Faculty of Electrical and Electronic Engineering Technology

DEVELOPMENT OF SIGN LANGUAGE INTERPRETER USING COMPUTER VISION TECHNIQUE

NASHA ATHILAH BINTI ZAINAL

Bachelor Of Computer Engineering Technology (Computer Systems) With Honours

2022

DEVELOPMENT OF SIGN LANGUAGE INTERPRETER USINGCOMPUTER VISION TECHNIQUE

NASHA ATHILAH BINTI ZAINAL

A project report submitted
in partial fulfillment of the requirements for the degree of
Bachelor Of Computer Engineering Technology (Computer Systems) With Honours

Faculty of Electrical and Electronic Engineering Technology

UNIVERSITI TEKNIKAL MALAYSIA MELAKA

2022

UNIVERSITI TEKNIKAL MALAYSIA MELAKA
FAKULTI TEKNOLOGI KEJUTERAAN ELEKTRIK DAN ELEKTRONIK

BORANG PENGESAHAN STATUS LAPORAN
PROJEK SARJANA MUDA II

Tajuk Projek :

Sesi Pengajian :

Saya NASHA ATHILAH BINTI ZAINAL …………………………………….. mengaku membenarkan laporan Projek Sarjana Muda ini disimpan di Perpustakaan dengan syarat-syarat kegunaan seperti berikut:

1. Laporan adalah hakmilik Universiti Teknikal Malaysia Melaka.
2. Perpustakaan dibenarkan membuat salinan untuk tujuan pengajian sahaja.
3. Perpustakaan dibenarkan membuat salinan laporan ini sebagai bahan pertukaran antara institusi pengajian tinggi.
4. Sila tandakan (✓):

☐ **SULIT*** (Mengandungi maklumat yang berdarjah keselamatan atau kepentingan Malaysia seperti yang termaktub di dalam AKTA RAHSIA RASMI 1972)

☐ **TERHAD*** (Mengandungi maklumat terhad yang telah ditentukan oleh organisasi/badan di mana penyelidikan dijalankan)

☑ **TIDAK TERHAD**

Disahkan oleh:

NASHA

(TANDATANGAN PENULIS)

Alamat Tetap:

No. 31 JALAN MELUR 15,
TAMAN MELUR,
68000 AMPANG,
SELANGOR DARUL EHSAN.

(COP DAN TANDATANGAN

DR. SUHAILA BINTI MOHD NAJIB
Pensyarah Kanan
Jabatan Teknologi Kejuruteraan Elektronik dan Komputer
Fakulti Teknologi Kejuruteraan Elektrik dan Elektronik
Universiti Teknikal Malaysia Melaka

Tarikh: 17/1/2023

Tarikh: 23/2/2023

*CATATAN: Jika laporan ini SULIT atau TERHAD, sila lampirkan surat daripada pihak berkuasa/organisasi berkenaan dengan menyatakan sekali tempoh laporan ini perlu dikelaskan sebagai SULIT atau TERHAD.

DECLARATION

I declare that this project report entitled "Development of Sign Language using Computer Vision Technique" is the result of my own research except as cited in the references. The project report has not been accepted for any degree and is not concurrently submitted in the candidature of any other degree.

Signature : NASHA

Student Name : NASHA ATHILAH BINTI ZAINAL

Date : 17/1/2023

APPROVAL

I hereby declare that I have checked this project report and in my opinion, this project report is adequate in terms of scope and quality for the award of the degree of Bachelor of Computer Engineering Technology (Computer Systems) with Honours.

Signature	:	
Supervisor Name	:	Dr.Suhaila Binti Mohd Najib
Date	:	23/2/2023

DEDICATION

I want to express my sincere thanks to my parents Roslim Binti Ali and Zainal Bin Amat for their constant words of support as I finished my final year 1 project. They are a huge support to me as I strive to carry out my project. They gave me some advice on how to complete it on time and without stress. They also set up a comfortable place where I could get inspiration and ideas to finish my project. Not to mention my sibling and friends, who provided me with a wealth of advice on how to make my work better. Thank you for all of your help and guidance during all of my inquiries about the project, Dr. Suhaila, my supervisor.

