Automatic Handwritten Digit Recognition On Document Images Using Machine Learning Methods

Akkireddy Challa

Dept. Computer Science & Engineering
Blekinge Institute Of Technology
Se--371 79 Karlskrona, Sweden

This thesis is submitted to the Faculty of Computing at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Computer Engineering. The thesis is equivalent to 20 weeks of full-time studies. The authors declare that they are the sole authors of this thesis and that they have not used any sources other than those listed in the bibliography and identified as references. They further declare that they have not submitted this thesis at any other institution to obtain a degree.

Contact Information:
Author(S):
Akkireddy Challa
E-mail: akch17@Student.Bth.Se

University Advisor:
Dr. Huseyin Kusetogullari,
Senior Lecturer,
Dept. of Computer Sci. & Eng.

ABSTRACT

Context: The main purpose of this thesis is to build an automatic handwritten digit recognition method 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.

Objective: The main purpose of this thesis is to find out the recognition performance of the methods. In order to analyze the performance of the methods, data is needed to be used for training using machine learning methods. Then digit data is tested on the desired machine learning technique. In this thesis, the following methods are performed:

- Implementation of HOG Feature extraction method with SVM
- Implementation of HOG Feature extraction method with ANN
- Implementation of Deep Learning methods with CNN

Methods: This research will be carried out using two methods. The first research method is the ¨Literature Review¨ and the second ¨Experiment¨. Initially, a literature review is conducted to get a clear knowledge on the algorithms and techniques which will be used to answer the first research question i.e., to know which type of data is required for the machine learning methods and the data analysis is performed. Later on, with the knowledge of RQ1, Experimentation is conducted to answer the RQ2, RQ3, RQ4. Quantitative data is used to perform the experimentation because qualitative data which obtains from case-study and survey cannot be used for this experiment method as it contains non-numerical data. In this research, an experiment is conducted to find the best suitable machine learning method from the existing methods. As mentioned above in the objectives, an experiment is conducted using SVM, ANN, and CNN. By considering the results obtained from the experiment, a comparison is made on the metrics considered which results in CNN as the best method suitable for Documents Images.

Results: Compare the results for SVM, ANN with HOG Feature extraction and the CNN method by using segmented results. Based on the Experiment results it is found that SVM and ANN have some drawbacks like low accuracy and low performance in the recognition of documented images. So, the other method i.e., CNN has greater performance with high accuracy. The following are the results of the recognition rates of each method.
- SVM performance - 39%
- ANN performance - 37%
- CNN performance - 71%.

Conclusion: This research concentrates on providing an efficient method for recognition of automatic handwritten digits recognition. Here a sample training data is treated with existing machine learning and deep learning methods like SVM, ANN, and CNN. By the results obtained from the experimentation, it clearly is shown that the CNN method is much efficient with 71% performance when compared to ANN and SVM methods.

ACKNOWLEDGMENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS AND SYMBOLS

AHDR	Automatic Handwritten Digit String Recognition
HOG	Histogram of Oriented Gradients
SVM	Support Vector Machine
ANN	Artificial Neural Network
CNN	Convolutional Neural Network
SRM	Structural Risk Minimization principle
ERM	Empirical Risk Minimization
IP	Interconnection Point
BP	Base Point
CD	Connected Digits
OD	Overlapped Digits
DJD	Disjoint Digits

CONTENTS

CHAPTER 1

1.1 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 representation phenomenon, 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 research done in this area to improve offline handwriting recognition.

The main task in offline handwriting recognition is to recognize the character of words. There are different approaches for recognizing the characters of a word in offline handwriting. The following are the processing steps:

```
┌─────────────────┐
│   Input Data    │
└─────────────────┘
          │
          ▼
┌─────────────────┐
│  Preprocessing  │
└─────────────────┘
          │
          ▼
┌─────────────────┐
│  Segmentation   │
└─────────────────┘
          │
          ▼
┌─────────────────┐
│    Feature      │
│   generation    │
└─────────────────┘
          │
          ▼
┌─────────────────┐
│ Classification  │
└─────────────────┘
          │
          ▼
┌─────────────────┐
│   Recognition   │
└─────────────────┘
```
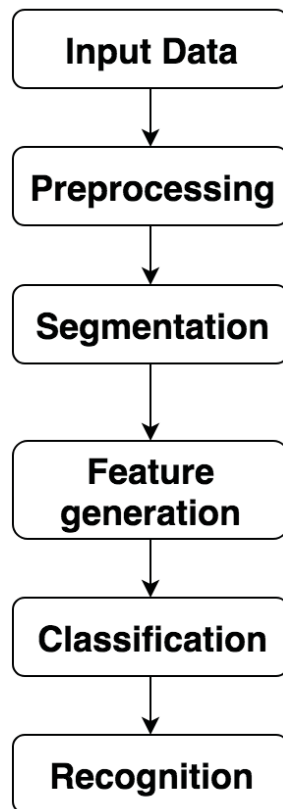
Figure 1.1. Steps of the typical character recognition system

Initially, a dataset is given as input. This is followed by preprocessing, where an image is subjected to various operations like noise reduction, document skew correction, slant correction, normalization, smoothing, and skeletonization[7]. The result of this preprocessing can be given as an input to feature generation. Segmentation of an image is done to isolate the characters of an image into different sub-images. Each sub-image is considered as one individual character[8].

The next phase is feature generating, in which various extraction techniques are used to represent an image as a vector feature in the feature generator[9]. To keep it clear without any noise, an algorithm is implemented to reduce the size of the image, which in turn reduces the noise in that image. The feature generation is followed by classification. There are a large number of classifiers, which reduces the performance of classification at each step. There are also classifiers for recognition including the statistical, the structural, the stochastic classifiers and finishing on a combination of classifiers. At each step, selecting the appropriate parameters could affect the final classification performance, which results in the complete recognition of an image[10][11].

Our main purpose is to find out the rules to be used in the automatic handwritten digit recognition for document images using machine learning methods. The field studied in this thesis work is to recognize the corrupted handwritten digits. Since handwritten digits recognition contains a wide variety of options the way the digits are written, the main concern of this study is to recognize handwritten digits. Although the performance of machine learning algorithms on handwritten digits are very good when the digits are segmented well, the segmentation performance of the existing algorithms is poor, which in turn reduces the recognition performance when digit strings are considered. Therefore, reliable handwritten digit string recognition methods are necessary in order to increase the recognition rate of handwritten digit strings[12]. The success of water reservoir method in handwritten digit recognition

motivated us to use them also to decide vertical cuts for segmentation of digit strings by training the regions between the digits[13].

1.2 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.

1.3 Research Questions

The following research questions were posed:

RQ1: What data is used to train and test the machine learning methods?
> **Motivation:** The motivation behind adopting the RQ1 is to gain knowledge towards the data sets in the machine learning and find about different types of machine learning methods can be adopted while training the data set for solving the problem for the connected digit, overlapping and disjunction(non-linear).

RQ2: What are the best parameters of machine learning methods to recognize digits?
> **Motivation:** The motivation for framing this research question is to find out the best parameters suitable for the machine learning methods to recognize digits.

RQ3: What is the recognition performance of machine learning methods for the handwritten digits which cannot be segmented into document images?
> **Motivation:** The motivation for framing this research question is to estimate the recognition performance of the Machine Learning methods.

RQ4: Which machine learning method provides the best recognition performance for digit recognition for those which cannot be segmented?
> **Motivation:** The motivation for framing this research question is to evaluate the best recognition performance of the machine learning method from SVM, ANN, and CNN.

1.4 Outline of the Thesis

In chapter 1, the purpose of the thesis with motivation and contributions to the field of study has already been introduced.

Chapter 2 describes Background and Related work. Describes the Literature survey and background information about the concepts used. This chapter includes an overall review of the literature about handwritten digits recognition and the techniques used in the steps of the handwritten digits recognition. The techniques used for preprocessing, segmentation and recognition parts are given respectively.

Chapter 3 includes the details of the Methods and the datasets used.

Chapter 4 includes the details of the Experimental results and Analysis. The experiments were done in order to observe the performance and accuracy of the particular method. Then the results of the experiments are presented.

Chapter 5 consists of Discussion of the proposed methods, expected outcomes and contribution or overcome all the work.

Chapter 6 provides the conclusion and future work of this thesis with a brief summary. The experimental results are discussed with some theoretical background and possible future works are explained.

CHAPTER 2

BACKGROUND AND RELATED WORK

2.1 Background

Prior to the experimental setup and classification of algorithms, one should have a clear knowledge about 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.

2.1.1 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. Machine learning is used where there is difficulty in programming tasks instead machine learning algorithms are used to achieve the task. Some of these tasks include Identity Fraud Detection, computer vision, population Growth Prediction, email filtering, Weather forecasting, OCR (optical character recognition), Diagnostics, real-time decisions 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 output can 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 the given 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].

2.1.2 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.

2.1.3 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].

2.1.5 Support Vector Machine (SVM)

Support Vector Machine (SVM) was introduced by introduced by Boser, Guyon, and Vapnik in 1992[25]. It is a core part of machine learning methods.

A support vector machine (SVM) is a supervised learning algorithm that can be used for binary classification or regression[18] and even it belongs to the family of a linear classifier. In other words, SVM is a popular application and it is used in natural language processing, speech recognition, image recognition, and computer vision. It constructs an optimal hyperplane for the decisions and the two margins separate action between the two classes in the data is maximized[26]. It refers to a small subset of the training observation and used to support the optimal location of the decisions[27]. In the process regression and classification prediction tools and take care that the algorithm does not lead to overfitting. While implementing the algorithm on the training data and the results are observed on the test data with much efficiency[27].
The main reason for the better performance of the SVM than any other algorithms[28]. Previously, SVM is developed to overcome the classification problems now recent studies are made on the use of SVM to overcome the regression issues[29].

2.1.6 Artificial Neural Network (ANN)

The artificial neural network is basically a mesh of a large number of interconnected cells[30]. The arrangement of cells is such that each cell receives an input and drives an output for subsequent cells. Each cell has a pre-defined. The diagram below is a block diagram that depicts the structure and workflow of a created Artificial Neural Network. The neurons are interconnected with each other in a serial manner[30]. The network consists of a number of hidden layers depending upon the resolution of comparison of inputs with the dataset[3].

2.1.7 Conventional Neural Network (CNN)

Convolutional Neural Network[85] is a family of multi-layer neural networks it is particularly designed for use on two-dimensional data, such as images and videos[86]. Basically, it is influenced

by earlier work in time-delay neural networks, which reduce learning computation requirements by sharing weights in a temporal dimension and are intended for speech and time-series processing [87]. It has many hierarchy layers to train in a robust manner. This architecture that leverages spatial and temporal relationships to reduce the number of parameters which must be learned and thus improves upon general feed-forward backpropagation training. It is proposed as a deep learning framework that is motivated by minimal data preprocessing requirements[31]. In CNN, small portions of the image are treated as inputs to the lowest layer of the hierarchical structure. The fully connected layers form a network in this first layer is named as "input layer" and the last layer are named as "output layer" and between these two layers and remaining all are known as "hidden layers". In the hidden layer, the inputs were passed and the output layer calculates the class probabilities for the classification. A regular neural network performs high computation if the size of data is increased or if the number of layers are increased. in these many parameters were calculated without overfitting results for accuracy[32]. On the other hand, CNN not fully connected in all levels of layers. The neurons in these layer connected to a small region. This encourages the local spatial relationship in the data and the hierarchy features to increase the abstraction from low-level to high-level when multiple were layers are stacked[28]. In these first layers can see only a small portion of the input data and the last layers can see the whole of the input data and draw conclusions from it. In CNN there are three types of layers were used in a convolutional network:

      1) Convolutional layer
      2) Pooling layer
      3) Fully-connected layer

Convolutional layer: In this layer few parameters like a number of filters, size of filters, stride, etc. Small window filter slid along with the dimensions of input data and performs dot products between the values stored in the filter and the input data points[33].

Pooling layer: This layer reduces the dimensionality of the input data which reduces the computations, number of parameters and therefore reduces overfitting. Typically, the pooling layer is inserted between convolutional layers. It discards the activations of previous layers and hence forcing the next convolutional layers to learn from a limited variety of data [33].

Fully-connected layer: In this layer neurons that are connected to all neurons of the previous layer as explained it[33].

