COMPUTER VISION-BASED BANGLA NUMERICAL SIGN LANGUAGE RECOGNITION

BY

MD. TOUHIDUR RAHMAN
ID: 152-15-6232

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

Supervised By

Prof. Dr. Monzur Morshed
Adjunct Professor
Department of CSE
Daffodil International University

Co-Supervised By

Prof. Dr. Md. Hasanuzzaman
Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY

DHAKA, BANGLADESH

MAY 2018

APPROVAL

This project titled "**Computer Vision-Based Bangla Numerical Sign Language Recognition**" submitted by Md. Touhidur Rahman, ID No: 152-15-6232 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 May 2018.

BOARD OF EXAMINERS

_____

Dr. Syed Akhter Hossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____

Dr. Sheak Rashed Haider Noori — **Internal Examiner**
Associate Professor and Associate Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____

Md. Zahid Hasan — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

_____

Dr. Mohammad Shorif Uddin — **External Examiner**
Professor
Department of Computer Science and Engineering
Jahangirnagar University

DECLARATION

I hereby declare that, this project has been done by me under the supervision of **Professor Dr. Monzur Morshed**, Adjunct Professor of Department of CSE of Daffodil International University. It is also declared that neither this thesis paper nor any part thereof has been submitted anywhere else for the award of any degree, diploma or other qualifications.

Supervised by:

_____

Prof. Dr. Monzur Morshed
Adjunct Professor
Department of CSE
Daffodil International University

Co-Supervised by:

_____

Prof. Dr. Md. Hasanuzzaman
Professor
Department of CSE
Daffodil International University

Submitted by:

_____

Md. Touhidur Rahman
ID: 152-15-6232
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First, I express my heartiest thanks and gratefulness to almighty Allah for His divine blessing makes us possible to complete the final year thesis successfully.

I am really grateful and wish my profound my indebtedness to **Professor Dr. Monzur Morshed**, **Adjunct Professor**, Department of CSE Daffodil International University, Dhaka. Deep Knowledge & keen interest of my supervisor in the field of "*Sign Language Recognition System*" to carry out this thesis. His endless patience, scholarly guidance, continual encouragement, constant and energetic supervision, constructive criticism, valuable advice, reading many inferior draft and correcting them at all stage have made it possible to complete this thesis.

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

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

Finally, I must acknowledge with due respect for the constant support and patients of our parents.

ABSTRACT

Sign Language is the mode of communication among the deaf and dumb. However, integrating them into the main stream is very difficult as the majority of the society is unaware of their language. So, to bridge the communication gap between the hearing and speech impaired and the rest in Bangladesh, I conducted a research to recognize Bangla sign language using a computer-vision based approach.

Sign language not only help for the people who can't speak or hear, it's also help for human computer interaction system or robotics system. To achieve my goals, I used Convolutional Neural Networks (CNNs) to train individual signs. In the future, this research, besides helping as an interpreter, can also open doors to numerous other applications like sign language tutorials or dictionaries and also help the deaf and dumb to search the web or send mails more conveniently.

It has two parts. One is Train part and second is sign detection part. Train part set by deep learning method using CNN network and make train dataset. The detection part takes sign from webcam and detect the Bengali numerical value by classifying from the train dataset.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURES **PAGE NO**

CHAPTER 1

Introduction

1.1 Introduction

Sign language is a communication medium by which people express their opinions and feelings. Communication between hearing impaired and mute community is not easy. Mute people need to learn sign language for effective communication. For this reason, it is necessary to develop an interactive, automatic and robust system that can be used as an interpreter among those communities. Bengali Sign Language is made from several core one-handed static sign alphabet and number signs [16]. Many researchers have been working on sign language recognition systems for various sign languages [16].

1.2 Motivation

Sign language is the mode of communication among the deaf and dumb. A common misconception about sign language is that it is universal, which is however not the case. In fact, sign languages, just like spoken languages are unique to a culture and have evolved over time. Moreover, they feature their own grammar and vocabulary and are generally acquired by deaf children as their mother tongue [1].

Sign languages use manual gestures and body language to convey meaning. Static signs are generally used for alphabets and numbers where hand shapes define each sign. On the other hand, words and sentences are generally expressed through a combination of hand shape, orientation and movement of hands and arms. Additionally, facial expressions exhibit emotions and sometimes head movement, shoulder position, body posture and lip patterns are important parameters in expressing the meaning of a sign. Hundreds of sign languages are in use around the World today, some of which have not yet gained any legal acceptance [1]. In Bangladesh, the Centre for Disability in Development (CDD) has developed a formal sign language for the Bengali deaf and dumb community, which is followed by schools for the speech and hearing impaired countrywide [1].

Incorporating the deaf and dumb into the mainstream is difficult, mainly due to a lack of knowledge about sign language by the rest of the society. So, as to bridge this communication gap, scientists have been researching on methods to develop automatic sign language recognition systems. This field of research is still far behind and struggling. Moreover, research on recognition of Bangla sign language has not prevailed as it has for some other sign languages. So, our goal is to conduct a research to recognize Bangla sign language.

1.3 Rationale of the Study

According to the manual of Bangladeshi Sign Language, there are approximately 5000 sets of gestures for alphabet, numbers and common words that are used by deaf and dumb people to communicate [8]. Some signs are expressed as static gestures while others incorporate some dynamic hand, face and body movements. For static gestures, the prominent sign is captured within a specific time frame. For dynamic gestures, a sequence of finger and hand positions needs to be identified and analyzed in order to recognize [8]. To interact with sign user the machine should recognize and learn sign characters and language. Automatic gesture and sign language recognition has been attracting computer vision researchers for a long time [8].

Unfortunately, in Bangladesh there is a tremendous lack of people who have an in-depth knowledge of sign language, which leads to the social isolation of the Deaf community [8].

So, I think that we have need more knowledge about sign language. That's why I am trying to develop sign language as computer understandable language.

1.4 Research Questions

How to interact with computer system for deaf and dumb people?

1.5 Expected Output

There are several technologies that can be and has been employed in sign language or gesture recognition. Bangladeshi sign language that can serve as an interpreter. For the purpose of this

thesis I am using a computer-vision based approach with the help of webcam and CNN to recognize signs.

For now, I am recognition of Bangla numeric signs only and also try to recognition real time feature for my experiment. The system is try to represent ten digits (০-৯). Finally, it has given output result that are shown the Figure 3.4.3 (a): Final output.

