

**Bangla Sign Language Conversation Interpreter Using Image Processing  
(CNN, Tensorflow and Keras)**

**BY**

**Prosenjit Roy**

**ID: 151-15-4815**

**Md. Arifur Rahman**

**ID: 151-15-5109**

**And**

**Md. Radwan Naeem Manik**

**ID: 151-15-5039**

This Report Presented in Partial Fulfillment of the Requirements for the Degree  
of Bachelor of Science in Computer Science and Engineering

Supervised By

**Fahad Faisal**

Senior Lecturer

Department of CSE

Daffodil International University

Co-Supervised By

**Saiful Islam**

Lecturer

Department of CSE

Daffodil International University



**DAFFODIL INTERNATIONAL UNIVERSITY**

**DHAKA, BANGLADESH**

**DECEMBER 2018**

## **APPROVAL**

This Project/internship titled “**Bangla Sign Language Conversation Interpreter Using Image Processing**”, submitted by **Prosenjit Roy** ID No: 151-15-4815, **Md. Arifur Rahman** ID No: 151-15-5109 and **Md. Radwan Naeem Manik** ID No: 151-15-5039 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on **December 9, 2018**.

### **BOARD OF EXAMINERS**

---

**Prof. Dr. Syed Akhter Hossain**  
**Professor and Head**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Chairman**

---

**Narayan Ranjan Chakraborty**  
**Assistant Professor**

Department of Computer Science and Engineering  
Faculty of Science & Information Technology  
Daffodil International University

**Internal Examiner**

---

**Md. Tarek Habib**  
**Assistant Professor**

Department of Computer Science and Engineering  
Daffodil International University

**Internal Examiner**

---

**Dr. Mohammad Shorif Uddin**  
**Professor**

Department of Computer Science and Engineering  
Jahangirnagar University

**External Examiner**

## DECLARATION

We hereby declare that this research has been done by us under the supervision of **Fahad Faisal, Senior Lecturer of Department of CSE**, Daffodil International University. We also declare that neither this research nor any part of this research has been submitted elsewhere for the award of any degree.

**Supervised by:**

**Co- Supervised by:**

---

**Fahad Faisal**  
Senior Lecturer  
Department of CSE  
Daffodil International University

---

**Saiful Islam**  
Lecturer  
Department of CSE  
Daffodil International University

**Submitted by:**

---

**Prosenjit Roy**  
ID: - 151-15-4815  
Department of CSE  
Daffodil International University

---

**Md. Arifur Rahman**  
ID: - 151-15-5109  
Department of CSE  
Daffodil International University

---

**Md. Radwan Naeem Manik**  
ID: - 151-15-5039  
Department of CSE  
Daffodil International University

## ACKNOWLEDGEMENT

First we would like to express our heartiest thanks and gratefulness to Almighty Allah for his divine blessing which has made us possible to complete the project successfully. We are grateful to our parents for their motivation and support.

We are grateful and indebted to our respected supervisor **Fahad Faisal**, Senior Lecturer, Department of CSE Daffodil International University, Dhaka. His endless patience, scholarly guidance, continual encouragement, constant and energetic correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Professor Syed Akhter Hossain**, Head Department of CSE, for his kind help to finish our project and also to other faculty member and the staff of CSE Department of Daffodil International University.

We would like to thank our entire course mate in Daffodil International University, who took part in this discuss while completing the course work.

Finally, we must acknowledge with due respect the constant support and patience and believe of our parents which help us to complete this long journey till childhood.

## **ABSTRACT**

This project about sign language conversation interpreter system. It is made for speak disable people. Here a man who don't know the sign language can make conversation with a speak disable people. In this project we used efficient methods to convert sign language into text. We used our own dataset and collected dataset of Bangla Sign Languages using hand gestures. Inputs will take by webcam and recognize by neural network system and give text output by using gestures. We have created a CNN which is similar model of the MNIST classifying model. It can use both Tensorflow and Keras. This model was trained using Keras by video streaming. In a video stream sign language, interpreter can turn it fast into text. We used histogram back projection method to recognize people's hand as object. For better result we calculated the grey level of hands.

