INFORMATICS INSTITUTE OF TECHNOLOGY

In Collaboration with

ROBERT GORDON UNIVERSITY ABERDEEN

Interactive Sign Language Learning Platform: SignIt

Group 03 Thesis Report by

Ms Jathusharini Basker - 20210452

Ms Anuttara Rajasinghe - 20210216

Ms A. B. Duweeja De Lima - 20210522

Mr Mathushan Manorangithan - 20200785

Supervised by

Ms Ganesha Thondilege

Submitted in partial fulfilment of the requirements for the BSc in Artificial Intelligence and Data Science degree at the Robert Gordon University.

February 2023

Table of Contents

List of Tables

List of Figures

List of Abbreviations

Acronym	**Description**
SSL	Sinhala Sign Language
BSL	British Sign Language
ASL	American Sign Language
CNN	Convolutional Neural Network
ANN	Artificial Neural Network
DNN	Deep Neural Network
RCNN	Region Based Convolutional Networks
SVM	Support Vector Machine
OOAD	Object Oriented Analysis and Design
SSADM	Structured System Analysis and Design Method
API	Application Programming Interface

KPI	Key Performance Indicator
CI / CD	Continuous Integration / Continuous Deployment

CHAPTER 01 : INTRODUCTION

1.1. Chapter Overview

The "SignIt" Sign language learning platform is an online web application that can be used to learn sign language which may be useful in day-to-day life of hearing impaired people. The goal of the website is to increase awareness and acceptance of Sinhala sign language, particularly within the deaf and hard-of-hearing community around the nation. Currently, there are no implemented systems to learn Sinhala Sign Language(SSL) in such an interactive and innovative way using new machine learning techniques which includes the latest AI technologies and real-time video gesture recognitions models to verify the user's input to the system, that way users are given the best way to learn a a sign for a word at their own pace. Additionally, the website also helpfully provides methods of tracking the user's learning progress so that the user is aware of how far they have come in their educational journey.

1.2. Problem Domain

Hearing-impaired people communicate using a different language known as Sign Language. Different sign languages are used in different countries or regions. The most widely used is the American Sign Language (ASL) and British Sign Language (BSL). ASL letters are signed in one hand while both hands are used in BSL. ASL is widely used in the United States and Canada, whereas BSL is widely used in Australia, New Zealand, India, and Sri Lanka. Sri Lankan hearing impaired People use SSL, derived from BSL.

As of 2022, a population of 430 million globally is a part of the hearing-impaired community, according to the World Health Organisation (WHO) Hearing loss affects over 20% of the world's population or more than 1.5 billion people. Furthermore, more than 389,000 individuals in Sri Lanka are affected by the same problem. More than 74% of the people of Sri Lanka speak Sinhala as their first language. Despite the fact that schools for the deaf exist, there are no any platforms to study SSL in Sri Lanka. There are students in remote regions of the country who do not have access to or afford these particular schools and equipment. There are also those who progressively lose their hearing, which implies they understand how their spoken word works.

This area of research in Sri Lanka is not really a popular area, not much awareness has been raised on the situation for the education of sign language. This disability needs much more attention than what is received as of now and we aim to raise awareness about the severity of this situation in Sri Lanka. This research is going to be risky as well as not a lot of known datasets are available for words in SSL. We believe that it is possible to aid this community and have a positive impact on the deaf and mute society to take a step forward to better learning using the newest technologies and utilising better accessibility features.

For these reasons, our primary objective is to provide an effective technique for assisting the deaf and mute community and other interested parties to learn sign language through the use of an online sign language learning platform.

1.3. Problem Definition

The deaf and hard-of-hearing community in Sri Lanka often encounter difficulties in communicating with the broader society due to a pervasive misconception that sign languages are inferior to oral languages. Our project aims to bridge this communication gap by promoting the use of Sinhala Sign Language, the most prevalent sign language in Sri Lanka, and elevating awareness of the language and the challenges faced by the deaf and hard-of-hearing community. By doing so, we hope to foster a more inclusive society where the deaf and hard-of-hearing community can communicate effectively with the rest of society and be recognized for the unique linguistic capabilities they possess.

By reducing the communication barrier between the deaf and hard-of-hearing community and the rest of society, we hope to create a more inclusive and equal society, where everyone has equal access to information and communication.

Artificial Intelligence and machine learning technologies are the centrepiece for this kind of problem. However, utilising AI tools in the situation could be a matter of great debate. Additionally, there are limited amounts of datasets available for SSL to train the AI, which is a common problem in most underdeveloped and developing countries like Sri Lanka.

