



**ADDIS ABABA UNIVERSITY**  
**SCHOOL OF GRADUATE STUDIES**

*Amharic Sign Language Recognition based on Amharic*

*Alphabet Signs*

Nigus Kefyalew Tamiru

A Thesis Submitted to the Department of Electrical and Computer Engineering in Partial Fulfillment for the Degree of Master of Science in Computer Engineering

Addis Ababa, Ethiopia

March 16, 2018

**ADDIS ABABA UNIVERSITY**  
**SCHOOL OF GRADUATE STUDIES**

*Nigus Kefyalew Tamiru*

*Advisor: Menore Tekeba*

This is to certify that the thesis prepared by *Nigus Kefyalew*, titled *Amharic Sign Language Recognition based on Amharic Alphabet Signs* and Submitted in partial fulfilment of the requirements for the Degree of Master of Science in Computer Engineering compiles with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by Examining Committee:

	<u>Name</u>	<u>Signature</u>	<u>Date</u>
Advisor:	<i>Menore Tekeba</i>	_____	_____
Examiner:	_____	_____	_____
Examiner:	_____	_____	_____

I, the undersigned, declare that this thesis is my original work and has not been presented for a degree in this and in any other university, and that all source of materials used for the thesis have been fully acknowledged.

**Declared by:**

Name: Nigus Kefyalew Tamiru

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

**Confirmed by my advisor:**

Name: Menore Tekeba

Signature: \_\_\_\_\_

Date: \_\_\_\_\_

## Abstract

Sign language is a natural language mostly used by hearing impaired persons to communicate with each other. At present day, sign language explainers are used to eliminate the language obstacles between people who are hearing impaired and non-impaired one. However, they are very limited in number. So, automatic sign language recognition system is better to narrow the communication gap between hearing impaired and normal people.

This thesis work deals with development of automatic Amharic sign language translator, translates Amharic alphabet signs into their corresponding text using digital image processing and machine learning approach. The input for the system is video frames of Amharic alphabet signs and the output of the system is Amharic alphabets.

The proposed system has four major components: preprocessing, segmentation, feature extraction and classification. The preprocessing starts with the cropping and enhancement of frames. Segmentation was done to segment hand gestures. A total of thirty-four features are extracted from shape, motion and color of hand gestures to represent both the base and derived class of Amharic sign characters. Finally, classification models are built using Neural Network and Multi-Class Support Vector Machine.

The performance of each models, Neural Network (NN) and Support Vector Machine (SVM) classifiers, are compared on the combination of shape, motion and color feature descriptors using ten-fold cross validation. The system is trained and tested using a dataset prepared for this purpose only for all base characters and some derived characters of Amharic. Consequently, the recognition system is capable of recognizing these Amharic alphabet signs with 57.82% and 74.06% by NN and SVM classifiers respectively. Therefore, the classification performance of Multi-Class SVM classifier was found to be better than NN classifier.

**Key Words:** - *Amharic Sign Language, Fourier descriptor, Neural Network and Support Vector machine.*

*Dedicated to*  
*My Grandmothers*  
*And*  
*My Grandfathers*

## **Acknowledgment**

First of all, I would like to thank almighty God and his Mother for giving me the strength, peace of my mind, and good health to achieve whatever I have achieved so far and for guiding me all the way through.

I would like to express my sincere gratitude to my advisor Mr. Menore Tekeba for his consistent follow-up and his willingness to offer me his time and knowledge from the inception to the completion of this thesis. My sincere thanks also goes to staff of Menelik the II Preparatory School who have dedicated their time for supplying different materials or equipment when they were needed. Some special thanks goes to Mrs. Habtam and her some students for capturing the data set. I also wish to express my gratitude to the IT staff members for providing a training class.

My thanks extend to Daniel Zemene who has spent many hours for proof, experiment, reading and discussion. Thank you Hassen Seid for your constructive comments on my report.

The final word of thanks goes to people who are not mentioned in name but whose support helped me complete the study successfully. Thanks to all.

# Table of Contents

<b>CHAPTER ONE: INTRODUCTION .....</b>	<b>1</b>
1.1 BACKGROUND.....	1
1.2 STATEMENT OF THE PROBLEM .....	2
1.3 OBJECTIVES .....	3
1.4 RESEARCH METHODOLOGY .....	3
1.5 SIGNIFICANCE OF THE STUDY.....	4
1.6 SCOPE AND LIMITATIONS.....	5
1.7 ORGANIZATION OF THE REST OF THE THESIS .....	6
<b>CHAPTER TWO: LITRATURE REVIEW .....</b>	<b>7</b>
2.1 SIGN LANGUAGE .....	7
2.1.1 Manual Signs .....	7
2.1.2 Non-Manual Signs .....	8
2.2 ETHIOPIAN SIGN LANGUAGE (ETHSL) .....	8
2.3 SIGN LANGUAGE RECOGNITION (SLR).....	13
2.4 PREPROCESSING .....	14
2.5 SEGMENTATION.....	15
2.6 FEATURE EXTRACTION .....	16
2.7 PATTERN CLASSIFIERS.....	17
2.7.1 Artificial Neural Network .....	18
2.7.2 Support Vector Machine .....	22
2.8 MODELS OF NN AND SVM CONSTRUCTION .....	23
2.9 SUMMARY .....	23
<b>CHAPTER THREE: RELATED WORKS.....</b>	<b>24</b>
3.1 INTRODUCTION.....	24
3.2 SIGN LANGUAGE RECOGNITION SYSTEMS FOR FOREIGN LANGUAGES .....	24
3.3 AMHARIC SIGN LANGUAGE RECOGNITION SYSTEM .....	28
3.4 SUMMARY .....	30

<b>CHAPTER FOUR: DESIGNING AMHARIC SIGN LANGUAGE RECOGNITION SYSTEM.....</b>	<b>31</b>
4.1 INTRODUCTION.....	31
4.2 THE PROPOSED ASLRS ARCHITECTURE.....	31
4.3 SIGN LANGUAGE VIDEO ACQUISITION .....	33
4.4 VIDEO TO FRAME CONVERSION .....	33
4.5 IMAGE PREPROCESSING .....	34
4.6 SEGMENTATION OF SIGN CHARACTER.....	35
4.7 FEATURE EXTRACTION .....	36
4.8 TRAINING.....	41
4.9 MODEL CONSTRUCTION.....	43
4.10 TESTING.....	44
4.11 SUMMARY.....	44
<b>CHAPTER FIVE: EXPERIMENTATION AND RESULT DISCUSSION .....</b>	<b>46</b>
5.1 INTRODUCTION.....	46
5.2 DATA SETS.....	46
5.3 IMPLEMENTATION .....	48
5.4 EVALUATION.....	49
5.5 TEST RESULTS.....	51
5.6 DISCUSSION .....	58
<b>CHAPTER SIX: CONCLUSION AND FUTURE WORK .....</b>	<b>60</b>
6.1 CONCLUSION.....	60
6.2 FUTURE WORK.....	61
<b>REFERENCES.....</b>	<b>63</b>
<b>APPENDIX.....</b>	<b>68</b>
<b>APPENDIX A: SAMPLE DATA USED FOR SYSTEM DESIGN .....</b>	<b>68</b>
<b>APPENDIX B: MATLAB CODE.....</b>	<b>69</b>

## List of Tables

<i>Table 2.1: The six forms of ‘<math>\nu</math>’ and its Trajectories with types of motions.....</i>	<i>11</i>
<i>Table 3.1: System Efficiency Using Three Different Classifiers .....</i>	<i>26</i>
<i>Table 4.1: Model Equations for Amharic Sign Character .....</i>	<i>40</i>
<i>Table 4.2: Extracted Shape, Motion and Color features for the sign ‘<math>\nu</math>’. .....</i>	<i>41</i>
<i>Table 5.1: The collected ETHSL manual alphabets .....</i>	<i>47</i>
<i>Table 5.2: Test Result for Each Fold, NN and SVM.....</i>	<i>51</i>
<i>Table 5.3 : Accuracy, Precision, Recall and F-Score measure for NN Model. ....</i>	<i>53</i>
<i>Table 5.4 : Accuracy, Precision, Recall and F-Score measure for SVM Model. ....</i>	<i>55</i>

## List of Figures

<i>Figure 2.1: Amharic Sign Language Finger Spelling</i> .....	10
<i>Figure 2.2: Signs for the letter ‘ሀ’</i> .....	11
<i>Figure 2.3: ETHSL Word created using Manual Signs</i> .....	12
<i>Figure 2.4: Approaches in SLR</i> .....	14
<i>Figure 2.5: Architecture of a Back Propagation Neural Network</i> .....	20
<i>Figure 4.1: Architecture of the Proposed Amharic Sign Language Recognition System.</i>	32
<i>Figure 4.2: Original Video Acquired</i> .....	33
<i>Figure 4.3: Frames Collected from Video</i> .....	34
<i>Figure 4.4: Preprocessing Image procedures</i> .....	34
<i>Figure 4.5: Contour of Amharic Sign Character ‘ሀ’</i> .....	37
<i>Figure 4.6: Fitting centroids to best adjusted R-square for ‘ሂ’ sign Character.</i> .....	38
<i>Figure 4.7: Fitting centroids to best adjusted R-square for ‘ሃ’ sign Character.</i> .....	39
<i>Figure 4.8: Fitting centroids to best adjusted R-square for ‘ሄ’ sign Character</i> .....	39
<i>Figure 4.9: Fitting centroids to best adjusted R-square for ‘ሀ’ sign Character</i> .....	39
<i>Figure 4.10: Neural Network Model with One Hidden Layer</i> .....	42
<i>Figure 4.11: Multi-Class Support Vector Machine Network Model</i> .....	43
<i>Figure 4.12: Constructed Model for NN</i> .....	44
<i>Figure 5.1: Screen shot of the running prototype</i> .....	48
<i>Figure 5.2: Process of 10-fold cross validation experiment</i> .....	50
<i>Figure 5.3: NN and SVM classification accuracy results with bar chart</i> .....	52

*Figure 5.4: Model Comparison Using Different Performance Metrics ..... 57*

## **List of Algorithms**

<i>Algorithm 4.1: An algorithm to prepare the binary frame .....</i>	<i>36</i>
<i>Algorithm4.2: An Algorithm to calculate Fourier Descriptors .....</i>	<i>37</i>
<i>Algorithm 4.3 : Color Difference Measure .....</i>	<i>41</i>

## List of Acronyms

<b>AMSL</b>	Amharic Sign Language
<b>ANN</b>	Artificial Neural Networks
<b>ArSL</b>	Arabic Sign Language
<b>ASL</b>	American Sign Language
<b>BASL</b>	Bangla Sign Language
<b>BSL</b>	British Sign Language
<b>CHSL</b>	Chinese Sign Language
<b>ENAD</b>	Ethiopian National Association for the Deaf
<b>ETHMA</b>	Ethiopian Manual Alphabet
<b>ETHSL</b>	Ethiopian Sign Language
<b>FD</b>	Fourier Descriptor
<b>HCI</b>	Human Computer Interaction
<b>ISL</b>	Indian Sign Language
<b>KNN</b>	K-Nearest Neighbor
<b>MS</b>	Manual Sign
<b>NMS</b>	Non-Manual Sign
<b>NN</b>	Neural Network
<b>RGB</b>	Red Green Blue
<b>SASL</b>	South Africa Sign Language
<b>SLR</b>	Sign Language Recognition
<b>SVM</b>	Support Vector Machine
<b>2D</b>	Two Dimensional

# Chapter One: Introduction

## 1.1 Background

Language is a communication tool that is useful for different people. People express their ideas to the community in which they are living. One of the types of language is sign language which is a manual form of communication language used in the hearing challenged (hearing impaired) community [1]. When we say hearing challenged community, it comprises hearing challenged people, their families and other people who communicate with them. Sign language enhances communication and transfer information among the members of hearing impaired community. Sign Language is also needed for formal teaching and learning process for the hearing challenged students. In addition to these issues, those people use sign language to transfer their culture from one generation to the other.

Usually, sign language is identified by the country where it is used e.g., American Sign Language (ASL), British Sign Language (BSL), Indian Sign Language (ISL) and Ethiopian Sign Language (ETHSL) [2]. By its nature, it is not universal because each country has its own sign language. Even we may obtain differences in sign languages among regions of the same country since like spoken language, it is unique to a culture and has evolved over time. In other idea, some people may think that sign language is a signed version of spoken language, but this is completely a miss assumption because sign language by itself is a language that has its own finger spelling alphabet, grammar and vocabulary structure [2].

In Ethiopia, there are a lot of hearing impaired people [3]. These people can communicate each other and the hearing people by using Amharic sign language (AMSL), with reading and writing techniques [1, 2]. Relative to communication with reading and writing techniques, communication through AMSL is more preferable for hearing challenged to hearing challenged people, hearing challenged to hearing people and hearing people to hearing people. However, there is communication gap among those people because they don't have the skill of Amharic sign language. Therefore, to reduce the short coming of the communication based on Amharic sign language, the research work about AMSL may be very useful since it would make communication to be good among them.

Including hearing impaired people into the common work was very difficult in the world, mainly due to lack of communication. However, the gap is partially solved by different researchers who have been researching on methods to develop different applications about different sign language [4]. However, less developments and research works have been done in Amharic sign language. Moreover, the field is still hot research area and is not matured well. We believe that working on this issue will contribute to the growth of the research and development of Amharic sign language tools.

## **1.2 Statement of the Problem**

Nature has gifted voice for human beings. The voice allows them to communicate with each other. Therefore, spoken language becomes the key language of humans. Unfortunately, some of human beings don't possess this skill due to the lack of hearing [5]. Due to this, sign language is the alternative language for hearing impaired people. It is challenged for most people who are not familiar with it because they are not enable to communicate with the others without the signer.

According to [6], there are 8% disabled people in Ethiopia. Out of these, 2% counts for hearing impaired people. Irrespective of this number, still there is a clear communication gap between hearing impaired and hearing people. In order to fill this gap, different measures should be considered; among those measures sign language training to hearing people and/or coming up with a system that facilitates the communication of theirs.

Concerning to the second option different research works have been conducted to come up with a system that converts sign language to text or vice versa for different sign languages all over the world. For example, American Sign Language (ASL) [7, 8, 9], Indian Sign Language (ISL) [10, 11], South African Sign Language (SASL) [12], Chinese Sign Language (CHSL) [13] and Bangladesh Sign Language (BASL) [14]. In our country, Legesse Zerubabel [15] has attempted to develop a recognition system for Amharic alphabet signs which translates a given alphabet sign into text. The drawback of this work is that it only focuses on the recognition of selected ten basic alphabet signs from static images; and hence further research work should be done at least to recognize all the basic and some derived alphabet signs from video of Amharic sign language. Even though different required resources for sign language recognition in Ethiopia are in their infant stages, the main aim of our work is to come up with a recognition system that translates

the basic and some derived Amharic alphabet signs into equivalent alphabet. So, the system tried to investigate the fundamental problems that exist between the hearing impaired and hearing people. To this end, our work discusses the answer to the following research question:

- Can we have a system that correctly recognizes the basic and some derived Amharic alphabet signs and translate it to the alphabet from?

### **1.3 Objectives**

#### **General Objectives**

The general objective of this study is mostly to implement and test Amharic alphabet signs recognition system using NN and SVM for all the basic and some derived Amharic alphabets from their corresponding signs.

#### **Specific Objectives**

The specific objectives are:

- To prepare the data set or corpus from different signers for this thesis and future researchers.
- To implement and test Amharic sign alphabet recognition system using NN and SVM machine learning tools.
- To provide part of a solution to lessen the communication gap between hearing and hearing challenged people.
- To enhance the computational resources of Amharic sign language recognition system.

### **1.4 Research Methodology**

In order to conduct this research work, the methodologies mentioned below will be used to select and implement appropriate methods and techniques.

#### **Literature Review**

Literature review is the basic methodology to study related works. Before starting the actual work, a deep study was made in the literature written on this area to have a clear picture about the work. Different previous papers written on sign language were reviewed to understand the various techniques and methods of Amharic sign language recognition. We used these reviewed papers which will be conducted for the recognition and classification of other sign languages

such as American Sign Language, Indian sign language, South African sign language, Bangladesh sign language and Chinese sign language. Based on the information that is obtained from these papers, the tools and the Algorithms were selected to develop the application.

In addition, to get broader understanding and also to insure the recognition or classification accuracy, we studied Amharic alphabet signs in detail.

### **Data Collection**

In the course of this study, the first task is to follow how the signers spell Amharic alphabet signs one by one so that we have an understanding of the Amharic alphabet signs. Hence, this task is carefully employed in order to address the research objective and come up with the recognition system. Then we have used a mobile device to record videos of the Amharic alphabet signs which are used as data set to train as well as test the recognition system.

### **Tools**

Unlike the video capturing part of the proposed system, other parts such as, preprocessing, segmentation, feature extraction and classification process were done by MATLAB built function.

Micro soft visio 2010 and Microsoft office 2016 software were used for designing the system architecture and various tables or diagrams to prepare the document respectively.

### **Prototype Development and Evaluation**

To evaluate the system, the samples data which are collected from different signers were feed into the developed prototype. Accordingly, the system is evaluated by comparing its output against the actual classes of Amharic alphabet signs. After the prototype was tested with sample signs the accuracy of the classification and verification was calculated. The result will be analyzed and evaluated from which a conclusion and further works will be recommended.

## **1.5 Significance of the Study**

The development of Amharic sign language is still in an infant stage because it is not well studied. In order to address the special needs of the hearing challenged people, several researches need to be done.

This work successfully incorporates the first order and some derived Amharic alphabet signs translation into their corresponding alphabet texts. The system anticipated to provide various benefits for both hearing challenged and normal people. Some of the benefits include:

- The key aim of this research work is to come up with a recognition system for ETHSL. This has an important contribution towards reducing the communication gap that happens between hearing challenged and hearing persons.
- The system supports the hearing challenged individuals to be participated in community services for using such kind of applications, as it did for the hearing individuals.
- This current work supports the students to quickly learn Amharic alphabet signs because sign language training process in the school consumes a long time.
- As an academic exercise, this research increases the experience of other researchers about ETHSL in the research area.
- The corpus collected and prepared can help future researches in this direction.

## **1.6 Scope and Limitations**

### **Scope**

This application is a software solution to recognize all the first Amharic alphabet signs and the derived Amharic alphabet signs of ‘ሀ’, ‘ለ’, and ‘ሐ’ only. The second, the third, the fourth, the fifth, the sixth and the seventh orders of other basic Amharic alphabet signs recognition is out of the scope of this research except the orders of some selected characters.

### **Limitation**

This thesis work has the following limitations:

- This application recognizes only the basic family and the derived Amharic alphabet signs of ‘ሀ’, ‘ለ’, and ‘ሐ’.
- This application depends on SVM and ANN models only for recognition (classification) of sign characters.
- This research will not cover the recognition of most of the derived alphabet signs, bastard alphabet (ዲቃላ ሆህያት) and local number signs.

## **1.7 Organization of the Rest of the Thesis**

The remaining part of the thesis is organized as follows. The basic theory, concepts and the different issues related to Sign Language (Ethiopian Sign Language and non-Ethiopian Sign Language) and other relevant topics of its classification and recognition for better understanding of our research domain are discussed in Chapter two. A general overview characteristic of Ethiopian Sign Language is discussed in the same chapter. Finally, this chapter presents about the main parts of Amharic sign language recognition system such as, preprocessing, segmentation, feature extraction and classification.

Chapter three reviews related research works that has been done on the classification and recognition of ASL, ISL, CHSL, SASL and BASL by using Machine Learning, Specially SVM and NN. Lastly, the recognition systems that were done on Ethiopian Sign Language are reviewed.

Chapter four gives a detailed description of the architecture and design issues of our System. The basic activities of the system, their main operation and the specific activities are presented in this chapter.

In Chapter five, the implementation of the proposed system architecture and experimental results are discussed.

Finally, the conclusion was presented from the research discussion (study) and the recommendation, and some future works are provided in Chapter six.

## **Chapter Two: Literature Review**

In this chapter points related with sign language in general, Amharic sign language in particular, sign language recognition systems using image processing methods, the tools and techniques implemented in each phases of the recognition system represented.