1.6 Report Layout

The following report consists of five chapters. Chapter 1 is the introduction, which highlights the motivation and goals behind the thesis. It has six sections. Section 1.1, 1.2, 1.3, 1.4, 1.5 and 1.6, titled introduction, motivation, rationale of the study, research questions, expected output and report layout. Chapter 2, titled background, prerequisites information relevant to the thesis and is divided in five sections. Section 2.1, 2.2, 2.3, 2.4 and 2.5, titled introduction, related work, research summary, scope of the problem and challenges respectively.

Chapter 3 illustrates the details of my thesis experiment spanned over five subsections. Section 3.1, 3.2, 3.3, 3.4 and 3.5 explains each phase of the recognition, i.e. Introduction, research subject and instrumentation, data collection procedure, statistical analysis and implementation requirements.

In chapter 4 have mentioned the experimental results and discussion of that result. It has four sections. Section 4.1, 4.2, 4.3 and 4.4 titled introduction, experimental results, descriptive analysis and summary.

Finally, chapter 5 have also four sections that titled summary, conclusion, recommendation and implication for future research.

An appendix, which lists all signs I used for my experiment, follows the five sections and at the end there is a reference of reading materials I have referred to during my research.

CHAPTER 2

Background

2.1 Introduction

A sign language is a language that instead of using sound to communicate uses manual gestures and body language to convey meaning. The hearing or speech impaired generally uses this mode of communication and today it is extensively studied by linguists [1].

Despite the common misconception that sign languages are not real language, they are as rich and complex as any spoken language. Sign language is a natural language and is not made up [1]. The acquisition process of sign language is similar to that of spoken languages as they are also adopted by deaf children in school from friends and teachers or at home from parents [1].

2.2 Related Work

Many researchers work for sign language. For example, American sign language, Bangla sign language, Indain sign language etc. are working many researchers. Sign language is important for the dumb and deaf people. It also important for human computer interaction.

Salma Begum and Md. Hasanuzzaman proposed a system that recognizes Bangla sign language using a computer-vision based approach [8].

Bikash Chandra Karmokar and many other are proposed a system that name is Bangladeshi sign language recognizer (BdSLR). In this regard BdSLR has been developed that can interpret Bangladeshi sign language into Bengali text and vice versa [9].

Muhammad Aminur Rahaman and many other are proposed a system that name is Real-Time Computer Vision-Based Bengali Sign Language Recognition [10]. The system is able to recognize 6 Bengali Vowels and 30 Bengali Consonants [10].

Zhou Ren and many other are proposed a system that title is Robust Hand Gesture Recognition with Kinect Sensor. They are proposed Hand gesture-based Human-Computer-Interaction (HCI) is one of the most natural and intuitive ways to communicate between people and machines, since it closely mimics how human interact with each other [11].

Syed Tauhid Ahmed and M. A. H. Akhand are also work for Bangladeshi sign language. They are gives name of it Bangladeshi Sign Language Recognition using Fingertip Position. The method considered relative tip positions of five fingers in two-dimension space and position vectors are used to train artificial neural network (ANN) for recognition purpose [12].

G. Anantha Rao and many other are proposed a new system for sign language. The title is Deep Convolutional Neural Networks for Sign Language Recognition. This proposed system extraction of complex head and hand movements along with their constantly changing shapes for recognition of sign language is considered a difficult problem in computer vision [13].

2.3 Research Summary

In this research I am wanted to recognize only the Bangla numeric sign. There are many researchers in Bangladesh are working in this field. Because this field is very complex and hard. They are always try to improve this field. For this purpose, they are used different technique. Its continuation firstly I have detected the segmentation of skin color and detect the largest blobs from the binary image. After that I am makes the training set and test set using deep learning method of convolutional neural network (CNNs).

Finally, when completed the largest binary blob image then it has classifying with the training data set that are shown the Figure 3.4.3 (a): Final output.

2.4 Scope of the Problem

In Bangladesh a formal sign language has been established only recently. In the year 2000, Center for Disability in Development (CDD) took the initiative to standardize communication with sign languages in this country [1]. Bengali Sign Language is structurally different from sign languages from other countries. To perform Bengali Sign Language, two hands are used generally but, in this system, I am only used one hand. Many researchers within computer vision communities have attempted to develop sign language recognizers for sign languages used in different countries [10]. The rule is also different in country basis that are shown the Figure 2.4 (a): Bangla numeric sign.

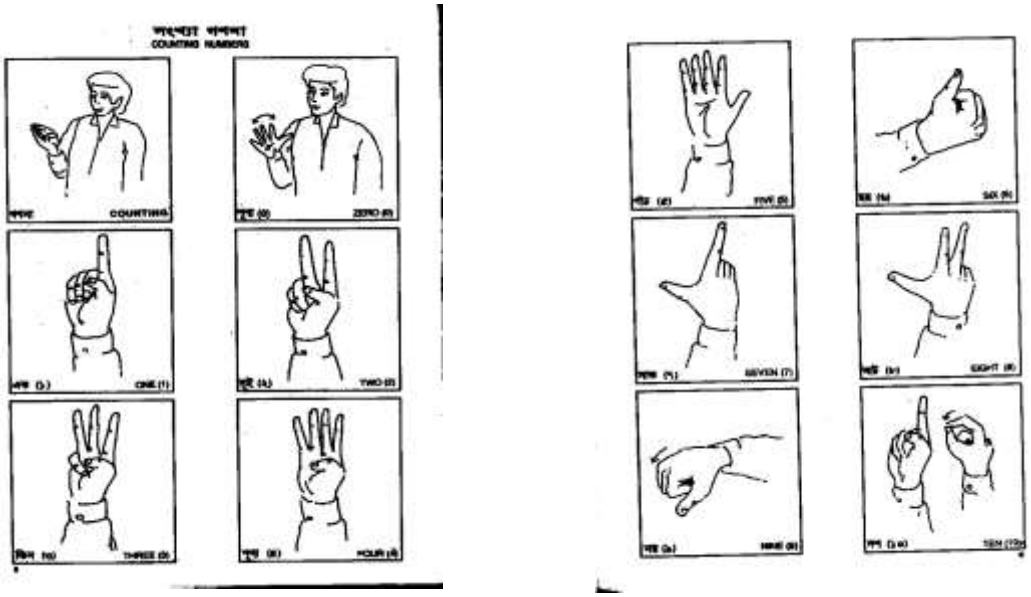

Figure 2.4 (a): Bangla numeric sign (collected).