## TABLE OF CONTENTS

<b>CONTENTS</b>	<b>PAGE</b>
Acknowledgement	III
Abstract	IV
List of Figures	VII
List of Tables	VIII
<b>Chapter</b>	<b>PAGE</b>
<b>Chapter 1: Introduction</b>	<b>1-2</b>
1.1 Introduction	1
1.2 Motivation	1
1.3 Objective	1
1.4 Expected Outcome	2
1.5 Report Layout	2
<b>Chapter 2: Background Study</b>	<b>3-6</b>
2.1 Introduction	3
2.2 Related work	3
2.3 Comparative study	3
2.3.1 First ever sign language	4
2.3.2 Real time hand gestures recognition	4
2.3.3 Bengali sign language and recognition	4
2.3.4 Neural network based	5
2.4 Scope of problem	5
2.5 Challenges	6
<b>Chapter 3: Research Methodology</b>	<b>7-12</b>
3.1 Introduction	7
3.2 Research subject and instrumentation	7
3.2.1 Research subject	7
3.2.2 Instrumentation	7
3.3 Data collection procedure	8

3.3.1 Collection images	8
3.3.2 Removing background	9
3.3.3 Creative gestures	9
3.3.4 Training of dataset	10
3.3.5 Testing gestures and results	11
3.4 Statically analysis	12
3.5 Requirements to implement	12
<b>Chapter 4 : Experimental results and discussions</b>	<b>13-16</b>
4.1 Introduction	13
4.2 Experimental results	13
4.3 Descriptive analysis	14
<b>Chapter 5: Conclusion and Future works</b>	<b>17</b>
5.1 Conclusion	17
5.2 Future work	17
<b>References</b>	<b>18-19</b>

## List of Figures

Figure 2.3.1.1: Some American sign language	4
Figure 2.3.3.1.1: তুমি	5
Figure 3.3.1.1: Data collection	8
Figure 3.3.2.1: Image conversion	9
Figure 3.3.3.1: Hand set up	10

## **Lists of Tables**

Table 3.3.5.1: Data set	11
Table 4.2.1: Accuracy for train set classes	13
Table 4.3.1: Accuracy measure based on performance matrix	15

# Chapter 1

## Introduction

### 1.1 Introduction

In this modern era, we have advanced technology to utilize human ability more and more. But some people are physically disabled by birth or for unexpected accidents. In this world about 15% of people are speech disabled (according to WHO reports). It's about 8% of the Bangladesh's population. They have to lead their life as a burden in society. This burden word is applied because they can't make conversation with normal people and that's why they can't gain skills of any work.

Basically people who are hearing disabled and speech disabled, they need to learn sign language for communication. At this matter, for communication, sign language is the preferred form. But most of the common people do not understand sign languages. For making communication between the people who use sign language and the people who do not understand sign language, sign language interpreter's service can be arranged. In this way of communication, it is important to know how to make use of the services of the system effectively.

Every country has their own sign languages. Our country has also well-formed sign language. Many researchers are working on sign language recognition systems. In our country most researchers worked on Bangla sign language, a system developed by Pavel et al. [1] analyzes video clips of different gestures of sign languages taken as input and gives audio output.

### 1.2 Motivation

We have a huge population who are speaking and hearing impaired. In our development country they have a very little chance to make them self-dependent as they can't communicate with others. If they can communicate, they can do all works. They have to face under humiliation situation in every step of life.

### 1.3 Objectives

Now-a day's people mostly depend on technology. We want to speech disabled and hearing disabled people can lead a modern life through this advanced technology. Our main goal, they

can communicate all common people and take parts in all of work of our life. We want to make this people our assets not burden in our society.

### **1.4 Expected Outcome**

The expected outcomes of the project are given below:

1. Converting sign languages into text where sign language will captured by webcam.
2. First recognizing sign languages word by used Neural Networks.

### **1.5 Report Layout**

The project report contains five chapters. Summarization is given below:

The motivation, objectives and expected outcome have been discussed in details in chapter one.

The background of our project which is covered with related works of the application and discussion of the problem and challenges of the system is discussed in chapter two.