ABSTRACT

To assist the social interaction of deaf and hearing-impaired people, efficient interactive communication tools are expected. Gesture recognition forms the basis for translating sign languages where gesture recognition plays a critical role in Sign Language Recognition (SLR). The purpose of this project is to provide a sign language interpreter to ease the interaction with the hearing-impaired person based on the computer vision approach. The dataset from the website Roboflow Universe will be used which consists of 26 different hand sign gestures which include A-Z alphabet gestures. All the labeled images will be set as a training set. For the test set, the Python OpenCV library will be used to capture sign gestures from the computer's webcam. The training and test set images will be compared and classified based on the Support Vector Machine (SVM). The output of the system will predict the accuracy of the hand gesture captured by the webcam.

ABSTRAK

Untuk membantu interaksi sosial orang pekak dan bermasalah pendengaran, alat komunikasi interaktif yang cekap adalah diharapkan. Pengecaman gerak isyarat membentuk asas dalam menterjemah bahasa isyarat di mana pengecaman gerak isyarat memainkan peranan penting dalam Pengecaman Bahasa Isyarat (SLR). Tujuan projek ini adalah untuk menyediakan penterjemah bahasa isyarat untuk memudahkan interaksi dengan orang cacat pendengaran berdasarkan pendekatan penglihatan komputer. Set data daripada tapak web Roboflow Universe akan digunakan yang terdiri daripada 26 gerak isyarat tangan berbeza yang termasuk gerak isyarat abjad A-Z. Semua imej berlabel akan ditetapkan sebagai set latihan. Untuk set ujian, perpustakaan OpenCV Python akan digunakan untuk menangkap gerak isyarat tanda daripada kamera web komputer. Imej set latihan dan ujian akan dibandingkan dan dikelaskan berdasarkan Mesin Vektor Sokongan (SVM). Output sistem akan meramalkan ketepatan gerak isyarat tangan yang ditangkap oleh kamera web.

ACKNOWLEDGEMENTS

I would like to sincerely thank my supervisor, Dr. Suhaila Binti Mohd Najib, for her helpful guidance, words of advice, and patience although this project was being conducted.

I thank Universiti Teknikal Malaysia Melaka (UTeM) and my family for helping me with the project by providing me with financial support. In addition, I would like to thank my twin sister Nasha Athirah Binti Zainal, and my friends Nur Najihah Binti Mohd Musa, Ilya Iziana Binti Kamaruzaman, Nur 'Izzati Binti Khalid, Nur Alysha Binti Norazizan, Noor Shafina Binti Abdul Ghani, Chan Yoke Lin, and Nur Hanis Binti Rohaimi for their the willingness to share

Finally, my parents, friends, and family have my deepest gratitude for their support and prayers during this time of study. Many thanks for your inspiration and tolerance

TABLE OF CONTENTS

		PAGE
DECLARATION		
APPROVAL		
APPROVAL		
DEDICATIONS		
ABSTRACT		**i**
ABSTRAK		**ii**
ACKNOWLEDGEMENTS		**iii**
TABLE OF CONTENTS		**iv**
LIST OF TABLES		**vi**
LIST OF FIGURES		**vii**
CHAPTER 1	**INTRODUCTION**	**1**
1.1 Background		1
1.2 Problem Statement		2
1.3 Project Objective		2
1.4 Scope of Project		3
CHAPTER 2	**LITERATURE REVIEW**	**4**
2.1 Introduction		4
2.2 Malaysian Sign Language		4
2.3 American Sign Language		6
2.4 Indian Sign Language		12
2.5 Table of Comparison		13
2.6 Review of Comparison Table		29
2.7 Summary		29
CHAPTER 3	**METHODOLOGY**	**30**
3.1 Introduction		30
3.2 Methodology		30
3.2.1 YOLO (You Only Look Once)		33
3.2.2 Image Acquisition		33

3.2.3 Image Pre-processing 33
3.2.4 Image Segmentation 34
3.2.5 Feature Extraction 34
3.2.6 Classification 35
3.2.7 Dataset 35
3.2.8 Epoch (training cycle) 38
3.2.9 Testing Flowchart 39
3.3 Equipment 40
2.4 Summary 42
3.5 Gantt Chart 43
CHAPTER 4 RESULTS AND DISCUSSIONS 45
4.1 Introduction 45
4.2 Training, Validation, and Testing Result 45
4.2.1 Loss Function of Training Model 47
4.2.2 Evaluation of Trained Model 48
4.2.3 Testing Results 50
4.3 Summary 53
CHAPTER 5 CONCLUSION 54
5.1 Conclusion 54
5.2 Future Recommendation 55
5.3 Project Potential 55
REFERENCES 56