This major problem at the present time sparked our interest towards them and led us to try to find a solution to their difficulties and minimise them as much as possible. Proposing an innovative system that is best and the platform should be accessible, affordable and should be a best solution so that it can be effective for a developing country like ours. We plan to provide an online free learning platform for the education of sinhala sign language. This makes it accessible even for the most rural provinces of the country and also making this version of the project freely available adds to the accessibility of our website.

1.3.1. Problem Statement

Sri Lanka, as of now, doesn't have an online learning platform for SSL education and a reliable method to acquire study material for learning SSL, which we intend to rectify by developing a web application that meets these requirements using machine learning implementations.

1.4. Research Motivation

Real-world problem-solving objectives, a need for intellectual fulfilment on a personal level, and motives for professional advancement are the driving forces behind this research. We feel that we should leverage the satisfaction we derive from curiosity and the desire to solve problems to offer and implement a system that is beneficial, novel, and useful to the world that might be used for a highly humanistic cause. Here in Sri Lanka, we want to lower the information gap between the hearing community and the deaf community.With SSL as our selected research topic, our research paper has given us the chance to share our own answer to significant and humanitarian real-world problems. We did this by using an interactive learning platform. Future advances focused on assisting the head community, to shed light on subjects that are much or little discussed, may be interested in conducting an original study on a topic, without uncovering any previous work or inventions, and sharing its results with a community based on data science. Our activities generally strive to improve and support society in this endeavour. We have committed ourselves to this research as a result of these motivations.

1.5. Existing Works

Table 1 : Existing Works

Citation	Limitation	Technology/ Algorithm	Advantage
(Dissanayake, Maheshi & Herath, H.C.M. & W.A.L.V. Kumari, & Senevirathne, W.A.P.B., 2013) [Not a learning platform]	Machine learning techniques aren't used.	Image processing technologies. .	Highly accurate recognition system because of accurate algorithms and calculations used to identify the number of fingers and positioning of the hand.
(S. Dilakshan and Y. H. P. P. Priyadarshana, 2020) [Not a learning platform]	Not using video input. Only supports 12 basic signs.	CNN Architecture	Accurate classification system because the CNN system is used. Good performance levels and outputs are given fairly quickly
(S.D. Hettiarachchi and R.G.N. Meegama, 2020) [Not a learning platform]	Limited number of texts. Only works for the right hand	Image processing along with CNN	The application is able to generate the relevant letter by getting an input of a hand gesture within 1.75 seconds of average time.

	gestures.		
(R.M. Rishan, S. Jayalal and T.K. Wijayasiriwardhane, 2022) [Not a learning platform]	Leap Motion is unable to see through the fingers which causes inaccurate validation.	Image processing. Convolutional neural network. Leap Motion technology.	It describes the importance of visual language interpretation based on SSL and highlights the raw interpretation of gestures.
(L.L.D.K.Perera and S.G.V.S.Jayalal. 2021) [Not a learning platform]	Haven't touched upon dynamic sign language gesture recognition and combining facial expressions of sign language in the gesture recognition process.	Scale Invariant Feature Transform (SIFT) CNN Image processing	A combination of SIFT features with CNN, improved the robustness to scale variations in sign language recognition, was implemented at a low cost and improved the model accuracy compared to a single-channel CNN implementation
(Kumarawadu, P. and Izzath, M. 2022)	Has used a limited	CNN ANN	Capable of handling both static and dynamic Signs. Uses LMC to

[Not a learning platform]	Combination of the general position of hands and speed at which the hand signs are performed is not paid attention to.	DNN Leap Motion	capture hand gestures, processes captured information, identifies corresponding words and phrases reference to the communication in SSL and displays them as a message in Sinhala Language enabling non-SSL speaker to understand and interact with a hearing-impaired person
Kumar, D.M., Bavanraj, K., Thavananthan, S., Bastiansz, G.M.A.S., Harshanath, S.M.B. and Alosious, J., 2020) [Not a learning platform] .	The same words and phrases can have different meanings according to the context of a sentence and many words by using natural language processing.	Image classification along with CNN and Linguistic Data Consortium to identify the word segments. Natural language processing and speech digital processing.	The proposed system has the ability to detect high-resolution images as well as low-resolution images of the hand signs. It's a highly accurate system which can also detect hand signs through live images.