In chapter three there will be Research Subject and Instrumentation, Data Collection Procedure, Statistical Analysis. Implementation Requirements.

Chapter four will contain Descriptive Analysis. Testing results.

Finally, chapter five have Conclusion, Recommendation and Implication for Future Research.

## **Chapter 2**

### **Background Study**

#### **2.1 Introduction**

Though there are quiet similar research and project about bangle sign language. Most of them about Bengali sign language alphabets interpreter. But with Bengali sign language alphabets, it's very tough to make conversation with common people. Without communication with people, hearing and speaking impaired persons will remain ignored in our society. They need a conversation interpreter about their Bengali conversation sign language.

#### **2.2 Related works**

When we selecting this project, we tried to find similar projects. We find some attractive related works. But all of them about sign language alphabet interpreter in our country. We will combine all Bengali sign language words to interpret it.

1. A project about Hand sign language recognition for Bangla alphabet using Support Vector Machine have done by MdAzher Uddin[2]
2. RahatYasir did a project about Two-Handed Hand Gesture Recognition for Bangla Sign Language using LDA and ANN.

Others either developed this systems or proposed new model to convert Bengali sign language alphabet have done by Rahat Yasir[3]

#### **2.3 Comparative Studies**

We studied about Bengali sign language alphabet existing project's details. We also studied about other sign language converting projects of different countries. In this present world, others countries researchers are trying to make a system for sign language word converting.

Our main goal is to make a easy and effective system to convert Bengali sign language words.

### 2.3.1 First ever sign language

In 17<sup>th</sup> century western countries sign language developing was started for a visual language where the conventional gesture, hand signs was, finger spelling also includes the positions to represent a meaningful line. In Europe, Laurent Clerc first taught French sign language. Then Thomas Hopkins Gallaudet brought Clerc to America. His plan was to start American school for deaf. After that from all over the world children came to attend this school and brought their home sign language. Before 19<sup>th</sup> century whole sign language system was confined only by fixed words using finger spelling system [13]. They probably figured and created signs to understand and speak with deaf people. Here is a figure of American Sign Language where it represent correspondingly “A”, “S”, “L”.



Figure 2.3.1.1:Some American sign language “A”, “S”, “L”

### 2.3.2 Real time based bare hand Gesture Recognition

Khedkhar Safaya and Prof. (Dr.). J.W.Bakal of Mumbai researched on Real time based hand gestures in 2013. They used a dynamic vision sensor which can recognize bare hand gesture. They figured out the delivery point track number for every frame and compared with given threshold values. They took this point to detect frame and find the hand position. They calculated the wrist point in the extraction phase used their own algorithm to define the hand sign.

### 2.3.3 Bengali sign language and Recognition

In our country our Bengali sign language is formalized in 2003 by CDC (Center for Disability in Development). Bengali language is the world’s 5<sup>th</sup> language in terms of population. There are 47 alphabets in Bengali language where is 11 vowels called “SOROBORNO” and 36 constant called “BENJONBORNO”. But it’s true that in conversation people don’t use alphabets. That’s why sign words should interpreted. American sign language recognition is

researched and developed in high procession. On the other hand Bengali sign language recognition field is still remain behind. In figure 2.3.3.1 a Bengali sign language which means “তুমি”

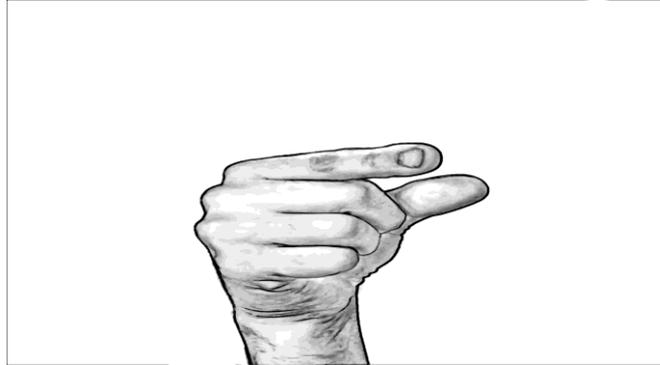


Figure 2.3.3.1: তুমি

### 2.3.4 Neural Network Based

Bikash Chandra Karmokar[14] and his team of Technology did a research employing Neural Network Ensemble. They took input from webcam. They also detected skin color of hand. After that captured image had been converted into its threshold value. In 30x33 scale pixels, images were converted by applying normalization process. Then feature extraction method has been applied and they used NCL algorithm to train these images.