2.5 Challenges

There are different Sign languages in different countries. One important fact to consider about sign languages is that, same hand shapes or same hand motion can be used to express different signs [1].

Figure 2.5 (a): Image digit one.

Figure 2.5 (b): Image digit six.

For example, in Bangla sign language, one and six digits have the near about same hand movement but the hand shape for the two signs are different. So, it is very challenges for every researcher to recognition the hand shape. That are shown the Figure 2.5 (a): Image digit one and Figure 2.5 (b): Image digit six.

CHAPTER 3

Research Methodology

3.1 Introduction

Sign language recognition is very difficult, because the computer system can't understand the sign easily. Every researcher tries hard and apply different technique to understanding the sign with the computer system. The proposed system is applied some technique to interact with computer system. This system only recognizes Bangla numeric sign (০-৯).

3.2 Research Subject and Instrumentation

This research topic is "**Computer Vision-Based Bangla Numerical Sign Language Recognition**". This is the field of image processing system.
This system uses Logitech c270 webcam for capture image. The system uses an Intel core i3 3.70 GHz PC with 4 GB RAM. It working very well if you can use GPU supported PC. I have used MATLAB as system development platform.

3.3 Data Collection Procedure

This system is applied some procedure for implementation. All the techniques have discussed in this section.

3.3.1 Skin Color Segmentation

Color space is a mathematical model to represent color information as three or four different color components. Different color models are used for different applications such as computer graphics, image processing, TV broadcasting, and computer vision [3]. Skin color detection means detect human skin color. This technique used for only detect human body. Either it can be hand or face. This system is used for this technique to detect hand for sign recognition that are shown Figure 3.3.1 (a): Original image.

Figure 3.3.1 (a): Original image.

It read RGB image and then capture the dimensions of the image. After converting the image RGB to YCbCr it detects the skin color. It can automatic segmentation of human face or hand in the complex background scene [3] that are shown Figure 3.3.1 (b): Skin color region.

Figure 3.3.1 (b): Skin color region.

Finally, it makes binary image and remove tiny regions from image that are shown Figure 3.3.1 (c): Tiny regions remove from the binary images.

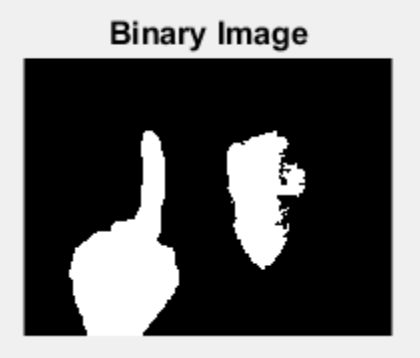

Figure 3.3.1 (c): Tiny regions remove from the binary images.

3.3.2 Extract N Largest Blobs

This feature helps to detect large blobs of object from binary image. Get all the blob properties from binary image and then get blob area. It selects all blob object.

It also creates boundaries in the original image and detect the largest object from the selected boundaries. I am using custom function for extract the N largest blobs from the binary image, that are shown **Figure 3.3.2 (a):** Binary image to extract the largest blobs.

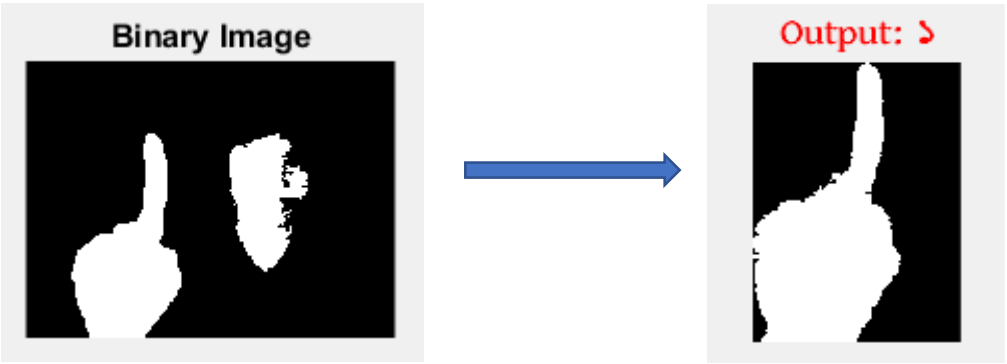

Figure 3.3.2 (a): Binary image to extract the largest blobs.

3.3.3 Convolutional Neural Network (CNNs)

Convolutional Neural Networks is also called CNNs or ConvNets, it is the part of deep learning neural network (Alex net). CNNs are compare the images piece by piece. The pieces are called features. Convolutional networks are powerful visual models that yield hierarchies of features [4].

Convolutional networks are driving advances in recognition. Convnets are not only improving for whole-image classification, but also making progress on local tasks with structured output. These include advances in bounding box object detection, part and keypoint prediction, and local correspondence [4].

3.3.4 Types of Architecture of CNNs

A convolutional neural network (CNN) contains one or more convolutional layers, pooling or fully connected, and uses a variation of multilayer perceptron's discussed above. Convolutional layers use a convolution operation to the input passing the result to the next layer. This operation allows the network to be deeper with much fewer parameters [5].

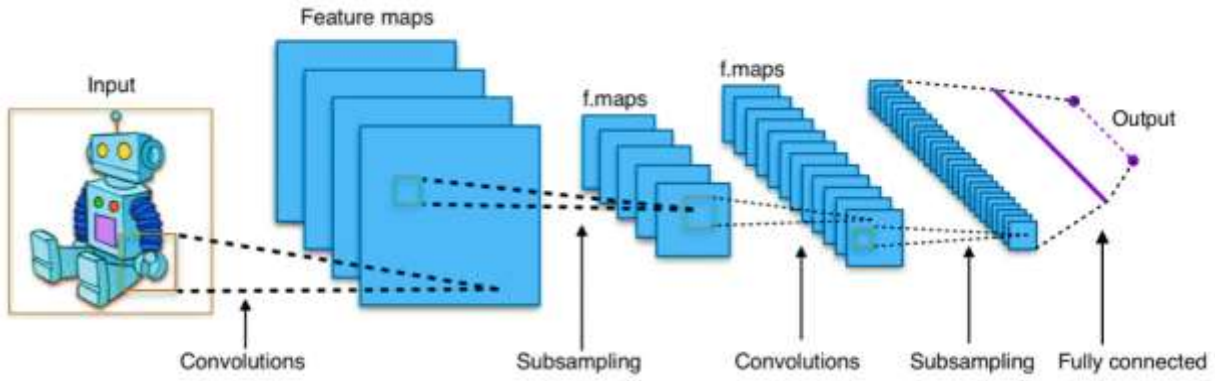

Figure 3.3.4 (a): Typical CNN architecture [2],[5].

