SRI LANKAN SIGN LANGUAGE TO TEXT & SPEECH AND VISE-VERSA USING MACHINE LEARNING & IMAGE PROCESSING

Project ID: 2020-077

Project Proposal Report

D. Manoj Kumar – IT17050272

G.M.A.S. Bastiansz – IT17143950

K. Bavanraj – IT17032766

S. Thavananthan – IT17068192

BSc (Hons) Degree in Information Technology (specialization in Software Engineering)

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

February 2020

SRI LANKAN SIGN LANGUAGE TO TEXT & SPEECH AND VISE-VERSA USING MACHINE LEARNING & IMAGE PROCESSING

Project ID: 2020-077

Project Proposal Report

BSc (Hons) Degree in Information Technology (specialization in Software Engineering)

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

February 2020

DECLARATION

I declare that this my own work & this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning & to the best of my knowledge & belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

Name	Student ID	Signature
D. Manoj Kumar	IT17050272	
G.M.A.S. Bastiansz	IT17143950	
K. Bavanraj	IT17032766	
S. Thavananthan	IT17068192	

The above candidate is carrying out research for the undergraduate Dissertation under my supervision.

---------- ----------

Signature of the Supervisor Date

---------- ----------

Signature of the Co-Supervisor Date

ACKNOWLEDGEMENT

We would like to express our sincere gratitude to our supervisor Mr. Buddhika Harshanath and our Co-Supervisor Mr. Jesuthasan Aloysius for their continued support and advice for our research. They provided us with great insight and how to and what not to do when it comes to conducting professional research.

Apart from our supervisors, we would like to thank Dr Dharshana for his continuous encouragement and for providing useful information regarding the research domain.

An extreme thanks go to Mr. Jegan Jeevaraj, the CEO and General Secretary of the Young Men's Christian Association of Batticaloa who has extended his support to obtain data which would be useful in conducting this research.

Above all, I would like to thank my parents who actively encouraged me to excel in my academics and providing a great motivation to take this research as an opportunity to show my skills in academics.

ABSTRACT

A Language is a way of communicating thought between two parties. Effective communication exists when both parties actively engage in the session and respond. Just like speech and text that we use to communicate, sign language is also a method used by the deaf and mute community all over the world to communicate with each other and with other normal people. Not all of us are aware of the sign language and we do require a translation. As Sri Lanka is still known as a developing country, most of the people don't have a clear understanding of the sign language system which is used locally. Therefore, many ordinary people are refusing to communicate with disabled people. Only the people who engage with these community are willing to learn sign language. There are some basic translator systems which are used to translate sign language into normal vocal language. But there are no robust interpreters available in Sri Lanka and most of them lack in real-time translation. Cost of the existing systems are high, and the quantity is not enough for a wide communication system.

Through this report, we are proposing a solution to develop a translation system with a low-resolution camera, which should be able to convert Sri Lankan sign language into English and vice-versa in real-time. The research is divided into four major components. The first component does the data acquisition and feature extraction part. Then it moves to the translation part where extracted images are changed into letters. The third component will do the text and voice assistant part and the final component will convert the texts into sign images. Our main goal is to achieve a system that can perform in real-time. With the advancement of image processing and machine learning, we think we can attain our goal.

TABLE OF CONTENT

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES & ABBREVIATIONS

1. SSL – Sri Lankan Sign Language
2. ML – Machine Learning
3. NLP – Natural Language Processing
4. PWD – Persons With Disabilities
5. SVM – Support Vector Machines
6. ANN – Artificial Neural Network
7. YCbCr - A family of color spaces used as a part of the color
8. ROI – Region of Interest
9. FCC – Freeman Chain Code
10. R-CNN – Region based Convolutional Neural Network
11. CV – Computer Vision
12. RPN – Region Proposal Network
13. DTW - Dynamic Time Wrapping.
14. LMC - Leap Motion Controller.
15. MLP - Multilayer Perception algorithm.
16. HD - High Definition.
17. CNN - Convolutional Neural Network.

1 INTRODUCTION

1.1 Background

In Sri Lanka, 8.7% of the total population is considered as PWD (Persons with Disabilities) [1]. They possess any form of disability from hearing impaired to physical disability. When we take the portion of the population who cannot speak or hear, they usually communicate using Sign Languages. There are different types of Sign Languages that are practiced around the world. In the Sri Lankan context, the deaf & mute of the country also follows a Sign Language called 'The Sri Lankan Sign Language' [2] which comprises of Sinhala, Tamil and English alphabets and a set of pre-defined words which are for common use.

The literacy rate of Sri Lanka is well above 96% [3] and growing. This is mainly due to the introduction of the free education system in the early 1950s. This action has led to creative and innovative solutions for different problems faced by different sectors of the community. However, the contribution of the deaf and mute community is much lesser due to communication issues, lack of financial and provisional donations etc. They're reluctant to actively participate in the socio-cultural activities in the background that they cannot speak or hear like other normal people.

According to basic human rights, the needs and rights of the deaf and mute should also be listened and fulfilled [4]. In Sri Lanka however, this is followed up to some extent. The country has special schools, meal programs, government aids etc. to keep them running daily. But their right to be a helpful citizen just like the others is not taken from them but rather ignored.

Types of difficulties	Number of Persons	% to total Cases	Number of persons not possible at all	% to total Case	Difficulty published in census	% to total Cases
Hearing	354,871	22.8	28,674	20.3	389,077	24.0
Communication	133,623	8.6	47,210	33.5	180,833	11.2

Table 1.1: Population of Deaf & Mute

According to the above table, we can see that there is a considerable amount of deaf and mute. This is just a census report from 2012. At the current date (2020) there can be even more individuals who are deaf and mute.

1.2 Literature Survey

According to the Census of Population & Housing – in 2012, there were 21 people older than 5 years out of a sample of 1000 people were suffering from difficulties in deaf [5]. The deafness can occur to someone at any stage & this may impact on his/her abilities to function well.

People around the world are using different types of sign languages such as Sri Lankan Sign Language [6], American Sign Language [7], British Sign Language [8], Indian Sign Language [9] & French Sign Language [10]. All these sign languages are unique to each country.

