Date	**15 Novmeber 2022**
Team ID	**PNT2022TMID46276**
Project	**A Novel Method For Handwritten Digit Recognition system**

A NOVEL METHOD FOR HANDWRITTEN DIGIT RECOGNITION SYSTEM

ABSTRACT

The main purpose of this thesis is to build an automatic handwritten digit recognitionmethod for the recognition of connected handwritten digit strings. To accomplish the recognition task,first, the digits were segmented into individual digits. Then, a digit recognition module is employed to classify each segmented digit completing the handwritten digit string recognition task. In this study, different machine learning methods, which are SVM, ANN and CNN architectures are used to achieve high performance on the digit string recognition problem. In these methods, images of digit strings are trained with the SVM, ANN and CNN model with HOG feature vectors and Deep learning methods structure by sliding a fixed size window through the images labeling each sub-image as a part of a digit or not. After the completion of the segmentation, to achieve the complete recognition of handwritten digits.

Motivation and Objective

This thesis is conducted by using Machine learning concepts. Before going deep into the topic, we must know about some of these concepts. Machine Learning is a method which trains the machine to do the job by itself without any human interaction. At a high level, machine learning is the process of teaching a computer system on how to make accurate predictions when fed the data. Those predictions will be the output. There are many sub-branches in machine learning like Neural Networking, Deep Learning, etc[1]. Among these, Deep Learning is considered to be the most popular sub-branch of Machine Learning.Initially, the idea of Machine Learning has come into existence during the 1950s, with the definition of perception[2]. It is the first machine which was capable of sensing & learning. Further, there was multilayer perceptron in the 1980s, with a limited number of hidden layers. However, the concept of perceptron was not in usage because of its very limited learning capability. After many years, in

the early 2000s, a new concept called Neural Networks came into existence with many hidden layers[3]. After the emergence of neural networks, many machine learning concepts like deep learning came into

force with multiple levels of representation. Because of these multiple levels of representationphenomenon, it has become easy to learn and recognize machines. The human brain

is considered as a reference to build deep learning concepts, as the human brain similarly processes information in multiple layers[4].A human can easily solve and recognize any problem, but this is not the same in the case of a machine. Many techniques or methods should be implemented to work as a human. Apart from all the advancements that have been made in this area, there is still a significant research gap that needs to be filled. Consider, for example, online handwriting recognition vs offline recognition [5]. In online handwriting recognition of letters, an on-time compilation of letters is performed while writing because stroke information is captured dynamically[5]. Whereas, in offline recognition, the letters aren't captured dynamically. Online handwriting recognition is more accurate when compared to offline handwriting recognition because of the lack of information[6]. Therefore, there can be researchdone in this area to improve offline handwriting recognition. The main task in offline handwriting recognition is to recognize the character of words.

Problem Definition and Approach

Here, the HOG feature extraction method is used for both SVM and ANN to derive accuracy and performance. The deep learning method is used for CNN to derive accuracy and performance. The results obtained from the above methods are compared and method with high accuracy and high performance is considered as the best method for handwritten digits[14]. There is a huge number of studies conducted in the field of handwritten digits. Out of them, the parameters which are considered here i.e. SVM, ANN, CNN are the most popular ones. Usually, segmentation and classification phases are the most challenging and play a vital role in the handwritten digit recognition process. Because in segmentation the image is broken into multiple images, each described as an individual digit and proper classification is done to every individual digit [15, 80]. To overcome the challenges faced in segmentation and classification, some rules are implemented to increase the accuracy and performance of both segmentation and classification[16]. The following rules are implemented in segmentation and classification.

- **Water reservoir concept is used for segmentation.**
- **SVM, ANN, CNN concepts are used in classification.**

By following the above rules, segmentation and classification are achieved successfully. Water reservoir concept is considered for segmentation because it is with high performance and accuracy [17]. In the same way SVM, ANN, CNN is considered as a suitable method for classification.

Background

Prior to the experimental setup and classification of algorithms, one should have a clear knowledgeabout the concepts which are going to be used. A literature review is done to get a clear picture of the concepts or algorithms used. In this research, two types of research methods are selected. First is Literature Review and the second is an experiment. To answer the RQ1, Literature Review is used and to answer RQ2, RQ3, RQ4 experimentation method is used.