Convolutional neural networks are giving great results in image and speech applications. The pooling and subsampling is the hidden units of CNN. Pooling basically takes a set of hidden units in a feature map and somehow aggregates the activations to obtain a single number [6].
For instance, in max pooling, the aggregation will be computing the maximum among the activations on a certain neighborhood or receptive field. This is shown in Figure 3.3.4 (a) and the following computations [6].

$$y_{ijk} = \max_{p,q} xi, j + p, k + q \quad (1)$$

Where

- xi, j, k is the value of the *i*th feature map at position *j*,*k*,

- p is the vertical index in the local neighborhood,
- q is the horizontal index in the local neighborhood,
- and $yijk$ is the pooled and subsampled layer. [6]

3.3.5 Working Procedure

Convolutional neural network is work the step by step. In this section it has briefly discuss.

Step 1: Break the image into overlapping image tiles

It makes same square tiles from the image and save each result as a separate, tiny picture tile, that are shown Figure 3.3.5 (a): Overlapping image tiles.

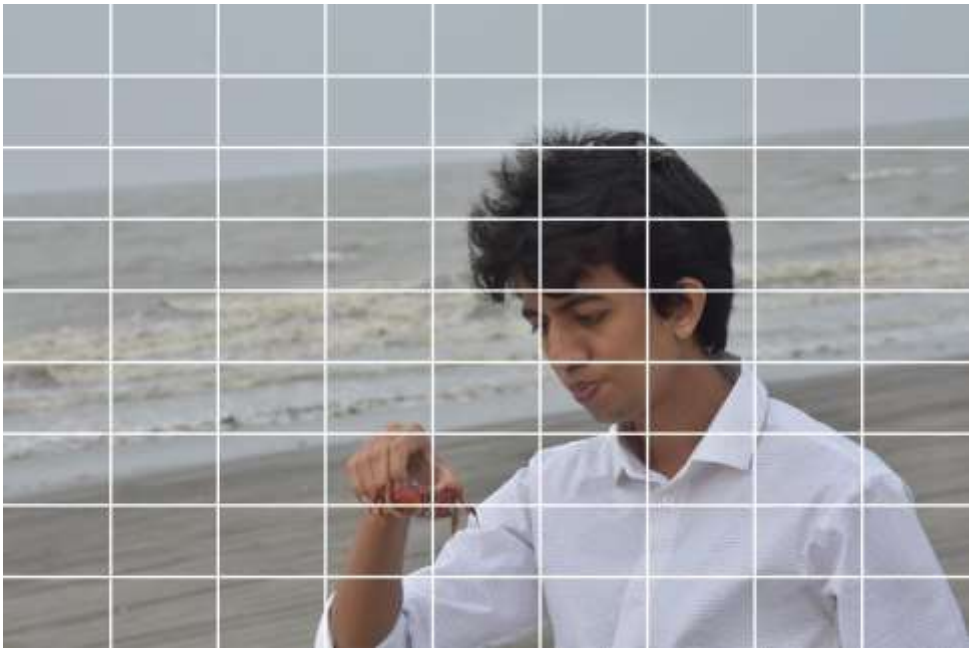

Figure 3.3.5 (a): Overlapping image tiles.

Step 2: Feed each image tile into a small neural network

We'll keep the same neural network weights for every single tile in the same original image.

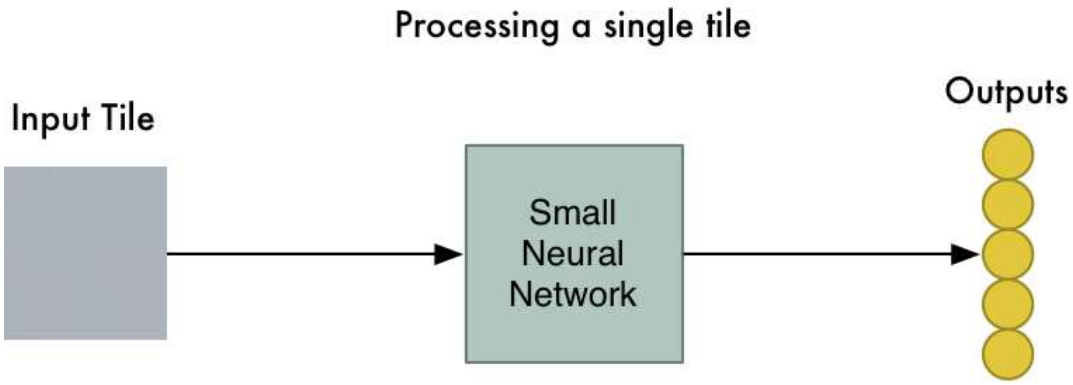

Figure 3.3.5 (b): Tile into a small neural network.

In other words, we are treating every image tile equally [7] that are shown Figure 3.3.5 (b): Tile into a small neural network.

Step 3: Save the results from each tile into a new array

We don't want to lose track of the arrangement of the original tiles. So, we save the result from processing each tile into a grid in the same arrangement as the original image [7]. It looks like this Figure 3.3.5 (c): Save result into new array.

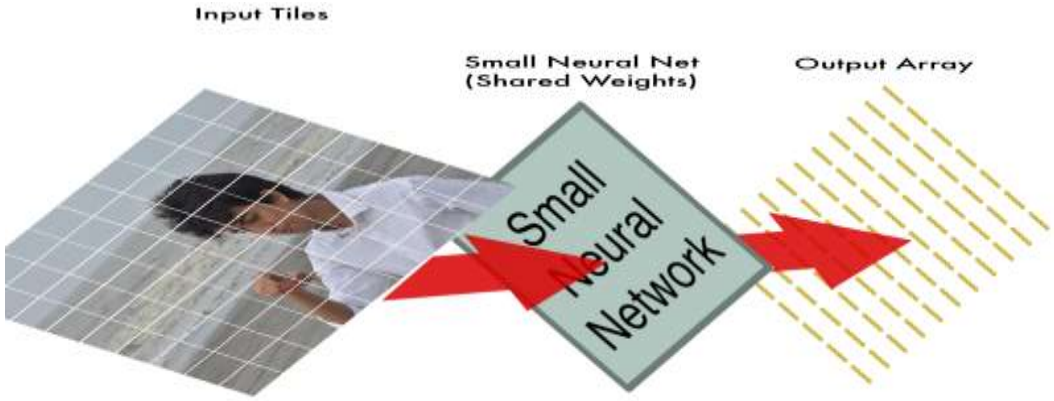

Figure 3.3.5 (c): Save result into new array.