## 2.4 Scope of the problem

This project is similar to other Bengali sign language projects and have a little bit difference. As we are working on Bengali sign language words . We are also working on video streaming process. We are to face some remarkable problems.

The scope of the problems are discussed below :-

1. As we want to give conversational output from video streaming, there can be problem to get input if the user move hands very fast.
2. Capturing images from video streaming can be face big problem because webcam has light issues. If light changes then system will not get clear image vision to recognize appropriate output.

3. As we working on sign language conversational words, there can be big problem in data collecting and analyzing. The sign language words sometimes vary district to district in our country.

## **2.5 Challenges**

In every sphere of life there are challenges but we are to overcome the challenges to make something good. The challenges we have faced are given below:

1. **Proper Dataset**

It was one of the biggest problems for us that we spent a huge time to collect proper dataset.

2. **Problem finding**

As some related applications have already been made, we tried heart and soul to find the bugs and missing features that should have been in the system.

3. **Model**

We faced some problems to choose which Model for image processing system.

4. **Time management**

It was one of the difficult parts to maintain. We always kept in our mind about the deadline. If we couldn't do it in time, it would be very dangerous for us. Thus we had to divide our time to develop the system part by part.

## **Chapter 3**

### **Research Methodology**

#### **3.1 Introduction**

Throughout this chapter we will discuss about our working procedure which will be divided into some steps. How we collect data, which method we used to processed those data, how the interface working, statistical analyzing of out puts and implementation requirements.

#### **3.2 Research Subject and Instrumentation**

##### **3.2.1 Research Subject**

Our main target is to capture image by our system and match it with our data set and show the outputs as Bangla language. For this work we take a lot of images of different peoples hand sign in different different specific gesture. Then we created 15 different gesture samples using OpenCV Library. We convert those captured images into 50x50 pixels and made our dataset. This system works in two steps. At first this system will take dataset images then it applies the skin detection algorithm and detect the skin and remove the background.

We use Convolutional Neural Network (CNN) in this system. And we created a CNN is looks a lots of similar to classifying model using both Tensorflow and Keras. If anyone want to add more gesture images in this project they need to make their own layer and some parameters. We trained this model using Keras on videos stream. Which trained by 4 Bangla font, 5 numbers and 6 words.

##### **3.2.2 Instrumentation**

For run this project in any platform you need some instruments and have to meet with some requirements. The instruments and requirements are given:

1. Install Python 3.x in your operating system
2. Import Tensorflow 1.5 using command
3. Need OpenCV 3.4
4. KerasAPI
5. CPU

### 3.3 Data Collection Procedure

For data collection and processing them we followed some steps. They are given below in the points.

#### 3.3.1 Collection Images

At this time Bangla sign language recognition is not still a widely researched topic so we did not find any dataset on any resources. So that, we made our own dataset with the help of our classmates. We took images of bare hands in different backgrounds. We collect some standard gesture from the internet for making our data set. Then capture images from hundred people bare hands for fifteen different signs.



Figure 3.3.1.1: Data collection

Those signs includes 4 Bangla font, 5 number and 6 words. Some of them are shown in fig: 3.3.1.1 . For unavailable of Bangla signs gestures data set we also use American English gestures for making our dataset. Make collection of hand gesture was very hard because normal people are not comfortable to make the hand gesture. Collecting those data was the first step for ours project.

### 3.3.2 Removing Background

As we use some similar dataset of American signs this step easy for us. But for our collected image we have to apply skin detection algorithm which detect the skin color pixel from it and identify the blob and remove all binary objects. Then it convert the image to inverted image. So the hand gestures turned into a black and white image. Example shown in fig 3.3.3.2.1



Figure 3.3.2.1: Image conversion

### 3.3.3 Creating Gesture

We make our collected images to gesture using some python script which makes our process easy to convert those images. First we set histogram for the light condition changes. The command is given below for this project.