LIST OF TABLES

TABLE	TITLE	PAGE
Table 3.1	Hardware Tool	40
Table 3.2	Software Tool	40
Table 3.3	Gantt Chart for PSM 1	43
Table 3.4	Gantt Chart for PSM 2	44
Table 4.1	Table of a different epoch using the testing dataset	50

LIST OF FIGURES

TABLE	TITLE	PAGE
Figure 2.1	(a) “Null” classification result, (b) “C” correct classification with 0.897 confidencescore (Khan *et al.*, 2021)	5
Figure 2.2	The gestures used for the system test (Shukor *et al.*,2015)	6
Figure 2.3	Speech to Sign module workflow(Patil *etal.*, 2019)	7
Figure 2.4	Sign Language interpreter flowchart(Apoorv, Kumar Bhowmick and SakthiPrabha, 2020)	8
Figure 2.5	Translation of gestures of ASL alphabets(Shivashankara and Srinath, 2019)	10
Figure 2.6	Background Subtraction Flowchart(Triyono *et al.*,2018)	11
Figure 3.1	Flowchart of a Sign language interpreter	32
Figure 3.2	Dataset Images from Roboflow Universe website.	35
Figure 3.3	Training flowchart of Yolov5	37
Figure 3.4	Summary of model training using 20 epochs	39
Figure 3.5	Testing flowchart of a sign language interpreter	39
Figure 4.1	Train using 20 epochs	46
Figure 4.2	Output of training batch	46
Figure 4.3	(a) validation label and (b) validation prediction	46
Figure 4.4	Loss Function of Training using 50 Epoch	48
Figure 4.5	Precision Bar Chart	48
Figure 4.6	Recall Bar Chart	49
Figure 4.7	Test using a webcam	52

CHAPTER 1

INTRODUCTION

1.1 Background

A sign language is a method of communication that involves the use of the hands and other body parts. It is not to be confused with nonverbal communication. Sign languages are an important means of communication for deaf individuals. They are frequently used by deaf individuals in place of spoken languages. Spoken languages are understood through the hearing and rely on sounds produced by the tongue. Sign languages are written with the hands and read with the eyes. Sign languages are easier for deaf and hearing persons to learn than spoken languages.

Communication is significantly more difficult when deaf and hearing people (who do not sign) meet. This problem is sometimes attributed to the deaf person, who is widely considered as impaired (unable to hear or speak clearly) and hence unable to communicate in the same way that hearing people do. Instead, the problem is one of a difference in comprehension between languages and cultures, as in spoken language interactions between different cultural groups. Deaf people exercise agency by using a variety of techniques to communicate with hearing people, such as iconic gestures, writing down words, and pointing.

During the COVID-19 epidemic, the world faced numerous obstacles in various sectors, including the education sector. Teachers and students have struggled to quickly adapt to distance learning at all levels and courses. For students with impairments, the experience may have been much more intense and difficult to adapt to a 'successful' online distance learning experience must involve a number of working components.

1.2 Problem Statement

Every human being need the ability to communicate. People with hearing and/or speech disabilities, on the other hand, require a method of communication other than vocal communication. However, understanding and learning Sign Language takes a lot of practice, and not everyone will grasp what the sign language movements represent. Because there is no good, portable technology for identifying sign language, learning sign language takes time. Hearing or speech-impaired people who know Sign Language will need a translator who also knows Sign Language to effectively communicate their views to others. This technology assists hearing or speech impaired people in learning and translating their sign language in order to help them overcome these issues.

1.3 Project Objective

The main aim project is to create a computer vision-based sign language interpreter to make interactions with hearing-impaired people easier. This is primarily for persons who are unable to communicate with others. Specifically, the objectives are as follows:

a) To recognize sign alphabets from the American Sign Language type using theYolov5 model computer vision technique.

b) To provide a real-time interface that allows a normal hearing-impaired person tocommunicate with a normal hearing-impaired person.