Dissanayake, I.S.M., Wickramanayake, P.J., Mudunkotuwa, M.A.S. and Fernando, P.W.N., 2020) [Not a learning platform]	The recurrent neural network (RNN) which is used for analysing video segments in the system will slow down the computational speed of this neural network and training activities can be difficult.	Image processing along with CNN and leap motion technology.	The unique feature of the proposed system is that it can interpret both static and dynamic signs using two separate machine learning models.The proposed system also accepts video feed as the input.

1.6. Research Gaps

Table 2 : Research Gaps

Citation	Theoretical gap	Performance gap	Empirical gap
(Dissanayake, Maheshi & Herath, H.C.M. & W.A.L.V.Kumari, &	Theoretically, this paper provides an advanced breakdown of the calculations	According to the paper, high levels of correlation were present in	Observationally, the system provided 100% successful

Senevirathne, W.A.P.B., 2013)	used to identify ratios and fingers. However, no neural networks were used in the final implementation.	the signed gesture and other gesture databases. Additionally, this system also takes 10 seconds on average to recognize a single gesture, which highly affects performance levels.	matches for only 11 out of 15 gestures. The testing and verification processes show that this system overall has a 93.3% success rate.
(S. Dilakshan and Y.H.P.P.Priyadarshan a, 2020)	A convolutional neural network is used in this research. They have utilised colour segmentation, gesture recognition techniques, convolution, pooling and feature extraction to gain the final output needed.	This system only utilises data for 12 basic gestures. Furthermore, this model has a good performance level because it doesn't use any motion-based recognition in its implementation, only identifying static gestures, also it mentions that no API service for video	Empirically, the Confusion matrix produced for this CNN model shows that the 3rd and 8th image classes have produced a miscalculation. This system provides an error rate of 10.1% which affects the accuracy of the overall system, only for 12

		streaming was provided.	gestures.
(L.L.D.K.Perera and S.G.V.S.Jayalal. 2021)	RGB is used to reinforce appearance. Preprocessing is finished by resizing representations and colour space change. HSV colour rooms are stronger than illuminations distinguished from RGB. Has also used SIFT descriptor and k-means clustering .	An average confirmation veracity of 86.5% was reached when the linked CNN-SIFT model was proven accompanying representations of 20 gestures of SSL. The most of the gestures accompanied accuracy, recall and F1 scores above 0.8 but any indication classes demonstrated reduced F1 principles. Out of the 20 sign classes, 13 displayed in addition 50% correct Sinhala quotation	Dynamic sign language gesture recognition and combining facial expressions of sign language in the gesture recognition process are not touched upon.

		predictions when the distance between help and camcorder were exchanged to 40cm and 60cm individually. An average veracity of only 68% was worked out when the SIFT physiognomy were not linked accompanying the CNN classifier	
(Kumarawadu, P. and Izzath, M. 2022)	The DNN model was executed in Python using Keras and TensorFlow open beginning athenaeums. The proposed DNN included individual input coating accompanying 23 input knots, two hidden tiers accompanying each 128 nodes and an	The system does not pay attention to the speed at which the hand signs are performed.	Has only used a limited Combination of general position of hands is not paid attention to.

	profit coating with 30 knots and provides an leading mishap of the calculations		
(S.D. Hettiarachchi and R.G.N. Meegama, 2020)	A convolutional neural network is used and a 2D convolutional layer is used for better validation. Softmax function is also used for the activation of output layers.	It was only able to identify 26 hand gestures using a convolutional neural network with 91.23% validation accuracy and 89.44% training accuracy.	In this project they limited the letters to 26 and when considering the Softmax activation function it is non-differentiabl e at zero and ReLU is unbounded.
(R.M. Rishan, S. Jayalal and T.K. Wijayasiriwardhane, 2022)	The image classification model utilised a CNN and an optimised ANN, while the Leap Motion method included an ANN.	The combined signs dataset consisted only of dynamic signs. Using a geometric template may affect the output timeline. validation	The research did not consider the output time and also the leap motion technology will not be able to see through the fingers that cause inaccurate values.
Kumar, D.M.,	A convolutional	Instead of	Large number of

