for the right class of future data parameters and delivers a fully accurate class overall performance over the training information[9]. As a result, it

always guarantees a set of parameter combinations based on a realistic subset of the data. In SVM, it is usually preferable to scale the statistics, since this greatly improves the outcomes.

5.2.Multilayer Perceptions

The handwritten digits are classified using a neural community-based fully classifier known as Multilayer perception (MLP). Enter layer, hidden layer, and output layer are the three special layers of a multilayer perceptron. Every node in a layer is linked to all other nodes in the next layer, and each layer may contain a certain number of nodes, also known as neurons. As a result, it's often referred to as a feed forward community. The number of nodes in the input layer is determined by the number of characteristics in the dataset. The number of nodes in the output layer is determined by the dataset's broad range of evident classifications. It's difficult to identify the most practical range of hidden layers or the most practical number of nodes in a hidden layer for a certain annoyance. But, in general, these figures are arrived at via trial and error. The connection between nodes in a multilayer perceptron is based on a weight. It essentially learns the proper weight adjustment that corresponds to each link throughout the educational system. It employs a supervised learning technique known as Back propagation set of rules for the goal of acquiring information[10].

5.2.1. Random Forest (RF) Algorithm

Random forest is an ensemble of unpruned regression or category trees generated from bootstrap samples of the training data using the tree imitation technique's random characteristic selection. The forecast is produced by combining the ensemble's predictions for each category using superiority balloting. It generates a generalization error charge and is more noise-resistant. RF, like maximum classifiers, may be plagued by the curse of learning from a highly unbalanced training data set. It will tend to concentrate more on the prediction performance of the bulk class, which consistently results in poor accuracy for the minority class, since it was designed to reduce the overall mistake price.

5.2.2. Bayes Net

Bayesian networks are a powerful probabilistic depiction, and its use to categorization has gotten a lot of attention. It represents the states of a simulated portion of the world and explains how those states are linked via probability. The Bayesian community is a graphical representation of the probability connections between various variables. The graphical form offers many advantages for records assessment when used in combination with statistical methods. One, since the version encodes all variables' dependencies, it can easily manage circumstances when certain data entries are absent. Two, a Bayesian community may be used to learn causal connections and, as a result, can be utilized to acquire knowledge in a problem area and anticipate intervention outcomes. The conditional probability of each characteristic given the class label is learned from education statistics by this classifier. The comparison of different methods is shown in Table 1.

TABLE 1: COMPARISON OF CORRECT AND INCORRECT INSTANCES OF VARIOUS ALGORITHMS

ALGORITHM	CORRECT INSTANCES	INCORRECT INSTANCES
MULTILAYER PERCEPTRON	90.47	9.53
SUPPORT VECTOR MACHINE	87.87	12.13
RANDOM FOREST	85.77	14.23
BAYES NET	84.30	15.70

2. DISCUSSION &CONCLUSION

It is hypothesized that image processing expertise coupled with machine learning may be utilized to extend a system to recognize different types of handwritten writing. This solution may be used to automate data input systems, among other things. This may be advanced for a variety of benefits and applications with appropriate training and a huge quantity of data sets. The purpose of this article is to find a representation of handwritten digits that may be used to recognize them. Various remarkable machine learning techniques have been used and applied in this work for the recognition of handwritten digits. The most difficult issue in any recognition system is dealing with feature extraction and proper type strategies. The suggested method, which uses machine learning algorithms, attempts to deal with each of the components accurately and quickly. The recognition technique using Multilayer Perceptron has the best overall accuracy of 90.47 percent.

REFERENCES

1. J. Watada and W. Pedrycz, “A Fuzzy Regression Approach to Acquisition of Linguistic Rules,” in *Handbook of Granular Computing*, 2008.
2. A. K. Seewald, “On the Brittleness of Handwritten Digit Recognition Models,” *ISRN Mach. Vis.*, 2012, doi: 10.5402/2012/834127.
3. N. Das, R. Sarkar, S. Basu, M. Kundu, M. Nasipuri, and D. K. Basu, “A genetic algorithm based region sampling for selection of local features in handwritten digit recognition application,” *Appl. Soft Comput. J.*, 2012, doi: 10.1016/j.asoc.2011.11.030.
4. E. Bhatia, “Optical character recognition techniques: a review,” *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, 2014.
5. R. F. P. Neves, A. N. G. Lopes Filho, C. A. B. Mello, and C. Zanchettin, “A SVM based off-line handwritten digit recognizer,” 2011, doi: 10.1109/ICSMC.2011.6083734.
6. A. Mittal, P. P. Roy, P. Singh, and B. Raman, “Rotation and script independent text detection from video frames using sub pixel mapping,” *J. Vis. Commun. Image Represent.*, 2017, doi: 10.1016/j.jvcir.2017.03.002.
7. A. Mosleh, N. Bouguila, and A. Ben Hamza, “Image text detection using a bandlet-based

edge detector and stroke width transform," 2012, doi: 10.5244/C.26.63.

8. F. Pereira, T. Mitchell, and M. Botvinick, "Machine learning classifiers and fMRI: a tutorial overview.," *Neuroimage*, 2009, doi: 10.1016/j.neuroimage.2008.11.007.

9. G. Orrù, W. Pettersson-Yeo, A. F. Marquand, G. Sartori, and A. Mechelli, "Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review," *Neuroscience and Biobehavioral Reviews*. 2012, doi: 10.1016/j.neubiorev.2012.01.004.

10. R. Kruse, C. Borgelt, F. Klawonn, C. Moewes, M. Steinbrecher, and P. Held, "Multi-Layer Perceptrons," 2013.