1.4 Scope of Project

In order to achieve the objectives of this project, the scope of project is:

a) The system's limitation is that it only enables the alphabet, not numbers.
b) The dataset will be used which consists of 26 different hand sign gestures.
c) The system will predict the accuracy of the hand gesture captured by the webcam.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

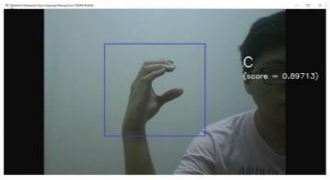
Those who have trouble hearing or speaking utilize sign language as a means of communication. They communicate well amongst themselves, but conversing with ordinary people is difficult. Malaysia Sign Language MSL was founded in 1998 by the Malaysia Federation of the Deaf (MFD). Approximately one million hard-to-hearing adults and nearly half a million hard-of-hearing children use by (Khan *et al.*, 2021) . The majority of hard hearing in town are unable to communicate via gestures. Several scholars proposed and created technologies for deaf or hard-of-hearing people in interacting with non-deaf or hard-of-hearing people. Furthermore, by teaching computers to understand human language, a user-friendly human-computer interface can be created.

In this chapter, I looked at some more similar studies that have been done in the field of sign language interpreters. The following are brief of the prproject'sarious works:

2.2 Malaysian Sign Language (MSL)

This research used a convolutional neural network (CNN) and a convolutional-based attention module (CBAM) to recognize Malaysian Sign Language (MSL) from images to tackle this challenge. For the project, CBAM-2DResNet (2-Dimensional Residual Network arebe used to "Within Blocks" and "Before Classifier" methods. The Python 3.6 programming language and Anaconda Spyder were utilized during he development period d Open CV(Khan *et al.*, 2021).

(a)



(b)

Figure 2.1 (a) “Null” classification result, (b) “C” correct classification with 0.897 confidencescore (Khan *et al.*, 2021)

As shown in figure2.2, to develop the real-time sign alphabets recognition application, the best trained CBAM-2DResNet “Before Classifier” was chosen as a classifier model. Using the OpenCV library, this real-time application provided a direct platform for evaluating the trained model using images taken from frames. Real-time sign images were retrieved from the blue box region for every four frames captured by a camera to feed as test inputs, and if the confidence score was greater than 0.5, the user was given the associated classification result.

Malaysia Sign Language (MSL) is a gesture-based communication system used by the deaf community in Malaysia. From this project, the data-glove approach, which used a specific glove-based device to extract hand posture and motion, is used for hand gesture identification. To identify hand poses that represent the alphabet, number, and numerous words from Malaysian Sign Language, the glove will use a microcontroller as the processor and tilt sensor and an accelerometer as the sensor. The gesture translation will be displayed on the phone(Shukor *et al.*, 2015).

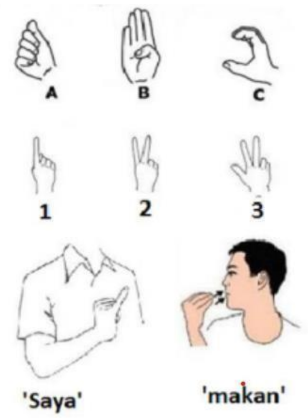


Figure 2.2 The gestures used for the system test (Shukor *et al.*, 2015)

As shown in figure 2.3, the glove also has an acceleration that detects gestures. The alphabets for the test are A, B, and C, selected from the standard Malaysian Sign Language database. 1, 2, and 3 are the numbers that need to be found. Three sets of gestures were examined for the motion: ‘saya,’ ‘Makan,’ and ‘Apa.’.

2.3 American Sign Language (ASL)

The purpose of this study was to describe methods for recognizing American Sign Language (ASL). The Support Vector Machine is a pattern recognition approach used to build sign language recognition systems. Furthermore, genetic algorithms are a sub-field of evolutionary computation, and the camera receives images of user input gestures, detection to determine whether it is hand or not, using a specific algorithm. As a result, the camera receives a user-input gesture image, and uses a specific algorithm, to determine whether it is hand-held or not. Image recognition is the following stage, where images collected from users are compared with photos in a data set to interpret the displayed gestures. The output is the following phase, the identified symbols are translated into text form (Dogra, Malik and Chowdary, 2018)

This study outlines modern ways to bridge the communication gap between the deaf and the hearing impaired. The Python programming language is used to create code. The data was collected from various ASL data set sites and included sign data for 26 letters and ten numbers. For each character input sample, 360 (10 samples per character) samples were used for testing various terms including the alphabet in them. This section outlines the "voice to sign" project proposal methodology. The system will receive speech input through a microphone, and the speech will be processed and converted to the appropriate text format(Patil *et al.*, 2019).