There are many researches going in sign language translators all over the world. Most of them use sensors to detect the signer's actions. The main advantage of the sensors are they get the clear view of the signs and they detect edges clearly. Commonly used sensors are Kinect and Leap motion controller. Kinect sensors are used to capture 3D depth images [14]. Leap motion controllers use near infrared rays to track hands and take grey scale images [15]. These high-end sensors can help to reduce the time taken in translations.

This research is mainly focusing on Sri Lankan Sign language which is used by deaf & mute community in our society. There are some applications which were developed based on SSL too. "Nihanda Ridma", "Ahanna" & "Kathana" are three examples for those applications based on SSL. Through this Literature Survey

section, we try to provide a brief summary of each system we mentioned above & a comparison with our proposed system.

1.2.1 Nihanda Ridma System

"Nihanda Ridma" is an application which can be used to convert dynamic gestures into text or voice & vice versa [11]. It is used as a game-based learning system for children with hearing-impaired. There are three components which are focused throughout the project. They are 3D model animating, Motion Tracking & Voice Recognition. 3D model animation used to output the result to a deaf person. Motion Tracking is used to detect & track the movement of the person & it acted as the key ingredient to recognize the sign language gesture in this application. "Nihanda Ridma" captures the text which entered by the user & finds out the relevant animation clip with the sign. Then the system displays the animation clip to the user in avatar mode. It is a web-based application.

1.2.2 Ahanna System

"Ahanna" system is mainly targeting teaching Sinhala Sign Language to the users who are willing to use the system. This is also a web-based application.

The main purpose of this application is to spread Buddhism through the deaf community in Sri Lanka. It is also providing many other benefits such as giving more valuable activities, innovative products & new thoughts in order to improve the knowledge, education of the deaf community.

1.2.3 Kathana, Speech Recognition System

"Kathana" is a Sinhala Speech Recognition System [12]. It uses to convert speeches on the Sinhala language captured by the microphone, to a group of words. Then the recognized words are used as commands, data entries for application.

1.2.4 Sanwadha System

"Sanwadha" is an intelligent mobile assistant application which was build targeting hearing-impaired people in Sri Lanka [13]. The core of this project is Instant Messaging (IM). Here the system gets the Sinhala text from an ordinary person & converts it into SSL. This output displays as a GIF. Also, through the system a user can convert SSL into text or voice format as he/she prefers. The main objective of this

application is that to reduce the digital divide between enabled & hearing-impaired users.

1.2.5 Mexican sign language recognition using Kinect

In this research they talk about the usage of the Kinect sensor. When a signer shows a sign, the system will store the color, depth and the skeleton tracking information using the sensor. Then using the Dynamic Time Wrapping (DTW) algorithm the gathered information is interpreted. For testing they have used the K-Fold Cross validation method. This testing has showed a mean accuracy of 99.1%. From these results we can understand that how accurate is the Kinect sensor in recognition [16].

1.2.6 Real time sign language recognition using the Leap Motion Controller

This system has used leap motion controller to detect and track signers' hands and fingers. The extracted and normalized features will be sent to the classification part. In classification part they have used Multilayer Perception algorithm (MLP) which uses features as input and convert them into specific alphabet letter. For training, authors have implemented Backpropagation algorithm. The total system has showed 96.15% of a recognition rate. This high percentage is achieved only because of the quality feature extraction through LMC [17].

1.2.7 Sign Language Recognition using Microsoft Kinect

This paper is focusing on identifying Sign Language using Microsoft Kinect [18] device. Using computer vision algorithms, they've developed a characteristic depth and motion profile for each sign language gesture. The feature matrix thus generated was trained using a multi-class SVM classifier and the results were compared with existing techniques [19]. The practical application of this system seems expensive since this system requires a Kinect device. The monetary provisions and technical expertise lack in this country. A mobile application or a web camera would be much more feasible since everyone has access to internet and especially mobile phones.

Features	Nihanda Ridma	Ahanna	Kathana	Sanwadha	Proposed System's Component
Text recognition system	✗	✗	✓	✓	✓
Convert English text to sign language	✗	✗	✗	✗	✓
Display the combination of signs as a GIF	✗	✗	✗	✓	✓
Display already-defined signs in a video format	✗	✗	✗	✗	✓
Low resolution web application	✗	✗	✗	✗	✓

Table 1.2: Comparison between Proposed System & Existing Systems

1.3 Research Gap

There is a communication gap between people with difficulties in hearing - speaking & ordinary people. Most of the time people use interpreters while communicating with hearing-impaired & mute people. But when there are no interpreters (translators), people face to a big problem because they cannot understand what each other says.

To avoid from this issue the countries like America, India found out various inventions what can easily convert sign language to textual format & vice-versa [20]. But those systems are not practical systems for Sri Lanka based on some reasons. The first reason is Sri Lanka still listed as a developing country. The tools & technologies which were used to build such systems by other countries are very expensive now-a-days. Sri Lankan government cannot be able to spend that amount of money for only one main purpose. The second reason is that, even we

buy one of those system, people cannot get the benefits of the system as each country has their own sign language. Another reason is that, the system which already implemented for specific language such as Sinhala. Due to this reason, if any person who doesn't have a proper practice in that language won't understand the system. Therefore, by-today the ordinary people who are willing to communicate with hearing-impaired & mute people have challenged with a huge communicating gap in their daily lives.

1.4 Research Problem

Sri Lanka is very hospitality country from the ancient times. Almost all the people in the country don't think twice to friend with another person who looks in pleasant manner. Therefore, in Sri Lanka "communication" acts a main role among the society.

As mentioned in Research Gap, there are different types of people in Sri Lanka. Among those categories, people with difficulties in speaking & hearing get a special attention. Those people use Sri Lankan Sign Language (SSL) for their communication purposes with people with same manner. But problems arise when there is no proper understanding about the SSL among the ordinary people. This issue may be occurred when having a poor guidance about the SSL. Even though many Sri Lankan universities, private institutions who are holding degrees & diplomas on SSL, people won't show any interesting because of above mentioned reason. Even there is a bit of interesting, only limited number of students get the chance to study. Due to this reason most of the ordinary people lose the chance to learn SSL. Therefore, they are refusing to be friendly & communicate with hearing-impaired & mute people frightening both parties could not be able to understand what each other tries to say.