### **2.1 Sign Language**

Communication is the means of exchanging information, views and expressions among different persons, in both verbal and non-verbal manner. Hand gestures are the non-verbal method of communication used along with verbal communication. A more organized form of hand gesture communication is the sign language [10]. Sign language is among many language types in which people with difficulty hearing can communicate with people of the same linguistic behavior and with the other hearing people.

In different countries, no sign language is the same. Each country developed its own sign language [1, 16]. Thus, different countries in the world have their own sign language, such as United States of America has American Sign Language (ASL), India has Indian Sign Language (ISL), United Kingdome has British Sign Language (BSL), South Africa has South African Sign Language (SASL), Austria has Austrian Sign Language (AuSL) and Ethiopia has Ethiopian sign language (ETHSL). This means that the communication among different countries' hearing impaired persons can be as difficult as the communication among different countries' Hearing persons.

The communication through spoken language is mainly conducted by verbal communication (the spoken words and their sounds). However, the communication through sign language proceeds by gestures. Therefore, sign language can be only understood by considering the gestures alone. Sign language has two basic components which are Manual Signs (MS) and Non-Manual Signs (NMS) [17, 18, 19].

#### **2.1.1 Manual Signs**

Manual Signs (MS) are the basic components that form a sign language [17, 18, 19]. These are performed using hand and arms only. Hand shape (hand-form), hand motion and hand position with respect to other body parts are manual signs. They can also be used to create words and/or sentences using isolated signing and continuous signing respectively.

### **2.1.2 Non-Manual Signs**

Even though manual signs include a large portion of sign language vocabulary, Non-manual signs (NMS) are significant to convey the information of signs. Unlike manual signs which are performed by hand and arms only, non- manual signs are conducted by facial expressions and also by body movements. These include head movement, facial expression (smile and anger), raised eye brow, eye shift and so on. Facial expressions can be categorized in to two parts which are lower and upper facial expressions. The former conveys information about a specific sign by using the mouth area such as, cheek and lips. The later used head and body movement for information expression and focuses on a sign or sentence types (i.e., negation, question, etc.) [17, 18, 19].

## **2.2 Ethiopian Sign Language (ETHSL)**

ETHSL was developed from Amharic language since Amharic is the official language of Ethiopia. As other known Sign Languages (SLs), it was acknowledged as minority language that coexist with majority languages which are spoken languages [17]. ETHSL shares many similarities with spoken language (sometimes called “oral language”). Both are considered as a natural language, but there are also some significant differences between them.

According to the 1994 Housing and Population Census of Ethiopia, there were 1, 90,220 hearing challenged and hard of hearing people and the most are youngsters [20]. These people live similarly to any other people within their given cultures; however, they are severing from good interaction with them. The most hearing challenged youngsters live in countryside areas where there are no schools for themselves. Because of this the majorities are uneducated and spend their lives in extreme isolation. They also consider themselves as mentally deficient and evil because of lack of speech [21, 22]. In towns, more awareness has been made regarding the hearing challenged people. Even though the resources available are not enough for the number of students, parents are eager to send their children to schools. This helps the hearing challenged people to be less isolated.

Initially [1, 6, 23, 15], ETHSL derived from American Sign Language. Even if it was derived from American Sign Language, there is also some influence of Nordic countries sign language such as Finish Sign language. Some local signs were created and used in specific hearing impairment people’ schools in the country and later incorporated into Ethiopian Sign Language.

Some of the examples of these words are ‘አንጂራ’, ‘ሀበሻ’, ‘ሬሳ’, etc. It was first taught in Ethiopia by American missionary and is based on American Sign Language (ASL) and signed English [21]. To be suited for Ethiopian culture, it has been modified but may still be intelligible with ASL. Then, after Ethiopia has developed its own sign language, which comprises finger spelling, notations, structure and signing hand from hearing challenged people community.

### **Amharic Sign Language Finger Spelling**

Ethiopian National Association for the Deaf (ENAD) [1, 23, 15] developed Ethiopian Finger Spelling in 1971 and later gained acceptance by the Ministry of Education. These are called Ethiopian Manual Alphabets (ETHMAs). According to [1], it has 33 base alphabet signs where each base alphabet has 6 other variations that are created with the same hand shape or form the base alphabet followed by unique hand movement for each variation. In 2009, additional one signed alphabet was set for the Amharic letter Ve/’ሸ’. So, now ETHSL has 34 basic Manual Alphabets.

Amharic finger spelling is different from American, Britain and Indian Figure spelling. In [24, 25], ASL finger spellings do not require motions for most of the letters to be performed by a dominated single hand. However, the two alphabet signs (‘J’ and ‘Z’) need movement. In British [26], finger spelling involves both hands but one letter (‘C’) conducted by a single hand. This is the opposite to the American Sign Language (ASL) and Ethiopian Sign Language, where letters’ signs are represented by single hand shape. Like ASL, British finger spellings do not need motion for most of the letters. Indian Sign Language [27] used two handed to show the signs of lower and upper case English Alphabets and hence it is more complex compared to single handed ASL. In ISL, the most alphabets do not require a unique hand motion but the rest of two alphabets (‘H’ and ‘J’) signs need movement.

In Ethiopia, the first orders Amharic alphabets are “Ge’ez” which do not need any movement and the other orders requires motion to show their own corresponding signs. Like American Finger Spelling, Ethiopian Finger Spelling uses a single dominated hand for all alphabet signs [2, 23, 15]. In Figure 2.1 [3], the Ethiopian finger spelling with the corresponding Ge’ez alphabet except ‘ሸ’ is depicted bellow.

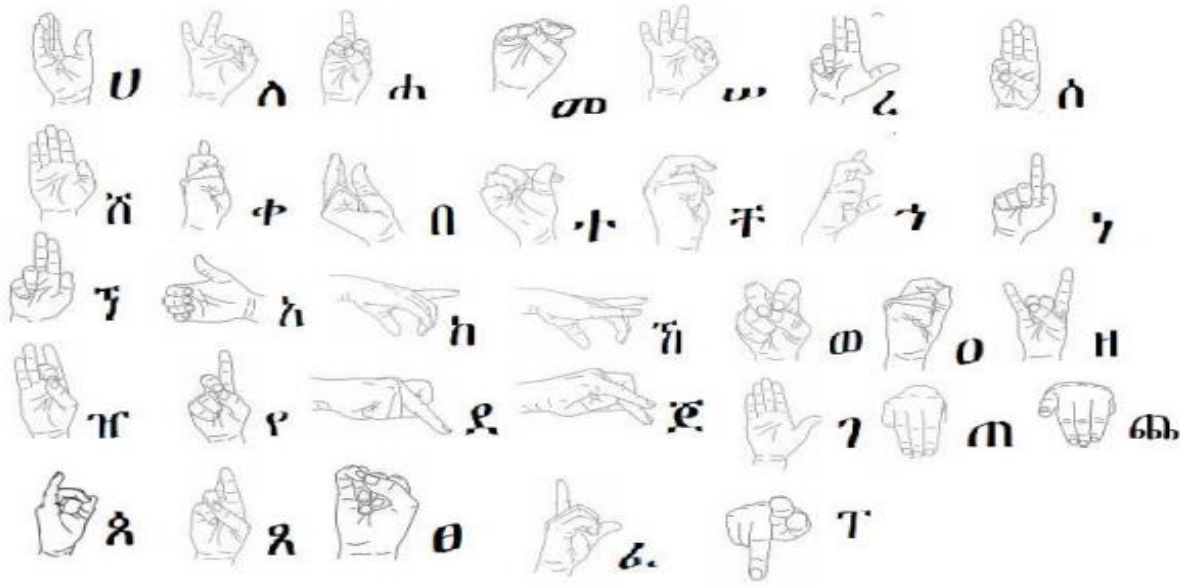


Figure 2.1: Amharic Sign Language Finger Spelling

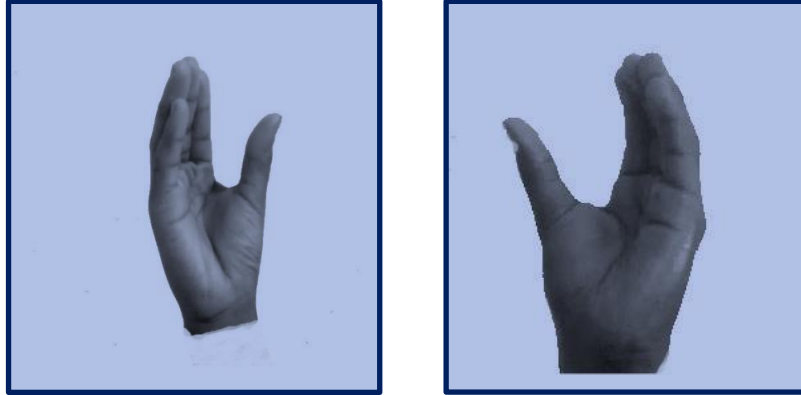
As we followed the training class about sign language in Menelik II preparatory school, finger spelling used to spell words that have no signs. If the word has three characters, it will be represented by the combination of three spelling signs, one for each letter. For example, for the word “ደረባ”, the signer firstly would perform the sign for the letter, ‘ደ’, then ‘ረ’ and finally ‘ባ’.

In many cases like names of persons, countries, cities and some other common words, their signs are first letter based. The sign for the name “አበበ” for instance, is the sign for ‘አ’ and touching part of our face. First alphabet-based signing is not limited to names only however the most common words like “አንጀራ” is also described by using the sign of ‘አ’ and adding some movement that shows the method of baking “አንጀራ” [1, 28].

### Amharic Sign Language Notations

Amharic sign language notations have four basic parameters: hand shape, hand movement, hand orientation, hand location [28].

**Hand shape:** The hand fingers particular orientation. When the signer is right handed, he/she conducts the signs by the right hand. On the other hand, when the signer is left handed, he or she performs the signs by the left hand. The right and left-handed signers are able to illustrate a sign on the same letter [28]. Figure 2.2 below depicts the first order Manual Alphabet (ETHMA) of ‘ሀ’ by the right and left dominated hand.









a. Right handed sign

b. Left handed sign

Figure 2.2: Signs for the letter ‘v’

**Hand movement:** The hand movement creates variation in the letters to be expressed by a provided sign. Amharic sign language uses seven orders of alphabet which have similar sign but differentiated by the movement of the hand shape to make an alphabet [28]. Table 2.1 shows the movement of hand to create the alphabet variants of ETHMA ‘v’.

Table 2.1: The six forms of ‘v’ and its trajectories with types of motions

Order	2 <sup>nd</sup>	3 <sup>rd</sup>	4 <sup>th</sup>	5 <sup>th</sup>	6 <sup>th</sup>	7 <sup>th</sup>
Form	v	v	v	v	v	v
Trajectory						
Direction	Left	Right	Down	Nearly Circle	Down Oscilatory	Rotation

**Orientation:** Orientation is the direction of signs. It is useful on signing because it may change the meaning. A sign can have one meaning when it moves one way and another meaning when it moves the other way [28, 29]. For example, to say “give” need the give sign from you to other body and to say “take” need the same with the reverse direction.

**Location:** This is the signing area when the sign takes place. The signs which are similar with other signs only distinguish in their location [29]. For example, signing the word "father" and the word "mother" in AMSL have the same hand shape except for the sign for "father" is signed

locating hand at the fore-head and that of "mother" at the chin. Figure [29] illustrates the Father's and Mother's signs.



a. Sign for "Father"



b. Sign for "Mother"

Figure 2.3: ETHSL Word created using Manual Signs

### Signing Hand

As explained earlier in the above, sign language is a visual language which uses our body parts to convey meanings. It combines hand shapes, orientation, movement of the hand, and facial expression. The main component in signing is hand shape which is constructed by changing the shape of our hands either using both or one hands.

Depending on the signer, the two hands can be classified as dominant and non-dominant hand. If the person is right handed, his right hand is dominant hand. On the other hand, if the person is left handed, his left hand is dominant hand [1, 21, 23]. There are three types of signs based on the use of hands. These are:

- **One handed signs:** are conducted by a single right or left dominant hand only. for example, Amharic alphabet signs.
- **Symmetric two handed signs:** Use both the right and left hands and these hands move the same way.
- **Asymmetric two handed signs:** Like symmetric two-handed signs, asymmetric two-handed signs use two dominant and non-dominant hands however, the dominant hand moves and the other non-dominant hand remains stationary.

## Amharic Sign Language Structure

Like Amharic spoken language, Amharic sign language has a common word order in sentence structure. However, the sentence structure in sign language is different from spoken language. Most of the time, the general syntax of a sentence in Amharic spoken language is “Subject” +” Object” +” Verb” but in Amharic sign language is “Time” (optional) +” Topic” +” Comment” where Time may indicate present, past and future tense. The comment is a word or a phrase that describe the topic. For example, “እኔ ዳቦ በላሁ”. This sentence is structured based on spoken language. On the other hand, the equivalent sentence in sign language is “እኔ ሙብላት ዳቦ”. In this example, the topic represents “እኔ ሙብላት” and “ዳቦ” is the comment [21, 23].

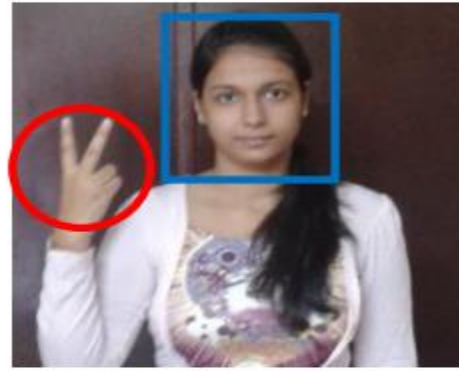
## 2.3 Sign Language Recognition (SLR)

Sign language recognition systems are being developed in order to provide an interface for the hearing-impaired persons. These allow the non-signers to interpret the meaning of what the signer wants to convey and therefore smoothing the communication between them [2, 23, 15]. The main approaches used in sign language recognition system can be classified as: Device based and vision-based approach [30].

- **Device based recognition:** In device-based approaches [30], devices such data gloves are needed. Data devices measure the hand shape. These approaches are difficult to signer because they need cumbersome devices to be worn by the signer and also limit their movement. However, this method reduces the computation of hand segmentation.
- **Vision Based recognition:** This technique [30] is mostly needed in the case of SLR. It uses image features like color, and shape. Even though, Vision Based technique introduces several challenges, such as the video pre- processing and segmentation of the hand, it provides more freedom to the signer by accommodating natural interaction. The examples of the two approaches are shown in Figure 2.4 [30].



a. Device based



b. Vision based

*Figure 2.4: Approaches in SLR*

The other main types of SLR are isolated and continuous sign recognition[30]. The former type is concerned with the recognition of single signs without continuation to the other sign. These signs can be either static or dynamic. No other sign is done before or after the isolated signs in continuation. Thus, this type of sign is not affected by the preceding or succeeding sign. In continuous signing, a complete sentence is recognized containing of various signs conducted one after the other. The objective is to identify different signs being conducted continuously.

There are varied techniques which used for recognition of sign language. Different authors have used different techniques according to the nature of sign language and the signs considered. In this work, Amharic sign language recognition is mainly handles four basic techniques: These are image preprocessing, Segmentation, feature extraction and classification methods. Let us see each method in detail below.

## **2.4 Preprocessing**

The first task before proceeding to any image preprocess is the collection of the input data in appropriate manner. We captured data set in a day that helped us to reduce computational cost and the noise of data. These are suitable to preprocess the data set easily [15]. After capturing the video database from different signers, the database should be transferred onto a computer. Then, the transferred video is automatically converted into the sequence of many frames. To be easy for further segmentation process, the frames are essentially cropped and enhanced since the quality of image was poor. Besides, smoothing and sharpening tasks should be performed on the sequence of frames in this step. Generally, the main aim of pre-processing is an improvement of

the image data that successfully removes unwanted distortions or enhances some image features relevant for further processing task.

## 2.5 Segmentation

Segmentation refers to the process of partitioning an input image into meaningful segments (regions/set of pixels) to be easily analyzed for further feature extraction [31]. This means it successfully separates the foreground object from background. The name of the operation comes from the result of the operation, which are segments. It is a process found between image pre-processing and image feature extraction techniques.

Image segmentation is one of the steps involved in the process of recognition in our case Amharic sign language recognition. For any vision-based application like sign language recognition, gray scale and binary images hold less information than color images. However, they need low computational cost, low computational complexity, and low computational time and also they do not require special hardware device relative to color images [15].

There are some of the well-established segmentation techniques [32]. Here we will discuss about threshold technique which is one of the important and the simplest technique in image segmentation. From a grayscale image, thresholding can be needed to form a binary image. This method replaces each pixel in an image with a black or white pixel if the image intensity  $I_{i,j}$  is less than some fixed constant  $T$  (that is,  $I_{i,j} < T$ ), or a white or black pixel if the image intensity is greater than that constant. The two important types of thresholding are global and local or adaptive [33]. The global thresholding uses a single threshold value for the entire image based on an estimation. The local or adaptive thresholding use multiple threshold values for each section of the image. It has been shown that Niblack's locally adaptive technique produces the best result. However, sometimes unnecessary noises such as holes created on the binary image. To successfully remove those, morphological operations are more appropriate [2]. The most basic morphological operations are dilation and erosion. Dilation adds pixels to the boundaries of objects in binary image, while erosion removes pixels on object boundaries.

## 2.6 Feature Extraction

The other most important step in recognition of sign language is feature extraction which is good to provide feature vectors as input to classifier for sign language recognition. In this step, important features are extracted after hand segmentation has been carried out successfully.

There are several types of image features that have been used for image classification. Shape, motion, color etc. are some of the basic image/frame features [32]. Selecting the right set of features is the crucial key in order to avoid uncertainty in all pattern recognition systems. The right features should discriminate between any patterns in the sample as an item of one category. In case of Sign language, it's necessary to choose features that uniquely identify the shape of Manual Alphabet Signs [34]. The description of shapes is generally categorized into two groups. These groups are:

***Region Based Shape Description:*** In a region-based approach [35, 36, 37], all the pixels within a shape are used to get the shape representation. For image recognition, region-based shape features extraction approach is better shape representation and be employed to describe non-connected or disjoint shapes but larger computational complexity than contour-based approach. Since it combines information across an entire object, rather than exploiting information just at the boundary points.

***Contour Based Shape Description:*** In contour or boundary-based approach [35, 36], only the contour information (pixels) of a shape is taken into account to get the shape representation. This technique is simple, or more popular because of less computational time. The most common contour-based shape descriptors wavelet and Fourier descriptors. Wavelet descriptors involve intensive computation due to the dependency of rotation. Fourier descriptors (FDs) improve the weak discrimination ability of Wavelets descriptors and also easily normalized [35, 36]. They are derived from the Fourier transform on a shape signature [35, 36, 37, 38]. Shape signature is a one-dimensional vector describing the outline of a two-dimensional shape. Even there are some types of shape signatures, the best results were obtained using the complex coordinates.

As the junior signer wants to illustrate signs in real time, problem of slight variation in angle (rotation), signer hand size (scaling) and translation (direction) need to consider while extracting the features. These problems are appropriately solved by Fourier descriptors. N points forming the boundary of a hand shape are required by taking all the pixels occupied by the

boundary or taking  $N$  samples from the boundary [37, 38]. This can be done by tracing the boundary counter clockwise. The coordinates of each point in the shape contour can be stated in the form of  $(x_k, y_k)$  where  $0 < k \leq N-1$ . The contour can be expressed by coordinate series as follows:

$$\mathbf{S}(\mathbf{k}) = [x(\mathbf{k}), y(\mathbf{k})], \text{ for } \mathbf{k} = \mathbf{0}, \mathbf{1}, \mathbf{2}, \dots, \dots, \mathbf{N} - \mathbf{1} \quad (1)$$

And each coordinate is a complex number. One of the easiest ways to represent the two-dimensional (2D) contour as a signature is to represent each coordinate pair as a complex number such that

$$\mathbf{S}(\mathbf{k}) = x(\mathbf{k}) + jy(\mathbf{k}) \quad (2)$$

Where  $\mathbf{S}(\mathbf{k})$  is the shape signature and  $x(\mathbf{k})$  and  $y(\mathbf{k})$  are the point coordinates. The advantage of this representation is that it reduces a 2D into a 1D (one-dimensional) problem. The Discrete Fourier Transform (DFT) of this sequence is the FD of the contour. The DFT of  $\mathbf{S}(\mathbf{k})$  is:

$$\mathbf{A}(\mathbf{u}) = \sum_{\mathbf{k}=0}^{\mathbf{N}-1} \mathbf{S}(\mathbf{k}) e^{-j2\pi\mathbf{u}\mathbf{k}/\mathbf{N}} \quad (3)$$

Where  $\mathbf{u} = 0, 1, 2, 3, \dots, \mathbf{N}-1$  and  $\mathbf{A}(\mathbf{u})$  is the complex coefficient or Fourier Descriptor (FD) of the boundary.