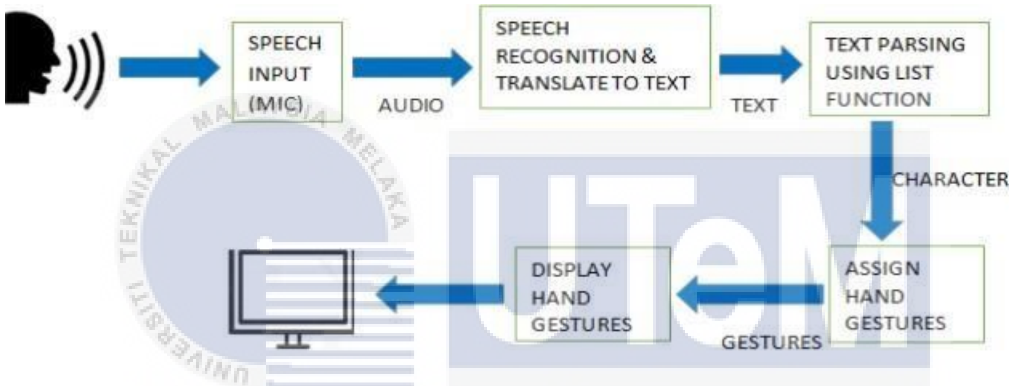


Figure 2.3 Speech to Sign module workflow (Patil *et al.*, 2019)

This report describes Sign Language Recognition using Media Pipes to recognize characters spelled in American Sign Language (ASL) for the deaf population, which can be used as a means of communication between the deaf and others. It provides a simple and accurate way to interact between humans and computers is one of the reasons for choosing a vision-based system. This research uses the Google Media tap. Media pipe solutions have improved their hand recognition models and now recognize 21 3D Palm Landmarks. Open Source Computer Vision (OpenCV) and the Python programming language are used to create the system. The method of taking pictures on the camera as touch data is used to collect sign language sign images. These sign language receivers will be able to recognize letters and detect hands as well as produce coordinators (A-Z). All alerts will be displayed in real-time (Gomase *et al.*, 2022).

This work reported the development of a dataset containing 26 English alphabets and described the deployment of a system that transforms Indian Sign Language into English. In this technique, a webcam is utilized to capture a still hand image frame. The image is then converted to grayscale before being converted to a binary image. The YCbCr model is used to identify skin color at the same time. Finally, to detect edges, the Canny edge detector is utilized. After that, the data was divided into 4800 training shots and 1200 testing photos. These images have been improved through post-production. Using feature extraction and classification algorithms, the sign language is subsequently translated into English text. This translation is converted to speech using the text to speech API (Apoorv, Kumar Bhowmick and Sakthi Prabha, 2020)

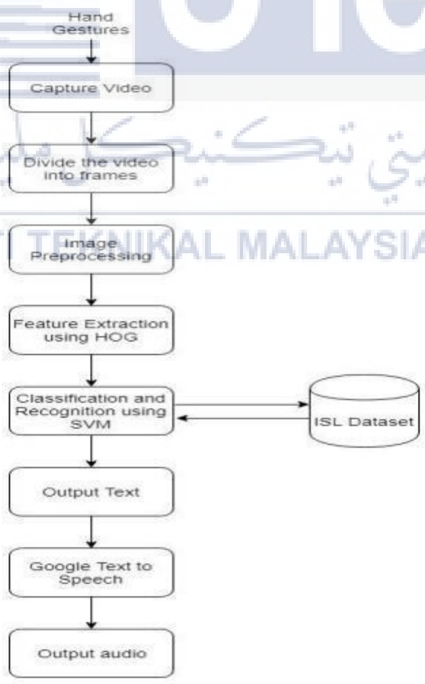


Figure 2.4 Sign Language interpreter flowchart (Apoorv, Kumar Bhowmick and SakthiPrabha, 2020)

The purpose of this paper is to show how to use Media Pipe open-source framework and machine learning algorithm to simplify Sign Language Recognition. Without any wearable sensors, real-time precise detection utilizing the Support Vector Machine (SVM) algorithm makes this technology more comfortable and simpler to use. On the American Sign Language (ASL), Indian Sign Language, and Italian Sign Language datasets, they constructed a simple image categorization model using logistic regression (Halder and Tayade, 2021)