2.2. Literature 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).
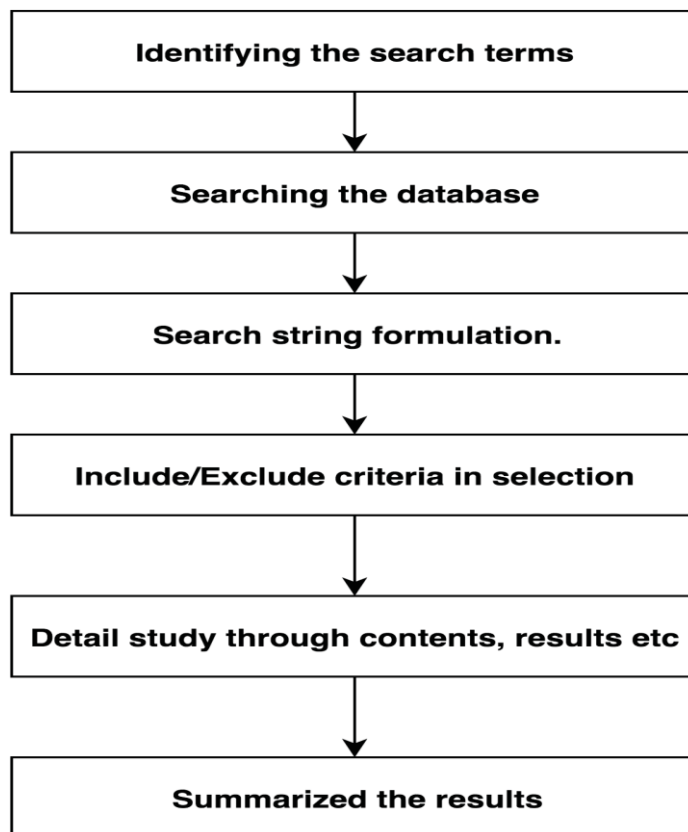
Figure 2.2 Steps of the literature review

Automatic handwritten digits extracts from images is a crucial role for creating documents and processing the systems[1]. The main purpose is to find out the rules to be used in the AHDR for document images using machine learning methods. The field studied in this work is to recognize the corrupted handwritten digits and increase the reliability of the result of the recognition process and to speed up the collecting training and test data from handwritten digit strings. The overall recognition process consists of preprocessing, segmentation, classification and finally recognition of given input data.
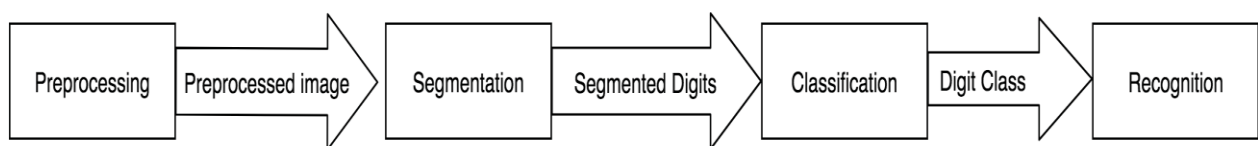
Figure 2.2 Modules of Handwritten Digit Recognition

2.2.1 Pre-Processing Method

Pre-processing and feature extraction are very important steps in automatic handwritten digit recognition(AHDR) for documented images[37]. The basic step is to improve the discriminating nature of the pixels or raw features being computed from input images. It has been taken a lot of work improving the preprocessing[38]. One of the problems in the recognition process is skew/slant detection and correction in the documented images, which introduces challenges for segmentation[39].

However, we can generally categorize the tasks into noise reduction, normalization, smoothing and skeletonization as the Fig.

Figure 2.2.1 Pre-Processing Method

At first, the Input image is in RGB format and huge saturation are discarded and intensity is used to obtain a grayscale image. Then the grayscale image should be turned into a binary form[40]. There are two main categories of binarization methods [77].

      a.   Global thresholding.
      b.   Local thresholding.

Global thresholding: In this algorithm uses a single threshold value for overall images. Twenty global thresholding techniques were compared by Sahoo et al. [41] based on uniformity and shape measures.

Local thresholding: In this algorithm uses different types of threshold values for each pixel using their spatial information[42] [73]. According to the comparison, Otsu's thresholding method[38] gave the best performance. There are various local thresholding techniques as well. Trier and Jain[43], Sezgin and Sankur[44] surveyed and discussed these thresholding techniques.

After binarization, the skew correction could be done in order to correct the angle of the digits and the X-axis. Knerr et al[45] determined the angle by computing the pixel densities between ±5 degrees with the help of horizontal guidelines. With the help of the histogram, we created the pixel densities, histogram with the longest peak is chosen as the angle of the text in the image. The problem with this method relies on the fact of horizontal guideline to realize the actual angle[46].

2.2.1.1 Noise reduction

Noise reduction is the process of removing noise from an image[47] [76]. There are many techniques to reduce noise. Basically, the noise filtering function is used to remove the noises and diminish spurious points in the image. Even we take an example, Symmetric Gaussian filter function is used for smoothing equally in all directions[48]. An alternative approach is to use Morphological operations, which are basically neighborhood operations. It can perform on the input images using the structure element[49].

2.2.1.2 Normalization

Normalization method is the most popular method used in character recognition. Because to reduce all types of variations and to obtain standardized data and it also gives excessive shape[13]. The characters for normalizing methods in the following:

- **Skew normalization:** It is used due to different types of writing style; the skew can hurt the effectiveness of recognition and therefore it is easy to detect and correct the baseline. Various methods have been used, which are the projection profile of the image, the Hough transform or the shape of the nearest neighbor clustering[50]. After skew detection, the character or word is translated to the origin and rotated until the baseline is horizontal.
- **Slant normalization:** The character inclination typically found in cursive script is called slant. Formally, it is defined as the angle between the longest stroke in a word and the vertical direction referred to the word slant[36], [50]. Slant normalization is used to normalize all

characters to a standard form with no slant. Many methods have been proposed to detect and to correct the slant of cursive words[6]. One of the used methods is based on the center of gravity, another method uses the projection profiles and some used a variant of the Hough transform[51].

- **Size normalization:** It is used to adjust the size, position, and shape (dimension) of the character image. This step is required for reducing the shape variation between images of the class to facilitate the feature generation and improve their classification[52].

2.2.1.3 Smoothing

Smoothing operation is done to regularize the edges in the image, to remove small bits of noise and to reduce the high-frequency noise in the image[13] [80]. Furthermore, different preprocessing methods are used for the smoothing image to acquire a more accurate output image.

2.2.1.4 Skeletonization

Skeletonization is a morphological operation used for reducing the contours in a binary image to a skeletal, the connectivity of the original region is detected while destroying most of the original foreground pixels[49], [53]. These were divided into two methods and they were iterative and non-iterative

- Iterative method approaches peeling contours of process parallel or sequentially by erasing or removing the unwanted pixels in each iteration[53].
- Non-iterative approaches, the skeleton is straightforwardly extracted without examining each pixel individually. Unfortunately, these techniques are difficult to implement and slow as well. Thinning can be somewhat performed for skeletonization using methods like erosion or opening. In this mode, it is commonly used by reducing all lines to single pixel thickness[53].

2.2.2 Segmentation

Segmentation is the most challenging part in the recognition of handwritten digits process[54]. The main reasons to occur this problem is because of the size of each digit, a number of digits and the gap between the digits are unknown[55]. To overcome this problem a perfect algorithm of digit segmentation should be implemented. If there are some rules like box for each digit to if there is a gap between the strings it will be much easier to segment the handwritten digit. By using a trivial algorithm, these types of strings can be segmented by applying connected component analysis after removing noises [56].
The main goal of this research is to develop a perfect segmentation algorithm for handwritten recognition. By using touching strings, the goal can be achieved. There are five types of touching digit strings and these can be further categorized into single touching digit strings and multiple touching digit strings [57].

Category	Type	Style of touching	Examples	
Single-touching	I.	◄ ►	59	33
	II.	⋈	24	02
	III.	━ ━	23	52
	IV.	⋈	40	00
Multiple-touching	V.	⊗	78	38

Table 2.2.2 Types of touching digit strings.

Segmentation algorithms are categorized into two types, one is segmentation-then based, and another is recognition-based algorithms. Initially, segmentation-then based algorithms, segmented images are extracted where each segmented part is assumed as a single character[57]. This segmented character is given as input to the classifier. Whereas, in the recognition-based algorithm, all segmented images are made into a segmented list and the list is given as input to the classifier. Due to the classification of all the options in the segmentation list, the computational cost of recognition-based algorithms is very high. Instead of having a high computational cost they provide good results. Besides, this algorithm must classify fragments, isolated characters and connected characters[56].

Classification of recognition-based algorithms is made into two types as implicit segmentation and explicit segmentation. segmentation generates candidate characters for the recognizer. in the explicit recognition[33]. Whereas, segmentation and recognition are performed simultaneously in implicit segmentation.
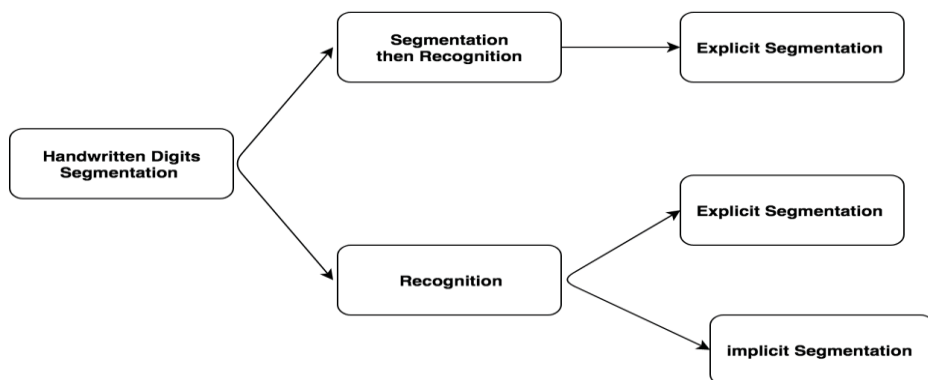
Figure 2.2.2(a) Full segmentation system for handwritten digit string recognition.

2.2.2.1. Segmentation using Water Reservoir Concept

In[17] proposed a new framework based on a concept named water reservoir for segmenting handwritten touching numerals. This concept is used to find the regions where two or multiple images combine or touch each other. Water reservoirs are formed by the reservoir points due to dropping

water from the top and bottom of an image and the locations. these waters accumulated locations are called as reservoir points. By using these reservoir points, segmentation points are decided without normalizing and thinning.

All the connected sub-images are extracted and by using reservoir point it is decided that the obtained image is connected or isolated. Segmentation is applied here, to the touching numerical. A large reservoir space is created when two digits touch each other. The cutting points are based on the reservoir. To get a precise cutting point attributes of reservoir like the center of gravity and height are considered[17].

Initially, by considering the structure of the image containing touching digits reservoirs are obtained. These reservoirs are grouped into two types as top and bottom reservoirs[24]. Touching positions of digits are decided based on the type of reservoir and base position of the reservoir. By considering the attributes like height, touching position, the center of gravity, closed loops of the reservoir the best cutting point is decided[17].
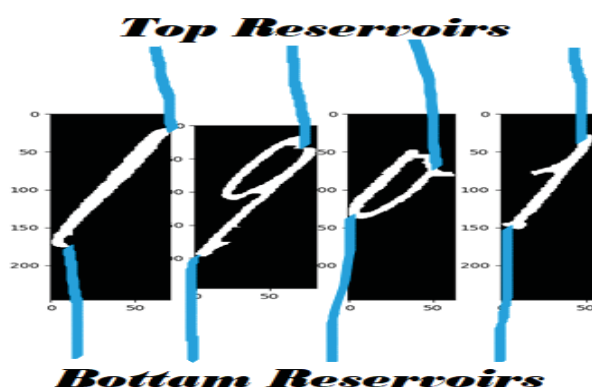

Figure 2.1.2.1 Segmentation using Water Reservoir Concept

2.2.3 Feature extraction

Feature extraction is a method of extracting features of characters from the sample image[46]. There are basically two types of feature extraction:
1. Statistical feature extraction
2. Structural feature extraction

1. **Statistical feature extraction:** In statistical feature extraction the feature vector is the combination of all the features extracted from each character and these association feature vector relative to positions of features in the character image matrix.
2. **Structural feature extraction:** In structural feature extraction extracts the morphological features of a character from the image matrix. It considers the edges, curvature, regions, etc. The functions that are used in feature extraction for indexing and labeling the dataset and it helps in the classification and recognition of handwritten digits.

2.3 Related Works:

The following are some of the terms and concepts used in this research. Our work performance of machine learning methods by using a support vector machine, artificial neural network and convolutional neural network on handwritten digits recognition is inspired by a few related works[55].