Here we used histogram back projection method for recognizing human hand as object. We count hand's grey level for recognizing object. Our histogram set up is needed to change in every light changing situation as it is light sensitive.

Here our histogram set up command for every light changing situation.

```
python set_hand_hist.py
```

In this set\_hand\_hist.py file we write some code to set make histogram in our camera position for 50x50 pixel from where the camera only hand gesture (in fig 3.3.3.3) for making real time gesture to compare with our dataset.

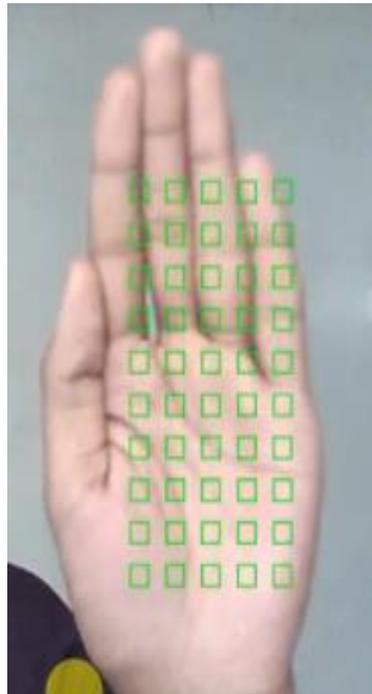


Figure 3.3.3.1: Hand set up

### 3.3.4 Training of Dataset and Model

For training our dataset we use two library of CNN model. One is Tensorflow and another is Keras. Keras can run over Tensorflow. So that, anyone can use either Tensorflow or Keras. Keras is a modified model of CNN. By using these we trained our dataset and we got good result to convert the sings to text which can make a conversation between deaf and numb in Bangla in future.

### 3.3.5 Testing Gestures and Results

For testing our model we make some gestures in front the camera to test our model. As we read a lot of papers and try some similarities projects, we understood that our model is working as well. Some of the results are given below in table 3.3.5.1

Table 3.3.5.1: Data set

১		
২		
৩		
৪		
৫		
খুব ভালো		

ক	
থামুন	

### 3.4 Statistical Analysis

The main thing is getting output. The more we train our system and give more dataset, the more we get best results. We worked here some number of characters and numbers. We tried to get best results and best output. So we make a statistics among the result and output. We discuss briefly in chapter 4.

### 3.5 Requirements to Implement

You need some instruments and have to meet with some requirements. The instruments and requirements are given below:

1. Install Python 3.x in your operating system
2. Import Tensorflow 1.5 using command
3. Need OpenCV 3.4
4. KerasAPI
5. CPU

After complete install and follow the processes that are written above chapters you can run this project and you can see results.

## Chapter 4

### Experimental Results and Discussion

#### 4.1 Introduction

In our proposed system to measure the performance we use fifteen types of signs and in future we want to increase the numbers of sign. We use 100 images from 15 datasets. Our first dataset contains 100 images and the average accuracy is 81%.

#### 4.2 Experimental Result

Total class of images =15;

In one class number of images = 100;

Our final dataset contain 1500 images.

Table 4.2.1: Accuracy for train set classes

Class	Train set average accuracy for each class	Class	Train set average accuracy for each class	Class	Train set average accuracy for each class
আমি	82%	১	81%	শুভকামনা	75%
তুমি	79%	২	78%	ক	78%
ভাল	80%	৩	80%	খ	82%
ভালোবাসি	83%	৪	80%	গ	80%
থাম	82%	৫	85%	ঘ	81%

## 4.3 Descriptive Analysis

### Get model report

i. To get classification reports about the model we have test\_images and test\_levels file which have generated load\_image.py .If we do not have them run load\_images.py file again. Then we have to run below file.

Python get\_model\_report.py

ii. We will get the confusion matrix, f scores, precision and recall for the prediction by the model.