This research looks at how deep learning can be used to classify American Sign Language into multiple classes (ASL). Every image in the ASL dataset will be assigned to one of 29 classes. On our ASL dataset, we design a fully convolutional neural network (CNN) to achieve this. The data set contains 87,000 photographs divided into 29 classes, with 26 for letters A-Z and three for SPACE, DELETE, and NOTHING, and they chose 21750 images, or 750 each class, to be used for the models. They built our model from scratch on a VGG16 network and used pre-trained weights to explore how transfer learning influences performance. They also wanted to test the effects of training the model with deeper and wider networks on performance, so we used InceptionNet and ResNet50 (Sood, 2022)

The goal of this project is to develop a vision-based application that provides sign language translation to text, allowing signers and non-signers to communicate more effectively. The suggested model extracts temporal and spatial characteristics from video sequences. Then, for identifying spatial characteristics, we utilize Inception, a CNN (Convolutional Neural Network). The RNN (Recurrent Neural Network) is then used to train on temporal information. The American Sign Language Dataset was used in this research (Bantupalli, 2018).

This research work shows an innovative context with the primary goal of converting 24 static motions from American Sign Language alphabets into human or machine read able English manuscript. More than ten thousand hand and face gesture signs exist to sign the various English words in ASL alphabets and numbers. This author also explains the system architecture, state of the art, data collecting for the proposed work, suggested system design, and thorough results evaluation by displaying a graphical representation of the proposed technique compared to existing techniques average recognition rate. Pre-processing the input image by an effective segmentation utilizing the BB Technique is distributed as part of the process of recognizing ASL Alphabets. Each data set comprises the 24 ASL alphabet movements. The alphabets 'J' and 'Z' have been left out since they need hand movement (Shivashankara and Srinath, 2019).

Figure 2.5: Translation of gestures of ASL alphabets(Shivashankara and Srinath, 2019)

The goal of this research is to create a sign language translation application utilizing OPEN CV on Android. The application will convert the movements of a finger into an alphabet letter using American sign language (ASL). Figure 2.8 depicts the flowchart. The application will first identify the background color on 7 section coordinates that are designated with 7 boxes. The threshold for acquiring binary imagery from RGBA data input is calculated using this color data. Following the acquisition of seven color data for the hand, seven upper and lower borders for the hand region are determined, which can be represented as a two-dimensional array. The binary images of the hand is then combined with the binary image of the background using the logic 'AND'. Next, is to use morphological operations (dilation and erosion) to remove noise (pixels that aren't in use) and achieve the best possible outcome. Following that, fingertip coordinates or Hu Moments values will be entered on an SVM data model with sequential labeling as part of the feature extraction process (Triyono *et al.*, 2018)

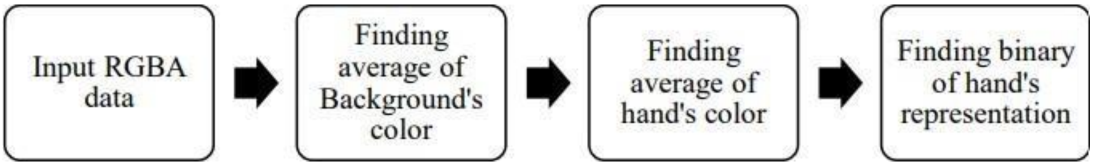


Figure 2.6: Background Subtraction Flowchart(Triyono *et al.*, 2018)

This difficulty is addressed by the system provided here. To recognize the sign made by a gesture, the proposed system leverages the American Sign Language (ASL) data collection. All of the hand shapes and actions are covered by these 70 samples of each symbol. All of the features are from the right hand. A camera is used in this system to collect various hand motions. After that, the image is processed using several techniques. Pre-processing of the image is the first step. The edges are then determined using an edge detection technique. Finally, the sign is identified and the text is displayed using a template-matching algorithm. Because the output is text, it is simple to decipher the meaning of a given sign. The system is implemented using OpenCV and Python language (Shrenika and Madhu Bala, 2020)

In this paper, this proposed system is to recognize ASL Alphabets and Numbers, which mainly depend only on hand and fingers. The process of identifying ASL Alphabets and Numbers is distributed as pre-processing the input image, computation of the region properties of the pre-processed image, and transliteration from treated image to text. 70.83 percent of the time (ASL alphabet movements M, N, Q, V, W, X, and Y are not recognised) and 97.5 percent of the time (ASL alphabet motions G, P, and Q are not recognised in some data sets (S and S, 2018).