While, applying this three classifier SVM, ANN, and CNN to recognizing digits with noise. It demonstrated that SVM, ANN and CNN system can achieve high accuracy on recognition of handwritten digits on documented images[39]. However, these methods are used in this work to find the best algorithm for handwritten digits recognition. They were few drawbacks identified by the research area, by this, we can say that it is important to conduct a pre-study in order to understand the work that has been already done on classifying the methods and to understand the limitations of existing machine learning methods[10]. The results from the literature review give us a lot of existing research area on preprocessing, segmentation, feature extraction with specific techniques and classification to recognize the digits

In the paper [93], the authors have conducted research related to "Handwritten Word Recognition Using Multi-view Analysis". The major contribution of this research is a solution to the problem of efficiently recognizing handwritten words from a limited size lexicon. The authors developed a multiple classifier systems, that analyzes the words from three different approximation levels, in order to get a computational approach inspired by the human reading process.

The authors of the paper [94] have conducted research related to "Handwriting Recognition On Form Document". The author used Freeman Chain Code, with the division of a region into nine sub-regions, histogram normalization of chain code as feature extraction and Artificial Neural Networks, to classify the characters on the form document.

In the paper [95], the authors have conducted research related to "Neural Networks for Handwritten English Alphabet Recognition." They have developed a system to recognize handwritten English alphabets by using neural networks. In this system, each alphabet has been represented by binary values that are used as an input to a simple feature extraction system, whose output is fed to the neural network system.

In the paper [96], The authors have extracted the features of numeral and mathematical operators. They have used SVM for classification as well as to remove the noise from the dataset. A feature extraction method has been used on NIST dataset which consists of uppercase, lowercase, and merger of uppercase and lowercase.

The authors of the paper [97] "Sunspot drawings handwritten character recognition method based on deep learning", presented a deep learning method for scanned sunspot drawings handwritten characters recognition. A Convolution Neural Network, which is a type of deep learning algorithm and is truly successful in the training of multi-layer network structure, is used to train the recognition model of handwritten character images. The advantages of the proposed method by Chinese Academy Yunnan and the experimental results show that the proposed method achieves a high recognition accuracy rate.

The authors of [98] "New approach for segmentation and recognition of handwritten numeral strings" have proposed a new system for segmentation and recognition of unconstrained handwritten numeral strings. The proposed system uses a combination of foreground and background features for segmentation of touching digits.

In this paper [99], the authors have proposed a directional method for feature extraction on English handwritten characters. The collected data has been classified based on the similarity between the vector feature of data training and the vector feature of data testing.

The authors of the paper [100] "New efficient algorithm for recognizing handwritten Hindi digits", have presented a new algorithm for recognizing handwritten Hindi digits, which is based on using the topological characters combined with statistical properties of the given digits in order to extract a set of features that can be used in the process of digit classification.

In this paper [101] "Post-processing for offline Chinese handwritten character string recognition", an offline Chinese handwritten character recognition has been done with the help of pattern recognition. A free writing style, large variability in character shapes and different geometric characteristics of recognition are the challenging problems that have been taken up in this research. To solve this problem, post-processing has been used to improve the recognition accuracy of the characters.

CHAPTER 3

METHODOLOGY

In this section we are describing the various steps and accepts such as methods, tools, datasets used, how the models are created and how the models were trained are tested. In this section, we are discussing how algorithms used and presented the block diagram of the proposed system.
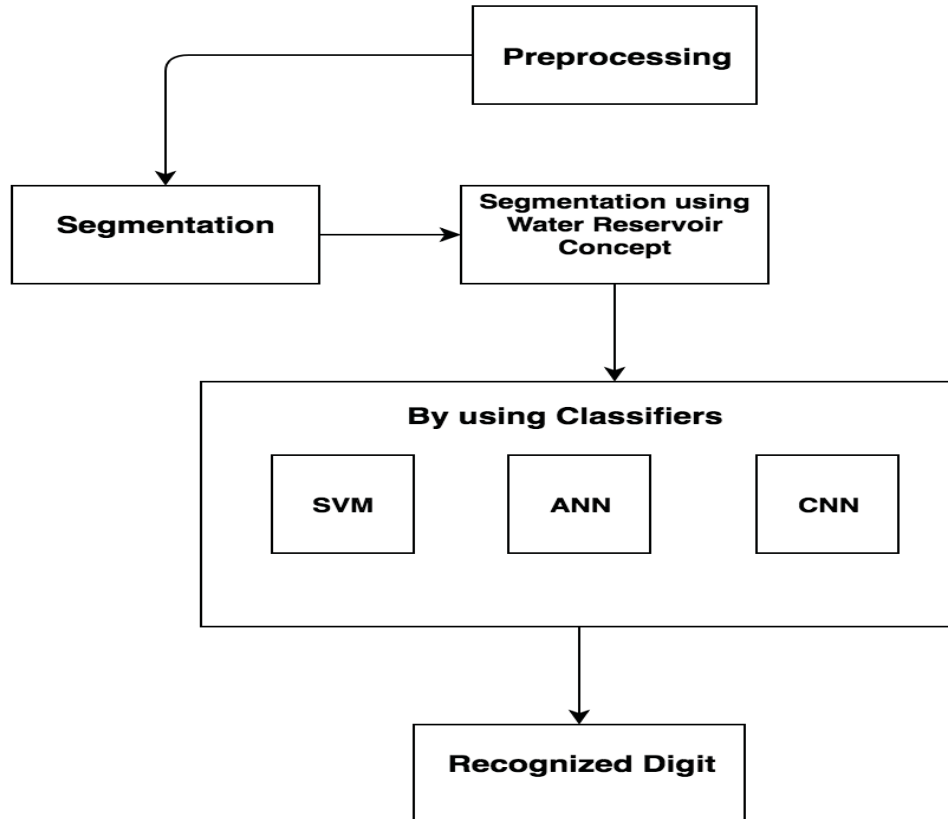

Figure 3.0 Block diagram of a Proposed method

3.1 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, segmentation and various machine learning methods to be adopted in the study.

Steps followed while performing the literature review:

Step 1: "Handwritten Digit Recognition", "Handwritten Digit Segmentation", "Handwritten Digit Classification", "Machine Learning Methods", "Deep Learning, Image processing on document images", "Support Vector Machine ", "Artificial Neural Networks", "Conventional Neural Networks", "Preprocessing Handwritten Digits". These keywords were identified before formulating the search string.

Step 2: Based on the listed keywords primary keywords were selected and used for the search string formulation.

Step 3: The following search strings were formed to perform the search in different digital libraries:

Search string 1 - "Automatic handwriting recognition system and method"

Search string 2 - "Handwritten digit recognition"

Search string 3 - "Handwritten digit recognition neural network"

Search string 4 - "Classifier methods handwritten digit recognition"

Search string 5- "Handwritten digit segmentation and recognition"

Search string 6 - "Handwritten digit recognition using Deep learning methods"

Search string 7- "Best efficient methods to recognize handwritten"

Step 4: After obtaining the list of articles, journals, and conference papers, inclusion and exclusion criteria was implemented to   and limit the results.

Inclusion Criteria:
- Papers over the past 20 years have been selected.
- Title and the abstract of the article match with the problem domain.
- Article available in full text in English.

Exclusion Criteria:
- Full-text article is not available.
- The language of the article is not in English.
- Articles falling in the domain other than computer science are rejected.
- Papers related to single digits recognition.
- Biologically inspired the hierarchical temporal memory model applied to the handwritten digit recognition.
- Handwritten digits by using mathematical morphology and various articles related to using rule-based decision fusion are excluded.

Step 5: Different machine learning methods that can be adopted for our study are listed in the results section of the experiment. While the knowledge about different data sets, preprocessing and segmentation has been utilized while documenting the related and background work.

Once the required data has been obtained from the literature review, then data analysis is performed[58]. 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[59]. The results gathered through this data analysis were documented and these were used for the experimental research method.

Experiment: The motivation for selecting experiment as a research method is to answer the RQ2, RQ3, RQ4 because quantitative data is required and descriptive methods like case-study and survey cannot provide us with the required data. In this research, two experiments are conducted. The first experiment is conducted to find the best machine learning methods for the recognition of digital images. The algorithms compared here are SVM and ANN. This experiment is designed in a way that, initially the input data which is the image data is divided into two parts which are training data and testing data. The training data is binarized. Binarization is nothing but converting the normal pixel image into a binary image[60]. After binarization, the using HOG, features are selected for training the algorithms. HOG stands for the histogram-oriented gradient which is used for feature selection in image processing[61]. Using the selected features, both machine learning algorithms are then trained with the training images. Once the trained model of SVM and ANN are created, then using the test

data images prediction is done. The performance of the algorithms SVM and ANN are estimated using the Accuracy metric and computational time for training.

The second experiment is conducted to evaluate the performance of the deep learning methods in recognition of digital images. Here for this experiment, the same amount of training data is used for experiment 1. But, instead binarization in this experiment, the training image is converted into grey scale images. Now, using these grey scale images, the deep learning algorithm which is CNN is trained. Once the model is trained, then using the same amount of testing data from experiment 1 is used for testing. The performance of the algorithm CNN is calculated using the Accuracy metric and training time.
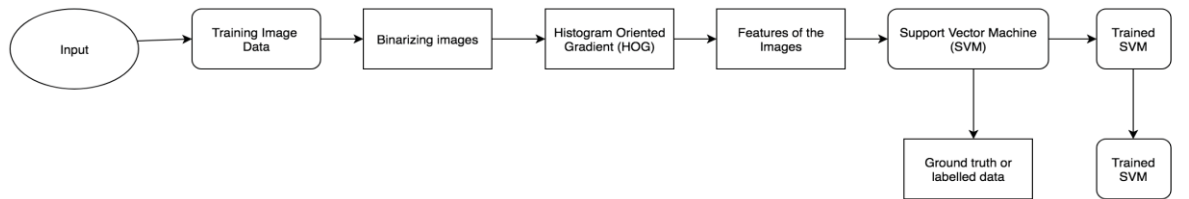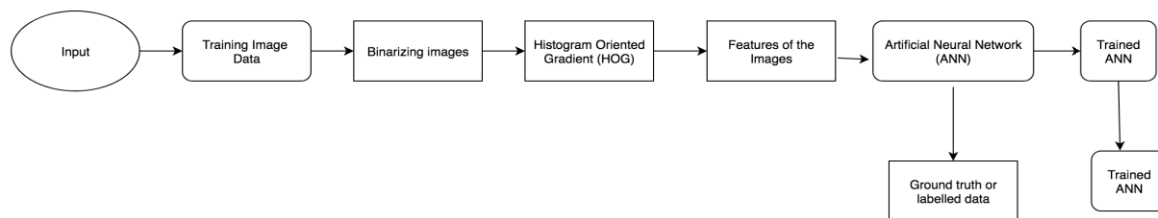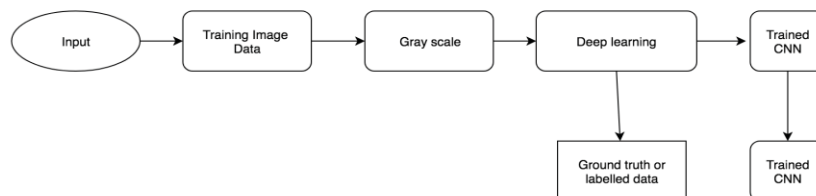
Figure 3.1 Block diagram classifiers used to recognize handwritten digits

To begin the experiment at first, we need to select the independent variables and dependent variables for this work.

Independent variables: The dataset and algorithms used in this study were Segmentation algorithms, Classification algorithms, i.e. Support vector machine, Artificial neural network, and Conventional neural network are the Independent variables.

Dependent variables: The performance algorithms, i.e. Segmentation algorithms, Classification accuracy, and Training classification algorithms are the dependent variables.

3.2 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.

3.3.1 Training new images on the constructed model

After preprocessing of the dataset, we will train the data from the top layer of the network and step by step outputs were:
- Training accuracy
- Validation accuracy
- Cross-entropy.

Training accuracy: While training accuracy we can find the percent of the images used in the current training batch was labeled with the correct class.

Validation accuracy: In validation accuracy, we will find the precision on a randomly-selected group of images from a different set.

Cross-entropy: The Cross entropy is a loss function which gives a glimpse into how well the learning process is progressing. [63] While training the dataset, we feed data to get predictions and these predictions are compared with the actual label data to update the layers weights through the back-propagation process. This process continues until accuracy increases. The trained model will perform on the classification task trained data to classify and generate a model.

3.3.2 Testing new images on the constructed model

After training model is completed with the selected algorithms or classifiers then we will perform the testing model by using the selected algorithms or classifiers[63]. While testing the dataset, initially all the steps used to train the model are followed except feeding the data to the classifier and generating a model. Here instead of feeding the data and training the model, we use the saved model to predict the class of the image i.e.[64] After predicting the classes of the test dataset, they are compared with the actual classes to check the accuracy of the algorithms or classifiers to check how well an algorithm works.