Bavanraj, K., Thavananthan, S., Bastiansz, G.M.A.S., Harshanath, S.M.B. and Alosious, J., 2020)	neural network is used as the system uses it for skin-colour based modelling and to feed the images into the relevant model. Leap motion technology is used to classify dynamic sign identification.	showing the accuracy value of the system it shows values such as precision value and recall value for both static and dynamic hand gestures. It achieves high precision and recall value for both static and dynamic signs (over 0.90).	images and high dimensionality of the data can be a problem for having two classifiers (static and dynamic) at the same time in the respective system.
Dissanayake, I.S.M., Wickramanayake, P.J., Mudunkotuwa, M.A.S. and Fernando, P.W.N., 2020)	A region-based convolutional neural network is used in the system. In the image classifier the images will be sent through an application programming interface for other classification.	The system shows an accuracy of more than 80% for almost all the scenarios and the responding speed was in milliseconds. In this project the hands or the position of hands are not considered.	In the project the region-based convolutional neural network had poor performance while the live feed of the images. In case if they need to change or modify the existing code to add some extra features by using

			an application programming interface they have to go through so many files to make changes.

1.7. Contribution to the body of knowledge

1.7.1. Technological Contribution

Artificial Intelligence technologies, such as video recognition and Convolutional Neural Networks (CNNs), are utilised in our system..Video recognition is used to identify and enhance the captured video and use it in further processes. CNN algorithm is used in the process of classifying the signs.

1.7.2. Domain Contribution

- Our contribution provides a novel form of learning SSL in an online learning platform to the Sri Lankan community.

- This will be a much better method to learn SSL because users will not have to travel to and pay for physical sign language lessons. Users can save time and effort with ease by obtaining knowledge in this manner.

- All that is necessary is a functional webcam and a stable internet connection for the user to fully benefit from this online learning experience. This is especially useful for users in rural regions of the country where some cannot afford to pay for special education.

- Users are allowed to progress with lessons at their own pace and are also allowed to track their progress to see how far they have gotten with their study journey. They might also repeat activities in which they were unsure for additional practice, therefore, boosting the student's confidence in the lesson and what they have learnt so far.

- This website strives to be age-friendly and completely accessible to all sorts of SSL learners. The users can choose to learn in both English and Sinhala Languages as some users are more comfortable with one language than the other. The user will be provided with an option to choose their preferred language to use on the website.

1.8. Research Challenge

It is believed that the research we are conducting for this humane project will have a positive impact on the Sri Lankan community by raising awareness of the hearing-impaired population and, eventually, easing the communication barrier. Our study was thorough and concise, and it was mostly focused on the currently existing proposed SSL solutions. We intend to improve these systems and provide a more user-friendly manner of learning sign language for people of all ages and levels of experience. In order to provide a robust and well-supported foundation for our project, it is crucial that we engage in thorough research and data collection efforts. This requires frequent examination and evaluation of existing literature reviews and data sources. Additionally, it is imperative that we stay informed of current advancements in knowledge management and technology to ensure the originality and high quality of our work. Despite limited prior research on SSL, we faced the challenge of unavailable datasets. As a result, we have taken it upon ourselves to create the necessary datasets for our project.

1.9. Research Questions

RQ1: How would using a CNN architecture elevate the performance of the system?

RQ2: How do spoken languages influence sign languages and How do sign languages influence spoken languages?

RQ3: How can sign language be used for deaf people with moderate intellectual disabilities?

RQ4: Which CNN architecture is best for the proposed system and how can we increase the efficiency of that model?

1.10. Research Aim

The objective of our research is to design and develop an Interactive Sign Language Platform that facilitates the learning and acquisition of SSL in a user-friendly manner. The platform will enable users to learn basic SSL gestures, test their comprehension through a guessing game, and expand their SSL vocabulary. The camera feature will allow users to practise the gestures they have learned and receive feedback on their accuracy. The platform will track dynamic signs using image processing, video recognition, and pattern recognition technologies, including a CNN model for sign language classification and OpenCV architecture for capturing user input. Our aim is to enhance communication between hearing-impaired individuals and the broader society and bridge the language barrier through the development of this innovative and comprehensive platform.