But it is important to build a communication bridge between ordinary people & people with difficulties with speaking & hearing. Hearing-impaired & mute people always communicate physically & visually. But those people too refuse to talk with ordinary people because ordinary people can't recognize the language if there are no interpreters for translation purposes.

Even there are some applications in Sri Lanka which are capable of translating SSL into natural vocal language or translating natural local language into SSL, it is hard to find out a system which can do both translation modes using one interface. Also, most of the all the implemented system uses Sinhala language for translating purposes. Even Sinhalese is the main language in Sri Lanka, some people cannot even understand Sinhalese. In cases like that, ordinary people or deaf people or mute people who are not familiar with Sinhala face to many problems while using existing systems.

Another problem is that using only one type of data set for translating purposes in the system. In many research papers we can see that the researchers use only Sri Lankan Sign Language alphabet “or” the set of pre-defined words in their research projects. It is not a best practical to use only one type of data sets in a translating system due to many reasons.

If there is a proper converting application instead of interpreters which can be used as a self-guidance, ordinary people as well as people with difficulties in hearing & speaking will be able to use them whenever they want. Then, it may be help them to break the communication gap & be friendly with each & everyone in the society.

This is the problem address in our research component. Some researchers have put their effort to address this fact. Also, they have come up with applications. But unfortunately, none of those application can fulfill all the requirements, which are expecting from an application such as converting day-to-day words into SSL, presenting converted answer as a proper identifiable output & vice-versa.

2 OBJECTIVES

2.1 Main Objective

- The main idea of this research is to improve the communication bond between the ordinary people & people with difficulties in hearing & speaking.
- To successfully fulfil the above idea, this system should be able to convert SSL into text format or audio format & converting English text format into SSL using low-resolution web application

The sub-objectives which are specified to each component can be listed as below:

2.2 Specific Objectives

- To verify that the product is reliable for ordinary people who are willing to keep better communication bond with people with difficulties in hearing & speaking and vice-versa.
- To capture an image using low resolution camera and train models through those low-resolution images. This would facilitate the system being used in a web interface of a laptop camera.
- To use R-CNN, in order to identify series of hand gestures with the aid of pre trained models of low-resolution images
- To prepare an optimal dataset to create the effective data model
- To apply a proper machine learning algorithm for translation
- To optimize the overall model to real time translating
- To convert alphabet into a textual format
- To Translate text to speech local languages
- To identify & extract the information, emotion get from the keyboard inputs.
- To investigate already-defined Sri Lankan signs for day-to-day words.
- To prepare the given text message to a GIF to convert the message to SSL with more accurately & eye-pleasant manner.
- To emerge the application in English language to reach the Sri Lankan society in effective way.
- To determine the usage of low-resolution web application for hearing-impaired people, mute people as well as ordinary people can be observed as an effort,

which allows those communities to clarify any need of learning & communication at anytime, anywhere.

- All the above objectives should be accomplished in real time. This is because we are proposing to build a mobile application in which the application response speed is of utmost importance

3 SYSTEM METHODOLOGY

The proposed system comprises of four major components which are namely: "**Data Acquisition & feature extracting**", "**Sign recognizing & translating**", "**Converting the identified sign into textual format or audio format**" & "**Converting textual format into SSL**". There are specific methodologies for each 4 components. The whole system will use many techniques & algorithms in ML, AI, NLP areas in order to fulfill the workload in each component.

3.1 Requirements

- Should be able to identify hand gestures in low resolution imagery provided.
- All the intended operations of the component should happen in real time regardless of device nature (Mobile, Web etc.)
- Getting the keyboard inputs
- Analyzing & extracting information from source text
- Classifying the input-text as already-defined common word or a new text input while studying the system database
- Converting a collection of signs into a GIF
- Sending the already-stored video clips, if the input-text matches with the data record in the system database

3.2 System Overview

3.2.1 Data acquisition & Feature extraction in real time

To implement a system using R-CNN, we must first understand the concept of Computer Vision (CV). CV is used for many kinds of image recognition and classification. Features such as edge detection, Object detection, Object Segmentation and Semantic Segmentation are some of the key features of CV. In this component however, we are going to identify hands and hand gestures in order to take the imagery.