Machine Learning

According to Arthur Samuel, "Machine learning is a subfield of computer science which gives computers the ability to learn without being explicitly programmed"[18]. This study helps in predicting and learning from the data imported with the help of algorithms implemented. Machinelearning is used where there is difficulty in programming tasks instead machine learning algorithmsare used to achieve the task. Some of these tasks include Identity Fraud Detection, computer vision,population Growth Prediction, email filtering, Weather forecasting, OCR (optical characterrecognition), Diagnostics, real-timedecisions etc.[18].

Machine learning concepts are classified into three categories:

- **Supervised Learning**
- **Unsupervised Learning**
- **Reinforcement Learning**

Supervised Learning:

Consider, a dataset is given as input and assumptions can be made on the output data how it looks like.In supervised learning, there's a relationship between the input data and the output data. The outputcan be predicted with the input given[18] [19].

Unsupervised Learning:

Unsupervised learning is an approach where the algorithm has to identify the hidden patterns in thegiven input. So, the algorithm works without any guidance as the input data is not labeled or classified[18].

Reinforcement Learning:

Reinforcement learning is a suitable action to maximize reward in a particular situation. It is to find the best possible behavior or path it should take in a specific situation[18].

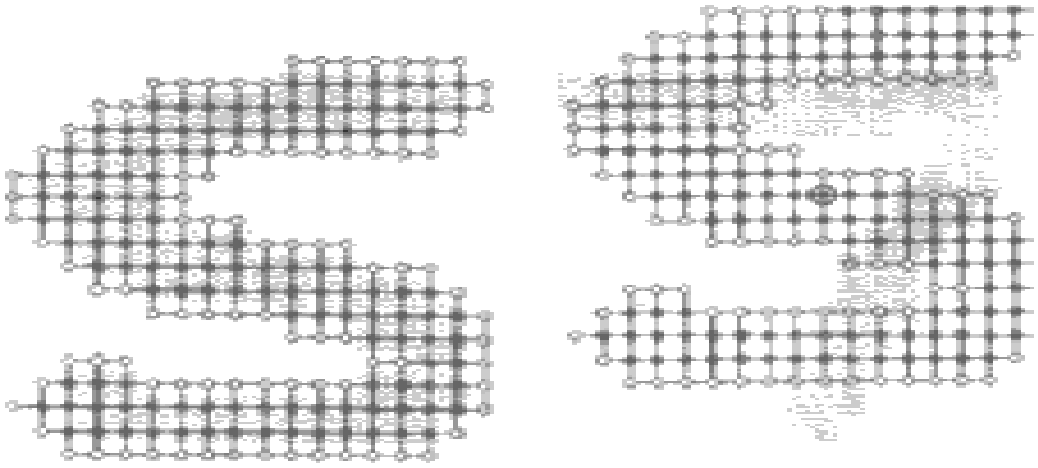


Figure 4.1 Numeral ―5‟ is mapping to numeral ―7‟.

Deep Learning

According to Arthur Andrew Ng, “Deep Learning is a superpower. With it, you can make a computer see, synthesize novel art, translate languages, render a medical diagnosis, or build pieces of a car that can drive itself. If that isn’t a superpower, I don’t know what is”[88]. Deep learning is a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms[89]. Learning can be supervised or unsupervised[90][91]. It is a set of algorithms in machine learning to learn multiple levels of representation, corresponding to different layers of abstraction that help to make sense of data [20]. Many layers are used to compute nonlinear functions with highly complex data. Each layer gets its input from a preceding layer, then it computes and transforms the data and sends it to the further layers. Each layer in a network consists of

neurons and has various modes of connections to other neurons in the same layer as well as to those of other layers depending on the type of network [3], [21]. The whole idea of deep learning is using brain simulations, helping to make learning algorithms more efficient to use and revolutionary advances in machine learning and Artificial Intelligence [4]. Nowadays deep learning gets more attention with development of modern technologies and easy execute it.

Histogram of Oriented Gradient (HOG)

HOG was proposed by Dalal and Triggs[22] for human body detection but it is considered as one of the most successful and popular used descriptors in recognition and computer vision[23]. It divides the input image into small square cells and then computes the histogram of gradient directions or edges of the image based on the differences. To improve the accuracy, local histograms have been normalized based on the contrast and it is the reason that HOG is stable on illumination variation.