1.11. Research Objective

Table 3 : Research Objectives

Research Objectives	Explanation	Learning Outcomes
Problem Identification	To identify the real-world issues in Sri Lanka and	LO1

	narrow down the scope to one topic	
Literature Review	RO1: Figure out existing Sinhala Sign Language Systems RO2: To build and finalise a vocabulary of signs and words to work with in out system RO3: To determine how the final output in the website will look like to have an idea about the design layout	LO1
Data Gathering and Analysis	To visit deaf community institutions to gather data from repositories and validate the needed data	LO2, LO3
Research Design	Gather information regarding the layout of the final design of the learning platform in comparison to the other available platforms in use.	LO3, LO4
Implementation	Developing the model and training with the necessary data, making edits to the system. Developing the front-end and back-end to lace it together as one functioning system.	LO2, LO3, LO4
Testing and Evaluation	Using test data apart from the training data to see whether the needed results are achieved, and using metrics to evaluate and rank the results.	LO4

1.12. Project Scope

1.12.1. In - Scope

Table 4 : In - Scope

No	Description
1	Identify signs that the user does using image processing and AI.
2	Obtain relevant descriptions, pictures and a video of how to do the sign.
3	Creating a web application tool which will be user friendly when learning sign language.
4	Contains the basic categories of vocabulary for the beginner users.
5	Updatable with new vocabulary and signs; the system will detect it and inform the user of it.
6	Make sure, for images or videos with incorrect specifications outlined under the user manual, it gives an error by specifying the reason.

1.12.2. Out - Scope

Table 5 : Out - Scope

No	Description
1	Identifying only small phrases.
2	The user will receive a responsive web application.
3	Not introducing the alphabet to the user.
4	Used a limited set of vocabulary which will give a basic knowledge about SSL

1.12.3. Prototype Diagram

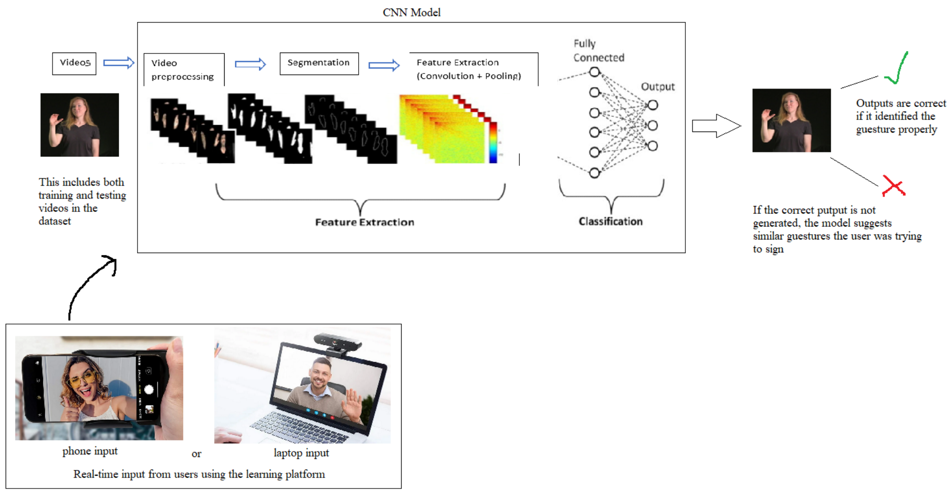


Figure 1 : Prototype Diagram

1.13. Resource Requirements

1.13.1. Hardware requirements

● Intel Core i5 10th generation processor or high - To be able to perform training the model and other high-priority tasks.

● 8GB RAM or high - To image processing the images and pre-process the images stored.

1.13.2. Software requirements

● Python – Main language to build the proposed system.

● Figma – To prototype the model architecture.

● Google Drive – To store data sets and other documents which were used in the process.

● MS Word – To write documents and drafts.

● Node js - we used node js to design the website.

● TensorFlow - Image pre-processing and training of the model are done using this.

- Keras - For image pre-processing and training the model of the system.
- Windows operating system – We used this to handle computational functionalities.
- Draw.io/star UML – To design wireframes for the system.

1.13.3. Data Requirements

- The amount of dataset available for this project is limited so that we have to create our own dataset and that has to be validated.

1.13.4. Skill Requirement

- Programming skills
- Design skills
- Mathematical application skills
- Time management skills
- Teamwork skills

1.14. Chapter Summary

Communication between hearing-impaired people and non-hearing-impaired people is narrowing because of the mode of communication. It is critical for non-hearing impaired people to learn sign languages in order to bridge the communication gap between these two parties. So a SSL learning platform would help them learn sign languages. The platform should be easy to use, convenient, and, most importantly, have an effective system to learn sign language.