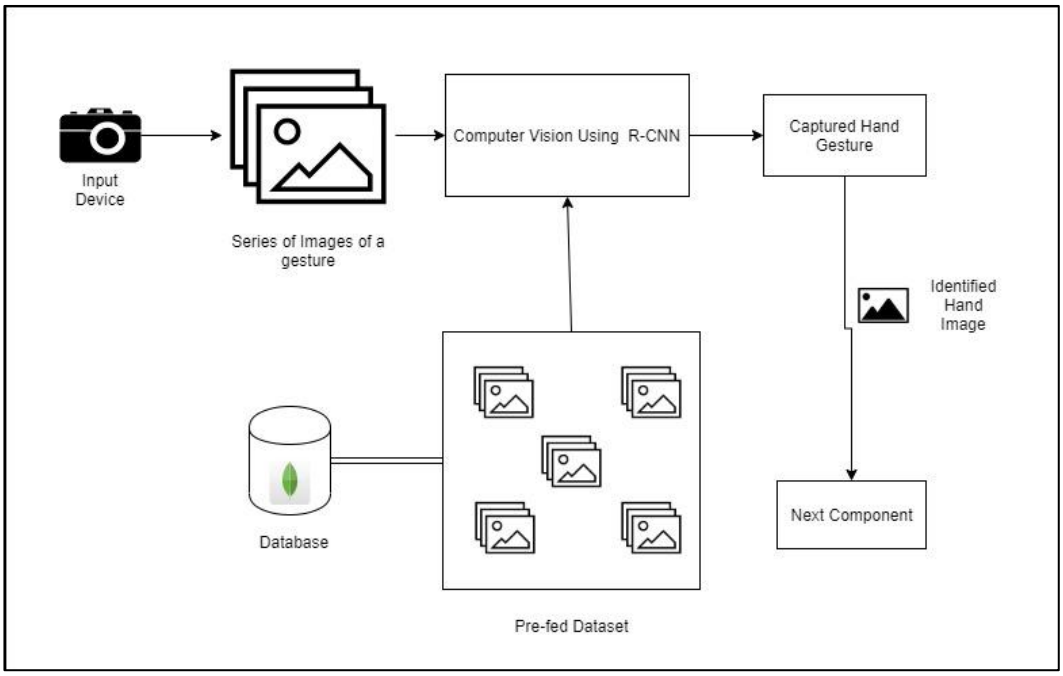

Figure 3.1: System overview diagram for Data Acquisition & Feature Extraction

Initially, the system is fed with a set of images for training. These images should be low-resolution images. We are planning to use 100 such low-resolution to train the model to detect hands. The reason for using low-resolution images is to make the application available on devices that have less camera resolution such as mid-range mobile phones and webcams.

The heart of the component is the place where images are classified and identified. R-CNN is a deep neural network for instance segmentation. This consists of two parts [21].

- RPN – Region Proposal Network which proposes candidate object bounding boxes.
- Binary mask classifier to generate mask for every object

By the term ‘Object’, we mean the hand gestures. First the features are extracted from the image using an architecture called ResNet101 architecture [22] and sent to the RPN. The RPN then generates a Region of Interest on where the object will be separated with bounding boxes.

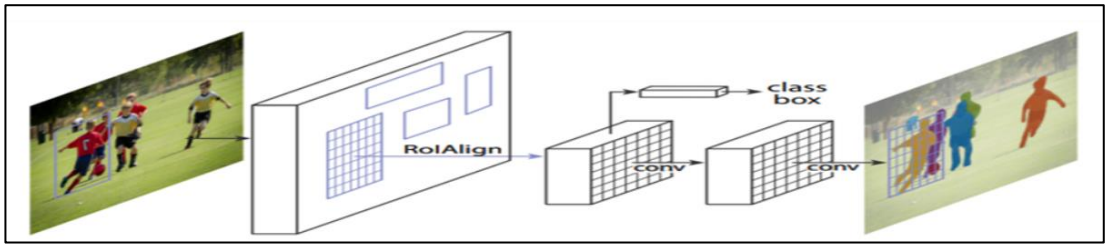

Figure 3.2: Generating ROI using RPN

In this way, the component can identify the hand gestures using this R-CNN and will produce a set of vectors to transfer it to the next components for translation and interpretation

3.2.2 Sign Recognition & Translation

To build this research component we must go through several steps. The major part for this component is the dataset. As the system is going to develop through a machine learning algorithm, we must choose the optimized dataset for the process. At first, we are planning to take 20 images for a sign. As there are 26 letters in the English alphabet, we are expecting to gather 520 images to train the module. Multiple signers show the signs of English alphabet in Sri Lankan sign method. These images are going to be taken by the web camera in different lightning conditions and backgrounds. From these images we select the most optimal images to create the effective data model.

As the next step the dataset must be trained using a machine learning algorithm such as Convolutional Neural Network (CNN). CNN algorithm can take images as an input and can learn different objects in the images. The image will be recognized as an array of pixels and it depends on the image quality. At the end the algorithm will process the image and classify it under given categories. So, in our context we can categorize the set of images to the A - Z alphabet. Compared to other algorithms CNN doesn't need much pre-processing for the image. It increases the processing speed of the component, which is a main advantage of using this algorithm [23].

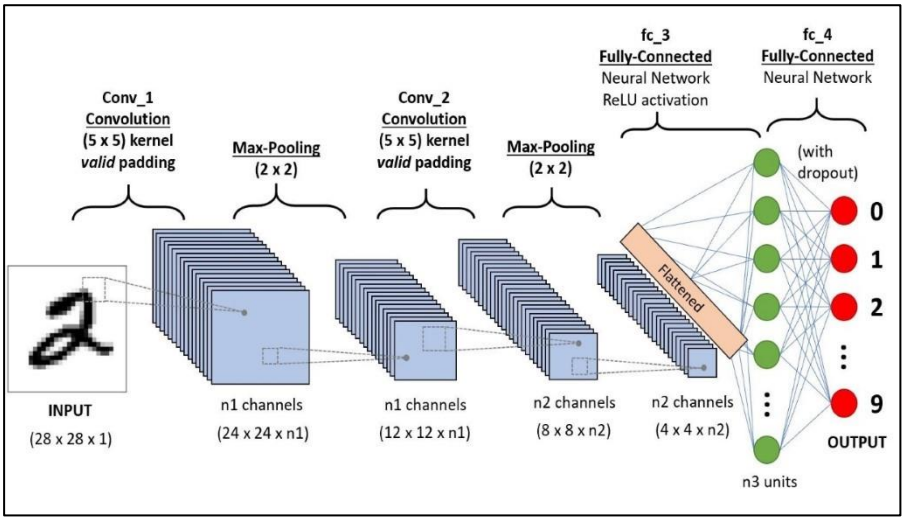

Figure 1.3: CNN Model

For better results we must optimize the selected machine learning algorithm. After training the dataset and optimizing the algorithm we can predict the signs using the image flow. Once the sign is identified, it will be passed to the next research component as an English alphabet letter.

Figure 3.4 shows the system diagram of the component. The extracted sign images are fetched from before component and will be send to the classification module. In the classification module the images are compared with the pretrained dataset and the letter according to the sign will be identified. The identified letters will be sent to the next component so that it will be made as a meaningful word. Optimising the whole component is the final part. Translations must happen in real time. For that optimisation must happen in the classification part. Time optimisation will happen in two parts. The first part is measuring the total time that has been taken by the method to build the training model. Second part is measuring the total time that has been taken to find the result, based on the test data. By using these results, we can optimize the time in each part and achieve the “real time processing” goal.