Step 4: Down sampling

The result of Step 3 was an array that maps out which parts of the original image are the most interesting. But that array is still pretty big shown at Figure 3.3.5 (d): Down sampling.

Figure 3.3.5 (d): Down sampling.

To reduce the size of the array, we down sample it using an algorithm called max pooling.

Final step: Make a prediction

So far, we've reduced a giant image down into a fairly small array. So, we can use that small array as input into another neural network. This final neural network will decide if the image is or isn't a match. To differentiate it from the convolution step, we call it a "fully connected" network [7]. So, from start to finish, our whole five-step pipeline looks like this Figure 3.3.5 (e): Overall design.

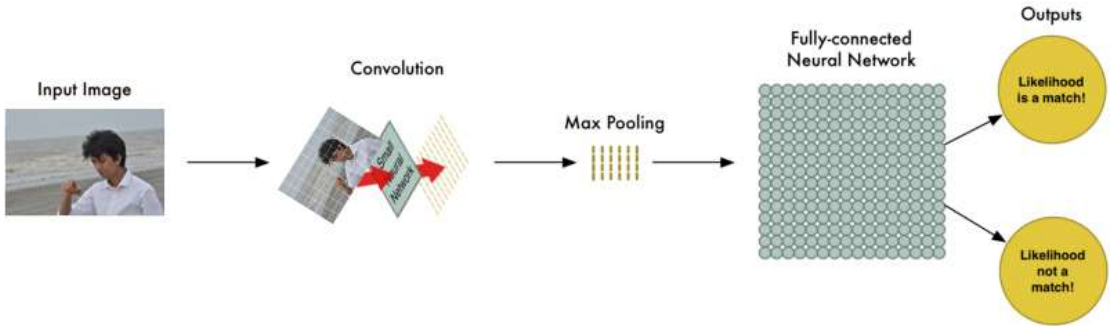

Figure 3.3.5 (e): Overall design.

Adding Even More Steps:

Our image processing pipeline is a series of steps: convolution, max-pooling, and finally a fully-connected network.

When solving problems in the real world, these steps can be combined and stacked as many times as you want! You can have two, three or even ten convolution layers. You can throw in max pooling wherever you want to reduce the size of your data [7].

3.4 Implementation Requirements

There are three categories in implementation requirement. All are described following:

3.4.1 Design Methodology

This proposed system takes snapshot from a video by using webcam. The system captures the dimension of the image and it convert RGB to YCbCr. Now the system detects the skin color from the image and it make into binary image. It is called segmentation. After completion the binary image the system removes the tiny regions from the binary image. Then the system extracts the largest area using Image Analyst's custom function.

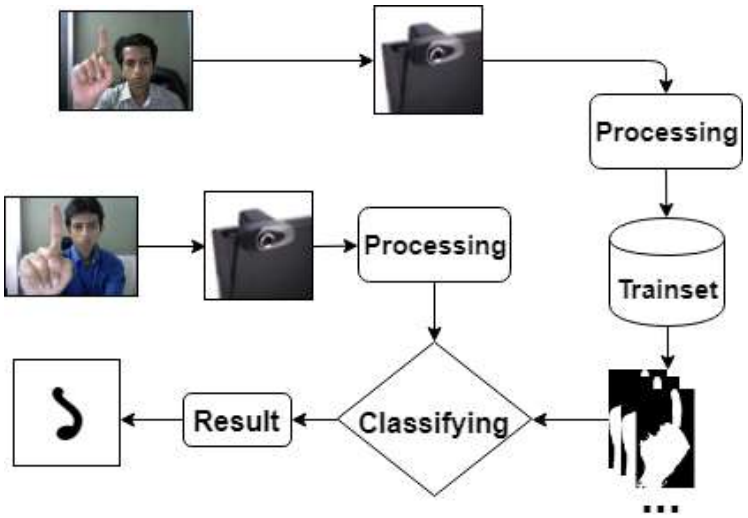

Figure 3.4.1 (a): Architectural view of proposed system.

The system also creates boundaries in the original image and detect the largest object from the selected boundaries. The system resizes the largest image blobs into [28x28] and classifying with the convolutional neural network.

After completed this part the proposed system gives the result from the given input. Figure 3.4.1 (a). Architectural view of proposed system.

3.4.2 Generate Train Network

To make train network this system also used a webcam but, preprocess the webcam images as the same process that are described in design methodology. That means this system used largest binary blob of image for train. It first takes multiple images for same sign. It takes multiple images for several signs. The proposed system saves the different sign images into different folder. For making category it saves the images into different folder. This system makes train network using 10,000 images for better accuracy. After completed to takes images, this system using convolutional neural network technique for makes train dataset that are shown Figure 3.4.2 (a): Random selected Images.