The high frequency descriptors contain information about finer details of the shape while the low frequency descriptors contain information about the general or global features of the shape [38, 39]. However, it is found that all the Fourier coefficients are not necessary to reconstruct the boundary because the computational process will be complex. Thus, the optimal number of FDs be selected for hand shape feature extraction.

### **Motion Features**

In addition to hand shape features, different activities were conducted to extract motion information by following the trajectory of the hands' centroid. As indicated by Samuel Teshome [17], a gesture path pattern which holds centroid points  $(x_{\text{hand}}, y_{\text{hand}})$ . So, the trajectory or curve is determined between two consecutive points from hand gesture path.

## **2.7 Pattern Classifiers**

Pattern classifiers can be defined as the ability to classify a physical object to one of several pre-specified classes [40]. Pattern classification which has an application in system like hand detection, face detection and sign recognition [15]. In the recognition of patterns [32], there are

two types of classification techniques which are basically supervised and unsupervised classification.

***Supervised classification:*** It is a commonly used technique on the given application and trying to predict the results for known examples (samples) [41]. It compares its prediction result to the target answer and learns from its error or mistakes. The aim of this supervised learning is to generate a classifier model to perform a right prediction for any provided input data.

***Unsupervised classification:*** Unsupervised learning is the most effective technique for expressing data rather than predicting it [41]. The basic aim of unsupervised learning is to find the structure in the available data and build the needed cluster for classification. There are no labels on the input data to find the suitable classification because of this it is harder to get the right classification.

In order to classify the signs' features which are given as an input to the system into different classes, a machine learning classifier is needed [42]. It is a branch of artificial intelligence. The methods or techniques of a system can be developed for enabling the computer to learn through machine learning. To perform this work, there are different machine learning classifiers. These are Hidden Markov Model (HMM), Bayesian Classifier, Template Matching, Boosted Classifier, K-Nearest Neighbor (KNN), Artificial Neural Network (AAN), Support Vector Machine (SVM) etc. In this Section, we will only discuss about Artificial Neural Network and Support Vector Machine pattern classifiers in detail.

### **2.7.1 Artificial Neural Network**

Artificial neural networks(ANNs) [43] are derived from the observation of basic building block of biological neural networks (BNNs) where biological neurons include soma, dendrites and axon components; on the other hand an artificial neuron handles inputs, weights, transfer function, bias and outputs. In biological neuron, information comes into the neuron via dendrite; soma processes the information and passes it on via axon. In the case of artificial neuron, the information comes into the body of an artificial neuron via inputs that are weighted. The body of an artificial neuron then sums the weighted inputs, bias and processes the sum with a transfer function. At the end an artificial neuron passes the processed information via outputs.

ANNs are highly distributed interconnections of adaptive non-linear processing elements which mean they are large set of interconnected neurons. The set of interconnected neurons execute in parallel to compute the task of learning. Therefore, ANN seems like human brain into two respects. The former respect is that knowledge is acquired by the network through a learning process. On the other hand, interneuron connection strengths known as weights are used to store knowledge, i.e., the weights on the connections encode the knowledge of a network. The neurons are modeled after the biological neurons and hence they are termed as neural networks (NNs) [32, 44].

Distributed computation of ANN has the advantages of reliability, fault tolerance, high throughput (division of computation tasks) and cooperative computing. The adaptation is the ability to alter a system's parameters according to some rule (normally, minimization of an error function). Adaptation enables the system to search for optimal performances. The ANN property of nonlinearity is also important to produce more powerful computation when compared to linear processing [32]. There are two basic phases in neural network operation. They are training (learning) phase and testing (recall or retrieval) phase. In the learning phase, data is repeatedly presented to the network, while weights are updated to obtain a desired response. In testing phase, the trained network with weights is applied to data that it has never seen.

ANN [43] has three simple sets of rules: multiplication, summation and activation. At the entrance of artificial neuron, every input value is multiplied with individual weight. In the middle Section of artificial neuron is sum function that sums all weighted inputs and bias. At the exit of artificial neuron, the sum of previously weighted inputs and bias is passing through activation function that is also called transfer function.

Although there exist many representations of ANNs, each one of these networks possesses four attributes  $\langle N_c, W, \sigma, \delta \rangle$ , where  $N_c$  is a finite set of highly interconnected neurons with outputs  $n_1, n_2, \dots, n_k$ ;  $W$  denotes a finite set of weights which represents the strength  $w_{ij}$  of the interconnection between neurons  $n_i$  and  $n_j$ ;  $\sigma$  is a propagation rule which shows how the input signals to a neuron  $n_i$  propagates through it. A typical propagation rule may be  $\delta(i) = \sum_j w_{ij}$  and  $\delta$  is an activation function which is usually a nonlinear function like sigmoid function [32]. The most popular neural network topology is the Multilayer Perceptron (MLP), which is an extension of the single layer perceptron proposed by Rosenblatt [44]. Multilayer perceptions, in general,

are feed forward network, having distinct input, output, and hidden layers. The architecture of multilayered perceptron with error back propagation network is shown in Figure 2.5 [32].

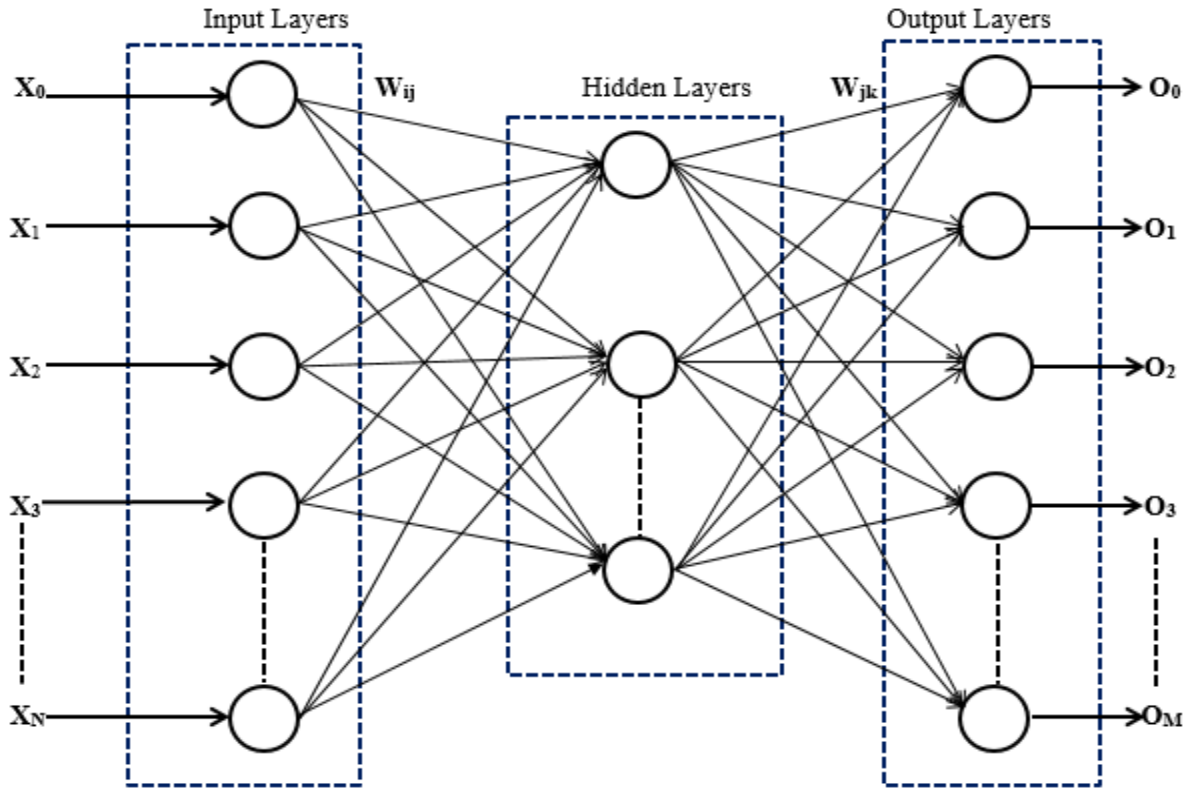


Figure 2.5: Architecture of a Back Propagation Neural Network

In an  $M$ -class problem where the patterns are  $N$ -dimensional, the input layer contains of  $N$  neurons and the output layer contains of  $M$  neurons. There can be one or more middle or hidden layer(s). The above Figure illustrates a single hidden layer case, which is extendable to any number of hidden layers. The output from each neuron in the input layer is fed to all the neurons in the hidden layer. No computations are done at the input layer neurons. The hidden layer neurons sum up the inputs and pass them through the sigmoid non-linearity and fan out multiple connections to the output layer neurons.

In feed forward activation, neurons of the first hidden layer compute their activation and output values and pass these on to the next layer as inputs to the neurons in the output layer, which produce the networks actual response to the input presented to neurons at the input layer. Once the activation proceeds forward from the input to the output neurons, the network's response is compared to the desired output corresponding to each set of labeled pattern samples belonging to

each specific class, there is a desired output. The actual response of the neurons at the output layer will deviate from the desired output, which may result in an error at the output layer. The error at the output layer is used to compute the error at the hidden layer immediately preceding the output layer and the process continues [32].

In view of the above, the net input to the  $j^{\text{th}}$  hidden neuron is expressed as:

$$I_j^h = \sum_{n=1}^N X_n W_{ij}^h + \theta_j^h \quad (4)$$

The output of the  $j^{\text{th}}$  hidden layer neuron is:

$$O_j = f_j^h(I_j^h) = \frac{1}{1+e^{-I_j^h}} \quad (5)$$

Where  $x_1, x_2, \dots, x_n$  is the input pattern vector, weights  $w_{ij}$  denotes the weight between the hidden layer and the input layer, and  $\theta_j^h$  is the bias term associated with each neuron in the hidden layer. These calculations are known as forward pass. In the output layer, the desired or target output is set as  $T_k$  and the actual output gained from the network is  $O_k$ . The error ( $T_k - O_k$ ) between the desired signal and the actual output signal is propagated backward during the backward pass. The equations governing the backward pass are used to correct the weights.

Thus, the network learns the desired mapping function by back propagating the error and hence the name error backpropagation. The average error  $E$  is a function of weight as shown below:

$$E(W_{jk}) = \frac{1}{2} \sum_{k=1}^M (T_k - O_k)^2 \quad (6)$$

To minimize the error  $E$ , we have to find the root of the partial derivatives

$$\sum_{k=1}^M \frac{\partial E}{\partial W_{jk}} = 0 \quad (7)$$

Hence, from this we can obtain the value of updated weights as follows

$$W_{jk}^{(new)} = W_{jk}^{(old)} + \eta \delta_j O_j \quad (8)$$

where  $\eta$  is the learning rate of the hidden layer neurons.

In summary, artificial neural networks can be regarded as an extension of many classification techniques, which have been developed over several decades. These networks are stimulated by the idea of the biological nervous system and have verified to be robust in dealing with the ambiguous data and the kind of problems that require large amounts of data. Instead of

sequentially performing a program of instructions, neural networks explore many hypotheses simultaneously using massive parallelism. Neural networks have the potential for solving problems in which some inputs and corresponding output values are known, but the relationship between the inputs and outputs is not well understood or is difficult to translate into a mathematical function [32, 44].

### **2.7.2 Support Vector Machine**

Support Vector Machine (SVM) is a machine learning algorithm which is used for binary classification. It was proposed by Vapnik [9, 30, 42, 45]. The basic idea of SVM is to find an optimal separating hyper plane (a separating plane of dimension  $n - 1$  where  $n$  is the number of features defining a data point) between positive and the negative classes. So, it correctly separates the classes even for large data set with small training samples and also makes a good generalization. Due to this, SVM gains popularity among world's people. The parameters of SVM are the orientation of hyper plane and the distance from the origin of the hyper plane.

In a binary classification, given linear separable training set  $\{x_1, x_2, \dots, x_n\}$  with their labels  $\{y_1, y_2, \dots, y_n\}$ ,  $y_i \in (-1, 1)$ , the SVM binary classifier is trained and the optimal hyper plane is yielded, which separates the data by a maximal margin. The optimal hyperplane divides the data points into two groups. Points lying on negative side are labeled as -1, and the other points on positive side are labeled as 1. When a new example is inputted for classification, a label (1 or -1) is issued by its position with respect to the hyperplane [37, 45].

For the case of data that is inappropriate (non-linearly separable), the SVM first maps the original data to a higher dimensional space by a kernel function  $K$ , such as, Radial Basis Function (RBF) and polynomial kernels, and then it linearly separates them [37, 45, 46, 47].

Basically, SVM was developed for binary classification and later on it was extended to solve multi class problems using “one-against-one” or “one-against-all” (one-versus-rest) strategy. The one-against-one strategy which builds one SVM for each pair of classes. Another popular method is one-against-all strategy (standard method) that comprises of constructing one SVM per class, which is trained to distinguish the samples of one class from the samples of all remaining classes and it is significantly more accurate for classification. In multi class SVM, multiclass labels are decomposed into several two class labels and it trains the classifiers to solve

the problems and the solution of multi class problem is reconstructed from the outputs of classifiers. Finally, multi-class SVM model generated from known samples [9, 37].

## **2.8 Models of NN and SVM Construction**

The two most important models are constructed or generated by Neural Network and Support Vector Machine with feature vectors during training phase. After the Models were generated or constructed, different cross validation techniques such as hold out, K-fold etc. are implemented to illustrate the correctness of them. Among these different cross validation techniques, k-fold cross validation ( $k = 10$ ) is the most common used for training and testing the two models [48].

## **2.9 Summary**

In this chapter, general overview regarding to Sign language have been presented. The notations, finger spelling, signing hand and structure of Amharic sign language are also discussed. In addition to this, the two basic approaches (device and vision based) of sign language recognition are introduced. Theoretical backgrounds about digital image processing that are applied to Amharic sign language recognition system such as preprocessing, segmentation, feature extraction, and classification are also presented.

Finally, ANN and SVM classifiers which have been used in Amharic sign language recognition system are described in detail and also their model's construction is presented.

## **Chapter Three: Related Works**

### **3.1 Introduction**

Researchers have been trying to design and develop Sign Language Recognition Systems (SLRSs) for different languages. This is due to its importance for narrowing the communication gap that exists between hearing and hearing challenged individuals. In this chapter, related works on sign language recognition will be presented. The review focused on recognition of American, Indian, South African, Chinese, Bangladesh and Ethiopian sign languages. Some of the recognition systems were done on static images while the others were done on videos.

### **3.2 Sign Language Recognition Systems for Foreign Languages**

#### **American Sign Language Recognition**

The authors in [7] proposed the Performance Analysis of KNN, SVM and ANN Techniques for Gesture Recognition System. This designed system can recognize exact human gestures in ASL. Five thousand gesture data sets were captured by utilizing cameras along with uniform background for this research. The system is capable of recognizing only numerical ASL static signs with 97.10%, 92.093% and 94.4% accuracy by SVM, KNN and ANN classifiers respectively. From this accuracy results, we can observe that the SVM method gives better performance related to other two classifiers.

The other researchers attempted to implement a vision based recognition of finger spelled American Sign Language alphabets in an automated manner [8]. It was designed based on static gestures which are the selected ten ASL alphabets namely A, B, C, D, G, H, I, L, V and Y for recognition. The objective of finger spelling recognition system is to describe a simple and efficient mechanism to translate the sign into text. This proposed system did not use the user to wear any data gloves for simplification of image processing and it gives human-machine interaction only through bare hand and also considered white and black background. In the work of this, feature extraction was the main task. For feature extraction process, Hough Transform that is an optimal edge detector is engaged since it finds both the locations and the quantity of features existing in the image and also it is reliable or efficient feature extractor. The other main task was gesture modeling and classification that was conducted through Support Vector Machine (SVM). The system is tested on 180 ASL dataset for recognition of individual alphabets

and achieved the overall recognition accuracy of 93.88%. Finally, the researchers recommended that the system to be extended for recognizing the hand gestures from complex background in real time.

### **Indian Sign Language Recognition**

The work in [10], is aimed to design Indian Sign Language Recognition System for hearing challenged people to convey their thoughts or ideas. The basic methods for classifying hand gestures for Human Computer Interaction (HCI) comprises Glove based and Vision based techniques. The system that uses bar hand gestures for recognition in the vision-based setup because it is suitable for both single and double handed gestures. As the method is implemented completely by using digital image processing technique so the user did not wear any special hardware device to get features of hand shape. Thus, a system recognizes Indian sign language (ISL) based on hand gestures and it allows the users to interact with the system in natural way. Then, the system can recognize 36 hand gestures of an Indian sign language which represents the English alphabets from A to Z and numbers from 0 to 9 via artificial neural network with feed forward back propagation algorithm which is used to train the network and translate them into text and voice. Finally, the system has been experimented with different gesture images collected by a web camera and succeeded a good result with an accuracy of 91.66%.

Additional paper was presented on the performance comparison of three different classifiers for human computer interaction (HCI) using hand Gestures in ISL [11]. This was developed based on Zernike moments (ZMs) with three classifiers KNN, ANN and SVM. For feature extraction, Anand Kulkarni and Sachin Urabinahatti selected Zernike moments which is region based feature extraction method because ZMs are translation (direction) and size invariant. The proposed approach is a robust hand gesture recognition system which is presented for recognizing static alphabet gestures irrespective of the angles in which the alphabet hand gestures were captured. Thus, the system developed here is real time hand gesture recognition. The main problem of gesture recognition depends in the complexity of the classification Algorithms, especially when using high dimensional feature vectors which become necessary in order to be able to distinguish several of gestures. Comparative study was carried out to show which classifier works better in reorganization of gestures among KNN, ANN and SVM. The comparative results of three classifiers are depicted in Table 3.1 [11] below.

Table 3.1: System Efficiency Using Three Different Classifiers

Classifier	System Efficiency
KNN	77.5%
ANN	82.5%
SVM	91%

As we see the above Table 3.1, the results show a good significant accuracy in this real time recognition system and we can conclude that the system works with better efficiency considering SVM as a classifier either than ANN or KNN classifiers.

### **South African Sign Language Recognition**

Vision-Based Static Hand Gesture Recognition System using SVM were proposed by S. Naidoo, C.W. Omlin and M. Glaser [12]. The major aim of their research is to develop a system that will enhance the interaction between hearing challenged and hearing communities in South Africa. South Africa has two groups of hand gestures, namely static and dynamic. A static gesture is a specific hand shape and represented by a single image. A dynamic gesture is a moving gesture, represented by a sequence of images. However, their research work focused on the design of a system that will recognize static hand images against complex backgrounds based on South African Sign Language (SASL). SVM classifier was used to classify hand postures as gestures, due to its high generalization performance without the need to add a priori knowledge, even when the dimensions of the input space are very high. The experimental results showed that the system has produced the recognition rate above 90%.

### **Chinese Sign Language Recognition**

In [13], the researchers developed Chinese Sign Language Recognition for a Vision-Based Multi-Features Classifier System. They focused on the Chinese manual alphabet which is composed by 30 hand gestures. At first, extracting features of letter images is done, and then classification method of SVMs for recognition is brought into use. Fourier descriptors and other multi-features were introduced in this work. The five image descriptors to describe numerous visual (color histogram and 7Hu moments) and geometrical (48 dimensional Gabor wavelet, 128 Fourier descriptors, scale invariant feature transform (SIFT)) properties of images. The first two

descriptors are computed from every pixel of the whole image, while the others descriptors are computed from small localized interesting image.

The researchers collected 195 images for each letter, 5850 in all by using a camera device. Experimentation with 30 classes of the Chinese manual alphabet images was conducted and the results proved that the features, such as Fourier descriptors, are simple, efficient, and effective to distinguish hand shape, and the SVMs method has excellent classification and generalization ability in solving learning problem with small training set of samples. This system was able to recognize images with 95.0256% accuracy when trained with 1500 images and tested with 4350 images. Finally, researchers recommended the idea of using videos as an input to a recognition system; obtaining more complete feature extractions for sign language image similarity characterization; and using Multi-kernel SVMs as classifier [13].