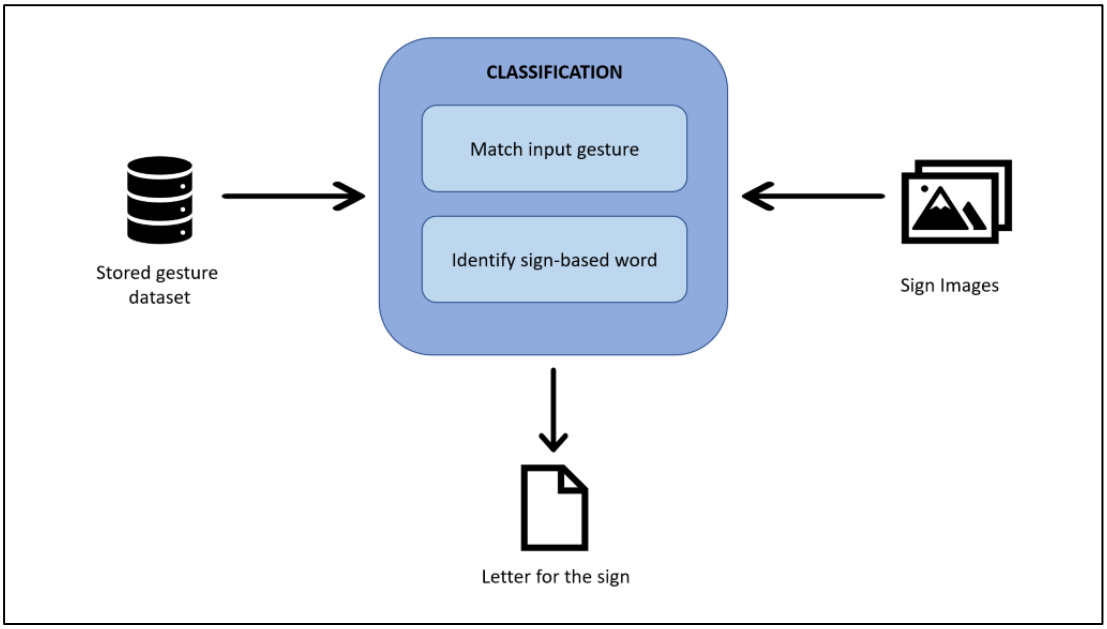

Figure 3.4: Component Diagram

3.2.3 Text & Voice assistant for SSL

3.2.3.1 Convert alphabet into textual format

Receive each alphabet from the previous component and then make the English sentences from the collection of alphabets, integrate them and the basic elements of translation programs are words and rules for combining them to form sentences, paragraphs and complete texts then we are using tokenization. In this phase, the input sentence is decomposed into tokens.

3.2.3.2 Translate text format into local speech languages

Certain tokens are assigned n to the POS stacker function to indicate tokens with their respective type. A partial speech (POS tagger) is a software that reads the text in some language and assigns parts of speech to each word (and another token) like noun, verb, adverb. Computational applications use very good POS tags such as 'noun-plural'. Searching for the source page of the corpus for 'close' matches and their translations. When applicable, each token marked with the POS tag is searched in the dictionary and if a match is found, the word is sent to the next step. The "classical" similarity measure suggested by Nago (1984) and used in many early EBMD systems is the use of vocabulary or similar methods to identify word

similarity in terms of meaning or usage. Here, matching words are allowed when words in the input string are replaced with similar words in the example sentences. At this point, after matching and rejoining, the matching words are matched to the English corpus. If it finds a fit, the English word is substituted. Consequently, substitute voice and speech according to be output.[26]

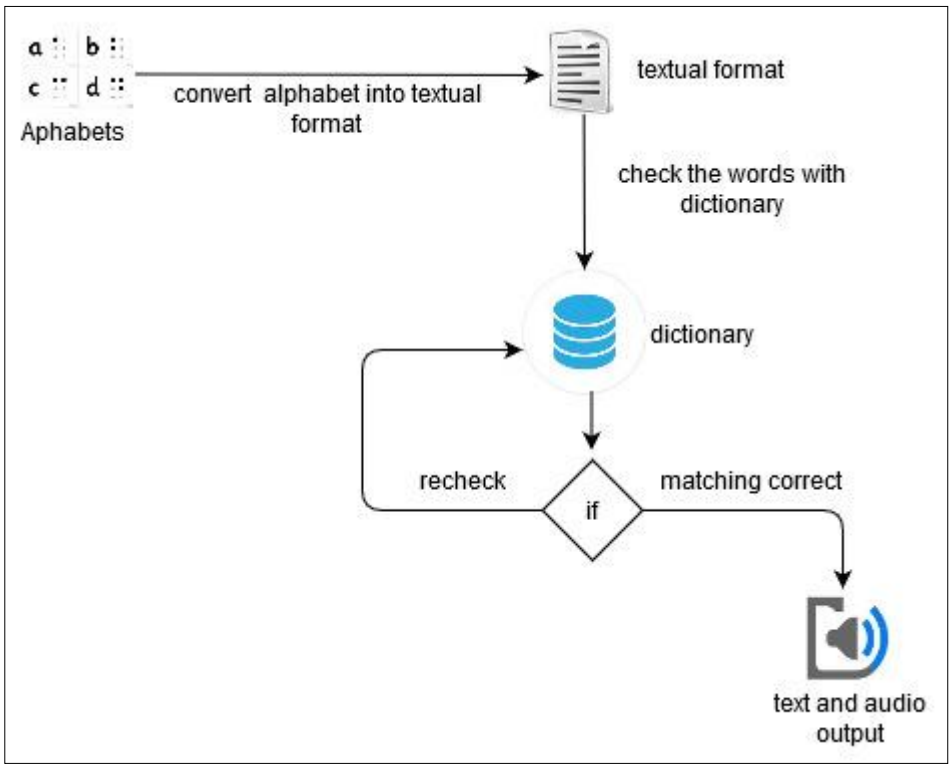

Figure 3.5: Text & Voice Assistant Processing Steps

3.2.4 Conversion on text format into SSL in real time

User can input the text using his/her keyboard to the system. After entering the text, the system will identify the text & pre-processes it to analyze whether the input is a text or an image. If the input is a text, the system will continue the pre-processing in order to identify whether the entered text is displayed in English format or not. If the text is not displayed in English, then the system will translate the text into English while respecting the grammar rules. After pre-processing gets over, the analyzed text will send to the system database to start the converting process.