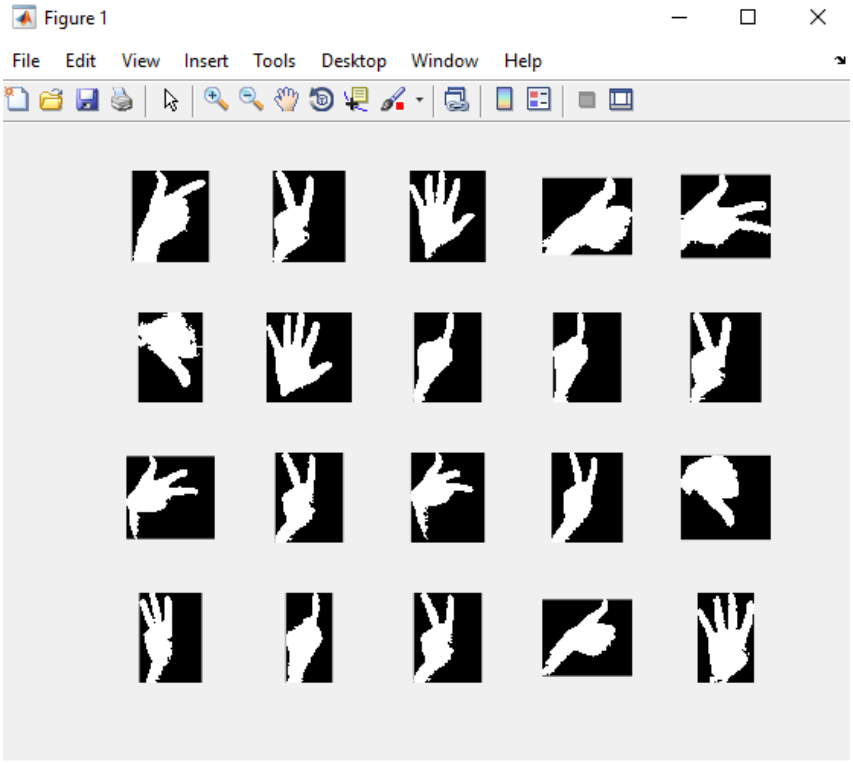

Figure 3.4.2 (a): Random selected Images.

The proposed system used alexnet of deep learning method for train network. Specifically, I using the Transfer Learning method of convolutional network.

It resized the binary images into [28x28] for the convolutional layer input. After that it makes trainset and testset using the random method.

Trainset means training database that used for comparing and testset means testing image. It has used 750 images for trainset and 250 images for test set. I used l Initial Learn Rate = 0.0001 Max Epochs = 20 and Mini Batch Size = 64.

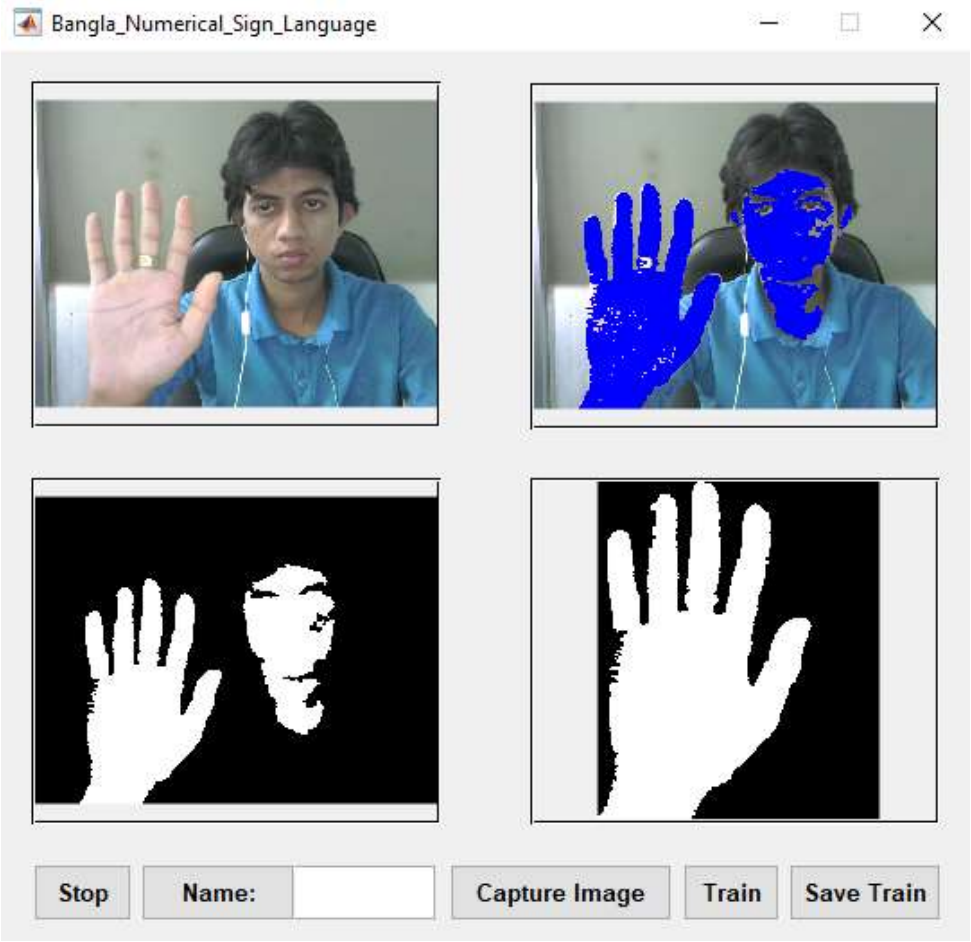

Figure 3.4.2 (b): Training window.

Finally, it makes my custom network and save as the name of Dataset.mat, that are shown Figure 3.4.2 (b): Training window.

3.4.3 Classifying with Train Network

This system loaded the training data set at first, then the sign has been taken when the figure window is opened. This system can detect Bangla numerical sign from one hand.

The proposed system classifying the webcam image. At first it has preprocessing the image using the technique that are described early in the design methodology.

When completed the largest binary blob image then it has classifying with the custom training data set.

Finally, it has given output result that are shown the Figure 3.4.3 (a): Final output.