CHAPTER 02 : LITERATURE REVIEW

2.1. Introduction

According to the World Health Organisation (WHO) as of 2022, 430 million people around the world are suffering from deafness. Nearly 20% of the world's population, or more than 1.5 billion individuals, live with hearing loss (World Health Organization, 2022). Additionally, there are more than 389,000 people in Sri Lanka who unfortunately experience the same issue. In Sri Lanka, more than 74% of the population speaks Sinhala as their native language. Despite the fact that there are schools dedicated to teaching deaf people, there isn't an adequate online learning environment set up for learning Sinhala sign language. Our goal is to develop an effective method of aiding the deaf and mute community by bringing a fresh approach to learning sign language. Our project is broken down into 4 main components that we will focus on introducing machine learning to enhance this website. These are 1) Image processing and Skin colour modelling, 2) Gesture Recognition, 3) Convolution and Classification and 4) A helpful chatbot.

2.2. Relevant Work

2.2.1. Image Processing and Skin Modelling

There are various methods of skin colour modelling and image processing used in the existing published papers on Sinhala Sign Language (SSL) systems. In an article about Image Processing used for Sinhala Sign Language (Dissanayake et al., 2013) , they have used Image processing technology to generate the image to be processed. They use a green background for the dynamic video input, break it down into frames and create a binary image of the hand by adjusting the RGB matrix of the skin colour. This allowed their model to have faster mapping but the accuracy of the binary image highly depended on the lighting conditions of the image captured (Perera & Jayalal, 2021). The article on using CNN architecture for SSL (Dilakshan & Priyadarshana, 2020) , used a very similar approach where they converted the RGB image to grayscale and then to a binary image before being fed to the gesture database. With this vision-based approach come challenges such as illumination change, background clutter, etc (Kumarawadu &

Izzath, 2022). To overcome some of the lighting issues, the Gaussian Mixture Model (GMM) is used.

In a system called EasyTalk (Kumar et al., 2020) , the abstract sign image serves as the input for the image classifier.The API machine learning model is used and it is created using CNN. It processes the image and classifies it under different categories.This component was developed utilising a Faster (Region-Based Convolutional Neural Networks) RCNN setup on top of the TensorFlow model trainer. In another similar system (Dissanayake et al., 2020) First, the background is removed. In this first the images are converted into grayscale images and thresholding is done to perform edge detection. Dilation and erosion methods were used to make the edges sharp. Next, the largest contour was taken as the person and the rest was taken as the image background. So the background is not focused and the accuracy is increased.

Sign Language gesture recognition follows stages of data acquisition, data preprocessing, segmentation, feature extraction, and classification of hand gestures (Perera & Jayalal, 2021). Gesture recognition can be achieved by using (Kumarawadu & Izzath, 2022) either vision-based approaches (camera images/videos) or sensor-based approaches.

Image preprocessing is done by resizing the images to 48x48 size, converting colour space from RGB to HSV and applying a skin colour mask to separate the hand region from the background. The HSV colour ranges used are HSV_min (0, 40, 30) and HSV_max (43, 255, 255)(Perera & Jayalal, 2021).

Research on SSL recognition and translation has been carried out focusing on (Kumarawadu & Izzath, 2022) appearance-based approaches, skeletal based approaches and 3D module based approaches. Research on image processing based SSL recognition using skin colour filtering and centroid finding approach for the (Perera & Jayalal, 2021) development of a still gesture mapping prototype has been conducted considering 15 gestures of the SSL alphabet.

Hand tracking devices with arrays of sensors have been used recently in sign language recognition research. After collecting hand gestures (Kumarawadu & Izzath, 2022). A depth sensor based real time hand-pose estimation framework has been proposed to recognize the first ten digits in American sign language.

In the article (Hettiarachchi & Meegama, 2020), they have only considered 26 letters which have static hand gestures having green as the background colour. A database of hand gestures is created and those digital images are processed. The images are taken under identical parameters such as background colour and the same side of the hand (Perera & Jayalal, 2021). When the user shows a sign from the hand to the web camera window in the computer, it proceeds 200 frames and the final frame will be captured to be used for further tasks. The selected images have a width and height as per a selected scaling factor. The image classification model utilised a CNN and an optimised ANN and Leap Motion method in gesture recognition (Rishan et al., 2022).

2.2.2. Gesture Recognition

The article concentrating on the image processing-based system (Dissanayake et al., 2020) employs a special technique to monitor the user's motions. In this instance, they allow the user to wear a wristband to distinguish the motion made in isolation from the complete hand. After it is complete, a contour map is created around the hand that has been filtered using a set of equations to position the hand and fingers (Kumarawadu & Izzath, 2022).