3.4 Handwritten Digit Segmentation

Segmentation of digit on documented images depends upon the link between adjacent digits. In this study, we are working with three situations which are Connected digits, Disjoint digits, Overlapped digits, respectively. In this part, describing the segmentation methods used in our system for which we propose a new segmentation method performed specifically for connected digits[55].
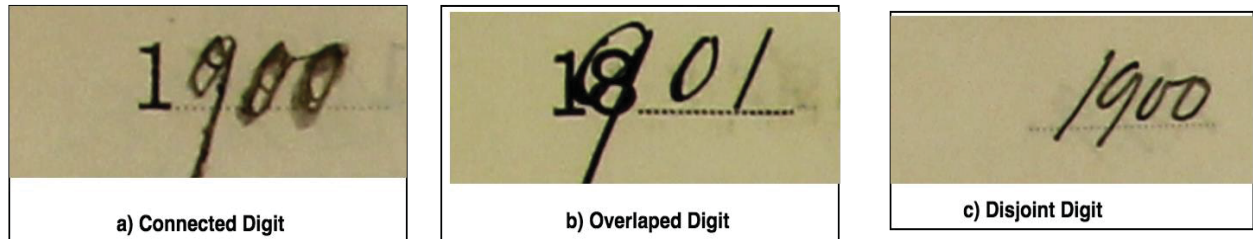

Figure 3.4 Different types of digits

Segmentation of Connected Digits

Segmentation of Connected digits various techniques were used to detect the connected digits by using the countermeasure skeleton technique[65]. We used split touching digits using thinning processes and it is based on segmentation strategy for a handwritten connected digit in this process to generate all possible conditions for segmentation. finally proposed a method for finding the Base Points (BPs) and Interconnection Points (IPs) on the contour and the skeleton of the connected digits according to the connection configuration. After that, a crossing-oriented window is set around IP for finding correctly the cutting path can be performed according to the following steps:
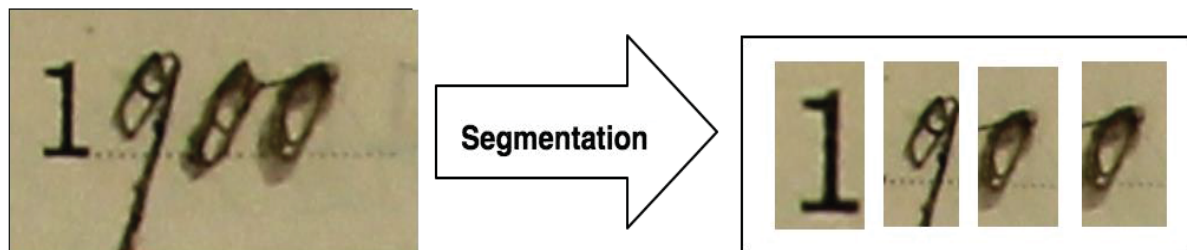

Figure 3.4 Segmentation of Connected digits

1. Apply a contour measure detection to detect all possible BPs from the local extrema
2. Perform the skeleton algorithm in order to detect all possible IPs.
3. If IP is detected, with a sliding window having the same height of the images and a fixed width is set on IP in the middle of the width. Then optimal orientation angle is used to match the right angle inter-digit it also allows reducing the number of segmentation cuts.
4. Segmentation hypotheses are helped to find the best cutting path. The hypothesis is performed using the digit recognition. In this case, the classifier plays an important role in detecting the overlapped or/and connected digits.

Segmentation of overlapped digits

Segmentation of overlapped digits is based on the contour analysis using the contour detection, which is extracted from the binary image using the morphology[56]. Hence, two adjacent digits are then separated using the fixed distance[66]. In some cases, broken parts of an overlapped

component are detected by examining the intersection with the median line of each component image.
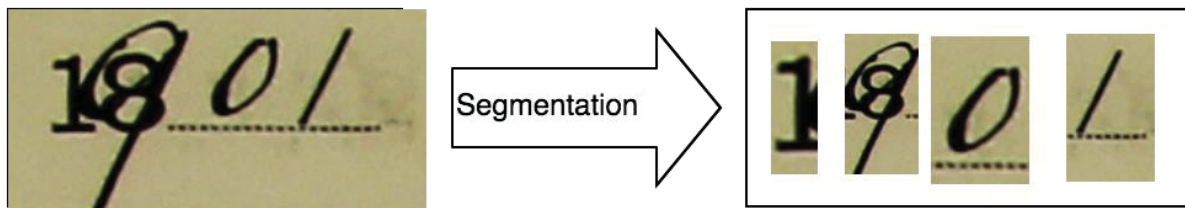

Figure 3.4 Segmentation of Overlapped digits

Segmentation of Disjoint Digits

Segmentation of disjoint digits by using the histogram of the vertical projection (HVP)[67]. The HVP is used to perform on the binary digits with a simple count of the black pixels in each column is running in order to detect the white space between successive digits[65]. It determines the location of each component of the image. During this method, the advantages of the ability of segmented digits with the unknown-length. In HVP we can split only digits comprising either it is a single digit or more than two digits.

3.5 Handwritten Digit Recognition

The handwritten digit recognition system is based on the input image, pre-process the image, segment the image, extract the feature of the image and classify the digit that is based on feature extracted module, and the proposed handwritten digits recognition[11], [51]. In the following, we briefly describe feature generation method, classifiers and the technique of recognition the digits.
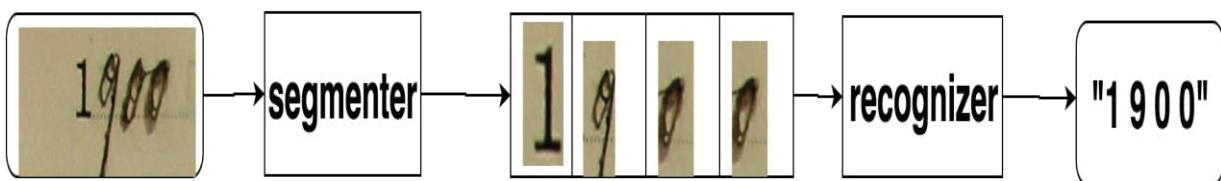

Figure 3.5 Handwritten digit recognition system

Feature generation

In feature generation, there are various feature generation methods have been proposed[68]. In the proposed system, we use a combination of multiple features for improving the recognition rate of handwritten digits by minimizing the intra-class variability and maximizing inter-class variability. These features were used in the study to include some global statistics and projection based features and features computed from the contour and skeleton of the digit.

Figure 3.5 Feature generation

Connected digits recognition

In connected digit recognition, all the images split into a sequence of the segment and each one of them is considered as a segmentation hypothesis, which is expected to contain a digit or a fragment of a digit[69].
If GC is accepted, then it is considered as a digit otherwise it is considered as a non-digit. The grouping of segments using conjointly GCA and DRV is performed according to the following heuristic rules:

X_{GC} {Accept if GC intersects median line and $f_{max}(X_{GC}) >= t_f$

X_{GC} {Reject $f_{max}(X_{GC}) >= t_f$

Overlapped digit recognition

Overlapped digit recognition is performed when the segmented component is less than a threshold. In this case, segmented components are considered as non-digits[70]. Therefore, contour detection is performed on segmented components for separation using a fixed some specific rules. The results of sub-components are identified by segmented component analysis for deciding if each one is a digit or non-digit. When it is detected as a digit, digit recognition is used for accepting or rejecting.

Disjoint digit recognition

Disjoint digit recognition is performed by using successively the Histogram of the vertical projection, segmented component analysis, and Digit recognition. HVP allows producing multiple segmented components and it is analyzed by segmented component analysis to find out it is a digit or non-digit. When a segmented component is detected as a digit, then digit recognition recognize digit with the help the of classifiers[71].

3.6 Statements of correct and incorrect segmentation-recognition

To recognize the handwritten digits whether the segmentation and recognition working correctly or incorrectly of our system Table.3.6 presents with some examples of correctly and incorrectly recognized digits.

	Digit string	Class labels	Segmented digits	Assigned classes
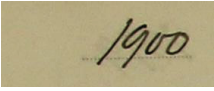
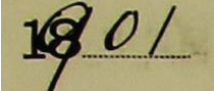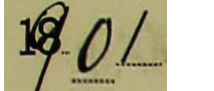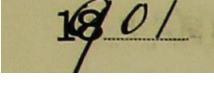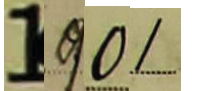|				
Correct segmentation and recognition		1900		1900
Incorrect segmentation		1901		1901
Incorrect recognition		1901		1701

Table 3.6: correct and incorrect segmentation-recognition

3.7 Algorithm for Touching Characters

Few rules have been added to the existing algorithm of touching characters for segmentation and the steps followed in the process are mentioned below:

Step 1) Separate the contour points into the upper boundary and lower boundary and this separation is done at the leftmost and rightmost points, giving two lists of contour points:

$C_U = \{(x_{up}, y_{up}) \mid u=1,2,\ldots,U\}$

$C_L = \{(x_{lp}, y_{lp}) \mid l=1,2,\ldots,L\}$

Step 2) Apply a vertical marginal operation for the upper and lower contours. At some x, more than one contour point exists. For such points, only one contour point is selected. For the upper contour, the lowest point is selected, and for the lower contour, the highest point is selected.

In parts (b) and (d) of Fig. 2.5, selected portions of the contour are shaded. As a result, the two lists of upper/lower contour points are converted to single-valued functions, $H_U(x)$ and $H_L(x)$, each of which represents the y coordinate of the contour point at x.

Step 3) Calculate an approximate measure of vertical width $H(x)$:

$H(x) = abs [H_L(x) - H_U(x)]$

Note here that the direction of the y coordinate is downward. In the intervals where $[H_L(x) - H_U(x)]$ is negative, $H(x)$ does not represent the real width of a stroke because of the approximation in step 2. It is an undefined interval in a sense.

Step 4) Search for candidate locations x^*_n, n = 1.... N, where the stroke should be cut. This is done by comparing the vertical width $H(x)$ with a threshold h_t. As shown in Fig. 2.5 (c), this comparison is made in the $[X_1, X_2]$ to limit the search range. The interval is defined to be the one where touching occurs. Candidate locations are where the curve $H(x)$ and line h_t cross each other. Let the candidate cutting contour points be

$\{(x^*_{up}, y_{up}), (x^*_{lp}, y_{lp}) \mid n=1,2,\ldots,N\}$

where subscripts U and 1 denote upper and lower, respectively. Examples of detected candidates of touching positions are depicted in Fig. 2.5 (a).

Step 5) If N is not zero, cutting contour points in (7) are shifted such that they are closer to the point where the width of the stroke expands abruptly. Otherwise, go to step 6. The width of the stroke at x^*_n can be calculated as $y_{ln} - y_{un}$, and that in the neighborhood can be determined by following the contour point chains in (5). Let the resulting modified candidate points be

$\{(x^*_{up}, y_{up}), (x^*_{lp}, y_{lp}) \mid n=1,2,\ldots,N\}$

Step 6) If N is zero, and if there is more than one inner loop, go to Algorithm Part 11.

Step 7) If N is zero, and if $H(x)$ is less than a threshold ht in the whole interval $[X_1, X_2]$, the middle of the interval is set as a candidate touching position and N is set as 1.

CHAPTER 4

RESULTS AND ANALYSIS

In this chapter contains all algorithms are applied for the digitization of image to segmentation and recognition of digital images. After the digitization, a grayscale image is calculated. Then thresholding technique is applied on Otsu's method to obtain binary images and it is resized shape of the image in certain aspect ratio is fixed so all the images have the same height and width. Finally, the morphological operator used to remove the noise in the image. After completing the preprocess segmentation part is applied. In segmentation, we have used water reservoir algorithms for comparison. That merges the concept of a drop-fall algorithm is applied as a baseline for comparison purpose. Then digit recognition module is used in preprocessing module in order to get compatible images for the classifier. In this module mainly consists of a center of mass extraction and normalization. Finally, the digit classifier module is used to recognize segmented digits. In this part, support vector machines, artificial neural networks and conventional neural networks are used and compare between them and to find the best-suited algorithm for recognition performance with high success rate. In this chapter, the modules mentioned earlier will be described step by step until the conclusion of the experiments.

4.1 Custom Dataset for Digits Segmentation

In this work, segmentation is done by using a water reservoir, and also for the validation of the segmentation algorithms, a dataset of digit and non-digits were used. Therefore, extraction of sample images CVL-Strings library is used for establishing the custom dataset. We have been worked with Python Script to make processing time essay. In this module, we use the input images and applies all for preprocessing algorithms and it will be applied in the main algorithms of segmentation and recognition module. Then a window which has 10 pixels horizontally slides through the image. Therefore, a window with h=100 pixels and w=10 pixels is used since all the digit string images are normalized to height h=100. While sliding, it asks the user the label of the image that is spanned by the window.