**Precision:** precision is the piece of related instances among the retrieved instances. High precision means that an algorithm returned substantially more relevant results than irrelevant ones.

**Recall:** Recall is the part of relevant instances that have been retrieved over the total amount of relevant instances. High recall means that an algorithm come back most of the relevant outcome.

**F-measure:** f-score is a measure of exam's accuracy by considering both precision and recall. it is a harmonic average of precision and recall.

**Accuracy:** accuracy indicates to the similarity of the measured value to a known value.

**False Positive Rate:** False positive rate is refer to the proposed method predict the gesture as another gesture and show wrong output.

**Specificity:** Specificity refers to the proposed method predict the gesture result perfectly as output.

Table 4.3.1: Accuracy measure based on performance matrix

Class	Recall %	False positive rate %	Specificity %	Precision %	F-measure %	Accuracy %
আমি	82	11	81	78	87	90
তুমি	81	9	78	83	90	91
ভাল	80	12	91	81	81	92
ভালোবাসি	78	11	79	84	78	87
থাম	79	10	93	82	78	86
১	82	7	82	86	88	83
২	82	9	82	82	79	86
৩	80	10	81	79	90	82
৪	83	15	89	81	87	83
৫	86	01	79	82	88	86
শুভকামনা	78	11	81	89	82	90
ক	79	12	82	88	88	84
খ	80	12	83	78	78	81
গ	82	0	81	81	78	91
ঘ	82	11	78	82	89	81

Average	81	9	83	82	84	86
---------	----	---	----	----	----	----

### Points scored

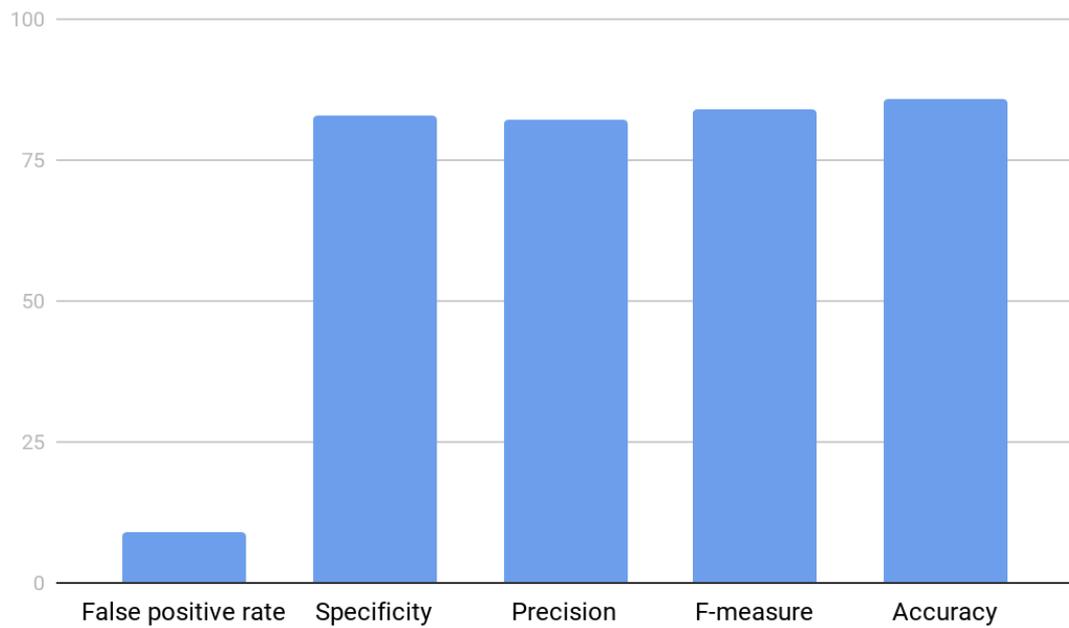


Figure 4.3.1: Statically analysis of accuracy, F-measure, Precision, Specificity, false positive rate.