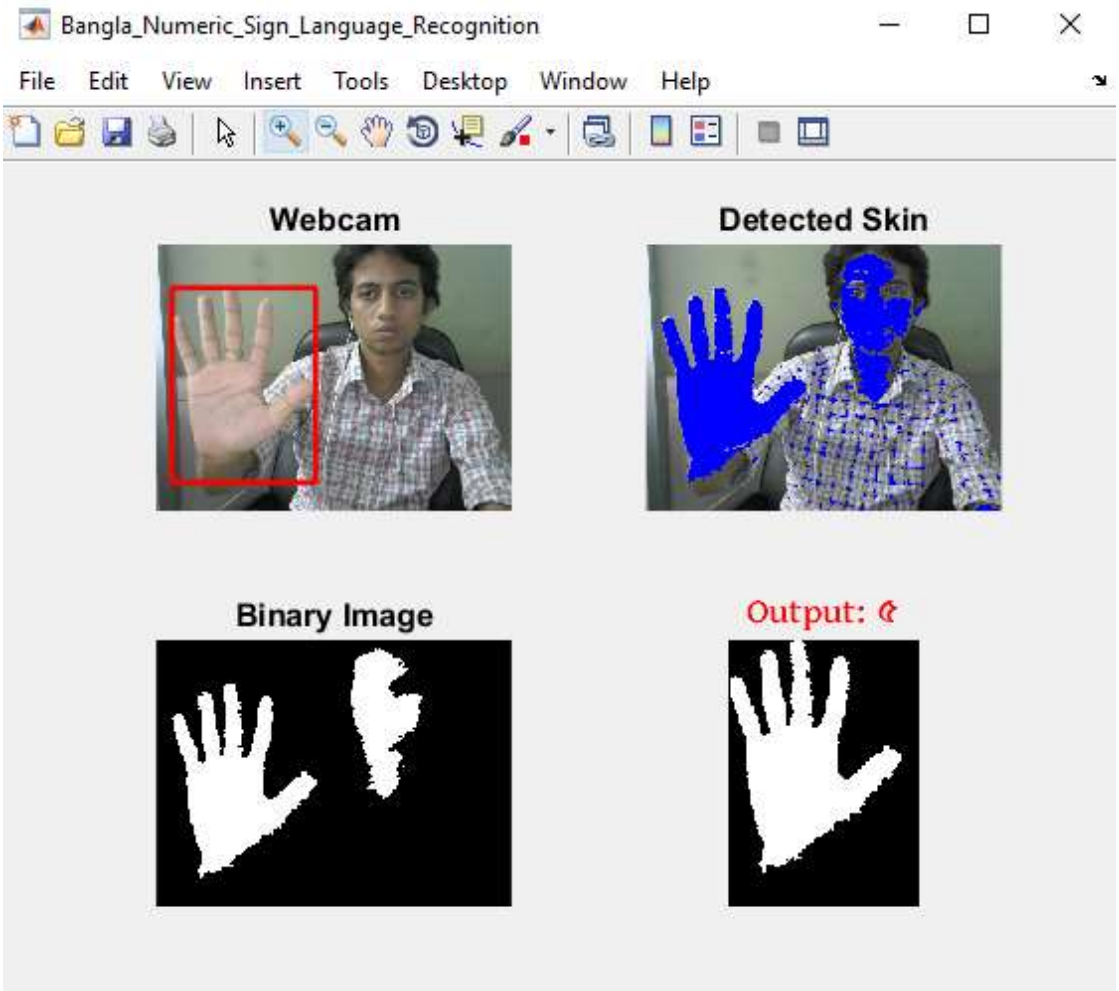

Figure 3.4.3 (a): Final output.

CHAPTER 4

Experimental Results and Discussion

4.1 Introduction

The proposed system has taken different technique to improve the sign language recognition field. This system tried hard for makes better accuracy. For experimental results I have needed to calculate the accuracy of the recognition of each sign. The accuracy of the system was calculated as described in section 4.1.

4.2 Experimental Results

I am calculated accuracy using the following variables:

Precision:

The precision is defined by the ratio of the numbers of correct recognition to the total numbers of recognition for each hand signs.

Recall:

The recall rate is defined by the ratio of the numbers of correct hand sign identification to the total number of hand signs.

Recall and Precision can be defined using the following variables:

TP (True Positive): True positive means number of correct results we are looking for.

TN (True Negative): Means number of correct detection of the incorrect input data.

FP (False Positive): Means wrong result in the system's output (Unexpected output).

FN (False Negative): Means missing expected outputs, i.e. The result should be included in system output but is not there.

As recall and precision are defined as follows:

$$\text{Precision} = \frac{tp}{tp+fp} \quad (2)$$

$$\text{Recall} = \frac{tp}{tp+fn} \quad (3)$$

Accuracy and True Negative rate are calculated as given below:

True negative rate = $\frac{tn}{tn+fp}$ (4)

Accuracy = $\frac{tp+tn}{tp+tn+fp+fn}$ (5)

The system takes among 10 test hand signs for every digit, where 9 are maximum correctly recognized and 6 are minimum correctly recognized and 2 are maximum incorrectly recognized and 1 are minimum incorrectly recognized as other numbers, that are shown **TABLE 4.2 (a)**.

TABLE 4.2 (a): PRECISION RATE, RECALL RATE AND ACCURACY FOR EACH SIGN.

Bangla Sign	**Precision**	**Recall**	**True negative rate**	**Accuracy**
০ 	100%	100%	0%	100%
১ 	100%	100%	0%	100%
২ 	100%	100%	0%	100%
৩ 	70%	100%	0%	70%
৪ 	100%	100%	0%	100%
৫ 	100%	100%	0%	100%
৬ 	100%	100%	0%	100%
৭ 	77%	87%	33%	70%
৮ 	66%	86%	25%	60%
৯ 	100%	100%	0%	100%

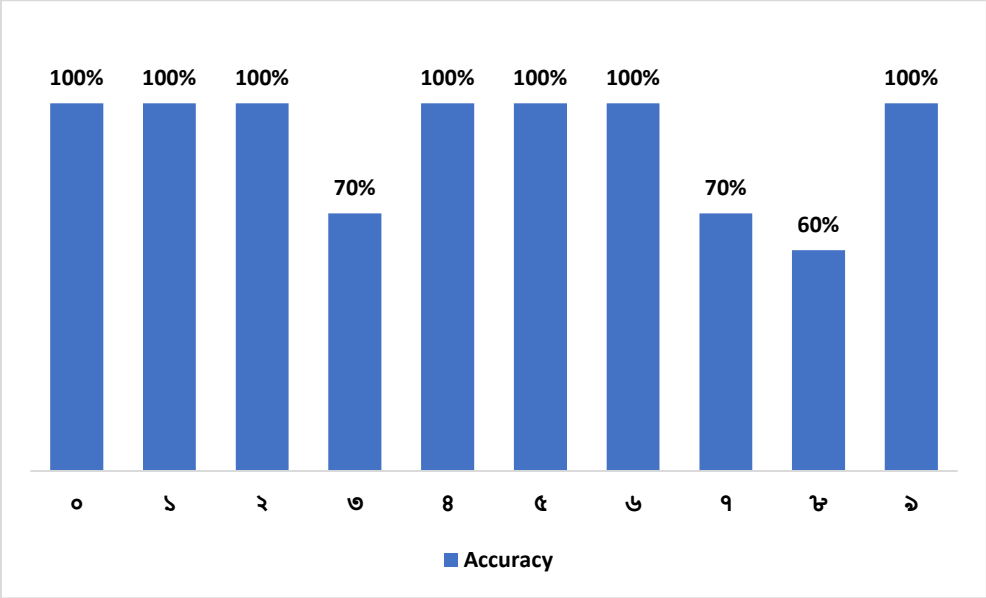

Figure 4.2 (a): Accuracy comparison of Ten different numeric sign.