4.2. Custom Dataset for Digit Recognition

For the digit recognition algorithms, the custom database is used for training purposes. However, since the tests are done with CVL Strings database, the recognition algorithms are decided to be tested with also classifiers trained with the individual digits from the CVL Strings database and it is a true comparison between SVM, ANN and CNN classifier results. Since the main database is a digit string library, correctly segmented and created this database. Therefore, the results from the segmentation module are passed to the digit recognition module trained with the custom database and classified the individual digits. Then correctly classified digits are, then, saved to a .mat file forming a new custom set of digits. For digit recognition module, SVM, ANN with HOG feature extracting and CNN is used, since they have the highest recognition rate.

4.3 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.
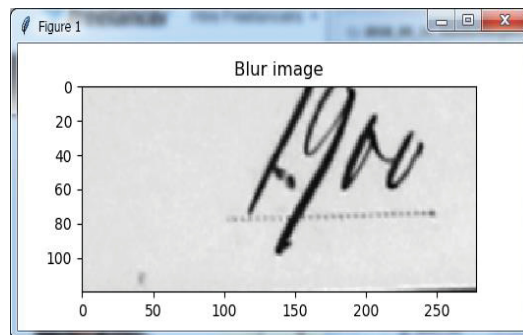
1. **Convert RGB to Gray image**
   We must convert RGB image to Gray image to get the binary image.
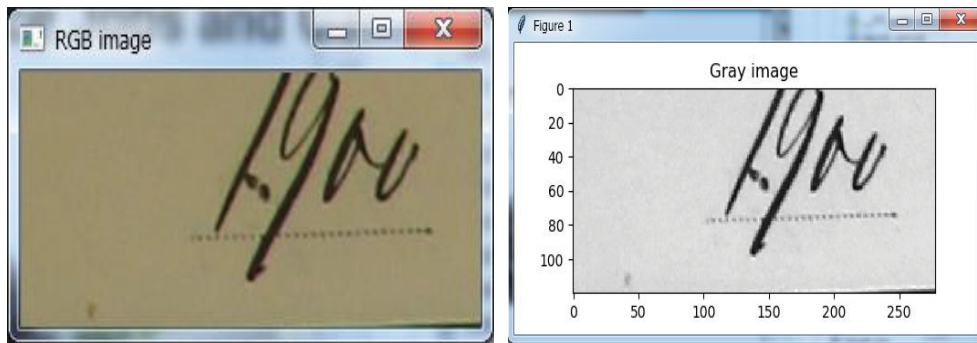
2. **Apply the blur filter with 3*3 size.**
   We must apply a blur filter to remove the Gaussian noise in the gray image.

3. **Convert the Gray image to a binary image.**
   We must convert Gray image to a binary image to foreground and background from image. Thus, the foreground will be the letters that we are going to get from the image. To convert Gray image to binary image. We can apply a thresholding method because it is the best method to make a binary image. So, we are applying the Adaptive thresholding method. The goal of thresholding an image is to classify pixels as either "dark" or "light". Adaptive thresholding is a form of thresholding that takes into account spatial variations in illumination. We present a technique for real-time adaptive thresholding using the integral image of the input.
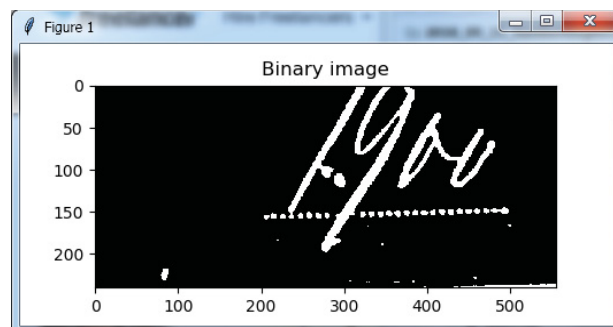
4. Remove regions that the area of the region is less than 150 pixels
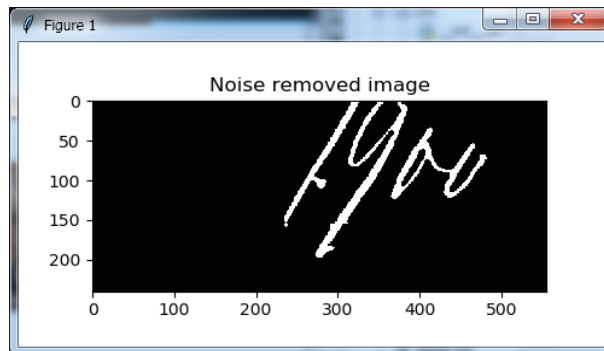
5. Merge the separated letters by expensing the pixels with block size 8*4
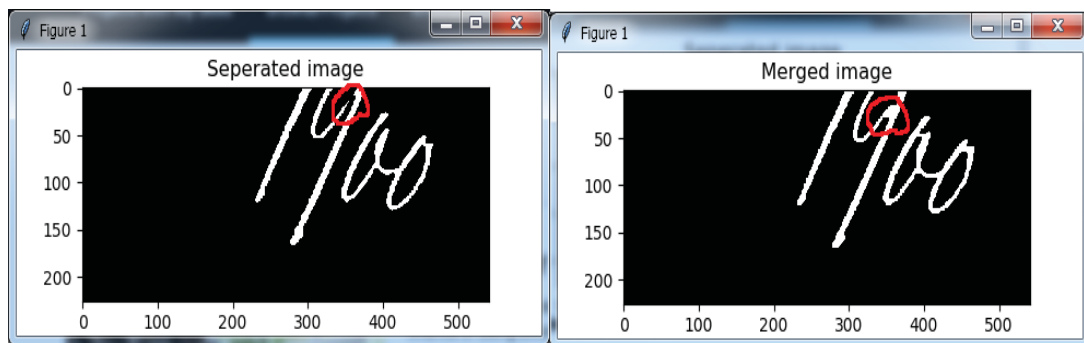

**6. Remove the several another noise like a line.**
To increase the accuracy of recognition, we have to obtain the best foreground regions, but the inputting images have a lot of the noise that is not the foreground as we can see from the testing images. Thus, we will remove those noises from images to get the good recognizing result like the following.

- Get the histogram according to with horizontal axis
- If the maximum value of the histogram is large than 80, we conclude that the image includes the line element and go until decrease the histogram value from a maximum index of the histogram by top direction.

- If the histogram value is larger or equal than previous histogram value, we select the point of the image that we are going to cut the image.
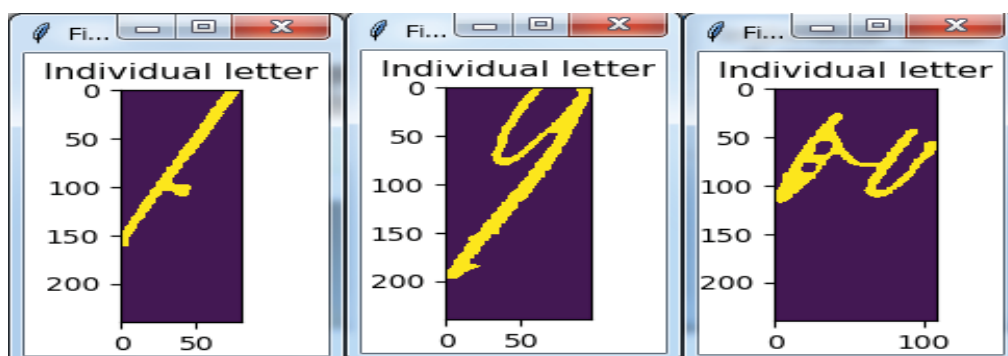- We crop the image by the obtained y-axis

The results can be shown in the following.

This plot is histogram according with horizontal direction
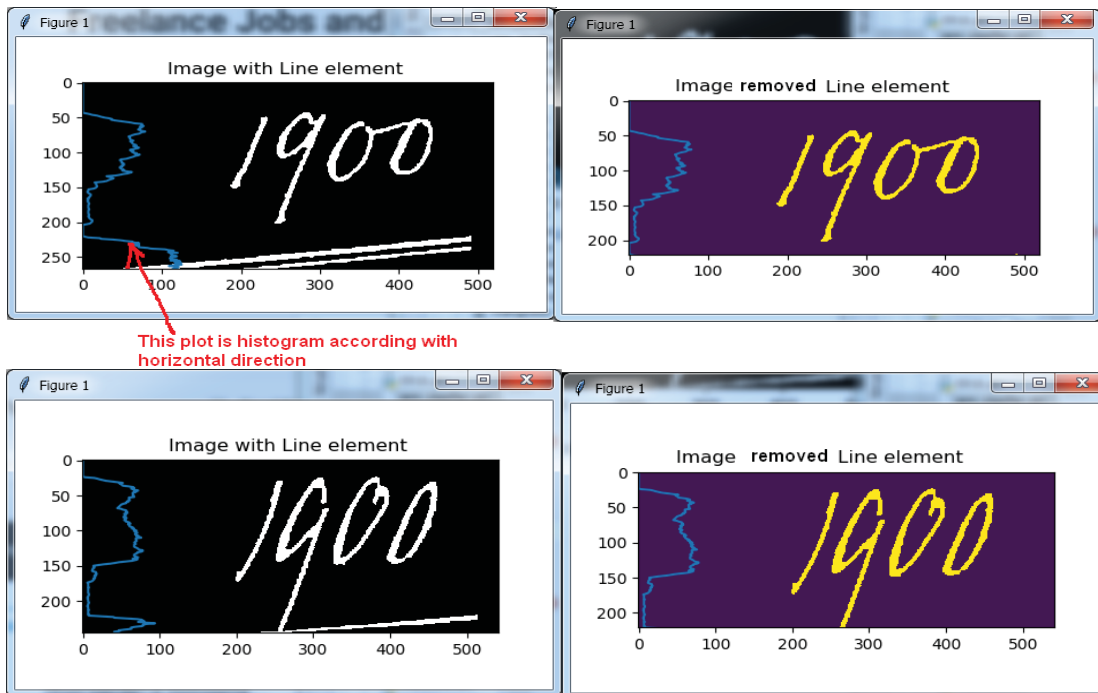
7. **Classify each letter by using 8-connection direction searching method.**
   As we can see the below figure, the letter classifies 3 letters.
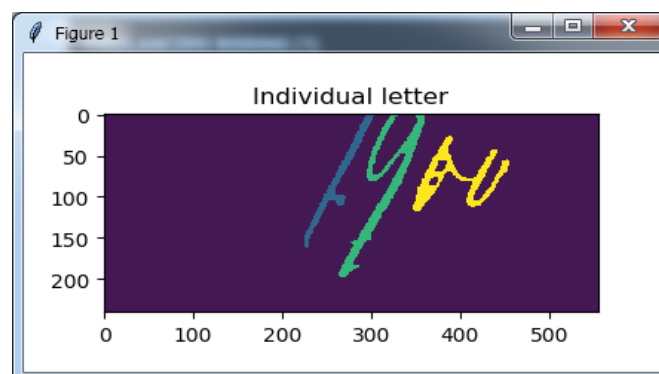
8. **Make classified each letter image into individual letter images**

Finally, re-array each letter images according to X-axis of each letter because the order of letters
may be changed when we classify the letters.

4.4 Segmentation

In this step, we have used the water reservoir concept algorithm. Segmentation using a water reservoir which is a novel approach proposed in this study[72].

4.4.1 Segmentation Using Water Reservoir Concept

By applying the water reservoir concept segmenting the connected pairs of the digits. At first connected component analysis is applied for input digits and therefore connected components are extracted and the bounding boxes of components were calculated[17]. Moreover, the bounding boxes are normalized to a constant column length which is equal to 100 pixels in this case. After finding the top and bottom reservoir of the classification algorithm is applied and deciding whether components of isolated digits or not touching types of the connected digits[75]. There are few drawbacks in this algorithm. The main drawbacks are unnecessary segmentation points are replaced in the single digits and it creates large reservoirs and when isolated digits are given to algorithm, and another is the wrong prediction during the segmentation when more than two digits are connected to each other, it fails to correctly segment all the digits[75]. However, since it is a rare possibility, the effect of this error is expected to be quite small in this study. After the segmentation is completed. 9096 components were obtained, and these are the digit recognition module as inputs.