### **Bangladesh Sign Language Recognition**

Among the global/international research reviewed, Md. Atiqur Rahman, Dr. Ahsan Ambia, Md. Ibrahim Abdullah and Sujit Kumar Mondal [14] worked a research work titled Recognition of Static Hand Gestures of Alphabet in Bangla Sign Language (BASL). The system has the ability to recognize 36 selected letters of BASL alphabet using ANN with the idea of it being popular in speech recognition as well as in handwriting recognition. ANN was trained with features of sign alphabet using feed-forward back-propagation learning algorithm. This recognition system used the images of the bare hand for the recognition rather than any gloves. There are 23 images of each sign of BASL used in the recognition. Therefore, a total of 828 signs are collected. Among 828 images, 540 samples were used for training and the remaining 288 images were used for testing. The average recognition accuracy of the proposed system is 80.902%. Using more images or samples for training ANN may improve the performance of the system. The limitation of this paper is that, the feature vectors should have integer values only. Future work would include extending the developed method to recognition of BASL with video based system [14].

Md Azher Uddin and Shayhan Ameen Chowdhury [49] proposed another framework for recognizing Bangla Sign Language (BASL) using SVM. Here, first the original RGB images are transformed to Hue, Saturation and Value (HSV) color space. Then, features are extracted from the segmented image by Gabor filter and then Kernel Principal Component Analysis (KPCA) is applied to reduce the dimensionality since KPCA is a nonlinear dimensionality reduction technique. Finally, SVM is used to identify alphabets of the sign. In the experiments, they have

used 4800 images with the size 480×360 pixels. The images were captured from different illuminations. From this database, 2400 images were used for training and 2400 images were used for testing. At the end, the system attained 97.7% accuracy result.

### **3.3 Amharic Sign Language Recognition System**

One of the local researchers who is Menelik Tesfaye developed Machine Translation System [6]. His system translates Amharic text to an equivalent Ethiopian Sign Language (ETHSL) by finger spelling representation with the help a 2D animating avatar which is used as a signer. Translating Amharic text to ETHSL benefits the hearing challenged people because it helps them to understand what is stated in Amharic. It also supports the Amharic speakers to express what they want to say to the hearing challenged people. The author used Macromedia Flash 8.0 and Action Script 2.0 as a tool to model and design the avatar. This model is tested by 10 hearing challenged people and its own overall performance rate is 66.6%. In this work [6], the researcher did not concern about signing (conceptual sign), but only about finger. In addition, he was also only focused on a single word which means the research reads only a single word to translate into ETHSL finger alphabet expression which is not applicable for others, such as phrases.

The other work was conducted by Tefera Gimbi [23] in 2014. Unlike the above work [6] which is rule based system, his work is a machine learning based system that is Isolated Signs Recognition System in ETHSL. The system receives videos of Amharic word signs and produce extracted frames. On the sequence frames, skin color detection algorithm was implemented and the equivalent binary image is created that has white value on the foreground for skin color and black value on back ground for another region. Based on the detected skin regions the hands and head were segmented from other parts of the body since they have very important role in the signing process of Amharic words. Significant features which help to get more shape information and get a better recognition result are extracted from the segmented body parts. Then, Hidden Markov Model (HMM) was trained by using these feature vectors. Three signers participated to capture the data set in this work and each signer performed a sign twenty times. Out of these twenty videos, Tefera Gimbi used fifteen for training and the rest five video for testing purpose. The system has been tested using the videos which were collected for learning purpose. He attained the overall recognition of 86.9% using the HMMs learned by eight features (i.e., Area, Centroid, Bounding Box, Major Axis, Minor Axis, Eccentricity, Orientation and Perimeter)

whereas by using only three basic features (i.e., Centroid, Area and Orientation), he got the recognition of 83.5%.

According to [20], the two authors attempted to design a hand gesture recognition system by using Gabor Filter (GF) together with Principal Component Analysis (PCA) for feature extraction and ANN for recognizing the ETHSL of 34 letters of ETHMA. For the proposed system, 170 images were captured from five ETHSL students for 34 Amharic alphabets. In order to reduce the difficulty of segmentation due to high variation of skin color among the signers, the system used white glove while capturing the total dataset. Moreover, their work describes that non- ETHEMA images were captured from websites and the system is capable of rejecting these non ETHMA images. Generally, the basic purpose of this work is translating those ETHSL alphabet signs to voice. The experiments performed and proved that with sufficient data, ANN approach produced good results and able to recognize an unknown input sign (non ETHMA) very fast. The experimental results show that the system has produced the recognition rate of 98.53% [20]. However, the paper could not provide full translation functionality because it could not be used to translate basic and specific Amharic alphabet signs to the corresponding alphabets.

In addition to the above research done by [20], Legesse Zerubabel [15] continued to propose Ethiopian Finger Spelling Classification System. The system architecture includes different components which are image preprocessing, feature extraction, segmentation and classification. The system worked by accepting Amharic alphabet signs as inputs and returning the corresponding Amharic alphabet text as an output.

Different experiments were performed to get the appropriate feature extractors and pattern classifier for hand detection and sign classification tasks. Besides, through experiments, the capability of principal component analysis and harr-like feature with neural network were tested. Basically, the work focuses on hand detection and sign classification that was used as input for Ethiopian fingerspelling recognition. According to experiment results, the overall recognition rate of 88.08% and 96.22% were obtained using neural network with PCA driven feature and neural network with harr-like feature respectively [15]. This developed system is only limited to the recognition of ten basic Amharic alphabet signs based on images. Therefore, it is necessary

to extend the system so that it handles specific and the remaining basic Amharic alphabet signs with video data set.

Later, this research work was extended by Abadi Tsegay [2]. He attempted to develop offline candidate hand gesture selection and trajectory determination system for continuous ETHSL. This recognition system used extracted candidate ETHMA frames from the video sequence and that can also determine hand movement trajectories. The system has two basic components namely:

- Candidate Gesture Selection (CGS) and
- Hand Movement Trajectory Determination (HMTD).

The CGS combines two metrics speed profile of continuous gestures and Modified Housdorff Distance (MHD) measure and obtained an accuracy of 80.72% for this module. The HMTD is done by considering each hand gesture centroid from frame to frame and using angle, x- and y-directions, and it returned a result with an accuracy of 88.31%. In addition to this the system as a whole has a performance of 71.88%. However, he did not focus on the first and the seventh orders of Amharic alphabet signs except the second, the third, the fourth, the fifth and the sixth orders. He hasn't also worked on recognition of these signs to text.

### **3.4 Summary**

Most of the research works were done in the recognition of Amharic sign language mainly focused on the translation of sign to its equivalent text but the systems couldn't be tested directly to all basic and/or some specific alphabets.

Besides, researchers did not use Support Vector Machine classifier for Amharic sign language recognition and also did not evaluate the performance between ANN and SVM based on this area. According to the above research reviews, in other countries like America, India, etc. different researchers used SVM and ANN for Sign Language recognition and also compare SVM performance with ANN and achieved good performance. Because of this, we selected ANN and SVM to be used as classifiers for our research work.

# **Chapter Four: Designing Amharic Sign Language Recognition System**

## **4.1 Introduction**

As described in different research works, recognition of Amharic sign language into the corresponding character using digital image processing and learning machines passes through different phases: sign language video acquisition, video to frame conversion, frame preprocessing, segmentation, and training and recognition. In this chapter, we propose a design for recognition of Amharic Sign Language into Amharic character. The Sections in this chapter present details of each component in the architecture. Section 4.2 presents the general overview of the proposed system architecture. Section 4.3 presents Sign Language video acquisition techniques and tools used. Section 4.4 explains techniques used to convert video into frame. Section 4.5 briefly explains frame preprocessing. Section 4.6 explains techniques and algorithms used in segmenting sign character from the background. 4.7 explains the tool and techniques used to extract features from the segmented character. After features are extracted, selected learning machines are trained from which models are constructed and tested. Finally, a summary for the chapter is briefly presented in the last Section.

## **4.2 The Proposed Amharic Sign Language Recognition System Architecture**

The system architecture shown in Figure 4.1 is the proposed architecture for recognition of Amharic Sign language. It has six components working together: Video to frame conversion, Image preprocessing, Segmentation of sign character, Feature extraction, Training and Testing.

The first component converts acquired sign language video to frame images. The frame images are preprocessed to remove unnecessary noise in the second component. The third component segments sign character from the background. Fourier descriptor, motion descriptors and color difference measure are used to extract features in the fourth component. In the training component, extracted features are feed into learning machines (NN and SVM) to construct their respective models. The Models constructed are used to recognize sign language into print text in the testing component. In this component the performance of the models constructed (for NN and SVM) are also tested.

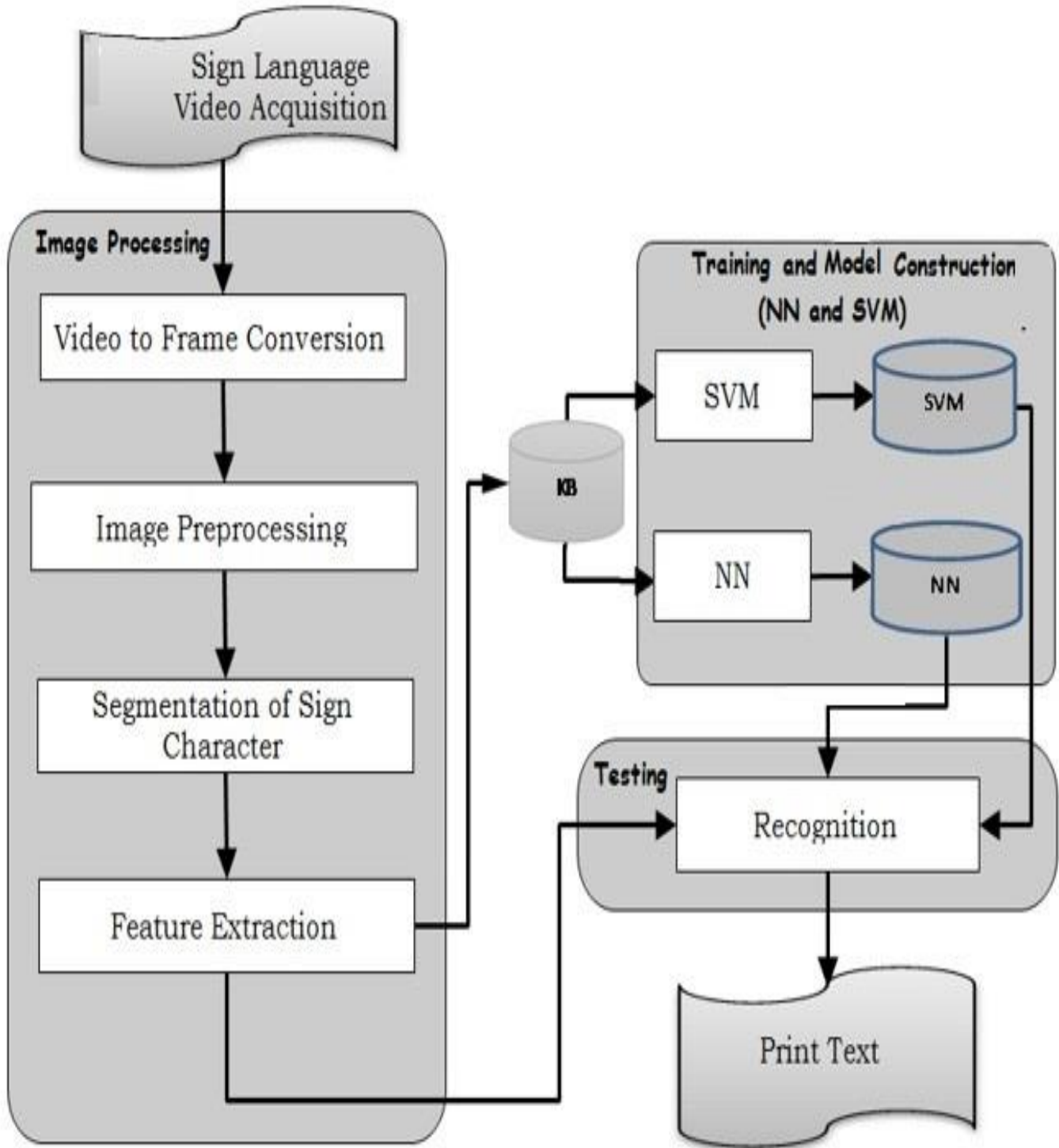


Figure 4.1: Architecture of the Proposed Amharic Sign Language Recognition System

### 4.3 Sign Language Video Acquisition

This is the first step in recognizing Amharic sign language. Before proceeding to any recognition processes, all the required videos are collected from which frames are extracted.

As discussed in Section 2.4, videos should be captured in lighting. Lighting is an important factor to be considered while capturing the video. The lighting should not be either too bright or too dim. Too bright produces reflective effect in the region of the hand on the other hand too dim produces the dark recording videos. The recording in dim light must require the lighting arrangement. Therefore, daytime is the most appropriate time to record the videos. In this research work, videos are collected by using uniform lighting condition which decrease the effect of bad illumination and uniform background (nearly white in color) which has great color difference measure with skin color, this helps us to segment the sign character from the background. Figure 4.2 shows a sample video for Amharic sign character 'ሀ'.

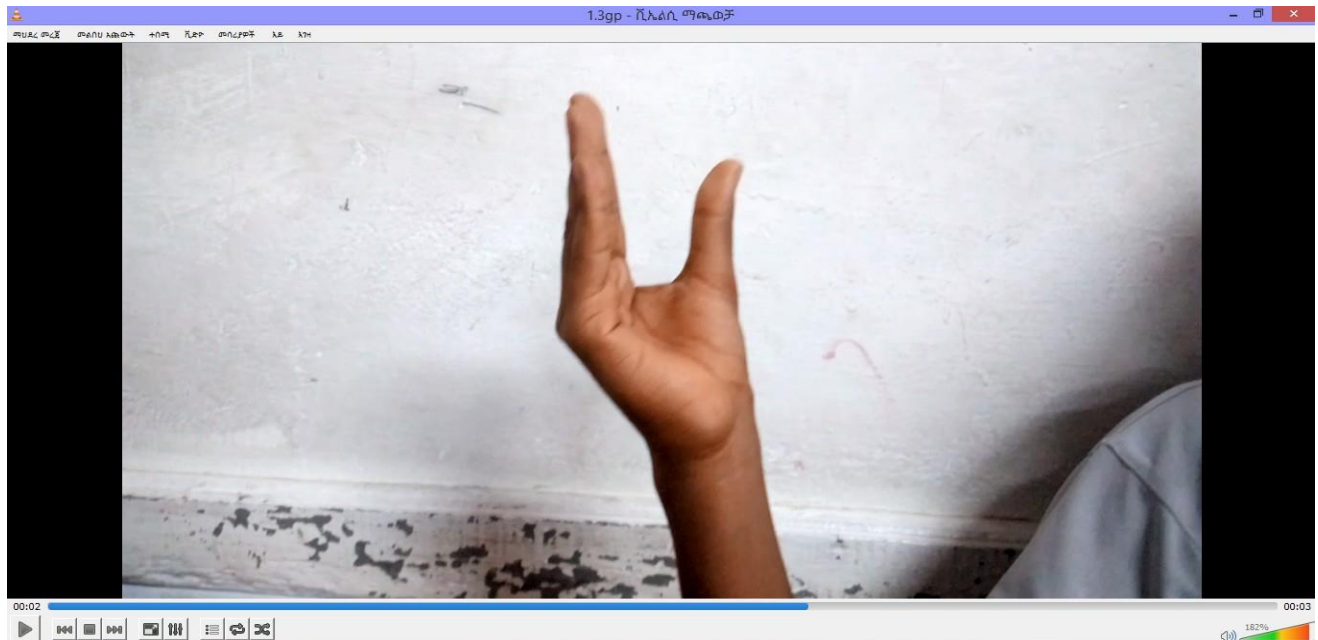


Figure 4.2: Original Video Acquired

### 4.4 Video to Frame Conversion

After we acquire the required videos from different junior signers, the sequence of frames are automatically extracted from the video using MATLAB builtin function. The number of frames automatically determined by the function depending on the play time of the video. In these

research work, we set the minimum frame to be not less than 50 frames (i.e play time greater than 3 seconds). This helps us to have enough number of centroids for each frames collected. The equation of motion of the frames: linear (left, right, down), sinusoidal and nearly circle is formulated from the position of the centroids of other frames in the video relative to the centroid of the first frame(reference frame). The positions (x,y) of the centroids of each frames are also autoamticaly detected by using the region property of the segmented frame image, from which the motion descriptors are drived as broadly described in Section 4.7. Figure 4.3 shows the number of frames (58 in number) detected from the video for a play time of 3.56 seconds for character ‘U’.

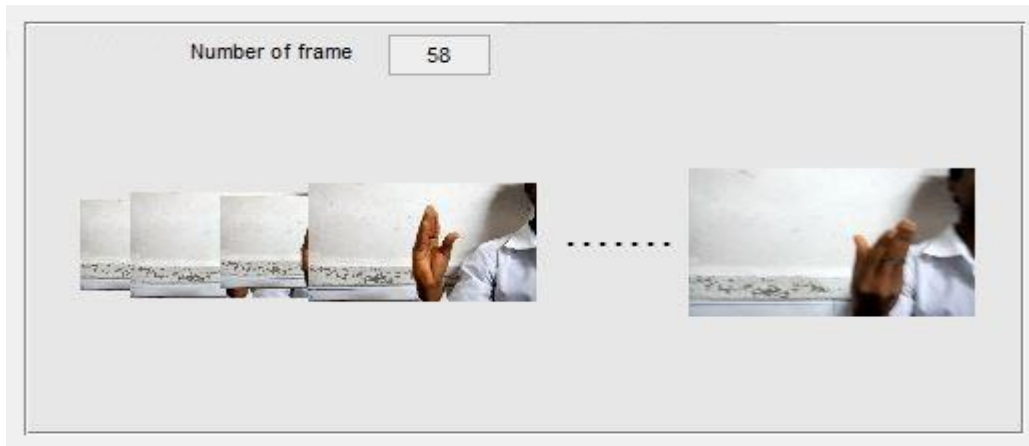


Figure 4.3: Frames Collected from Video

## 4.5 Image Preprocessing

In this component frame images obtained are pre-processed using different image preprocessing techniques since the original frame images are exposed to different noises during data capturing and then it proceeds computational complexity. To reduce these things, we apply the following image pre-processing tasks: Cropping, Converting the RGB frame into grayscale form, contrast adjustment and sharpning.

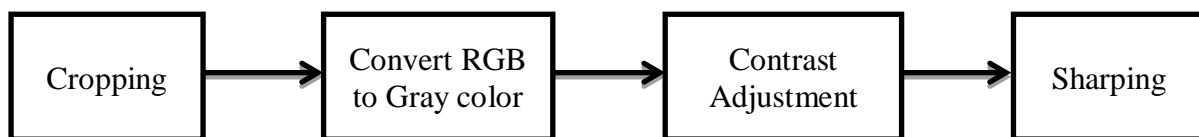


Figure 4.4: Preprocessing Image procedures

***Cropping:*** The frames which are extracted from the original video have huge size and unnecessary frame components. As a result, processing the sequence of video frames with huge dimension is computationally heavy for further processing. Hence, the height and the width of original frames are cropped.

***Converting RGB image into grayscale form:*** Firstly, the original extracted frame is RGB (24 bits color) form in which each pixel is presented by Red, Blue, and Green components. On the other hand, Grayscale image (8 bits image) only has shades of black and white. It doesn't have any color information. Processing the image in RGB form is computationally heavy and it takes more processing time when compared with grayscale form. Because of this, the sequence of video frames is converted into grayscale representation form for suitable processing.

***Contrast adjustment:*** After the original video frame in RGB color space is converted into Grayscale color frame, identifying the foreground and the background object by our eye is difficult. Enhancing the quality and adjusting the contrast of the video frame reduce the effect of this problem on the performance of the system. Basically, the idea behind this technique is to bring out detail that is obscured, or simply to highlight certain features of interest in a frame. A familiar example of enhancement is when we increase the contrast of an image because it looks better.