This stage is one of the most important section in whole research component. In system database, there are 2 type of data tables. One data table keeps the hand gestures which are related to the letters in English Alphabet. The sample for the hand gestures related to English Alphabet letters are shown in Figure 3.6

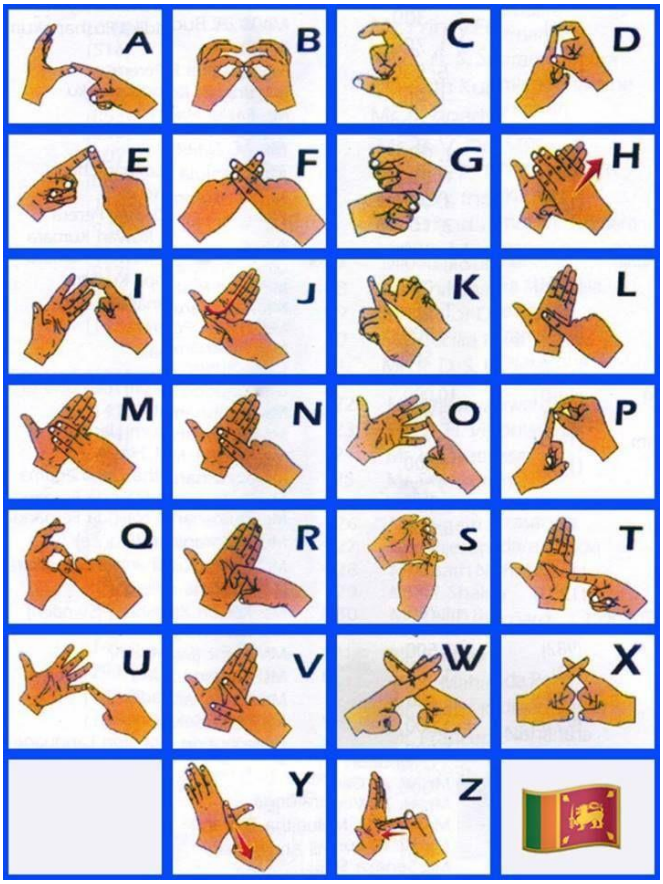

Figure 3.6: English Alphabet Using in Sri Lankan Sign Language

The other data table consists of already-defined Sri Lankan hand gestures. As below Figure 3.2 there are some sign which already give the meaning of one specific word like “Good morning” or “Thank you”.

Our system database will be stored with small video clips for each already-defined sign. When pre-processed text arrives to the system database, the system start will compare the text with the data records. If system finds there is already-defined sign for

the text, then chooses the related video clip. Otherwise, the system will select signs for each letter in the word & combine all the signs together. Then the combination of signs will be converted into a GIF & pass to the next section.

At the last stage, all the video clips & GIFs will be ordered according to the order of text entered to the system & displayed to the user.

3.2.4.1 Pre-processing the Input-Text

Pre-processing the input-text is a series of events happens between get an input-text from the user & Converting the text into a language which can be understood by the system. This series of events are displayed in Figure 3.7

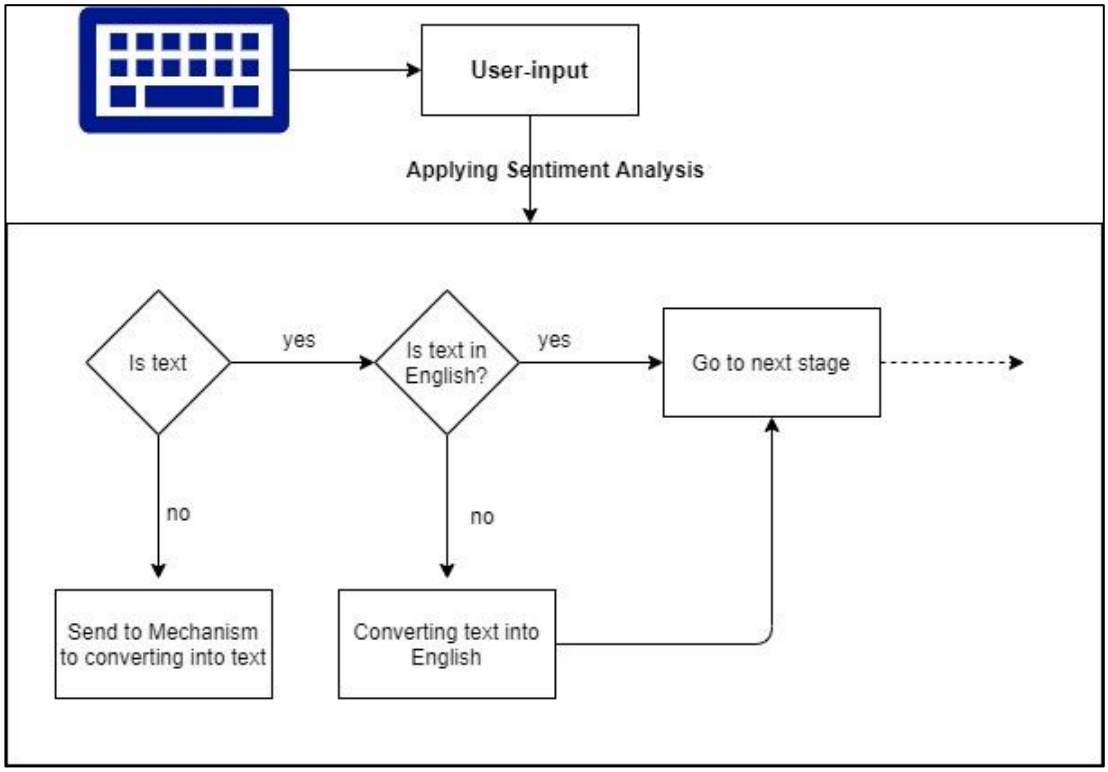

Figure 3.7: Pre-processing Input Text Diagram

As soon as the system get the input, which was entered by the user, it starts to apply the Sentiment Analysis on the input [24]. At the first step, the system tries to figure out whether the input is a text or an image. If it is a text, then the system tries to figure out whether the text is in English text format or not. If it is not is English, then the system will translate the text into English while following the grammar rules. Finally, the pre-processed text will send to the next stage.