Next, the CNN-based article (Dilakshan & Priyadarshana, 2020) discusses using a straightforward gesture recognition pipeline. Starting with sensor, Feature extraction, classification and finally obtaining the gesture class label. It is said that the gesture recognition system may be divided into two categories: a three-dimensional (3D) hand model-based method and an appearance-based system. The 3D method is built using the ResNet architecture and depth-based sensor devices.

The gesture recognition in the system EasyTalk (Kumar et al., 2020),Tensor flow model was used along with RCNN. For the accelerated training and storage efficiency, the photos were captured at a resolution of 800 x 600 pixels. Laptop webcams with lower resolution were used to capture the images. The detected images were divided into training and test datasets for a successful model. Next in the system UTalk (Dissanayake et al., 2020), the features were detected statically and dynamically. As micro movements cannot be detected easily, and also it is not a part of a particular gesture it has to be neglected. This problem was solved with a more extended time duration in the required frames.

SL gesture recognition follows stages of data acquisition, data preprocessing, segmentation, feature extraction, and classification of hand gestures (Perera & Jayalal, 2021). Gesture recognition can be achieved by using either vision-based approaches (camera images/videos) or sensor-based approaches

Authors have used a hand model with a hierarchical skeleton for sign recognition and a model using artificial neural networks support vector machine has been utilised as a pose classifier (Kumarawadu & Izzath, 2022). Kinect sensors with high performance 3D image capturing have been used for a gesture recognition system focusing on fingertip position. A glove based data acquisition technique has also been used as another approach wherein the position and orientation of hands were recognized using a special glove equipped with several sensors worn by a user (Kumarawadu & Izzath, 2022).

A database of hand gestures was created in the article and those digital images were recognized and classified by the Convolutional neural network (CNN). Then the authors identify the most appropriate design and then they implement a platform to develop the system to translate the gesture into Sinhalese (Hettiarachchi & Meegama, 2020).

Vision-based gesture detection and recognition systems are cheaper than sensor-based gesture recognition systems (Rishan et al., 2022). However, background effects, changes in light intensity, computational time against resolution and frame rate and background objects with similar skin colours or otherwise the hands will be a challenge for vision-based approaches.

2.2.3. Convolution

For the article that works on the Image Based processing system (Dissanayake et al., 2020), a convolutional approach is not taken on by the respective researchers. Instead, they segment the hand and crop the image using a series of equations to more accurately detect the user's gesture.

Contrarily, the publication that discusses CNN architecture (Dilakshan & Priyadarshana, 2020) utilised convolutional neural networks (CNN), as the name implies, is used in the field of image classification. The CNN architecture consists of hidden layers that can learn features by repeatedly performing three actions. These were termed "Convolutional Blocks" by the authors (Perera & Jayalal, 2021). Each layer of the convolution process uses matrices, and then a pooling mechanism is used to produce dense matrices.

In the system EasyTalk (Kumar et al., 2020),The Convolutional Neural Networks (CNN) is used in the process of building the machine learning API which gets the abstracted sign image as the input. When taking an image as an input those individual images are converted into Comma Separated Value files. The model is started to train using the TensorFlow model trainer and Faster RCNN Configuration utilising the CSV files. Since python is used in building the EasyTalk (Kumar et al., 2020) system's backend implementation, the *keras* library is used to build the CNN model. In the system UTalk (Dissanayake et al., 2020) the CNN is used to develop the static and dynamic sign classifier. The creators have used max pooling and some of the layers for the model building process. The trained data set was transformed into a numpy array, as such it can translate a 3D array into a 1D array.

The methodology consists of the major stages of data acquisition, image preprocessing, feature extraction, classification and displaying Sinhala text. The classification model consists of two input channels, one from CNN and the other from the SIFT (Perera & Jayalal, 2021). The final fully connected layer will concatenate both feature vectors from SIFT and CNN layers to generate the final output of the gesture recognition model.This comprises three major components: builder, interpreter and classifier (Perera & Jayalal, 2021). The system architecture was all managed and executed with the help of the leap service of LMC, which was connected with the Interpreter and the Builder through the API service of LMC (Kumarawadu & Izzath, 2022). The gestures were recognized using the builder application and they were saved in a dataset. The interpreter was responsible for recognizing a sign from the dataset and displaying the text output of the sign. The Builder created the Sinhala hand signs, trained the data, and saved the gestures in the dataset (Kumarawadu & Izzath, 2022). A utility application was used to help the classifier extract features from hand gestures. A classifier was used to train the dataset value and to normalise hand gesture data received from the leap motion controller using its leap service (Kumarawadu & Izzath, 2022). The dataset was the storage and this was used to store the recognized values from the leap motion controller (Kumarawadu & Izzath, 2022). This stored the raw data representing hand gesture values of the palm and fingertips in a numeric way.