***Sharpening:*** The Sharpening process increases the contrast between bright and dark hand regions to bring out good features. In applications where edges are very important, an image sharpening technique is used to enhance edges from the blurred image. In addition to this, it also differentiates the blurred open or close hand fingers. To increase the discrimination ability of the neural network and support vector machine model, the input frames should somewhat be sharp.

## **4.6 Segmentation of Sign Character**

As described in Section 2.5, Segmentation refers to the process of extracting required objects in our case hand signs only from the background. In this component, we applied adaptive threshold algorithm to segment hand signs from the background.

We select adaptive threshold algorithm because our data sets are affected by shadows, shading and lighting effects. Besides, we applied different morphological operators such as dilation,

erosion and morphological filling operations to remove tiny objects and re fill missed objects of the segmented sign.

```

Input: A sample Video File: V
Output: List of Segmented Frames (Binary)
    Load video files V
    Convert v into frames  $F_{i_{rgb}}$ ,  $i=1: N$ 
For each frame  $F_{i_{rgb}}$ ,  $i=1: N$ 
    Convert  $F_{i_{rgb}}$  RGB to gray scale;
    Apply adaptive thresholding for segmentation;
    Convert  $F_{i_{gray}}$  into binary:  $F_{i_{bin}}$ ;
Apply morphological operators on binary frames;
End
Return  $F_{i_{bin}}$ 

```

*Algorithm 4.1: An algorithm to prepare the binary frame*

## 4.7 Feature Extraction

This is the crucial stage in our research work. In this component basic features are selected from the shape of the frames and direction of motion of the frames. As discussed in Section 2.2, the fundamental Amharic sign characters, the Ge'ez (ግዕዝ) are different in shape. However, the derived Amharic sign characters' frames, from Kabi (ካዕብ) to Sadis (ሳድስ) keep the shape of the fundamental sign character. They differ in direction of motion only as shown in Table 2.1. For the last derived sign character or the Sabi (ሳብ) Amharic sign character, all the frames keep the shape of the fundamental character like the other family except the last frame takes the color of the dorsum which is opposite to the palm as shown in Figure 4.3. Thus, in this research work, we consider three major feature descriptors: shape feature, motion feature and color feature.

### Shape Feature Descriptors

These are features selected to identify the fundamental Amharic sign characters from each other. We used Fourier Descriptor (FD) which is a contour-based shape descriptor invariant of size, rotation and translation as discussed in Section 2.6. In this research work, we used 31 set of combined shape feature descriptors (fd1, fd2, fd3... and fd31) to represent all the 34 Amharic

sign characters. These shape features are taken after the original shape is resampled as shown in Figure 4.5 and they are extracted using Algorithm 3 illustrated below. As described in the Algorithm, segmented binary frame image is taken as an input, as discussed in Section 2.6 and the segmented image is resampled to reduce the computational time of the Fourier descriptor. On the resampled image, we applied a 1D complex signal on which Discrete Fourier Transform (DFT) is implemented which returns list of feature descriptors (FDs). These list of feature descriptors are normalized to 31 best feature descriptors.

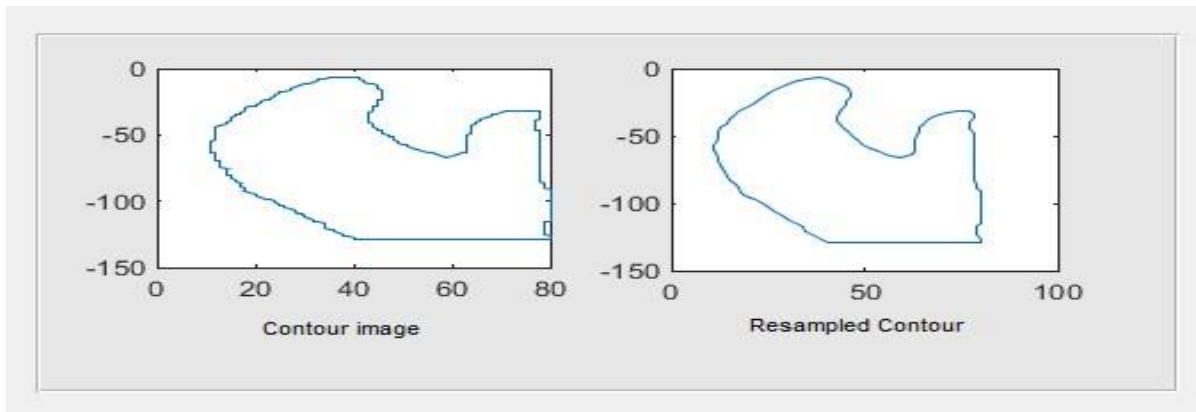


Figure 4.5: Contour of Amharic Sign Character 'v'

```

Input: -List of Segmented Frame Images (Binary), I
Output: - List of features, FDs
    Get segmented frame image I.
For each segmented frame I
    Extract boundary points from I: B.
        Resample B into K points: R.
    For each resampled point in R
        Make a 1D complex signal:  $x_k + i*y_k$ 
        Apply DFT on  $x_k + i*y_k$ : FDs.
        Compute the norm of FDs;
    End
End

```

Algorithm 4.2: An Algorithm to calculate Fourier Descriptors

## Motion Feature Descriptors

In addition to the shape features extracted from the captured videos, features conveyed by the hand motions were also considered for this research. As stated in Section 2.6, motion can be represented in various techniques. In the proposed system, the two motion features are found to be crucial to represent sign gestures namely, Direction and Angle to identify the type of trajectory or curve.

Direction describes the type of movement with reference to the four-coordinate plane. Hence, the direction of sign characters could be identified based on the centroid information, gathered from a sampled frame lists, followed by analyzing the direction of the last frame with respect to the first video frame. By this, it is meant that the last frame can be referred as either to the right, left, or down side of the first video frame.

The type of curve or trajectory might depend on the curve trajectory formed by drawing a line between each individual centroid and realized by employing different models. For example: A polynomial fitting with order one or two produce a good result to model for the first three derivatives (ካዕብ, ሣልስ and ራብዕ). Figure 4.6 and Figure 4.7 shows fitting the centroids, extracted from ten sampled frames for Amharic sign characters ('ሂ') and ('ሃ') respectively.

The other derivatives (the Hamis(ኃምስ)-ሄ and the Sadis-ዐ) are modeled by circle and sinusoidal models respectively as shown in Figure 4.8 and 4.9.

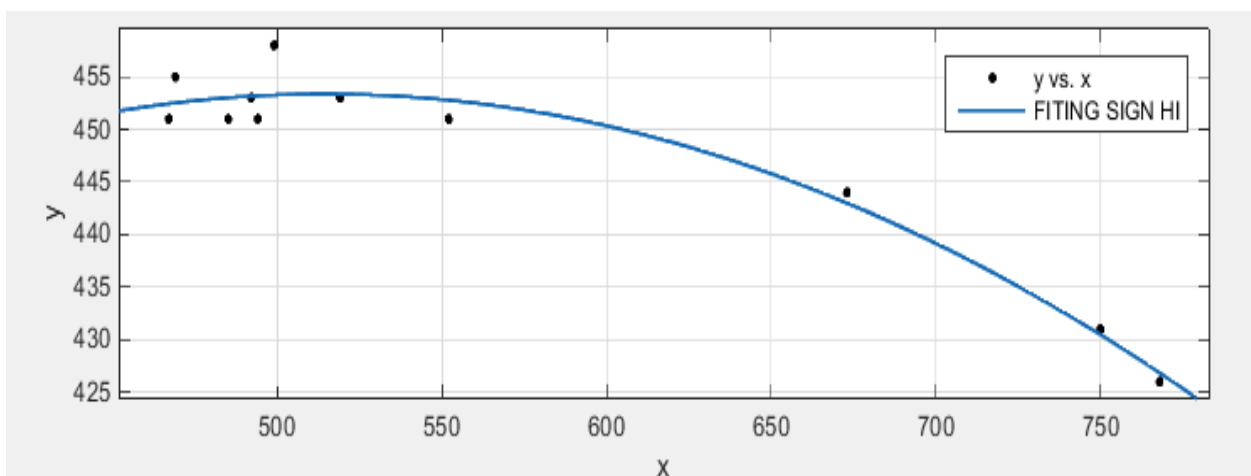


Figure 4.6: Fitting centroids to best adjusted R-square for 'ሂ' sign Character

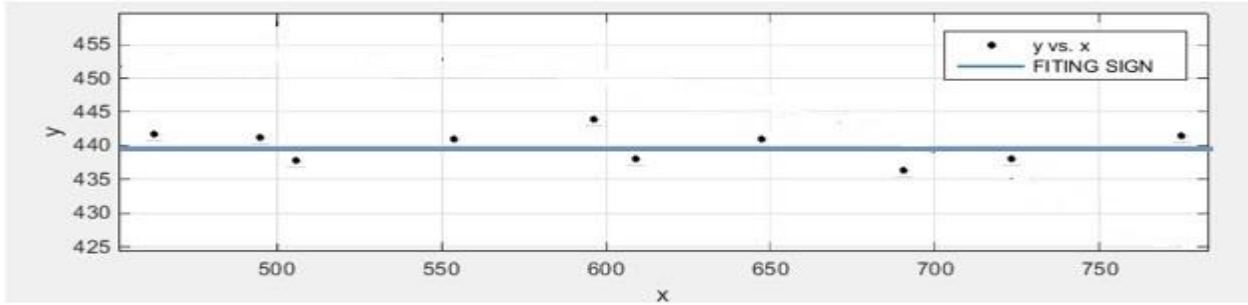


Figure 4.7: Fitting centroids to best adjusted R-square for 'Y' sign Character.

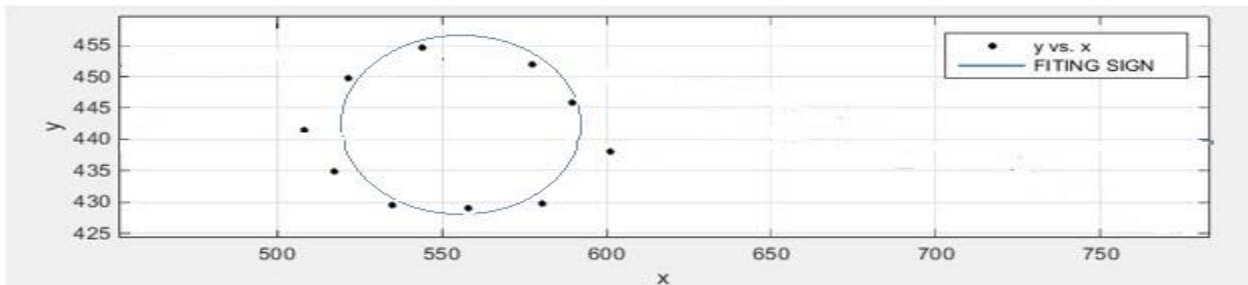


Figure 4.8: Fitting centroids to best adjusted R-square for 'Z' sign Character

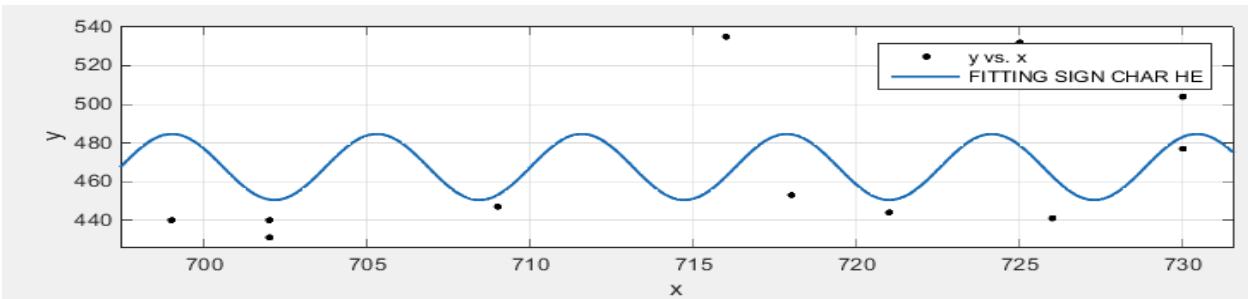


Figure 4.9: Fitting centroids to best adjusted R-square for 'v' sign Character

Accordingly, the five derived Amharic sign characters could be model by the following equations as shown in Table 4.1. From these models, we depict two features to represent the characters: the slope named as angle difference and direction. Angle refers the angle make by consecutive centroids of the frames with respect to the centroid of first frame. For the first three derived Amharic sign characters ( $\nu$ ,  $\zeta$  and  $\gamma$ ) the angle difference measure is measured to be less than  $30^\circ$  in our experiment. However, for the  $\zeta$  sign of the angle difference measure is found to be more than  $180^\circ$  and the rest  $\nu$ 's sign, the angle difference is less than  $180^\circ$  in our experiment with respect to a line drown from the first frame to the last frame in the direction of motion.

Table 4.1: Model Equations for Amharic Sign Character

Sign Character	Model Equation	Model Description	Angle Difference	Direction
ሁ	$F(x) = \{ax^2 + bx + c\}$	Parabolic	$0 \leq \Theta < 30$	Positive slope, increasing
ሂ	$F(x) = \{ax^2 + bx + c\}$	Parabolic	$0 \leq \Theta < 30$	Negative slope, decreasing
ሃ	$F(x) = \{ax + b\}$	Linear	Nearly Zero	Zero slope
ሄ	$\{ax^2 + by^2 = 1\}$	Nearly Circle	$180 < \Theta \leq 360$	Increasing and decreasing
ህ	$F(x) = \{a * (\sin(x)) + b\}$	Sinusoidal	$30 \leq \Theta < 180$	Increasing and decreasing

### Color Difference Measure Feature Descriptor

As we can see in Figure 4.3, the last frame of Sabi (ሳብ) character of Amharic sign language turns to the back side of the palm's color which has in difference with the first frame's color. Thus, once we identify the character from the shape of the first frame, we check the color of the last frame whether it has the color of the skin or not using color difference measure as stated in the following Algorithm 4.3. As implemented in the Algorithm 4.3, high in color difference measure value refers skin color. An HSV color was chosen to get color difference of the first and the last frames of Sabi character for this research due to the fact that HSV color is more popular when compared to RGB or YCbCr color since it is compatible with human color perception [17, 23]. Thus, the Sabi chatters (ህ for instance) are always represented by:

*Amharic Sign for Sabi (ሳብ) Character = Shape of the first frame + Skin color of the last frame*

<p><b>Input:</b> -colors of first frame and last frame around the centroids</p> <p><b>Output:</b> - color difference measure value, CDM</p> <p>Get collection of RGB colors [ ]c<sub>1</sub> from the first frame around its centroid.</p> <p>Get collection of RGB colors [ ]c<sub>2</sub> from the last frame around its centroid</p> <p>Convert both color collections from RGB to HSV collections.</p> <p>Calculate average value of each collection: C<sub>1avg</sub> and C<sub>2avg</sub></p> <p>Compute a reference color value from C<sub>1avg</sub> and C<sub>2avg</sub>: <math>C_{ref} = \frac{1}{2}(C_{1avg} + C_{2avg})</math></p> <p>Compute rms value of C<sub>1avg</sub> and C<sub>2avg</sub> relative to C<sub>ref</sub>: C<sub>d1</sub> and C<sub>d2</sub></p> <p>i.e. <math>C_{d1} = \sqrt{(C_{1avg} - C_{ref})^2}</math> and <math>C_{d2} = \sqrt{(C_{2avg} - C_{ref})^2}</math></p> <p>Compute color difference measure: <math>CDM =  C_{d1} - C_{d2} </math></p> <p>Return CDM</p>
---

*Algorithm 4.3 : Color Difference Measure*

Combining the above features, 31 from shape feature, 2 from motion feature and 1 from color feature, we come up 34 features to represent Amharic Sign Language Characters as shown in Table 4.3. These features are stored in the Knowledge Base (KB) for training.

*Table 4.2: Extracted Shape, Motion and Color features for the sign ‘v’.*

Fid	fd1	fd2	fd3	...	fd31	TetaDiff	ShiftDirection	ColorDiff	Class
1	0.007796297855576045	0.008938650165959244	0.02764011575515158	...	0.016965436532464788	0	0	0	1
2	0.0287854909223919	0.014282323343644336	0.013779690623821743	...	0.01959670219375705	0	0	0	1
3	0.02598301989447569	0.011905781217853188	0.015539306817221113	...	0.009819085120985806	0	0	0	1
4	0.01248367509245044	0.014145379301528687	0.006999666780245172	...	0.008413410929484172	0	0	0	1
5	0.022581586715767953	0.0221128682541622	0.01279251188235577	...	0.013919055793543486	0	0	0	1
6	0.008930162241472424	0.024603483283511095	0.03515261140511974	...	0.01837218039969434	0	0	0	1
7	0.01673406253165888	0.035778408416667774	0.03275662807966468	...	0.02438927465083855	0	0	0	1
8	0.01669119216197339	0.04870946144510985	0.03563323172776649	...	0.013997466845466097	0	0	0	1
9	0.014854231210061377	0.026792134710513088	0.08445494352977642	...	0.03981418055277261	0	0	0	1
10	0.03314870889069797	0.03632921782364478	0.07724712760823194	...	0.03860451705653878	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
41	0.015771833169043497	0.026487325322723593	0.017863312519031744	...	0.01598728568966026	0	0	0	1
42	0.020080112604308357	0.013927147784344512	0.012055596767133726	...	0.010367087548439945	0	0	0	1
43	0.015328582602856655	0.011859425134043212	0.007215824547779402	...	0.016912474174175185	0	0	0	1

## 4.8 Training

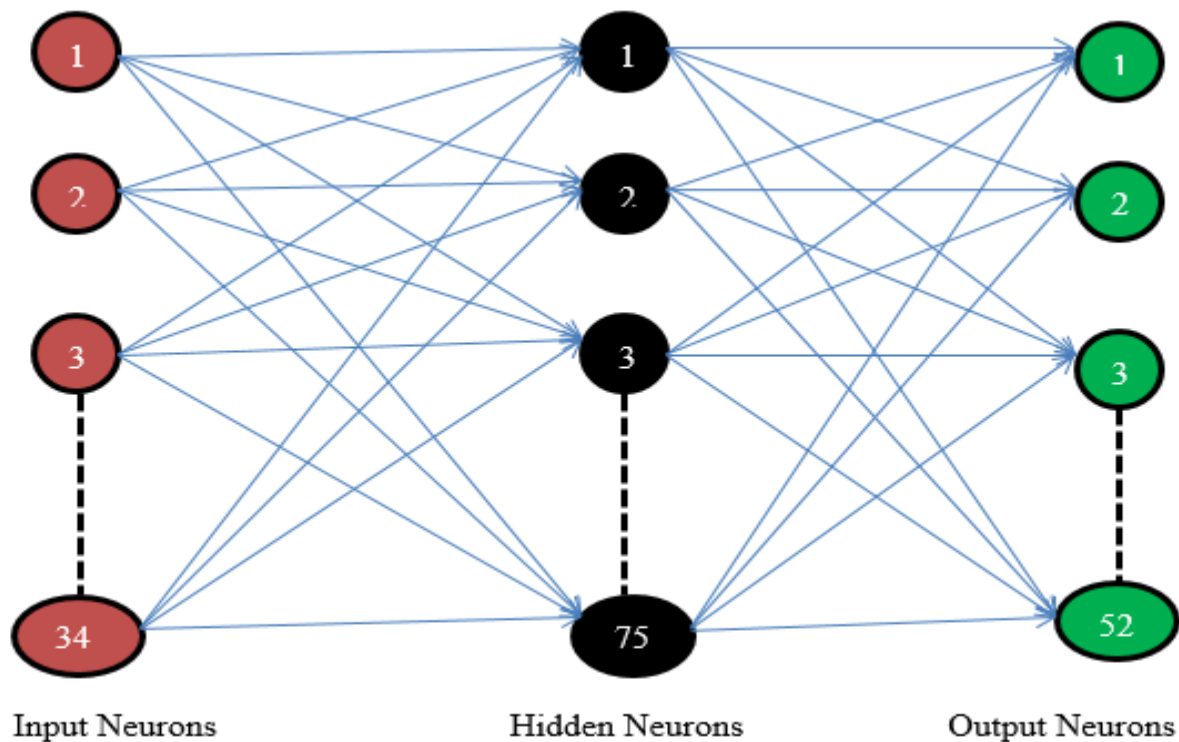
In this component extracted features obtained from the feature extraction is used as an input in performing classification through training. We have selected two classifiers: artificial neural network (ANN) and Support Vector Machine (SVM) as training machines to construct models which would be used as classifiers for Amharic sign characters into the corresponding Amharic print characters.