3.2.4.2 Mechanism of Converting Text to GIF

Mechanism of converting text into GIF is a series of events which happens between input text & GIF output. This series of events is known as “Text Adaptation” which displays in below Figure 3.8

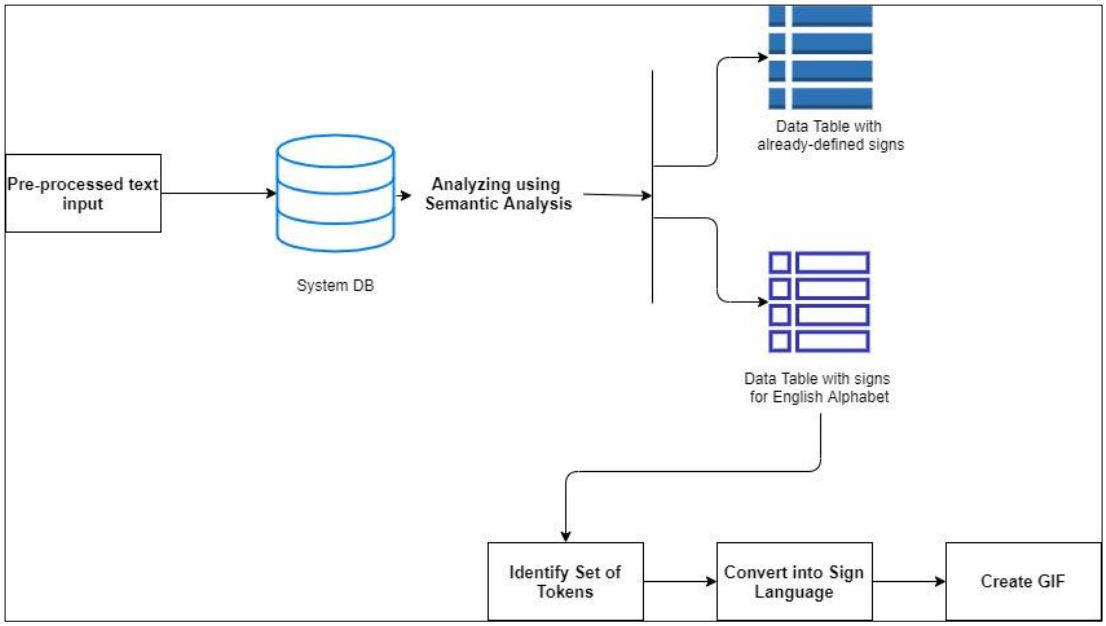

Figure 3.8: Mechanism to Converting Text to GIF

The pre-processed text is sent to the system database. In here, system tries to recognize one word at each time with the assistance of Semantic Analysis [25]. After the analyzing the system will recognize that the related sign for the entered word is existing in system database or not. If there is no sign related to the input text, then the system will create a GIF & display to the user.

4 DESCRIPTION OF PERSONAL & FACILITIES

4.1 Work Breakdown Structure

4.1.1 Data Acquisition & Feature Extraction in Real Time

4.1.1.1 Capturing Image and Identifying Hands

- Gathering a set of individuals who have different skin color, varying palm sizes etc.
- Take photos of both their hands on the front and the back sides. Everyone is taken 8 low resolution photos each (maximum 20 individuals)
- Those images are fed into the database for training the model to identify hands.
- Using those pre-fed images, implementing the system to identify hands using leading tech stack and algorithms

4.1.1.2 Real Time Operation

- All the operations should happen in real time
- Fine tuning the image recognition algorithm
- Improving the database schema

4.1.2 Sign Recognition & Translation

- Preparing an optimal dataset to create the effective data model.
- Applying a proper machine learning algorithm for translation.
- Translating images into representing English alphabet letter.
- Optimizing the overall model to real time translating.

4.1.3 Converting text format into SSL in Real Time

4.1.3.1 User Input Capturing & Analyzing

1 Capture the user inputs efficiently

2 Apply Sentiment Analysis technique to the captured input

- By applying Sentiment Analysis, system should be able to pre-process the user-inputs while checking whether the type of the user-input is text & if the user-input fulfill above fact then check whether it is entered to the system using English language.

4.1.3.2 Presenting the Solved Sign Language Conversion

1 Compare the pre-processed input-text with the data records in system database

- When comparing, first the system should check whether there is already-defined sign. To fulfill this task, the system should compare the input-text with the data table which contain the video clips of already-defined signs.
- If there is similar text, then system should get the video clip which related to the text & pass to next section.
- If system unable to find a similar text in the data table with already-defined video clips, then system should check with the data table which contains the letters of English alphabet & the sign related to each letter.
- Then system should convert each letter in the text into sign using the above-mentioned data table with English letter. (The order of the sign should like the order of the English letters in the text)
- Then convert the combination of signs into a GIF using Semantic Analysis & send to the last step.

2 Display the input-text with converted sign to the user

- In the last stage, the system should consist with all the video clips & GIF related to the user-input. (All the video clips & GIF should place in the similar order as the user-input)
- Then system should display the converted signs with the user-input in eye-pleasant manner.

5 BUDGET & BUDGET JUSTIFICATION

Without a budget, no system can function or do efficient analysis on a given domain. The table below shows the breakdown of the budget which might vary according to changing prices of the products and services.