- Calculate the threshold value whether the letter image is connected or not.
- First, we must take a certain threshold to decide whether the letter image is connected or not. To get this, we get the sum of the width of each letter images and then the value to get by dividing the sum with 4, called connected threshold.
- Therefore, if the letter image is larger than this threshold, we conclude that it includes the connected letters and implement the algorithm that reflects in my report to split them.
- Implement the algorithm to split the connected letters.
    1) Get the contours for the latter image.
    2) Get the Upper boundary and Lower boundary by re-array gotten contours according to with X-axis direction.
    3) Find the position that the difference between Upper boundary and Lower boundary is most big in a limited region.
    4) We must limit the searching region to get the good result for splitting.
    5) Thus, we took the searching region from 1/4 of letter width to 4/3 of letter width.
        a) Find the nearest points to the searched point from Upper and Lower boundary.
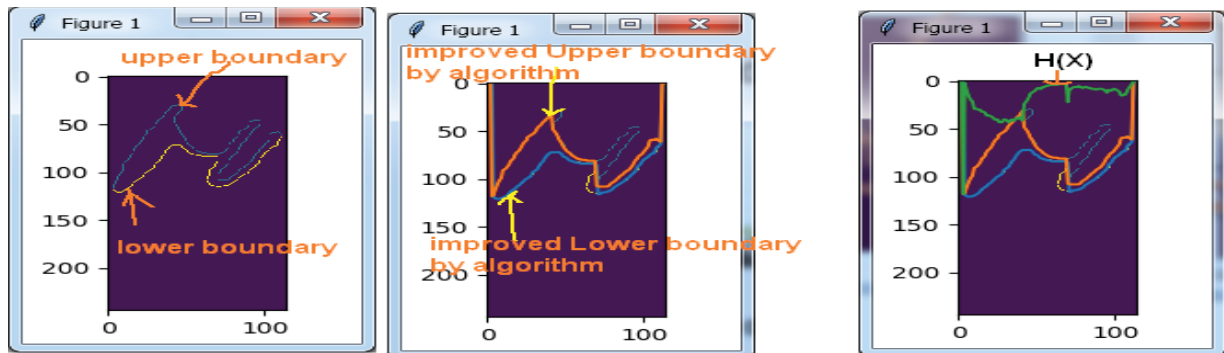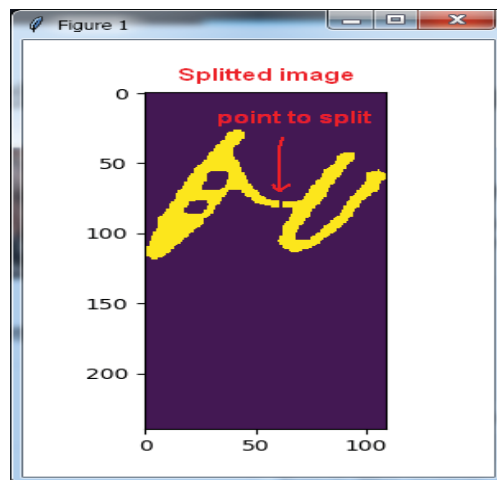   b) Split the letters by linking the gotten points with black color.

Figure 4.4 Implement the algorithm to split the connected letters

Through the processing until now, we can obtain the complete letter images.

4.5 Handwritten Digit Classifier

In this step, we will train and test the model to recognize the handwritten digits with the following algorithms or classifiers.

4.5.1 Train the SVM and ANN model with HOG feature vectors and the CNN model
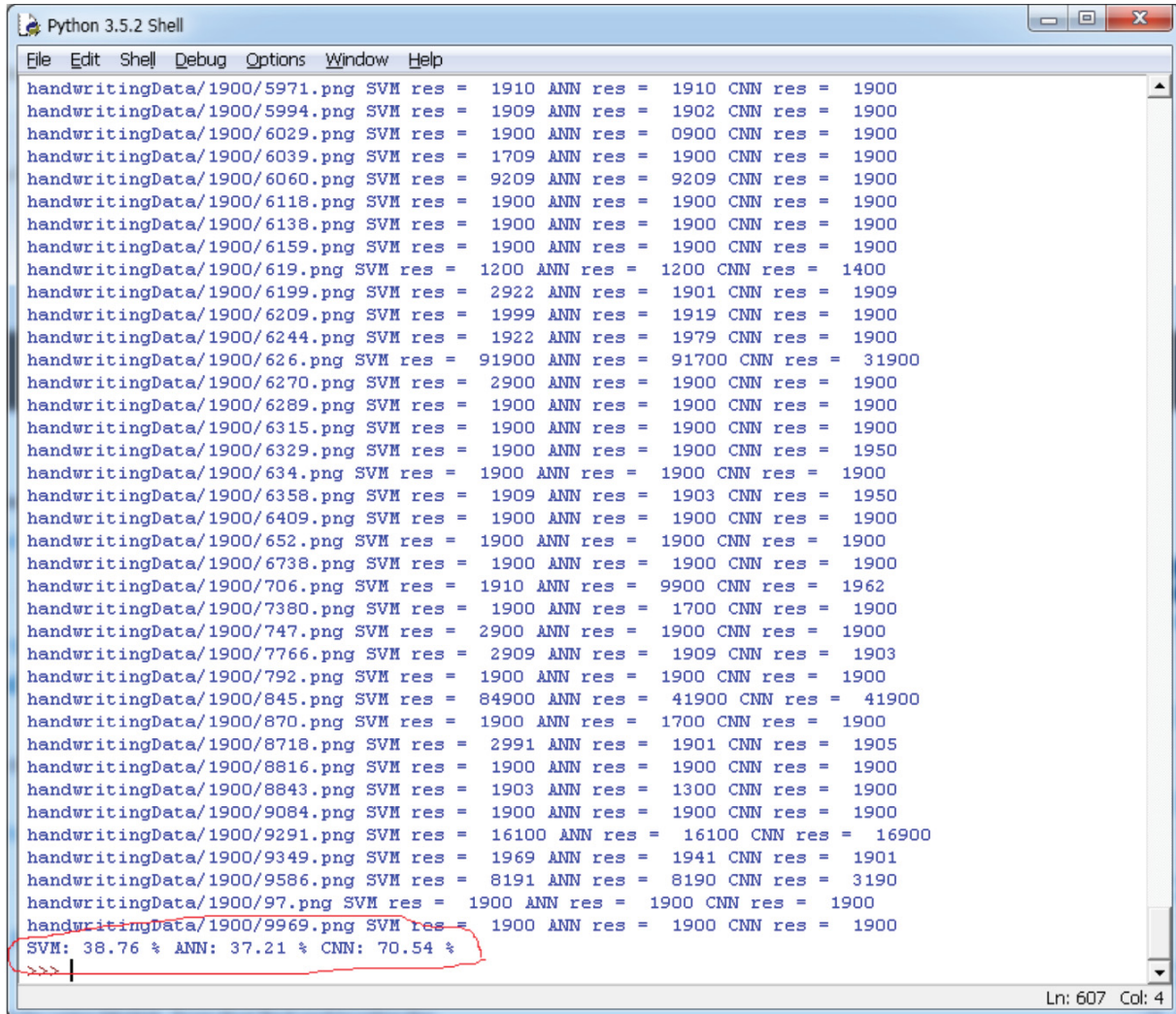
The histogram of oriented gradients (HOG) is a feature descriptor used in computer vision and image processing for object detection. The technique counts occurrences of gradient orientation in localized portions of an image.

- Resize letter images into 28*28 size image to detect the features of letters.
- Save resized image and output values by pressing the key for digit according to a letter from the keyboard.

- After repeating this processing until for the total images, we can see label.dat and traingData.dat, saved files, where label.dat is digit values and traingData.dat is letter images. We can train each model with these data. And then, we get the HOG feature vectors of letter images by using skimage lib and python library.
- We can use the sklearn lib and Python library for the SVM model and train the SVM model with the gotten HOG feature vectors.

- We can use sklearn lib, Python library for the ANN model and train the ANN model with the gotten HOG feature vectors.
    - We use the CNN model that made already and train the CNN model with the saved traingData.dat and label.data.

4.5.2. Recognize the handwriting with Trained SVM, ANN and CNN model

First, we must split each letter images of the image through preprocessing progressing. And then, we have to make HOG feature vectors by using python libs and recognize the handwriting by trained models.

Figure 4.5.2: Results

4.6 Comparison of Results of Conducted Experiments

Based on the experimental results and statistical tests, we conclude that CNN is the best-suited algorithm for recognition performance of machine learning methods. The recognition performance of CNN is 71%, SVM is 39%, and ANN is 37%.
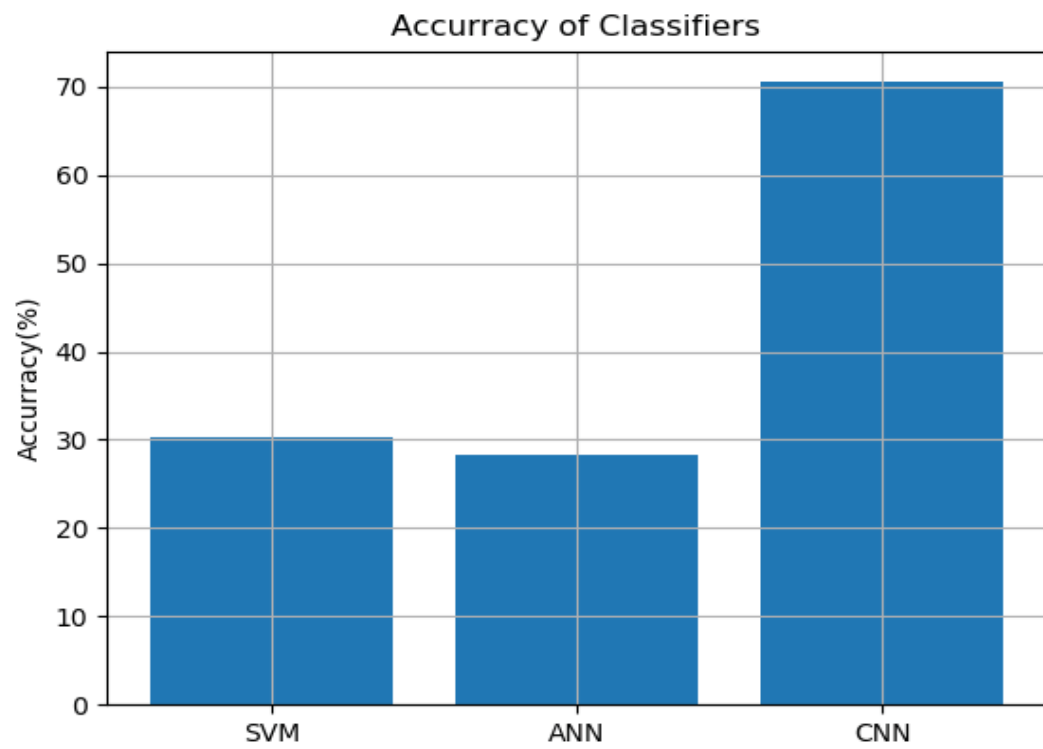
Figure: 4.6 Comparison of Results of Conducted Experiments

CHAPTER 5

5.1 Answers to Research questions

RQ1: What data is used to train and test the machine learning methods?

Answer: We have used the custom dataset to train and test. There are 9096-digit handwriting images. we get the separated letters by using the segment method to split connected letters as you can see the code and trained and tested the classifier models with them.

For example, we can see below figures for this image shows the individual letters. So, we can train the models with these images.
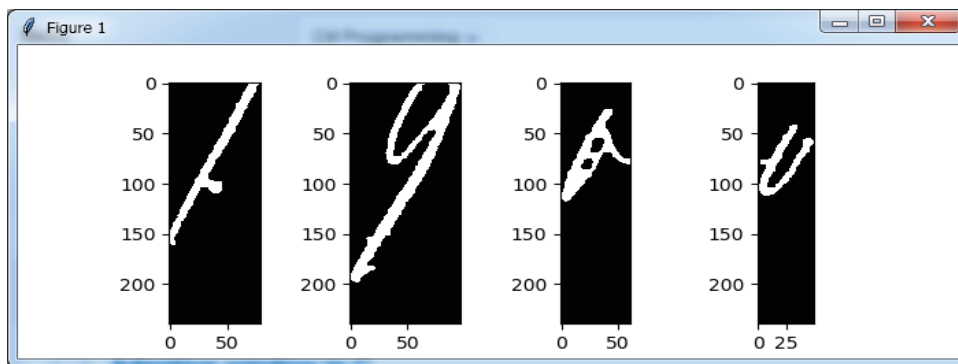
Figure 5.1 segment method to split connected letters

RQ2: What are the best parameters of machine learning methods to recognize digits?

Answer: After collecting the data from the test results, an experiment is conducted on the three algorithms which are SVM, ANN, and CNN.
- For SVM, best ML parameters are C = 1.0, penalty = 12, max_iter = 1000, tol=1e-4.
- For ANN, best ML parameters are solver = "lbfgs", hidden_layer_sizes = 30*30
- For CNN, best ML parameters are lr = 0.01. decay=1e-6, momentum=0.9.