It is a fast descriptor when compared to any other descriptor due to simple computations, it has been also shown that HOG is a successful descriptor for detection[22], [24].

2.2. Literatur2e Review

"Literature review (also referred to as a systematic review). A form of secondary study that uses a well-defined methodology to identify, analyze and interpret all available evidence related to a specific research question in a way that is unbiased and (to a degree) repeatable"[34]. The motivation behind adopting the literature review is to gain knowledge towards the data sets and the implementation of different types of classifiers to recognize the handwritten digits. A systematic literature review does not opt for this research as the results gathered through this were not used as the results. Once the required data has been obtained from the literature review, then data analysis is performed[35]. Narrative synthesis is adopted as our data analysis method for our literature review. During the
literature review, the data collected through the articles were gathered together and they summarized in a paragraph[36]. The results gathered through this data analysis were documented and these were used for the experimental research method. While conducting the literature review, towards recognition of handwritten digits a critical analysis is taken for the methods used solving this problem. In order to search relevant resources, the following steps have been and clearly overview was mentioned in the
Methods section(Chapter 3).

Method

For this research, two types of research methods have been selected and they were:

- **Literature review**
- **Experiment**

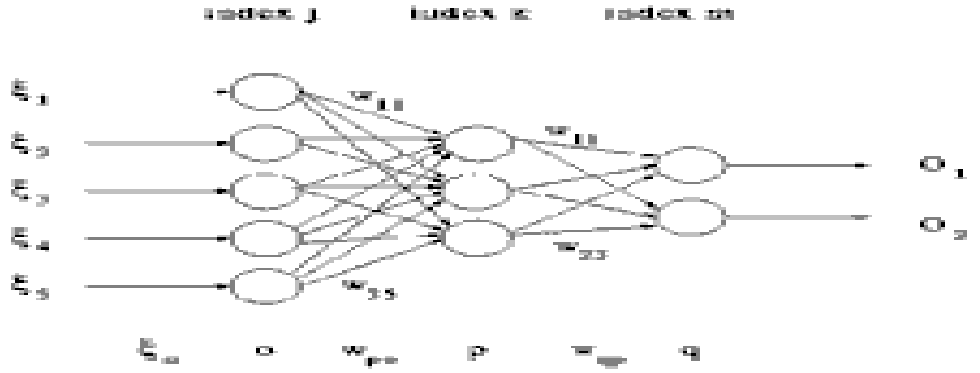
Literature review: Initially, the literature review has been conducted to answer RQ1 to know which type of data is required to train and test the machine learning methods. The motivation behind adopting the literature review is to gain knowledge towards the data sets in the machine learning and find about different types of machine learning methods that can be adopted while training the data set. A simple literature review was performed to gain knowledge about different datasets that can be used for training and testing the data. The author also gained knowledge of different processes of data preprocessing, gmentation and various machine learning methods to be adopted in the study. Steps followed while performing the literature review:

Tools Used

This study is to identify AHDR on documented images with the use of machine learning methods[19]. At first, we need to construct a suitable model or method for training and testing[54]. The program able to extract characters one by one to get target output for training & testing model. The implementation and the experimentation of the algorithm had been carried out by using Python and supported with the usage of Graphical User Interface (GUI). We have used the Python 3.5 version, TensorFlow backend, OpenCV, sklearn, Kera's it consists of the statistics and machine learning Toolbox which is used for training and testing the data using for different classifiers.

3.3 Dataset Used

The dataset is required for the training and testing[62]. The images of data are represented in datasets and it contain colored images. The dataset contains a total of 9096 images. From the available data, we have used 70% of the images for training the classifier and rest of the 30% used for testing.



re 3.1 Feed-forward Neural Network with 3 inputs

4.1 Preprocessing

In preprocessing the dataset were colored input images with the help of the handwritten digits recognizer, first the colored image is converted into a grayscale image. Then the image is resized keeping the aspect ratio. The input images varied in vertical length and all the images must be in the same height for the segmentation and classification modules to work correctly since the segmentation module creates a sequence of images by windows of the same size and equalizing heights to the same size would be enough. Therefore, the height is fixed as h=100 and horizontal length is adjusted according to the aspect ratio of the image. From the grayscale image, thresholding is applied based on Otsu Thresholding[23]. Finally, morphological operations are applied to remove noise and step by step
process we seen in below with figures.