4.3 Descriptive Analysis

When I give the training data as input the recognition rate is 100 percent, meaning that the training was successful. Then when a separate test data was given to the neural network, the accuracy rate was 90 percent. The same signers who signed for the training sample provided this data.

When the system gets the individual sign, it given 100% accuracy for the sign of ১,২,৪,৫,৬,৯ and ০. figure 4.2 (a) it shown that.

Thus, it can be seen that the system performance becomes worse when gives the input sign for ৩,৭,৮.

Over all, the proposed system accuracy is 90%.

4.4 Summary

In this chapter I have discussed the accuracy calculation and accuracy figure of the individual sign of my thesis project. I have also discussed the result briefly and shown the table for each sign in this chapter.

CHAPTER 5

Conclusion and Future Work

5.1 Summary of the Study

I have implemented the Computer Vision-Based Bangla Numerical Sign Language Recognition system successfully. There is always chance of upgrading the system further more. In this chapter I will like to discuss about the summary of the research, conclusions, and recommendations and further improvement ideas of this research.

5.2 Conclusions

This proposed system has been trying to recognize the Bangla numeric sign language. But it has some limitation. It cannot detect sign in the dark places. On the other hand, it detects the skin color, so if any color is equal with skin color it takes that.

This system accepts the largest binary blob, So, If the face is the largest from hand shape this system accepts the face. Future project I will try to improve the unpredictable error.

To develop the sign language field this type of interpreter is very important. So, I will try hard in my next research to improve the interpreter. I think this system will be helps to the deaf and dumb peoples to communicate with computer system.

5.3 Recommendations

The next step will be to increase the number of training samples and to try training methods other than neural network to analyze which one produces the best result for Bangla Sign Language recognition.

This project also has the scope of enhancements like:

• It helps for dumb and deaf people as an interpreter.

• It also helps to robotics system.
• Update with English version.

5.4 Implication for Further Study

This system has been developed with future improvement possibilities consideration. Next time this system, I think that will be more efficient and faster. It is very important to reduce the processing time of any system.

In future this system will be upgraded is much better from now. I want continued research in this area. My next objective is to improve the accuracy of the recognition system.

Next, I try to change training method for get better accuracy. Besides helping as an interpreter this system help the deaf and dumb to search the web or send mails more conveniently.

References:

[1] Golam Kayas, Sakhawat Hossain Himel, and Mahraj Hasan, "Automatic Recognition of Bangla Sign Language Using Artificial Neural Networks (Anns) for Deaf and Dumb to Bridge the Communication Gap," In: International Journal of Recent Advances in Multidisciplinary Research July, 2016, Vol. 03, Issue 07, pp.1649-1654.

[2] Yann LeCun, Koray Kavukcuoglu and Clement Farabet, "Convolutional Networks and Applications in Vision," In: Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on.

[3] Khamar Basha Shaik, Ganesan P, V.Kalist, B.S.Sathish, J.Merlin Mary Jenitha, "Comparative Study of Skin Color Detection and Segmentation in HSV and YCbCr Color Space," In: Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on3rd International Conference on Recent Trends in Computing 2015 (ICRTC-2015).

[4] Jonathan Long, Evan Shelhamer, Trevor Darrell, "Fully Convolutional Networks for Semantic Segmentation," In: IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume: 39, Issue: 4, April 1 2017).

[5] Learn about Medium, available at <<https://medium.com/@datamonsters/artificial-neural-networks-for-natural-language-processing-part-1-64ca9ebfa3b2>>, last accessed on 30-03-2018 at 12:00pm.

[6] Learn about Researchgate, available at <<https://www.researchgate.net>>, last accessed on 30-03-2018 at 12:20pm.

[7] Learn about Medium, available at <<https://medium.com/@ageitgey/machine-learning-is-fun-part-3-deep-learning-and convolutional-neural-networks-f40359318721>>, last accessed on 30-03-2018 at 1:00pm.

[8] S. Begum and M. Hasanuzzaman, "Computer Vision-based Bangladeshi Sign Language Recognition System," In: International Conference on Computer and Information Technology, MIST, Dhaka, Bangladesh, 2009, pp. 414 - 419.

[9] Bikash Chandra Karmokar, Kazi Md. Rokibul Alam, and Md. Kibria Siddiquee, "Bangladeshi Sign Language Recognition Employing Neural Network Ensemble," In: International Journal of Computer Applications 58(16):43-46, November 2012.

[10] Muhammad Aminur Rahaman, Mahmood Jasim, Md. Haider Ali and Md. Hasanuzzaman, "Real-Time Computer Vision-Based Bengali Sign Language Recognition," In: 17th International Conference on Computer and Information Technology (ICCIT) 2014.

[11] Zhou Ren, Jingjing Meng, Junsong Yuan, “Robust Hand Gesture Recognition with Kinect Sensor,” In: IEEE Transactions on Multimedia (Volume: 15, Issue: 5, Aug. 2013).

[12] Syed Tauhid Ahmed and M. A. H. Akhand, “Bangladeshi Sign Language Recognition using Fingertip Position,” In: Medical Engineering, Health Informatics and Technology (MediTec), 2016 International Conference on.

[13] G.Anantha Rao, K.Syamala , P.V.V.Kishore, A.S.C.S.Sastry, “Deep Convolutional Neural Networks for Sign Language Recognition,” In: Signal Processing And Communication Engineering Systems (SPACES), 2018 Conference on.

[14] Wei Wang, Jing Pan, “Hand segmentation using skin color and background information,” In: International Conference on Machine Learning and Cybernetics (ICMLC), 2012.

[15] Sayantan Thakur, Sayantanu Paul, Ankur Mondal, Swagatam Das, Ajith Abraham, “Face detection using skin tone segmentation,” In: Information and Communication Technologies (WICT), 2011 World Congress on.

[16] Angur M. Jarman, Samiul Arshad, Nashid Alam and Mohammed J. Islam, “Face detection using skin tone segmentation,” In: International Journal of Electronics & Informatics, 22 July 2015 Vol.4, No.1.