Component	Amount (USD)	Amount (LKR)
Variable Cost		
Travelling – Train (Batticaloa -YMCA)	7.92	1440.00
Travelling – Bus (Ratmalana)	1.12	200.00
Internet Charges	11.01	2000.00
Mongo DB cloud – 2GB	11.01	2000.00
Total	**31.06**	**5640.00**
Fixed Cost		
Domain Name Registration	4.40	800.00
Hosting	55.05	10000.00
Total	**59.45**	**10800.00**

Table 2.1: Budget Plan

6 REFERENCES

[1] Wikipedia, "Disability in Sri Lanka," 29 08 2017. [Online]. Available: https://en.wikipedia.org/wiki/Disability_in_Sri_Lanka. [Accessed 22 January 2020].

[2] Wikipedia, "Sri Lankan sign Languages," 24 February 2011. [Online]. Available: https://en.wikipedia.org/wiki/Sri_Lankan_sign_languages. [Accessed 25 January 2020].

[3] Wikipedia, "Education in Sri Lanka," 15 February 2018. [Online]. Available: https://en.wikipedia.org/wiki/Education_in_Sri_Lanka. [Accessed 26 January 2020].

[4] CBM Australia, "Sri Lanka Disability Information Sheet," 2014, pp. 4-5.

[5] Department of Census and Statistics, "Census of Population and Housing," Ministry of Finance and Planning, Government of Sri Lanka, 2012. [Online]. Available: http://www.statistics.gov.lk/PopHouSat/CPH2011/Pages/Activities/Reports/FinalReport/FinalReportE.pdf [Accessed 26 January 2020].

[6] Wikipedia, "Sri Lankan sign Languages," 24 February 2011. [Online]. Available: https://en.wikipedia.org/wiki/Sri_Lankan_sign_languages. [Accessed 25 January 2020].

[7] Wikipedia, "American Sign Language," [Online]. Available: https://en.wikipedia.org/wiki/American_Sign_Language. [Accessed 13 January 2020].

[8] Wikipedia, "British Sign Language," [Online]. Available: https://en.wikipedia.org/wiki/British_Sign_Language. [Accessed 13 January 2020].

[9] Wikipedia, "Indian Sign Language," [Online]. Available: https://en.wikipedia.org/wiki/Indo-Pakistani_Sign_Language. [Accessed 13 January 2020].

[10] Wikipedia, "French Sign Language," [Online]. Available: https://en.wikipedia.org/wiki/French_Sign_Language. [Accessed 13 January 2020].

[11] K. N, S. Wellage, H. Samarawickrama, W. Perera and J. Yasas, “The Rhythm of Silence” - Gesture Based Intercommunication Platform for Hearing-impaired People (Nihanda Ridma), Colombo: SemiScholar, 2014.

[12] M. Tharindu , R. Chamal, C. Maleen and Y. Randika, "Kathana - Sinhala Speech Recognition System," [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.462.1848&rep=rep1&type=pdf. [Accessed 20 January 2020].

[13] S. Perera, J. Jayalath, W. S. Tissera and A. Bandara, "Sanwada-App," 18 November 2017. [Online]. Available: http://project.sliit.lk/redmine/documents/206. [Accessed 23 January 2020].

[14] Microsoft Azure. “Azure Kinect DK Documentation”, [online]. Available at: https://docs.microsoft.com/en-us/azure/kinect-dk/ [Accessed 22 Feb. 2020].

[15] Leap motion developer. “Leap Motion Developer documentation”, [online]. Available at: https://developer.leapmotion.com/documentation/ [Accessed 22 Feb. 2020].

[16] Gibran García-Bautista, Felipe Trujillo-Romero, Santiago Omar Caballero-Morales, “Mexican Sign Language Recognition Using Kinect and Data Time Warping Algorithm”, Available at: ieeeexplore.ieee.org. [Accessed 13 Feb. 2020].

[17] Deepali Naglot, Milind Kulkarni, “Real Time Sign Language Recognition using the Leap Motion Controller”, Available at: ieeeexplore.ieee.org. [Accessed 13 Feb. 2020].

[18] A. Agarwal and M. K. Thakur, "Sign Language Recognition using Microsoft Kinect," IEEE, pp. 181-182.

[19] "1.4 Support Vector Machines," [Online]. Available: https://scikit-learn.org/stable/modules/svm.html. [Accessed 02 02 2020].

[20] Ahanna Foundation, "E – Sign Reader - Next age of Sri Lankan Sign Language…," [Online]. Available: https://application.isif.asia/projects/projects/view/1308. [Accessed 23 January 2020].

[21] R. Khandelwal, "Computer Vision — A journey from CNN to Mask R-CNN and YOLO -Part 2," 27 July 2019. [Online]. Available: https://towardsdatascience.com/computer-vision-a-journey-from-cnn-to-mask-r-cnn-and-yolo-part-2-b0b9e67762b1. [Accessed 18 01 2020].

[22] P. Jay, "Understanding and Implementing Architectures of ResNet and ResNeXt for state-of-the-art Image Classification: From Microsoft to Facebook [Part 1]," 2017 February 2018. [Online]. Available: https://medium.com/@14prakash/understanding-and-implementing-architectures-of-resnet-and-resnext-for-state-of-the-art-image-cf51669e1624. [Accessed 02 02 2020].

[23] Techopedia, “Convolutional Neural Network (CNN)”, [online]. Available at: https://www.techopedia.com/definition/32731/convolutional-neural-network-cnn [Accessed 22 Feb. 2020].

[24] Altexsoft, "Sentiment Analysis: Types, Tools, and Use Cases," 21 September 2018. [Online]. Available: https://www.altexsoft.com/blog/business/sentiment-analysis-types-tools-and-use-cases/. [Accessed 25 January 2020].

[25] R. M. Siegfried, "Compiler Construction," 2004. [Online]. Available: https://home.adelphi.edu/~siegfried/cs372/372l8.pdf. [Accessed 3 February 2020].

[26] Manish Rana, “Example Based Machine Translation Using Fuzzy”. 2016.[Online]. Available: at https://pdfs.semanticscholar.org/d0dc/e48493e1ffc42453f8335842f7dbef61bd71.pdf. [Accessed: 22 February 2020]