In the case of ANN, we need to select optimum architecture (topology) and training an algorithm that best matches to the proposed classification model. In our study, feed forward multi-layer

perceptron (MLPs) architecture is used. It is composed of three types of layers: input layer, output layer and hidden layer.

To compute the neuron final output state, a backpropagation algorithm needs continuous and differentiable activation function. For this study, in order to control the smooth relation between the input and output, a sigmoid transfer function is used.

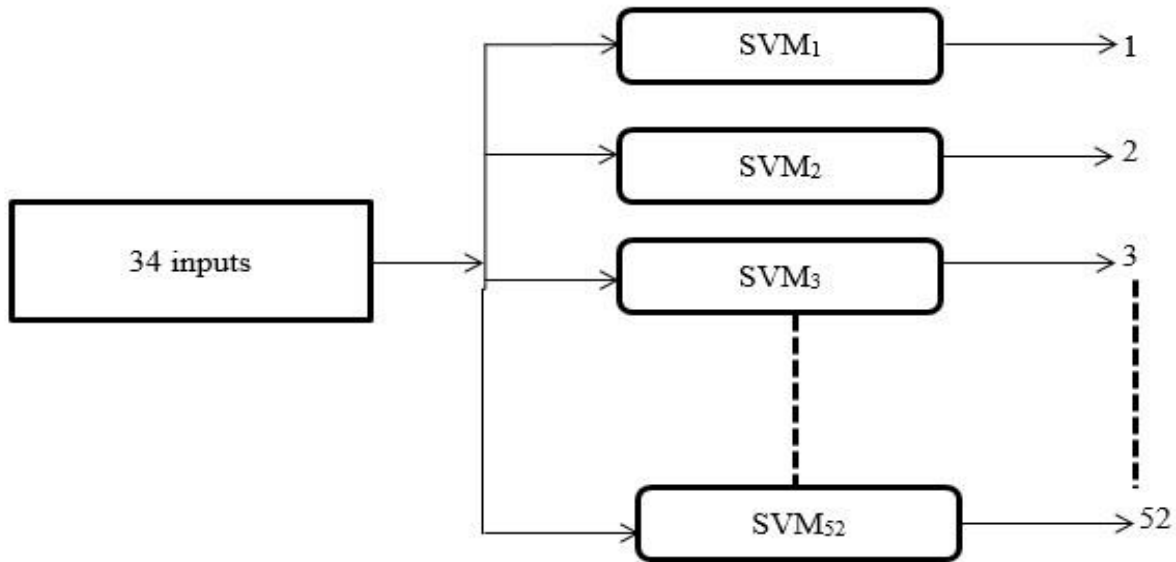
Generally, as depicted in Figure 4.10, we used feed forward multilayer perceptron with one hidden layer which has 75 neurons in addition to the input and output layers. The features: shape feature descriptors, motion feature descriptors and color feature descriptors computed are used as an input layer in NN architecture. Based on these selected features, the number of input neurons is 34 for the selected 52 Amharic alphabet signs, as a result NN gives us 52 generated output classes in the output layer as shown in Figure 4.10.



*Figure 4.10: Neural Network Model with One Hidden Layer*

To build a model using SVM which is basically binary classifier, we applied a multi class SVM classifier since the expected outputs generated from our model has to be more than two classes as shown in Figure 4.11. Though, there are various multi-class SVM with kernel functions which

are used to construct a model for classifying Amharic sign characters, one against all with radial basis function is found to be better classifier to identify actual multi-classes for Amharic sign language with large number of signs. We used different parameters and initial values to generate an optimal result.



*Figure 4.11: Multi-Class Support Vector Machine Network Model*

#### **4.9 Model Construction**

To construct the models namely NN and SVM, several steps are iteratively done until an optimal result is found as described in Chapter two, Section 2.8. These steps are applied differently for the training machines (NN and SVM) we had selected.

For NN, as we already illustrate in Figure 4.12, the optimal number of hidden layer neurons and the size of the layers in the architecture are determined by doing several training steps iteratively. As a result, we finally constructed a NN model with 34 different input neurons, 75 neurons on the hidden layer to determine one of the 52 classes. Figure 4.12 shows the final generated model of NN.

In the case of SVM, the training steps are selecting different parameters along with its corresponding initial values iteratively. We selected “One against All” method which is an optimal to construct 52 classes by taking 34 features as an input as can be seen in Figure 4.11.

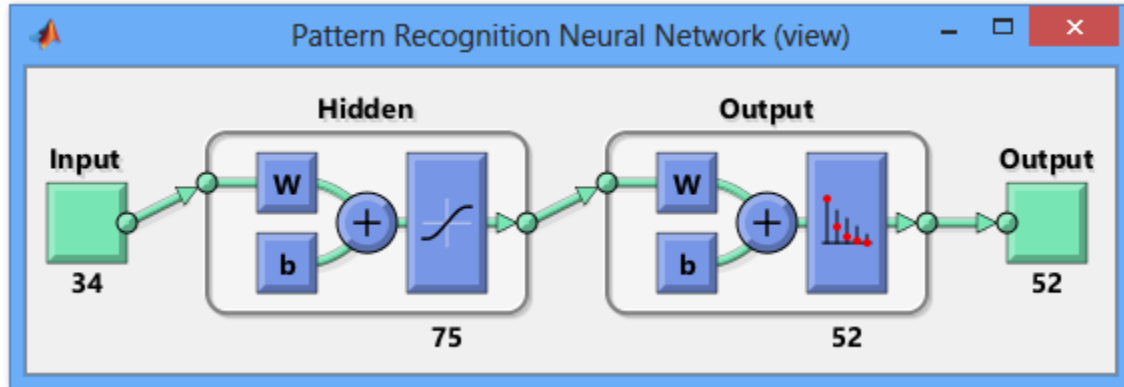


Figure 4.12: Constructed Model for NN

#### 4.10 Testing

In this component, the recognition or classification accuracy of both constructed models: NN and SVM are tested. As early presented in Section 2.8, k-fold cross validation is an appropriate method to test NN and SVM models. Hence, we select k-fold cross validation technique in this work because all data set are used for both training and validation. In our case the value of k is ten. The whole data set is partitioned into 10 equal parts and each partition is used for both training and evaluation. As described in Table 5.1, the total data that are 1,710 from which 58,140 set feature vectors are feed into the training machines.

#### 4.11 Summary

This chapter systematically went through the designing of Amharic sign language recognition system.

The proposed system has different components for systematically working together. Initially, for each Amharic sign language characters, sample videos are acquired. Collection of frames are taken from the videos. The frames are preprocessed and segmented from the background. Then, after preprocessing and segmentation, from the first frame of each character sign, we take shape features using Fourier Descriptor. From the other consecutive set of frames, motion and color features are derived in the feature extraction components. From this component, thirty-four

different features are extracted. Among these features 31 of them are shape features extracted from the first frame. Others are motion and color features extracted from the motion of other frames and the color of the last frame respectively. Finally, an artificial neural network and support vector machine are trained using these features from which 52 output classes are generated to classify or recognize the selected Amharic Sign language characters. The sign characters are the 34 basic Amharic sign language characters. And the others are derived Amharic sign characters of 'ሀ', 'ለ' and 'ሐ'. Through training, in this work two models (NN and SVM) are constructed and then tested. The models are tested using ten-fold cross validation testing technique.

## **Chapter Five: Experimentation and Result Discussion**

### **5.1 Introduction**

In the previous chapter, we described the design of the system. This chapter discusses implementation detail of the proposed design for recognition of Amharic sign language. Section 5.2 presents the datasets used in training and testing the system. Section 5.3 describes the tools used and the overview of the system. Section 5.4 presents the evaluation and test results found in Section 5.5. Finally, the discussion part is presented in the last Section of the chapter.

### **5.2 Data Sets**

The data sets were taken from one preparatory school's willingness students in Addis Ababa, namely Menelik the II Preparatory School. Totally, ten students (junior signers) were fully participated. From these, there were eight right handed and two left handed signers. Each signer is expected to perform a sign 160 times for only four basic alphabet signs from among thirty-four first order Amharic sign language characters. However, we selected 153 best signs. For eighteen derived Amharic sign language characters, only three voluntary signers were participated from among ten signers and we captured 60 data set from each. The captured video has third generation partnership (3GP) format and a play time ranging from 3 to 4 seconds. From the collected data set, all basic sign characters (34 in number) and some derived sign characters of 'ሀ', 'ለ' and 'ሐ' (18 in number) have been chosen to be part of the research. The videos were taken in the same controlled environment in order to avoid external effects of sunlight and other environmental conditions.

As depicted in the Table 5.1, total sample videos were collected for basic Amharic alphabet signs which are 1,530 and also total samples are 180 were captured for some derived Amharic alphabet signs. Accordingly, we have created our own data set that contain around 1,710 real videos of 52 different classes that are basic and some specific manual alphabets.

Table 5.1: The collected ETHSL manual alphabets

ETHSL manual alphabets	No of sample frames	ETHSL manual alphabets	No of sample frames
<b>Considered Amharic Sign Languages- Basics</b>			
ሀ	43	ዘ	50
ለ	49	ዠ	45
ሐ	49	የ	43
መ	47	ደ	50
ሠ	46	ጀ	38
ረ	50	ገ	50
ሰ	49	ጠ	48
ሸ	50	ጨ	50
ቀ	49	ጸ	28
ቦ	49	ጸ	50
ተ	37	ፀ	42
ቸ	35	ፈ	50
ኅ	46	ፐ	38
ነ	50	ሸ	48
ኘ	47	ኸ	44
አ	41	ወ	37
ከ	39	ዐ	43
<b>Considered Amharic Sign Languages-Derived</b>			
	<b>Total No of sample Videos</b>		<b>Total No of sample Videos</b>
ሀ's derived - ሁ, ሂ, ሃ, ሄ, ህ and ሆ.	60	ሐ's derived - ሑ, ሒ, ሓ, ሔ, ሕ and ሖ.	60
ለ's derived - ለ, ል, ሎ, ሏ and ሏ.	60		
<b>Total</b>	<b>1,710</b>		

In the process of data collection, a lot of challenges have been occurred to address about Amharic alphabet signs with different signers (teachers) in Menelik the II Preparatory School. Unfortunately, the reason that school's signers are not willing for giving information was the main challenges. The challenges have been made their own impact on the success of our work, especially, during recording. Besides, getting notes and books about ETHSL was also a very challenging task.

### 5.3 Implementation

We used MATHLAB 2014a and MySQL to implement the prototype. MATHLAB is a high-level language and interactive environment used to design and test the tools and techniques used in our approach. MySQL is used to persist the features extracted from Amharic sign language characters. Figure 5.1 shows the running prototype. The designed prototype is tested on Toshiba laptop core i3, processor speed of 1.6 GHz, 4.0GB RAM,700 GB hard disk capacities and 64 bits operating system.

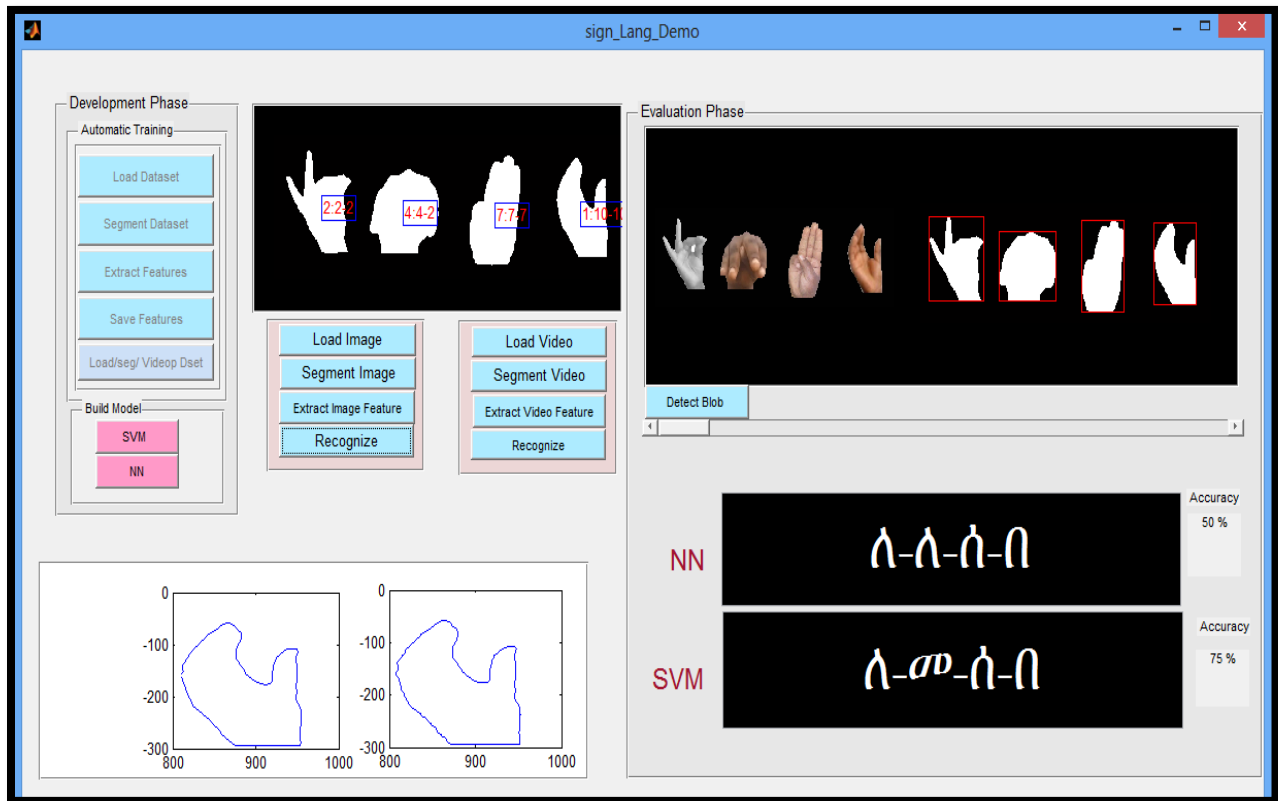


Figure 5.1: Screen shot of the running prototype

## 5.4 Evaluation

In many machine learning areas, a basic problem is obtaining an accurate estimate for the generalization ability of a learning algorithm trained on a given dataset. Hence, the basic concern in machine learning is to obtain an accurate estimate of the generalization error of a model trained on a finite dataset [50]. Simply splitting the corpus or dataset into a single training and testing set may not give the best estimate of future performance. Therefore, we used different evaluation techniques which is cross-validation with confusion matrix on our approach to clearly point out the accuracy of the model. This can also help to choose an algorithm from a variety of learning algorithms.

Most machine learning approaches are evaluated by using a methodology of cross validation which is believed to be a more reliable methodology [51]. It is a statistical method of evaluating and comparing learning algorithms or models by dividing data into two segments: one used to learn or train a model and the other used to validate the model. The basic form of cross-validation is k-fold cross validation which is the most appropriate technique. “In k-fold cross validation, the dataset  $D$  is randomly split into  $k$  subsets  $D_1, D_2, \dots, D_k$  of approximately equal size” [51].

K-Fold cross validation helps us to estimate performance of the learned models (NN and SVM) from our dataset to gauge the generalization of the algorithms for our Amharic sign language classification as it was described in Section 4.10. Thus, we can compare the performance of models constructed from NN and SVM and find out which one is the best. Generally, it generates an approximate measurement of how well the learned model will do on “unseen” data.

Our experiment is done using a 10-fold cross-validation on the available data. This means that the data is split in ten equal partitions, and each of these is used once as test set, with the other nine as corresponding train set. This way, all examples are used at least once as a test item, while keeping training and test data carefully separated, and the NN and SVM based classifiers are trained each time on 90% of the available training data. The process of our 10-fold evaluation is illustrated in Figure 5.2.

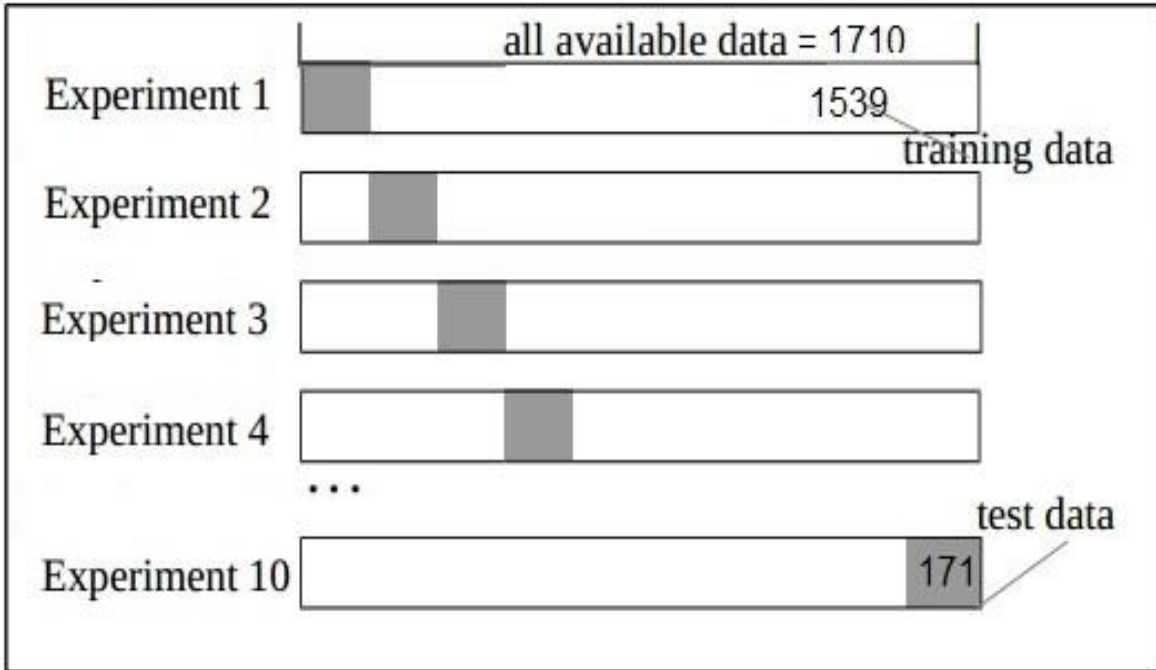


Figure 5.2: Process of 10-fold cross validation experiment

The lighter section of the data (1,539 Amharic sign characters' data set) are used for training while the darker sections (171 sign characters) are used for validation in each experiment.

In our work, we have used four performance matrices to analyze the classification performance of NN and SVM models using 10-fold cross validation. These performance matrices are: Accuracy, Precision (positive predictive value), Recall (True positive rate or Sensitivity) and F-score. Accuracy is the total number of correctly predicted signs to all test samples. Precision is the ratio of correctly identified instances of a class to the total number of positive observations identified a sign as that class. Recall is the ratio of correctly identified instances of a class to all the instances where the sign was truly that class and f-score is the harmonic mean of precision and recall [4, 8]. All of these parameters are given by the following equations (9), (10), (11) and (12) respectively:

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+FP+TN} \quad (9)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (10)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (11)$$

$$F_{Score} = 2 * \left( \frac{Precision * Recall}{Precision + Recall} \right) \quad (12)$$

In these equations, TP, TN, FN, FP present true positive, true negative, false negative, false positive respectively. A true positive occurs when the predicted class is the same as the actual class. A false positive occurs when a classifier classifies sign character as the incorrect class. A true negative is when the classifier correctly predicts sign character not part of an incorrect class and a false negative occurs when the classifier does not classify sign character into the correct class [4].

## 5.5 Test Results

The data set is partitioned into two parts (training part and testing part) as it is indicated in Section 5.4. The two models (NN & SVM) are used as classifiers, each of which is trained and tested. For evaluation purposes, the models obtained from training phase are tested using new Amharic alphabet signs in addition to the training set. In this way we compute the accuracy of the 10-fold cross validation examination by taking the average of the sum of each training and test set as illustrated in Table 5.2. The performance of the models is presented in Table 5.2 for each fold.