RQ3: What is the recognition performance of machine learning methods for the handwritten digits which cannot be segmented into document images?

Answer: Based on the experimental results and statistical tests, we conclude that CNN is the best-suited algorithm for recognition performance of machine learning methods. The recognition performance of CNN is 71%, SVM is 39%, and ANN is 37%.

RQ4: Which machine learning method provides the best recognition performance for digit recognition for those which cannot be segmented?

Answer: By the above RQ3 and from the experimental results, the best machine learning method is CNN.

5.2 Contribution

In this study, the main purpose is to recognize corrupted handwritten digits using recognition method. Image of connected digits, disjoint digits and overlapped digits was divided into sub-images. Then sub-images were used to train a classifier. Then the classifier was used to identify parts of an image. In this study, we showed that machine learning methods can be used for efficient classification in handwritten digits. Thus, our study serves as a reference for aspiring researchers in automatic handwritten digits or hand recognition and classification. Apart from this major contribution of this study is to implement some specific rules while segmentation and analyze the results. The experimental results of this study showed that conventional neural network is the best suitable algorithm or classifier for automatic handwritten digits on documented images accuracy rate is also high when compared with other algorithms like SVM and ANN

5.3 Threats to validity

The results of the research are helpful to the extent the various techniques for improving the performance of the segmentation and resolve issues for recognition. It is essential to consider the validity threats in this research and the validity of research is concerned with the question of how the conclusions might be wrong[78]. In quantitative research there are four types of validity threats, which are conclusion validity, internal validity, construct validity and external validity[79].

Internal Validity: The Internal validity threats affect the independent variables concerning causality[81]. In this paper, the independent variables are the dataset and algorithms used in this project. To mitigate this validity threat, the data required for the dataset is selected by discussing with the supervisor currently working on the image processing. Also, data preprocessing for algorithms is done by using previous research work.

Construct Validity: Construct validity threats concerns about the design of the experiment[82]. To mitigate this validity threat, the hypothesis is defined, and the experiment is designed upfront. Here, the experiment is designed based on the amount of data obtained and by evaluating other designs.

External Validity: The fact that the administration of a test may affect the performance of the participants in the study. Its results may not be generalizable to situations where pretesting will not occur[75], [78]. This threat is mitigated by employing an experimental design which does not have any pretests. The experimenters involved in the study alter the observations by creating expected outcomes by themselves. We have used thresholding technique to combat this bias.

5.4 Computation cost of the proposed system

The proposed system of automatic handwritten digits on documented images by using machine learning methods allows resoling different segmentation and recognition problems. Even this system can be deployed in the real environment, to compute the cost required for treating all situations. Hence, the computation cost is measured by considering three situations:

Connected digit: The total number of Segmentation Hypothesis is equal to the sum of all numbers of segmentation hypotheses and formula is $N_{SH} = N_{DS}(N_{HVP} + N_{CA})$.

Overlapped digit: Let N_{GBC} be the Number of Grouped Broken Component provided by CA when the broken component exists and N_{ID} the total Number of Isolated Digit provided by CA, then, the total Number of Segmentation Hypotheses from CA (N_{CA}) is $N_{CA} = N_{GBC} + N_{ID}$

Spaced digit: The total Number of Segmentation Hypotheses using HVP (N_{HVP}) is the same as the number of the spaced digits.

5.5 Limitations

The limitations were classifier robustness, no obvious segmentation point on connected digits and training time of the conventional neural networks is too high to perform statistical analysis of all possible combinations of the experiment variables.

5.6 Summary

The goal of this thesis is to present automatic handwritten digits on documented images by using machine learning methods and it is a combination of several techniques like HOG, Ostus, Segmentation and Recognition methods. While segmentation we have been used several methods like Histogram of the Vertical Projection, Component Analysis and used with Segmented Component Analysis and Digit Recognition for Recognition. The benefit of this system is the ability to get proper segment information while segmenting the handwritten digits like connected digits, overlapped and disjoint digits.

The obtained performance of the proposed automatic handwritten digits on documented images depends on mainly three different factors of accuracy SVM, ANN, CNN on Digit Recognition, while using segmentation methods, the combination problem is with some heuristic rules. In some cases, the unknown-length of the digits or images makes more difficult to recognize the task for classifiers[83]. However, by combining HVP and CA, the proposed system allows the confusion of the automatic handwritten digits on documented images by using machine learning methods.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1. Conclusion

The primary focus of this research is to build an Automatic Handwritten Digit Recognition on document images, the important problem is segmenting adjacent digits inherit to the disjoint digits, overlapping digits, connected digits, variability of the digits, unknown length of the string, over-segmentation and under-segmentation. the development of the above system is not a problem. Water a reservoir is used to for segmenting two connected handwritten digits by using SVM, ANN and CNN method to find the best possible approach for isolating two adjacent digits. recognizing the handwriting is designed in three steps. segmentation of connected digits by searching the interconnection points using the water reservoir method, generation of features on isolated digits, recognition of the segmented digits.

Two complementary structural features i.e. counter and skeletal points, are used by the segmentation two connected digits for finding the base points, interconnection and detecting the optimal path. selection of the appropriate orientation angle is the main constraint for the performance of system. The results obtained explains that the by using appropriate angle the recognition rate increases.

Selecting the appropriate optimal angle is the important step for cutting two digits optimally. there is a tradeoff between the number of segmentations cut and correct segmentation according to the state-of-the-art of the proposed framework.

There are limitations in the proposed framework when the following two constraints are involved:

The first limitation is because of the classifier robustness. Even though the segmentation is done precisely sometimes the classifier doesn't allow to recognize the segmented digits because of its shape. To mitigate this limitation, it is recommended to add this digit to the training set t increase the performance of SVM, ANN, CNN classifier.

The second limitation is due to "no obvious segmentation point". segmenting connected digits are applied by the assumption 3 (No IP). By this, acceptable results are not provided by the segmentation path in between. According to the configuration link between digits, a new design is proposed by combining several segmentation methods to recognize the digit of unknown length.

Based on the histogram vertical projection segmentation methods are combined. by using verification strategy, counter analysis for unknown length of the string can be found.

The primary target of the proposed framework is to avoid the under segmentation and over segmentation by using good cutting path.

The recognition module is used to configure the digit strings in difficult situations. The initial results are important because all the ways of segmentations are managed by the proposed framework.

By implementing few rules, the proposed framework has many advantages in providing the correct segmentation without having any prior knowledge in the research.

6.2. Future Work

A new method can be proposed to cut or segmenting the digit strings still there are some limitation for this method, where improvements has to be made. Thus, there is a place for some future work such as:

- Different classifications models can be used at a time to improve the performance of the segmentation.
- To reduce the complexity of the algorithm, it's better to reduce the number of hypothesis to function the algorithm faster.
- To reduce the computation time, better filters are to be used to eliminate the unnecessary segmentation hypothesis.

REFERENCES

[1] R. Jayadevan, S. R. Kolhe, P. M. Patil, and U. Pal, "Automatic processing of handwritten bank cheque images: a survey," *Int. J. Doc. Anal. Recognit. IJDAR*, vol. 15, no. 4, pp. 267–296, 2012.

[2] N. M. Nasrabadi, "Pattern recognition and machine learning," *J. Electron. Imaging*, vol. 16, no. 4, p. 049901, 2007.

[3] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.

[4] Y. Bengio, "Learning deep architectures for AI," *Found. Trends® Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.

[5] R. Plamondon and S. N. Srihari, "Online and off-line handwriting recognition: a comprehensive survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 63–84, 2000.

[6] R. M. Bozinovic and S. N. Srihari, "Off-line cursive script word recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 1, pp. 68–83, 1989.

[7] K. Marukawa, M. Koga, Y. Shima, and H. Fujisawa, "A High Speed Word Matching Algorithm for Handwritten Chinese Character Recognition.," in *MVA*, 1990, pp. 445–450.

[8] R. W. Smith, J. F. McNamara, and D. S. Bradburn, "Pattern Classification Techniques Applied to Character Segmentation," *TRW Financ. Syst. Berkeley CA*, 1998.

[9] F. Lauer, C. Y. Suen, and G. Bloch, "A trainable feature extractor for handwritten digit recognition," *Pattern Recognit.*, vol. 40, no. 6, pp. 1816–1824, 2007.

[10] L. Bottou *et al.*, "Comparison of classifier methods: a case study in handwritten digit recognition," in *Pattern Recognition, 1994. Vol. 2-Conference B: Computer Vision & Image Processing., Proceedings of the 12th IAPR International. Conference on*, 1994, vol. 2, pp. 77–82.

[11] G. S. Lopes, D. C. da Silva, A. W. O. Rodrigues, and P. P. Reboucas Filho, "Recognition of handwritten digits using the signature features and Optimum-Path Forest Classifier," *IEEE Lat. Am. Trans.*, vol. 14, no. 5, pp. 2455–2460, 2016.

[12] L. S. Oliveira, R. Sabourin, F. Bortolozzi, and C. Y. Suen, "A methodology for feature selection using multiobjective genetic algorithms for handwritten digit string recognition," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 17, no. 06, pp. 903–929, 2003.

[13] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[14] N. Arica and F. T. Yarman-Vural, "An overview of character recognition focused on off-line handwriting," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 31, no. 2, pp. 216–233, 2001.

[15] H. Murase, M. Shinya, T. Wakahara, and K. Odaka, "Segmentation and recognition of handwritten character string using linguistic information," *Trans Inst. Electron. Inf. Commun. Eng.*, vol. 69, no. 9, pp. 1292–1301, 1986.

[16] D. Lu and Q. Weng, "A survey of image classification methods and techniques for improving classification performance," *Int. J. Remote Sens.*, vol. 28, no. 5, pp. 823–870, 2007.

[17] U. Pal, A. Belaıd, and C. Choisy, "Touching numeral segmentation using water reservoir concept," *Pattern Recognit. Lett.*, vol. 24, no. 1–3, pp. 261–272, 2003.

[18] "Machine learning," *Wikipedia*. 14-Jan-2019.

[19] A. Talwar and Y. Kumar, "Machine Learning: An artificial intelligence methodology," *Int. J. Eng. Comput. Sci.*, vol. 2, no. 12, 2013.

[20] "Deep learning," *Wikipedia*. 14-Jan-2019.

[21] D. C. Cireşan, U. Meier, L. M. Gambardella, and J. Schmidhuber, "Deep, big, simple neural nets for handwritten digit recognition," *Neural Comput.*, vol. 22, no. 12, pp. 3207–3220, 2010.

[22] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, 2005, vol. 1, pp. 886–893.

[23] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst. Man Cybern.*, vol. 9, no. 1, pp. 62–66, 1979.

[24] C. D. James *et al.*, "A historical survey of algorithms and hardware architectures for neural-inspired and neuromorphic computing applications," *Biol. Inspired Cogn. Archit.*, vol. 19, pp. 49–64, Jan. 2017.

[25] "Support-vector machine," *Wikipedia*. 03-Jan-2019.

[26] C. J. Burges and B. Schölkopf, "Improving the accuracy and speed of support vector machines," in *Advances in neural information processing systems*, 1997, pp. 375–381.

[27] V. Vapnik, *Estimation of dependences based on empirical data*. Springer Science & Business Media, 2006.

[28] M. I. Fanany, "Handwriting recognition on form document using convolutional neural network and support vector machines (CNN-SVM)," in *Information and Communication Technology (ICoIC7), 2017 5th International Conference on*, 2017, pp. 1–6.

[29] S. R. Gunn, "Support vector machines for classification and regression," *ISIS Tech. Rep.*, vol. 14, no. 1, pp. 5–16, 1998.

[30] R. Beale and T. Jackson, *Neural Computing-an introduction*. CRC Press, 1990.

[31] "Convolutional neural network," *Wikipedia*. 05-Jan-2019.

[32] X.-X. Niu and C. Y. Suen, "A novel hybrid CNN–SVM classifier for recognizing handwritten digits," *Pattern Recognit.*, vol. 45, no. 4, pp. 1318–1325, 2012.

[33] L. Deng and D. Yu, "Deep Learning: Methods and Applications," in Foundations and trends in Signal Processing, 2014, pp 192-387.

[34] "Literature review," *Wikipedia*. 26-Nov-2018.

[35] ช. ı. ฏ. ฐ. พ. ศรี, "A Guide to Writing the Thesis Literature Review for Graduate Studies," *Exec. J.*, vol. 34, no. 1, pp. 11–22, 2014.

[36] T. Greenhalgh, G. Robert, F. Macfarlane, P. Bate, O. Kyriakidou, and R. Peacock, "Storylines of research in diffusion of innovation: a meta-narrative approach to systematic review," *Soc. Sci. Med.*, vol. 61, no. 2, pp. 417–430, 2005.