## **Chapter 5**

### **Conclusion and future works**

#### **5.1 Conclusion**

In Bangladesh about 4.8 million people who are disabled in hearing or speaking. This huge population can be our asset if they can make conversation with us. In this paper, we showed about 15 combined Bengali sign language where has Bengali sign language alphabets, some Bengali sign language Numeric numbers and individual signs too. Our main goal is interconnect this speaking and hearing impaired people with common people. Without this interconnection, We can not turn them into skillful and self-dependent population.

#### **5.2 Future Works**

In the future, Our goal is to make this system more efficient. We collected our own data-set. That's why It took huge time to collect data and train data. In future our target is to collect and combine all Bengali sign language alphabets , all Bengali sign language numeric numbers and other individual Bengali sign languages. This is modern era. Most of the people in our country use smart phones. We will work to make a phone-based application of this system for making easy communication with all.

## References

- [1] D. Shahriar Hossain Pavel, Tanvir Mustafiz, Asif Iqbal Sarkar, M. Rokonzaman, "Geometrical Model Based Hand Gesture Recognition for Interpreting Bengali Sign Language Using Computer Vision", ICCIT, 2003.
- [2] Md Azher Uddin ; Shayhan Ameen Chowdhury , "Hand sign language recognition for Bangla alphabet using Support Vector Machine", 2016 International Conference on Innovations in Science, Engineering and Technology (ICISSET), 2016.
- [3] Mahmood Jasim, Md. Hasanuzzaman, "Sign Language Interpretation using Linear Discriminant Analysis and Local Binary Patterns", 2014 International Conference on Informatics Electronics & Vision (ICIEV), pp. 1-5, 2014.
- [4] Rahat Yasir, Riasat Azim Khan, "Two-Handed Hand Gesture Recognition for Bangla Sign Language using LDA and ANN", 2014 8th International Conference on Software Knowledge Information Management and Applications (SKIMA), pp. 1-5, 2014.
- [5] Tanzila Ferdous Ayshee, Sadia Afrin Raka, Quazi Ridwan Hasib, Md. Hossain, Rashedur M Rahman, "Fuzzy Rule-Based Hand Gesture Recognition for Bengali Characters", 2014 IEEE International Advance Computing Conference (IACC), pp. 484-489, 2014.
- [6] Salma Begum, Md. Hasanuzzaman, "Computer Vision-based Bangladeshi Sign Language Recognition System", 12th International Conference on Computers and Information Technology (ICCIT) 2009, pp. 414-419, 2009.
- [7] Matheesha Fernando, Janaka Wijayanayaka, "Low cost approach for Real Time Sign Language Recognition", 2013 8th IEEE International Conference on Industrial and Information Systems (ICIIS), pp. 637-642, 2013.
- [8] Shreyashi Narayan Sawant, M. S. Kumbhar, "Real Time Sign Language Recognition using PCA", 2014 International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), pp. 1412-1415, 2014.
- [9] J.R. Movellan, "Tutorial on Gabor Filters", 2008, [online] Available: <http://mplab.ucsd.edu/tutorials/gabor.pdf>.
- [10] M. Haghghat, S. Zonouz, M. Abdel-Mottaleb, "Identification using encrypted biometrics" in Computer analysis of images and patterns (CAIP), Springer, pp. 440-448, 2013.
- [11] Milad Moghaddam, Manoochehr Nahvi, Reza Hassanzadeh Pak, "Static Persian Sign Language Recognition Using Kernel-Based Feature Extraction", 7th Iranian conference on Machine Vision and Image Processing (MVIP), pp. 1-5, 2011.
- [12] A. Kocsor, L. Toth, "Kernel-based feature extraction with a speech technology application", Signal Processing IEEE Transactions on, vol. 52, pp. 2250-2263, 2004.
- [13] Baker-Shenk, Charlotte Lee, and Dennis Cokely. American Sign Language: A teacher's resource text on grammar and culture. Gallaudet University Press, 1991.

[14] Bikash Chandra Karmokar, Kazi Md. RokibulAlam and Md. KibriaSiddiquee: “Bangladeshi Sign Language Recognition employing Neural Network Ensemble “,International Journal of Computer Applications (0975 – 8887) Volume 58– No.16, November 2012