*Table 5.2: Test Result for Each Fold, NN and SVM*

Experiment	1	2	3	4	5	6	7	8	9	10	Avg (%)
Total No of Dataset	1710	1710	1710	1710	1710	1710	1710	1710	1710	1710	1710
Accuracy-NN (%)	61.73	52.63	60.17	61.66	59.77	47.81	51.97	59.57	60.29	62.64	<b>57.82</b>
Accuracy-SVM (%)	73.5	74.93	72.95	75.42	72.71	74.26	72.74	73.6	74.51	75.98	<b>74.06</b>

As can be seen in the Table 5.2, we found that the recognition accuracy of Neural network is 61.73%, 52.63%, 60.17%, 61.66%, 59.77%, 47.81%, 51.97%, 59.57%, 60.29% and 62.64% on experiment one, experiment two, experiment three, experiment four, experiment five, experiment six, experiment seven, experiment eight, experiment nine and experiment ten respectively. The lowest recognition percentage (47.81%) is recorded on experiment six which indicates different

signs in this experiment are very similar in shape but the highest recognition percentage (62.64%) is recorded on experiment ten that means it includes signs in dissimilar shape. The average recognition accuracy of NN from those experiments is 57.82%.

The other recognition accuracy of Support Vector Machine is 73.50%, 74.93%, 72.95%, 75.42%, 72.71%, 74.26%, 72.74%, 73.60%, 74.51% and 75.87% on experiment one, experiment two, experiment three, experiment four, experiment five, experiment six, experiment seven, experiment eight, experiment nine and experiment ten respectively. In experiment five (fold 5), there is the lowest recognition percentage (72.71%). However, the highest recognition percentage (75.98%) is recorded on experiment one. The average recognition accuracy from those ten experiments is 74.06%. In each experiment, the accuracy result of SVM is greater than NN's accuracy result. The classification performance of NN and SVM Models in the above Table 5.2 are clearly illustrated with a Bar chart below.

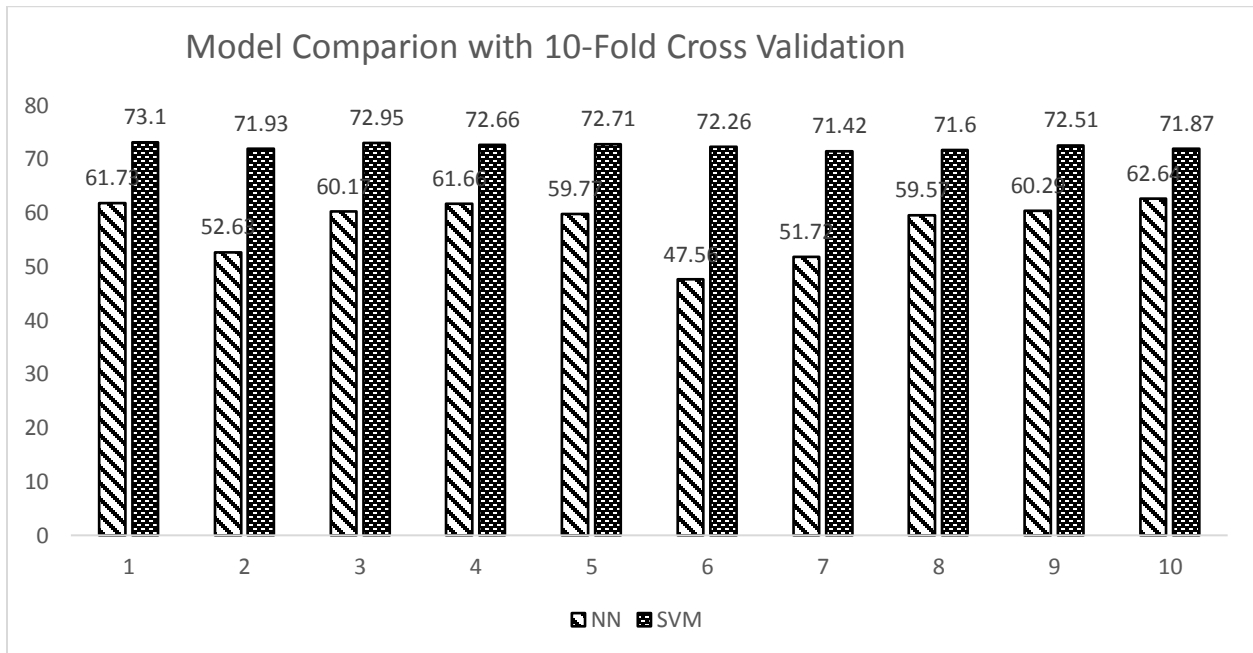


Figure 5.3: NN and SVM classification accuracy results with bar chart

The Bar chart in Figure 5.3 is a visual representation of the results above. The blue color represents the Neural Network's accuracy and the red color represents the Support Vector Machine's accuracy. This is a good chart on how well the performance of the NN and SVM is compared in each fold.

Each class result of accuracy, precision, recall and F-score based on NN and SVM models from the confusion matrix is presented in the Table 5.3 and Table 5.4 respectively as tabulated below.

*Table 5.3 : Accuracy, Precision, Recall and F-Score measure for NN Model.*

Signs	Class	TP	FP	FN	TN	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)
U	1	47	36	38	51	56.97674419	56.62650602	55.294118	55.95238095
Λ	2	36	38	41	57	54.06976744	48.64864865	46.753247	47.68211921
h	3	29	43	35	65	54.65116279	40.27777778	45.3125	42.64705882
ω	4	32	26	31	83	66.86046512	55.17241379	50.793651	52.89256198
ω	5	29	37	39	67	55.81395349	43.93939394	42.647059	43.28358209
ζ	6	31	29	39	73	60.46511628	51.66666667	44.285714	47.69230769
ł	7	36	36	39	61	56.39534884	50	48	48.97959184
ñ	8	31	32	34	75	61.62790698	49.20634921	47.692308	48.4375
φ	9	31	39	41	61	53.48837209	44.28571429	43.055556	43.66197183
π	10	30	39	41	62	53.48837209	43.47826087	42.253521	42.85714286
†	11	25	31	33	83	62.79069767	44.64285714	43.103448	43.85964912
ƒ	12	33	34	36	69	59.30232558	49.25373134	47.826087	48.52941176
γ	13	34	35	39	64	56.97674419	49.27536232	46.575342	47.88732394
ι	14	36	29	33	74	63.95348837	55.38461538	52.173913	53.73134328
Ƴ	15	35	31	35	71	61.62790698	53.03030303	50	51.47058824
κ	16	32	39	38	63	55.23255814	45.07042254	45.714286	45.39007092
h	17	35	34	32	71	61.62790698	50.72463768	52.238806	51.47058824
ñ	18	31	28	33	80	64.53488372	52.54237288	48.4375	50.40650407
ω	19	38	37	39	58	55.81395349	50.66666667	49.350649	50
o	20	31	24	26	91	70.93023256	56.36363636	54.385965	55.35714286
h	21	37	29	31	75	65.11627907	56.06060606	54.411765	55.2238806
ƒ	22	31	34	33	74	61.04651163	47.69230769	48.4375	48.0620155
Ƒ	23	28	37	36	71	57.55813953	43.07692308	43.75	43.41085271
Ɣ	24	29	36	34	73	59.30232558	44.61538462	46.031746	45.3125
Ɔ	25	37	35	36	64	58.72093023	51.38888889	50.684932	51.03448276

7	26	31	39	28	74	61.04651163	44.28571429	52.542373	48.0620155
ᆞ	27	38	25	28	81	69.18604651	60.31746032	57.575758	58.91472868
ᆡ	28	36	33	34	69	61.04651163	52.17391304	51.428571	51.79856115
ᆣ	29	32	32	33	75	62.20930233	50	49.230769	49.6124031
ᆥ	30	39	29	25	79	68.60465116	57.35294118	60.9375	59.09090909
ᆦ	31	34	38	25	75	63.37209302	47.22222222	57.627119	51.90839695
ᆧ	32	38	35	24	75	65.69767442	52.05479452	61.290323	56.2962963
ᆨ	33	33	34	29	76	63.37209302	49.25373134	53.225806	51.1627907
ᆩ	34	37	38	33	64	58.72093023	49.33333333	52.857143	51.03448276
ᆪ	35	15	39	35	83	56.97674419	27.77777778	30	28.84615385
ᆫ	36	12	34	36	90	59.30232558	26.08695652	25	25.53191489
ᆬ	37	11	36	38	87	56.97674419	23.40425532	22.44898	22.91666667
ᆭ	38	12	35	39	86	56.97674419	25.53191489	23.529412	24.48979592
ᆮ	39	11	35	36	90	58.72093023	23.91304348	23.404255	23.65591398
ᆯ	40	11	33	38	90	58.72093023	25	22.44898	23.65591398
ᆰ	41	10	39	41	82	53.48837209	20.40816327	19.607843	20
ᆱ	42	12	44	43	73	49.41860465	21.42857143	21.818182	21.62162162
ᆲ	43	15	56	54	47	36.04651163	21.12676056	21.73913	21.42857143
ᆳ	44	17	43	42	70	50.58139535	28.33333333	28.813559	28.57142857
ᆴ	45	12	56	47	57	40.11627907	17.64705882	20.338983	18.8976378
ᆵ	46	13	58	59	42	31.97674419	18.30985915	18.055556	18.18181818
ᆶ	47	18	53	55	46	37.20930233	25.35211268	24.657534	25
ᆷ	48	14	41	42	75	51.74418605	25.45454545	25	25.22522523
ᆸ	49	9	39	34	90	57.55813953	18.75	20.930233	19.78021978
ᆹ	50	11	31	34	96	62.20930233	26.19047619	24.444444	25.28735632
ᆺ	51	13	31	34	94	62.20930233	29.54545455	27.659574	28.57142857
ᆻ	52	11	32	28	101	65.11627907	25.58139535	28.205128	26.82926829
<b>Total</b>						<b>3006.976744</b>	<b>2124.926236</b>	<b>2124.0268</b>	<b>2121.604091</b>
<b>Average (%)</b>						<b>57.82647585</b>	<b>40.86396608</b>	<b>40.846669</b>	<b>40.80007866</b>

Table 5.4 : Accuracy, Precision, Recall and F-Score measure for SVM Model.

Signs	Class	TP	FP	FN	TN	Accuracy (%)	Precision (%)	Recall (%)	F-Score (%)
U	1	39	16	18	99	80.23255814	70.90909091	68.421053	69.64285714
Λ	2	32	28	21	91	71.51162791	53.33333333	60.377358	56.63716814
h	3	31	25	35	81	65.11627907	55.35714286	46.969697	50.81967213
ω	4	32	26	31	83	66.86046512	55.17241379	50.793651	52.89256198
ω	5	29	27	31	85	66.27906977	51.78571429	48.333333	50
ζ	6	38	18	17	99	79.65116279	67.85714286	69.090909	68.46846847
ñ	7	36	16	19	101	79.65116279	69.23076923	65.454545	67.28971963
ñ	8	33	12	14	113	84.88372093	73.33333333	70.212766	71.73913043
ϕ	9	28	19	21	104	76.74418605	59.57446809	57.142857	58.33333333
η	10	30	19	19	104	77.90697674	61.2244898	61.22449	61.2244898
†	11	25	21	23	103	74.41860465	54.34782609	52.083333	53.19148936
‡	12	29	21	23	99	74.41860465	58	55.769231	56.8627451
γ	13	34	18	19	101	78.48837209	65.38461538	64.150943	64.76190476
ι	14	29	21	24	98	73.8372093	58	54.716981	56.31067961
ϣ	15	33	13	16	110	83.13953488	71.73913043	67.346939	69.47368421
λ	16	32	14	17	109	81.97674419	69.56521739	65.306122	67.36842105
h	17	35	14	12	111	84.88372093	71.42857143	74.468085	72.91666667
ñ	18	31	18	16	107	80.23255814	63.26530612	65.957447	64.58333333
ω	19	38	17	19	98	79.06976744	69.09090909	66.666667	67.85714286
ο	20	28	11	12	121	86.62790698	71.79487179	70	70.88607595
η	21	29	19	18	106	78.48837209	60.41666667	61.702128	61.05263158
η	22	27	14	13	118	84.30232558	65.85365854	67.5	66.66666667
ρ	23	37	17	21	97	77.90697674	68.51851852	63.793103	66.07142857
ρ	24	41	25	24	82	71.51162791	62.12121212	63.076923	62.59541985
ξ	25	33	18	19	102	78.48837209	64.70588235	63.461538	64.0776699
γ	26	34	19	18	101	78.48837209	64.1509434	65.384615	64.76190476
η	27	28	21	21	102	75.58139535	57.14285714	57.142857	57.14285714

ᄁ	28	36	20	24	92	74.41860465	64.28571429	60	62.06896552
ᄂ	29	32	23	23	94	73.25581395	58.18181818	58.181818	58.18181818
ᄃ	30	29	20	19	104	77.3255814	59.18367347	60.416667	59.79381443
ᄄ	31	28	33	21	90	68.60465116	45.90163934	57.142857	50.90909091
ᄅ	32	39	31	15	87	73.25581395	55.71428571	72.222222	62.90322581
ᄆ	33	31	18	13	110	81.97674419	63.26530612	70.454545	66.66666667
ᄇ	34	27	18	14	113	81.39534884	60	65.853659	62.79069767
ᄈ	35	12	29	31	100	65.11627907	29.26829268	27.906977	28.57142857
ᄉ	36	13	24	24	111	72.09302326	35.13513514	35.135135	35.13513514
ᄊ	37	9	26	24	113	70.93023256	25.71428571	27.272727	26.47058824
ᄋ	38	12	25	19	116	74.41860465	32.43243243	38.709677	35.29411765
ᄌ	39	9	20	20	123	76.74418605	31.03448276	31.034483	31.03448276
ᄍ	40	8	27	18	119	73.8372093	22.85714286	30.769231	26.2295082
ᄎ	41	12	29	18	113	72.6744186	29.26829268	40	33.8028169
ᄏ	42	11	34	22	105	67.44186047	24.44444444	33.333333	28.20512821
ᄐ	43	11	47	24	90	58.72093023	18.96551724	31.428571	23.65591398
ᄑ	44	13	31	34	94	62.20930233	29.54545455	27.659574	28.57142857
ᄒ	45	9	22	28	113	70.93023256	29.03225806	24.324324	26.47058824
ᄓ	46	12	25	29	106	68.60465116	32.43243243	29.268293	30.76923077
ᄔ	47	11	23	38	100	64.53488372	32.35294118	22.44898	26.5060241
ᄕ	48	9	21	36	106	66.86046512	30	20	24
ᄖ	49	11	28	37	96	62.20930233	28.20512821	22.916667	25.28735632
ᄗ	50	12	26	34	100	65.11627907	31.57894737	26.086957	28.57142857
ᄘ	51	11	21	34	106	68.02325581	34.375	24.444444	28.57142857
ᄙ	52	9	32	20	111	69.76744186	21.95121951	31.034483	25.71428571
<b>Total</b>						<b>3851.162791</b>	<b>2638.429929</b>	<b>2644.5932</b>	<b>2629.803292</b>
<b>Average (%)</b>						<b>74.0608229</b>	<b>50.7390371</b>	<b>50.857561</b>	<b>50.57314023</b>

As indicated in Table 5.3, the summary results of neural network model showed that 57.82% were correctly classified and 42.18% were misclassified and we can see that the best cases (above 65%) are the letters ‘ᄁ’, ‘ᄈ’, ‘ᄍ’, and ‘ᄒ’, mainly because these signs have very few

similarities with the others. Conversely, the worst case is the letters ‘ላ’, ‘ሎ’ and ‘ሐ’ with only below 40% accuracy. In other way, it is observed that the precision, recall and f\_score results of NN model are 40.86%, 40.84% and 40.80% respectively.

On the other hand, the overall SVM Model of accuracy, precision, recall and f-score is 74.06%, 50.74%, 50.86% and 50.57% respectively as tabulated in the above Table 5.4. in which the highest recognition rate (above 80%) is obtained for ‘ሀ’, ‘ሸ’, ‘ገ’, ‘አ’, ‘ከ’ ‘ዐ’, ‘ዠ’, ‘ጥ’ and ‘ቨ’ because these alphabet signs are a lot in number. The lowest recognition rate is found for ‘ላ’ that has 58.72% accuracy because its shape complex. From the above two experimental results, it has been shown that SVM based system has given a promising recognition performance.

The classifiers achieve high precision, recall and f- score value for high frequency classes of Amharic alphabet signs but the low results of those metrics is the cause of the low frequency alphabet signs.

In general, the classification performance of support vector machine is by far better than neural network classifier as summarized in Figure 5.4 which shows the comparison of both NN and SVM model with respect to the four performance metrics namely Accuracy, Precision, Recall and F-score.



Figure 5.4: Model Comparison Using Different Performance Metrics

According to the above Figure 5.4, a high value of precision in NN and SVM indicates that the proposed models can correctly classify those samples as alphabet sign which really has alphabet sign and also the number of selected signs is more relevant. A high recall shows that the many of the sign data introduced were recognized and also high relevant signs selected while a low recall shows that the major of the sign cases were ignored. The other high value of f-score indicate that the best values are recorded at precision and recall performance metrics but the low value is the cause of worst value at precision and recall.

## 5.6 Discussion

The proposed Amharic sign language recognizer is evaluated from the capability of learning machines (SVM and NN) to recognize Amharic sign language into the corresponding text character. Learning-ability of the selected machines is evaluated by training the 90 % of the dataset and obtaining their accuracy by using the remaining 10% for testing in each experiment (fold).

The experimental result of SVM showed significant performance better than NN even though it uses more memory. The problem of speed and memory may not be a series problem for Amharic sign language recognition into text character because the fundamental concern is first obtaining a model that recognizes Amharic sign language with better accuracy. Then speed and memory may be the second issues when there is an optimization of the classifier. From this perspective, SVM classifies better than NN regardless of its high memory usage in storage instances.

Past attempts to recognize Amharic sign language achieved fairly good results, but this is mainly due to the use of few characters. One such attempt is that of Legesse Zerubabel who has attained the recognition rate of 88.08% and 96.22% using neural network with PCA driven features and neural network with harr-like features respectively [15]. However, he only focused on ten selected basic alphabets signs which are *ሀ*, *መ*, *ሠ*, *ረ*, *ሰ*, *ሸ*, *ቦ*, *ኀ*, *ኘ* and *አ*. In these characters' hand shape, there is very few similarities among them. Due to this, his work has scored better accuracy result. The other related work was done by Abadi Tsegaye [2] who used motion features (such as angle and direction) and the overall system performance is 71.88% but he only considered the second, the third, the fourth, the fifth and the sixth orders of signs and also Abadi's work only focused on trajectory identification not on recognition. The two mentioned works used sample data from limited number of signers. In our proposed approach, we covered

all the basic and some derived alphabet signs based on shape, motion and color features with NN and SVM classifiers as it is described earlier. As we can see from Tables 5.2, 5.3 and 5.4 our proposed system has less accuracy result than previous work that was done by Legesse Zerubabel. This is attributed due the number of classes used (52 vs 10) and the less similarity of the selected characters in Legesse's work.

The answer of the research question which was mentioned in Chapter one is the system correctly recognizes the basic family and some derived Amharic alphabet signs based on NN and SVM models. From the two models, SVM is found to be more preferable classifier than NN as the test result in SVM shows very good performance. It is worth mentioning that our result and findings also render positive response towards answering the two research equations.

## **Chapter Six: Conclusion and Future Work**

### **6.1 Conclusion**

In Ethiopia there are millions of people who are living with hearing problem. These people need ETHSL as their communication channel among themselves. However, they are living with other people who communicate with spoken language. The hearing challenged people do not have a skill to use spoken language well while the hearing people do not have a skill to interpret the ETHSL. This gap makes the life of hearing impaired people very challenging. Local researches efforts to fill the gap among those people are very useful. This study is part of the efforts that were applied to solve the communication problem of hearing impaired people with the hearing ones.

In this study, an attempt has been made to design and implement a system which is capable of recognizing Amharic sign language characters. The system has four main parts: Image preprocessing, Segmentation, Feature extraction and Classification. The first three parts use various image processing techniques and the last uses two classifiers (NN and SVM) to conduct the required task.

The system starts by accepting RGB video frame of an Amharic alphabet signs. A video was captured with a 1280x720 dimension. Therefore, handling each pixel one by one was computationally expensive and hence before starting the process of segmentation, first the input video passes through some preprocessing tasks, namely, cropping, gray scale conversion, contrast adjustment and sharpening. Then the subsequent outcome passes through segmentation which is implemented by adaptive thresholding to prepare the binary image.