[37] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," *ArXiv Prepr. ArXiv13126229*, 2013.

[38] R. F. Moghaddam and M. Cheriet, "AdOtsu: An adaptive and parameterless generalization of Otsu's method for document image binarization," *Pattern Recognit.*, vol. 45, no. 6, pp. 2419–2431, 2012.

[39] K. Bansal and R. G. Kumar, "Cleaning and Recognition of Numerals in a Handwritten Devnagari Document Consisting of Roll Numbers and Marks," PhD Thesis, 2012.

[40] M. Cheriet, R. F. Moghaddam, and R. Hedjam, "A learning framework for the optimization and automation of document binarization methods," *Comput. Vis. Image Underst.*, vol. 117, no. 3, pp. 269–280, 2013.

[41] P. K. Sahoo, S. Soltani, and A. K. Wong, "A survey of thresholding techniques," *Comput. Vis. Graph. Image Process.*, vol. 41, no. 2, pp. 233–260, 1988.

[42] F. Nie, C. Gao, Y. Guo, and M. Gan, "Two-dimensional minimum local cross-entropy thresholding based on co-occurrence matrix," *Comput. Electr. Eng.*, vol. 37, no. 5, pp. 757–767, 2011.

[43] Ø. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 12, pp. 1191–1201, 1995.

[44] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," *J. Electron. Imaging*, vol. 13, no. 1, pp. 146–166, 2004.

[45] S. Knerr, V. Anisimov, O. Baret, N. Gorski, D. Price, and J.-C. Simon, "The A2iA intercheque system: courtesy amount and legal amount recognition for French checks," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 11, no. 04, pp. 505–548, 1997.

[46] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, 2005.

[47] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Phys. Nonlinear Phenom.*, vol. 60, no. 1–4, pp. 259–268, 1992.

[48] B. Scholkopf *et al.*, "Comparing support vector machines with Gaussian kernels to radial basis function classifiers," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2758–2765, 1997.

[49] L. Vincent, "Morphological grayscale reconstruction in image analysis: applications and efficient algorithms," *IEEE Trans. Image Process.*, vol. 2, no. 2, pp. 176–201, 1993.

[50] A. Bagdanov and J. Kanai, "Projection profile based skew estimation algorithm for JBIG compressed images," in *Document Analysis and Recognition, 1997., Proceedings of the Fourth International Conference on*, 1997, vol. 1, pp. 401–405.

[51] C.-L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, "Handwritten digit recognition: investigation of normalization and feature extraction techniques," *Pattern Recognit.*, vol. 37, no. 2, pp. 265–279, 2004.

[52] S. Kahan, T. Pavlidis, and H. S. Baird, "On the recognition of printed characters of any font and size," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 2, pp. 274–288, 1987.

[53] "Skeletonization Algorithm for Binary Images - ScienceDirect." [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2212017313004027. [Accessed: 15-Jan-2019].

[54] S. Knerr, L. Personnaz, and G. Dreyfus, "Handwritten digit recognition by neural networks with single-layer training," *IEEE Trans. Neural Netw.*, vol. 3, no. 6, pp. 962–968, 1992.

[55] F. C. Ribas, L. S. Oliveira, A. S. Britto, and R. Sabourin, "Handwritten digit segmentation: a comparative study," *Int. J. Doc. Anal. Recognit. IJDAR*, vol. 16, no. 2, pp. 127–137, 2013.

[56] A. Elnagar and R. Alhajj, "Segmentation of connected handwritten numeral strings," *Pattern Recognit.*, vol. 36, no. 3, pp. 625–634, 2003.

[57] Y.-K. Chen and J.-F. Wang, "Segmentation of single-or multiple-touching handwritten numeral string using background and foreground analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 11, pp. 1304–1317, 2000.

[58] J. Webster and R. T. Watson, "Analyzing the past to prepare for the future: Writing a literature review," *MIS Q.*, pp. xiii–xxiii, 2002.

[59] J. R. Edwards, *Person-job fit: A conceptual integration, literature review, and methodological critique.* John Wiley & Sons, 1991.

[60] A. Gattal, Y. Chibani, C. Djeddi, and I. Siddiqi, "Improving isolated digit recognition using a combination of multiple features," in *Frontiers in Handwriting Recognition (ICFHR), 2014 14th International Conference on*, 2014, pp. 446–451.

[61] Q. Zhu, M.-C. Yeh, K.-T. Cheng, and S. Avidan, "Fast human detection using a cascade of histograms of oriented gradients," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, 2006, vol. 2, pp. 1491–1498.

[62] Y. Al-Ohali, M. Cheriet, and C. Suen, "Databases for recognition of handwritten Arabic cheques," *Pattern Recognit.*, vol. 36, no. 1, pp. 111–121, 2003.

[63] T. Blaschke, C. Burnett, and A. Pekkarinen, "Image segmentation methods for object-based analysis and classification," in *Remote sensing image analysis: Including the spatial domain*, Springer, 2004, pp. 211–236.

[64] J. Brownlee, "How to Train a Final Machine Learning Model," *Machine Learning Mastery*, 16-Mar-2017.

[65] "Segmentation-Verification for Handwritten Digit Recognition | Publish your master's thesis, bachelor's thesis, essay or term paper." [Online]. Available: https://www.grin.com/document/374168. [Accessed: 15-Jan-2019].

[66] S. Hua and Z. Sun, "A novel method of protein secondary structure prediction with high segment overlap measure: support vector machine approach1," *J. Mol. Biol.*, vol. 308, no. 2, pp. 397–407, 2001.

[67] "(PDF) Combining Multiple Segmentation Methods for Handwritten Digit Recognition of Algerian Bank Checks." [Online]. Available: https://www.researchgate.net/publication/282184214_Combining_Multiple_Segmentation_Methods_for_Handwritten_Digit_Recognition_of_Algerian_Bank_Checks?_sg=A_AFv47pMx-WN7PZIOuWrwN_qfaTPY3KHtna3P_LFM7F0rjsbMYqHdFlpakAHibUpzN_-WcKuzkIqpQuGkxriVnU8LENtHEjJQ. [Accessed: 15-Jan-2019].

[68] A. GATTAL and Y. CHIBANI, "Structural features generations for automatic recognition of isolated handwritten digits."

[69] A. Gattal and Y. Chibani, "Segmentation strategy of handwritten connected digits (SSHCD)," in *International Conference on Image Analysis and Processing*, 2011, pp. 248–254.

[70] G. L. Martin and M. Rashid, "Recognizing Overlapping Hand-Printed Characters by Centered-Object Integrated Segmentation and Recognition," in *Advances in Neural Information Processing Systems 4*, J. E. Moody, S. J. Hanson, and R. P. Lippmann, Eds. Morgan-Kaufmann, 1992, pp. 504–511.

[71] H. Fujisawa, Y. Nakano, and K. Kurino, "Segmentation methods for character recognition: from segmentation to document structure analysis," *Proc. IEEE*, vol. 80, no. 7, pp. 1079–1092.

[72] M. Suwa, "Segmentation of connected handwritten numerals by graph representation," in *Document Analysis and Recognition, 2005. Proceedings. Eighth International Conference on*, 2005, pp. 750–754.

[73] H. Kusetogullari, H. Grahn and N. Lavesson, "Handwriting image enhancement using local learning windowing, Gaussian Mixture Model and k-means clustering," *2016 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, Limassol, 2016, pp. 305-310.

[74] U. Pal, A. Belaïd, and C. Choisy, "Water reservoir based approach for touching numeral segmentation," in *Document Analysis and Recognition, 2001. Proceedings. Sixth International Conference on*, 2001, pp. 892–896.

[75] U. Pal, A. Belaid, and C. Choisy, "Water reservoir based approach for touching numeral segmentation," in *Proceedings of Sixth International Conference on Document Analysis and Recognition*, Seattle, WA, USA, 2001, pp. 892–896.

[76] H. Kusetogullari, A. Yavariabdi, Evolutionary multiobjective multiple description wavelet based image coding in the presence of mixed noise in images, Applied Soft Computing Volume 73, Pages 1039-1052, 2018.

[77] H. Kusetogullari, Unsupervised Text Binarization in Handwritten Historical Documents Using k-Means Clustering, Proceedings of SAI Intelligent Systems Conference (IntelliSys), pp 23-32, 2016.

[78] R. Feldt and A. Magazinius, "Validity Threats in Empirical Software Engineering Research - An Initial Survey," p. 6.

[79] I. Bjelland, A. A. Dahl, T. T. Haug, and D. Neckelmann, "The validity of the Hospital Anxiety and Depression Scale: an updated literature review," *J. Psychosom. Res.*, vol. 52, no. 2, pp. 69–77, 2002.

[80] T. Celik and H. Kusetogullari, "Self-sampled image resolution enhancement using dual-tree complex wavelet transform," *2009 17th European Signal Processing Conference*, Glasgow, pp. 2017-2021, 2009

[81] rodsazon, "Threats to internal and external validity," 03:21:44 UTC.

[82] Stephanie, "Statistical Conclusion Validity," *Statistics How To*, 01-Oct-2015. [Online]. Available: https://www.statisticshowto.datasciencecentral.com/statistical-conclusion-validity/. [Accessed: 15-Jan-2019].

[83] L. Xu, A. Krzyzak, and C. Y. Suen, "Methods of combining multiple classifiers and their applications to handwriting recognition," *IEEE Trans. Syst. Man Cybern.*, vol. 22, no. 3, pp. 418–435, May 1992.

[84] A. Cheddad, H. Kusetogullari and H. Grahn, "Object recognition using shape growth pattern," Proceedings of the 10th International Symposium on Image and Signal Processing and Analysis, Ljubljana, 2017, pp. 47-52.

[85] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proc. IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[86] F.-J. Huang and Y. LeCun, "Large-scale learning with SVM and convolutional nets for generic object categorization," in Proc. Computer Vision and Pattern Recognition Conf. (CVPR'06), 2006.

[87] A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. Lang, "Phoneme recognition using time-delay neural networks," IEEE Trans. Acoust., Speech, Signal Processing, vol. 37, pp. 328–339, 1989.

[88] www.deeplearning.ai

[89] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural networks, vol. 61, pp. 85–117, 2015.

[90]Bengio, Yoshua; LeCun, Yann; Hinton, Geoffrey (2015). "Deep Learning". Nature. 521 (7553): 436–444.

[91] Bengio, Y.; Courville, A.; Vincent, P. (2013). "Representation Learning: A Review and New Perspectives". IEEE Transactions on Pattern Analysis and Machine Intelligence. 35 (8): 1798–1828.

[92] I. G. Maglogiannis, Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in EHealth, HCI, Information Retrieval and Pervasive Technologies. IOS Press, 2007.

[93]J. J. De Oliveira, C. de A. Freitas, J. M. de Carvalho, and R. Sabourin, "Handwritten word recognition using multi-view analysis," in Iberoamerican Congress on Pattern Recognition, 2009, pp. 371–378.

[94] F. A. Hadi, T. A. B. W, and K. N. Ramadhani, "Freeman Chain Code yang Dimodifikasi Handwritten Numeral Recognition by application of modified Freeman Chain Code Abstrak Pendahuluan Optical Character Recognition (OCR) Freeman Chain Code," 2014.

[95] Y. Perwej and A. Chaturvedi, "Neural networks for handwritten English alphabet recognition," arXiv preprint arXiv:1205.3966, 2012.

[96] E. Juharwidyningsih, C. Fatichah, and N. Khotimah, "Pengenalan Karakter Tulisan Tangan Angka dan Operator Matematika Berdasarkan Zernike Moments Menggunakan Support Vector Machine," vol. 2, no. 1, pp. 1–5, 2013.

[97] S. Zheng et al., "Sunspot drawings handwritten character recognition method based on deep learning," New Astronomy, vol. 45, pp. 54–59, 2016.

[98] J. Sadri, C. Y. Suen, and T. D. Bui, "New approach for segmentation and recognition of handwritten numeral strings," in Document Recognition and Retrieval XII, 2005, vol. 5676, pp. 92–101.

[99] S. B. Hallale and G. D. Salunke, "Twelve Directional Feature Extraction for Handwritten English Character Recognition," no. 2, pp. 39–42, 2013.

[100] Y. El-Sonbaty, M. A. Ismail, and K. Karoui, "New efficient algorithm for recognizing handwritten Hindi digits," in Document Recognition and Retrieval IX, 2001, vol. 4670, pp. 68–74.

[101] Y. Wang, X. Ding, and C. Liu, "Post processing for offline Chinese handwritten character string recognition," in Document Recognition and Retrieval XIX, 2012, vol. 8297, p. 829709.