In addition to Amharic alphabet signs segmentation which discriminates the targeted object from back ground, extraction process that produces different feature vectors was done. The feature vectors were created by combining shape feature descriptors, motion feature descriptors and color feature descriptors altogether by applying different feature extraction algorithm which is Fourier descriptor algorithm.

The last part of the system is sign recognition. For sign classification, NN and SVM classifiers are employed. Experimental result illustrates that the NN and SVM classifiers achieved an overall accuracy of 57.82% and 74.06% respectively. The reason for the low overall system

performance is the non-existence of well-constructed corpus for experimentation. Thus, we captured videos of signs from different signers but a number of challenges occurred, including: Poor video quality, Mobile vibration, and lighting problem. The cause of these problems are due to the fact that we used junior signers which in return created inconsistency in hand flinging during signing.

In conclusion, our developed system that recognizes more alphabet signs than the existing or the previous works for Amharic sign language and it has promising recognition accuracy.

## **6.2 Future Work**

In this paper, we attained a result in recognizing Amharic sign language and it has its own impact on the Ethiopian sign language. However, there are gaps which should be filled in the future papers because the work cannot be used as a full translation system of Amharic Sign Language. The following are some of the recommendation that the researchers propose for future work:

- Improving this proposed design for better recognition by using other features that discriminate the alphabet signs.
- Enhancing Amharic sign language recognition system by implementing components that can handle the remaining specific, bastard Amharic alphabet signs and also Ethiopian number signs.
- This work has used SVM and NN and has a good evaluation result for recognition of Amharic sign language in SVM. Thus, extending this work to word, phrase or sentence level for the study of the language can be a good fit for future researches.
- Our work presented about one-way communication which means it only translates the sign to text. Therefore, we suggest that other researcher will design the system which should work like two-way communication to translate sign to text and vice versa.
- We used a uniform background (nearly white color) while we captured videos of the signs. However, in real world we are not able to find such uniform background. Therefore, future local researches should consider complex backgrounds on videos and images of signs.
- In this research work, the main challenge was obtaining enough data in ETHSL. There should be a database that includes all the Amharic alphabet signs with high frequencies

performed by different signers with great diversity in background and illumination. This will significantly help the training of classifiers not to be shallow.

## References

- [1] Masresha Tadesse, "Automatic translations of Amharic text to Ethiopian Sign Language," Master's Thesis, Addis Ababa Universty, Addis Ababa, Ethiopia, 2010.
- [2] Abady Tsegaye, "Offline candidate hand gesture selection and trajectory determination for continuous Ethiopian sign language," Master's Thesis, Addis Ababa University, Addis Ababa, Ethiopia, 2011.
- [3] Ethiopian National Association for Deaf, "BERTAT," Yearly Magazine, Ethiopia, 1997.
- [4] Najeefa Nikhat Choudhury and Golam Kayas, "Automatic Recognition of Bangla Sign Language," Bachelor's Project, Department of Computer Science and Engineering, BRAC University, (Knoxville - USA), 2012.
- [5] Daniel Mart´inez, "Sign Language Translator using Microsoft Kinect XBOX 360TM," Department of Electrical Engineering and Computer Science, University of Tennessee, USA, 2012.
- [6] Minilik Tesfaye, "Machine Translation Approach to Translate Amharic Text to Ethiopian Sign Language," Lecturer, Faculty of Informatics, St. Mary's University College, P.O.Box 18490, Addis Ababa, Ethiopia.
- [7] Pradeep Kumar and Manjunatha, "Performance Analysis of KNN, SVM and ANN Techniques for Gesture Recognition System," *Indian Journal of Science and Technology*, vol. 9(51), December 2016.
- [8] Nagarajan, Subashini and Balasubramanian, "Visual Interpretation of ASL Finger Spelling using Hough Transform and Support Vector Machine," *International Journal of Advanced Research in Computer and Communication Engineering* , vol. 4, no. 6, June 2015 .
- [9] Nagarajan and Subashini, "Static Hand Gesture Recognition for Sign Language Alphabets using Edge Oriented Histogram and Multi Class SVM," *International Journal of Computer Applications (0975 – 8887)*, vol. 82, no. 4, November 2013.
- [10] Tavari and Deorankar, "Implementation of Neural Network based Hand Gesture Recognition System," *International Journal Of Engineering And Computer Science*, vol. 3, no. 6, 2014.

- [11] Anand, Sachin and Urabinahatti, "Performance Comparison of Three Different Classifiers for HCI Using Hand Gestures," *International Journal of Modern Engineering Research (IJMER)*, vol. 2, no. 4, pp. 2857-2861, July-Aug 2012.
- [12] Naidoo, Omlin and Glaser, "Vision-Based Static Hand Gesture Recognition Using Support Vector Machines," *University of the Western Cape, South Africa*, on 22 May , 2014.
- [13] Yang Quan and Peng Jinye, "Chinese Sign Language Recognition for a Vision-Based Multi-features Classifier," in *International Symposium on Computer Science and Computational Technology*, Shaanxi Xi'an, P. R. China,, 2008.
- [14] Atiqur, Ahsan, Ibrahim and Sujit, "Recognition of Static Hand Gestures of Alphabet in Bangla Sign Language," *IOSR Journal of Computer Engineering (IOSRJCE)*, vol. 8, no. 1, pp. 7-13, 2012.
- [15] Legesse Zerubabel, "Ethiopian Finger Spelling Classification: A Study To Automate Ethiopian Sign Language," Master's Thesis, Addis Ababa University, Addis Ababa, Ethiopia, 2008.
- [16] Matthews and Baoill, " the Irish Deaf Community (Volume 2): The Structure of the Irish Sign Language," The Linguistics Institute of Ireland, Dublin, Ireland, 2000.
- [17] Samuel Teshome, "Isolated Word-level Ethiopian Sign Language Recognition," Master's Thesis, Addis Ababa University, Addis Ababa, Ethiopia, 2013.
- [18] Shujjat, Gourab, Donald, Serge and Chris, "Sign Language Analysis and recognition: A Preliminary Investigation," in *24th International Conference on Image and Vision Computing New Zealand*, (IVCNZ 2009).
- [19] Fikrte Shiferaw, "Challenges And Opportunities Of Teaching Signed Afan Oromo In Case Of Sebeta Special Needs Education Teachers' College," Master's Thesis, Addis Ababa University, Addis Ababa, Ethiopia, 2014.
- [20] Yonas Fantahun and Kumudha Raimond, "Ethiopian Sign Language Recognition Using Artificial Neural Network," in *10th International Conference on Digital Object Identifier*, PP 995-1000, 2010.
- [21] Daniel Zegeye, "Amharic Sentence To Ethiopian Sign Language Translator," Master's Thesis Addis Ababa University, 2014.

- [22] Ethiopian Sign Language Community, the deaf people of Ethiopia, people and language detail report, 2005.
- [23] Tefera Gimbi, "Recognition of Isolated Signs in Ethiopian Sign Language," Master's Thesis, Addis Ababa University, 2014.
- [24] Nicolas Pugeault and Richard Bowden, "Spelling It Out: Real-Time ASL Fingerspelling Recognition," in *IEEE Work-shop on Consumer Depth Cameras for Computer Vision*, Barcelona, Spain, 2011.
- [25] Ricco and Tomasi , "Fingerspelling Recognition through Classification of Letter-to-Letter Transitions," in *Proceedings of ACCV (3)*, pp. 214-225, 2009.
- [26] Susanna Ricco and Carlo Tomasi, "Finger Spelling Recognition through Classification of Letter-to-Letter Transitions," *In Proceedings of ACC*, vol. 3, pp. 214-225, 2010.
- [27] Joyeeta Singha, "Indian Sign Language Recognition Using Eigen Value Weighted Euclidean Distance Based Classification Technique," vol. 4, no. 2, 2013.
- [28] Daniel Assefa, "Amharic Speech Training For The Deaf," Master's Thesis, Addis Ababa University, Addis Ababa ,Ethiopia, 2006.
- [29] Alemayehu Teferi, "Ethiopian Sign Language Manual," User Guide on Ethiopian Sign Language, 2007.
- [30] Dr. Anand Singh, "Automatic Recognition of Dynamic Isolated Sign in Video for Indian Sign Language," Master's Thesis, GLA University, India, 2015.
- [31] Christine Fernandez-Maloigne, "Advanced Colour Image Processing and Analysis," Springer Science and Business Media New York, 2013.
- [32] Tinku Acharya and Ajoy K. Ray, *Image Processing Principles and Applications*, John Wiley , 2005.
- [33] Chaudhuri, Mandavya, Badela and Gosh, "Optical Character Recognition Systems for Different Languages with Soft Computing," *Springer International Publishing, Studies in Fuzziness and Soft Computing 352*, AG 2017.
- [34] Nadia, Albelwi and Alginahi, "Real-Time Arabic Sign Language Recognition," IEEE Paper, Research Gate, January 2012.
- [35] ZHANG Gang, MA Zong-min, NIU Lian-qiang and ZHANG Chun-ming, "Modified

- Fourier descriptor for shape feature extraction," in *Central South University Press and Springer-Verlag Berlin Heidelberg*, 2012.
- [36] Direkoglu, Mark and Nixon, "Shape Classification via Image-based Multiscale Description," *Pattern Recognition*, *ELSEVIER*, pp. 2134-2146, 2011.
- [37] Jonathan Rupe, "Vision-Based Hand Shape Identification for Sign Language Recognition," Master's Thesis, Rochester Institute of Technology, 2005.
- [38] Tahir, Hussain, Samad, Husain and Rahman, "Human Shape Recognition using Fourier Descriptor," *Journal of Electrical and Electronic Systems Research*, vol. 2, June 2009.
- [39] Corneliu Lungociu, "Real Time Sign Language recognition Using Artificial Neural Networks," *Studia Univ,Babes-Bolyai,Informatica*, vol. LVI, no. 4, 2011.
- [40] Amit Sharan, "Character Recognition Using Fourier Coefficients," Master's Thesis ,in Texas Tech University, December 1993.
- [41] Klimis Symeonidis, "Hand Gesture Recognition Using Neural Networks," Master's Thesis, UniS, On August 23, 2000.
- [42] Sanjay Meena, "A Study on Hand Gesture Recognition Technique," Master's Thesis, National Institute Of Technology, Rourkela, India, 2011.
- [43] Kenji Suzuki, *Artificial Neural Networks-Methodological Advances And Biomedical Applications*, In Tech Book, Ivana Lorkovic, 2011.
- [44] Coolen, Kühn and Sollich, *Theory of Neural Information Processing Systems*, New York: Oxford University Press, 2005.
- [45] George Awad, "A Framework for Sign Language Recognition using Support Vector Machines and Active Learning for Skin Segmentation and Boosted Temporal Sub-units," Master's Thesis Dublin City University(DCU), in 2007.
- [46] D.Karthikeyan and Mrs.G.Muthulakshmi, "English Letters Finger Spelling Sign Language Recognition System," *International Journal of Engineering Trends and Technology (IJETT)*, *M.S.University,Tirunelveli,India*, vol. 10 , no. 7, Apr 2014.
- [47] Nikhil, Shreyas, Sumesh, Gowranga and Bhakthavathsalam, "Implementation and Comparison of Machine Learning Algorithms for Recognition of Fingerspelling in Indian Sign Language," *Society For Science And Education United Kingdom*, vol. 5, no. 5, 20th

August, 2017.

- [48] Marijana, Sanja and Natasa, "A Comparison of Machine Learning Methods in a High-Dimensional Classification Problem," *Business Systems Research, DE GRUYTER*, vol. 5, no. 3, September 2014.
- [49] Azher Uddin and Shayhan Ameen, "Hand Sign Language Recognition for Bangla Alphabet using Support Vector Machine," *IEEE International Computer Science and Engineering Conference*, 2016.
- [50] Michael Kearns and Dana Ron, "Algorithmic Stability and Sanity-Check Bounds for Leave-One-Out Cross-Validation," AT&T Labs Research, Murray Hill, New Jersey, January 1997.
- [51] Walter Daelemans and Antal van den Bosch, "Memory-Based Language Processing," Cambridge University Press 978-0-521-80890-3, 2009.

# Appendixes

## Appendix A: Sample Data Used for System Design





## Appendix B: MATLAB Code

### I. Segmentation

#### **%Read from video source**

```
ff= fullpathname;
obj = VideoReader(ff);
nframes = get(obj, 'NumberOfFrames');
%Take sample points for hue difference measure
LF=read(obj, nframes);
FF=read(obj, 1);
[w,hi]=size(FF);
SZ=floor(nframes/2);
```

#### **%GetNumber of Bin to take sample frames**

```
DESCRIPTOR_QTY=floor(nframes/10);
index=1;
xcor=[];
ycor=[];
[xcor, ycor, LF, FF, DESCRIPTOR_QTY,BASE_FRAME, END_FRAME,
nframes, BF_IDX] = video_segmentation(obj);
```

```
[L,no]=bwlabel(BASE_FRAME,8);
BASE_FRAME=dispLabel(L,BF_IDX, BASE_FRAME);
BASE_FRAME=logical(BASE_FRAME);
SE=strel('disk',20);
BASE_FRAME=imerode(BASE_FRAME,SE) & BASE_FRAME ;
BASE_FRAME=bwareaopen(BASE_FRAME, 2000);
SE=strel('disk',30);
BASE_FRAME=imdilate(BASE_FRAME,SE) & BASE_FRAME ;
Return BASE_FRAME
```

```
[xcor, ycor, LF, FF, DESCRIPTOR_QTY,BASE_FRAME, END_FRAME,
nframes, BF_IDX]=Function video_segmentation(obj)
```

#### **%Segment only the selected frames**

```
for k=1:DESCRIPTOR_QTY:nframes
    col=read(obj, k);
    [h,s,v]=rgb2grey(col);
    Sframe=zeros(size(FF,1),size(FF,2));
    for i=1: size(FF,1)
        for j=1: size(FF,2)
```

```

    if (( h(i,j) >= .009 && h(i,j) <= .09 ) && s(i,j) >= .23)
        Sframe(i,j)=1;
    end
    end
end
    SE=strel('diamond',10);
    Sframe=imdilate(Sframe,SE) & Sframe ;
    Sframe=bwareaopen(Sframe,1500);
    Sframe=imfill(Sframe,'holes');
    stats = regionprops(Sframe, {'Centroid','Area'});
    areaArray = [stats.Area];
    [junk,idx] = max(areaArray);
    c = stats(idx).Centroid;
    c = floor(fliplr(c));
    xcor(index)=c(2);
    ycor(index)=c(1);
if k==1
        BASE_FRAME=Sframe;
        BF_IDX=idx;
end
    index=index + 1;
    END_FRAME=Sframe;
end
%End function
end

```

## II. Feature Extraction

```

% Get list of x-coordinates xcor,list of y-coordinates,
% ycor,First frame,FF,Bin size, DESCRIPTOR_QTY, index of base
%frame, and Last frame BF_IDX,LF
LF=handles.LF;
FF=handles.FF;
xcor=handles.xcor;
ycor=handles.ycor;
BF_IDX=handles.BF_IDX;
DESCRIPTOR_QTY=handles.DESSCRIPTOR_QTY;
BASE_FRAME=handles.BASE_FRAME;
[VFS, DES, CON,DESP, CONP,FE]= Extract_Video_Features (xcor,
ycor,LF, FF,DESCRIPTOR_QTY, BASE_FRAME, BF_IDX);

```

**Function** [VFS, DES, CON,DESP, CONP,FE]= Extract\_Video\_Features (xcor, ycor,LF, FF,DESCRIPTOR\_QTY, BASE\_FRAME, BF\_IDX)

```

[DES, CON,DESP, CONP,FE, VAL]=DESC_FOURIOR((BASE_FRAME));
VFS2=VAL;
%color difference measure Last frame (LF) and first frame
%segment the last channel and compare and apply color difference
% measure between the last and first channel

[h0,s0,v0]=rgb2hsv(LF);
[h,s,v]=rgb2hsv(FF);
len=numel(xcor);
coldiff=sqrt(abs(h(ycor(len), xcor( len )) - h0(ycor(1), xcor(1)
)) + abs(s(ycor(len), xcor(len))-s0(ycor(1), xcor(1))) +
abs(v(ycor(len),xcor(len))-v0(ycor(1),xcor(1))));

%Degree of coliniarity/ sinusoidal motion test metrice
teta=[]; index=1;sumteta=0;
for i=1:DESCRIPTOR_QTY-2
    teta(index) = tetaof3( [xcor(i),ycor(i)] ,
[xcor(i+1),ycor(i+1)] , [xcor(i+2),ycor(i+2)] );
    sumteta=sumteta + abs(teta(index));
    index=index + 1;
end
sumteta=deg2rad(sumteta);
%direction of object movment
movType=-1;%unpridictable
count=numel(xcor);

%compute eculidian distance for evaluating positional value
po=[xcor(count), ycor(1)];p1=[xcor(1), ycor(count)];
d1=distancePoints([xcor(count),ycor(count)], po);
d2=distancePoints([xcor(count),ycor(count)], p1);
if((xcor(count)-xcor(1)) >0 && (ycor(count)-ycor(1)) <=0 && d1
<= d2)
    movType=1;%move to the left
    elseif((xcor(count)-xcor(1)) >0 && (ycor(count)-ycor(1)) >0
&& d2 < d1)
        movType=3;%move down
    elseif((xcor(count)-xcor(1)) <0 && (ycor(count)-ycor(1)) >0
&& d2 <= d1)
        movType=3;%move down
    elseif((xcor(count)-xcor(1)) <0 && (ycor(count)-ycor(1)) )

```

```

        movType=2;%move right
    elseif((xcor(count)-xcor(1)) <0  && (ycor(count)-ycor(1)) >0
&& d2 > d1)
        movType=2;%move right
    elseif((xcor(count)-xcor(1)) >0  && (ycor(count)-ycor(1)))
        movType=1;%move left
end

```

```

VFS1(1)=movType;
VFS1(2)=sumteta;
VFS1(3)=coldiff;
VFS2(32)=VFS1(1);
VFS2(33)=VFS1(2);
VFS2(34)=VFS1(3);
VFS=VFS2;
end

```

**%A function to compute furious descriptor**

```

Function [xe,ye,xer,yer,Fe,RFV]=DESC_FOURIOR(img)
    img=logical(img);
    b_cell=bwboundaries(img);
    xe= b_cell{1}(:,2);
    ye= -b_cell{1}(:,1);
    xer=P_RES(xe,128);
    yer=P_RES(ye,128);
    %make a 1D complex signal out of the contour points
    fe=xer+i*yer;
    Fe=fft(fe,128); %and compute its FDs
    RFV=normFD(Fe,[-16,-15,-14,-13,-12,-11,-10,-9,-8,-7,-
6,-5,-4,-3,-2,-1, 2, 3, 4,5,6,7,8,9,10,11,12,13,14,15,16]);
return

```

### III. Recognition

```

%Get extracted video features
ExtractedVFS=handles.ExtractedVFS;
valNN={};
valSVM={};
input=[];
output=[];

```

```

%Load the dataset
DS=load('dataset.mat');
input=DS.input;
input=cell2mat(input);
output=DS.output;

%Convert the output in terms of row data
O=output;
val=unique(O);
for i=1:numel(unique(O))
for j=1: size(O,2)
    if output(j) == val(i)
        O(i,j)=1;
    else
        O(i,j)=0;
    end
end
end

    %Load each constricted model
ModelSVM=load('svmnet.mat');
ModelNN=load('nnNet.mat');

%% Testing the Neural Network
test=ExtractedVFS;
[a,b]=max(sim(ModelNN.net,test));
numNN=round(b);
nn_pridicted=numNN;
if( numNN > 0)
    valNN=numNN;
end
##### case II using multi class SVM
predict=svm.predict(ModelSVM.svmnet,test);
numSVM=round(predict);
if(numSVM > 0)
    valSVM=numSVM;
end
set(handles.editresult, 'string','');
strnn = valNN;
strsvm = valSVM;
if isempty(str{1,strnn})

```

```
    nndisp=strn;
    else
    nndisp=str{1,strn};
end
if isempty(str{1,svms})
    svmdisp=svms;
    else
    svmdisp=str{1,svms};
end
set(handles.editresult, 'string',nndisp);
set(handles.edit4, 'string',svmdisp);
```