The Leap Motion controller is a commercially available, inexpensive sensor for recognizing hand and finger motions, including bones and joints in a 3D interactive zone. The Leap Motion Controller can capture a user's hands and fingers by being surrounded by cameras with infrared light. Similar to the sign training model, a function is used to receive the performing sign recordings. Then, using another function, the data is recorded and the process begins (Rishan et al., 2022).

A 2D convolutional layer was used by them as it provides better validation accuracy than 3D convolutions (Hettiarachchi & Meegama, 2020). The main task of the convolution stage is to derive high-level features such as edges and quality from an input image. After inserting an image with 3 colours into the convolutional layer, it produces a 3 coloured image.

2.2.4. Classification

The "error in ratios" and the "error in fingers" were used to determine the "highest mark" in the image-processing system (Dissanayake et al., 2020). Each gesture in the sequence receives a determined score. The gesture that received the highest score is ultimately picked as the input gesture's best match.

The main aspect of the classification utilised is also part of the CNN architecture for the CNN-based article (Dilakshan & Priyadarshana, 2020). A callback list of matched motions is found using the final CNN architecture. A histogram will be generated and used for the segmentation process of the system. Each database includes categorised files for the Unicode characters that are used to correspond with a Tamil or Sinhala letter.

When considering the classification in the system (Kumar et al., 2020),the CNN classifier goes through a number of phases with an image. First the image passes through a series of convolutional layers. Then the pooling process takes place. Then, several convolutional layers are added, and pooling is continued until the desired level of filtration is achieved.

In the system (Dissanayake et al., 2020) , Static sign classification plays a major role in this system as at least one static sign supports getting a meaningful sentence.CNN classification technique is used to create the dynamic sign classification model. CNN classifies dynamic signs using input and output layers as well as numerous hidden layers (Perera & Jayalal, 2021). And also image entropy is used to decide which pixels spread constantly through each frame and which regions are highlighted by the appearance of the frame. With the value of the image of entropy it is decided whether it is a dynamic frame or a static frame (Kumarawadu & Izzath, 2022). Here VGG base 16 model is used for video classification.

The CNN consists of three convolution layers, three max-pooling layers followed by a dense layer. The combined feature map from both channels is used in the fully connected layer of CNN as an enhanced set of features for scale variations (Kumarawadu & Izzath, 2022). The concatenated output at the final dense layer is given as the output of the model. The training dataset is fed to the gesture classification model as input in the CNN channel and the SIFT feature vectors fed as the input for the SIFT channel to train the model (Perera & Jayalal, 2021).

In this study, the proposed DNN model was compared with a Naïve Bays model and a multi-class SVM model for prediction of Sinhala Sign Language (Kumarawadu & Izzath, 2022). The DNN model was implemented in Python using Keras and TensorFlow open source libraries. The proposed DNN consisted of one input layer with 23 input nodes, two hidden layers with each 128 nodes and an output layer with 30 nodes. The complete application consisted of three major system components: Interpreter, Builder and Classifier. Three classifiers which were experimented in study were (Kumarawadu & Izzath, 2022) DNN classifier, Naïve Bayes Classifier and SVM based classifier which were trained and tested to classify a number pre-identified Sinhala signs and words which have been pre-identified for feature extraction.

To classify the dataset in this project, an artificial neural network was added to the CNN (Hettiarachchi & Meegama, 2020). Basically, a fully attached layer looks at what high level features most firmly correspond to a particular division to produce an

effective output. Number of units was used, which is the number of nodes that should be present in a hidden layer to achieve non-linearity in the fully connected layer.
The Leap Motion and the Leap Trainer framework are also used in the sign recognition model (Kumarawadu & Izzath, 2022). In the article the Leap Motion controller-based sign language recognition, proposed a combined approach along with image and Leap Motion data classification for the identification of gestures (Rishan et al., 2022).

2.3. FAQ/ Progress Checker

The FAQ system and Progress checker are two novel systems we tend to introduce to our website implementation. The FAQ system will help the user navigate around the system with ease and help them clear their doubts in a much easier way than waiting for a response from the support site. The innovative progress checker introduces a new way of keeping track of the work being done by the user, by showing encouraging messages and milestones achieved by the user, the learner can know where they stand in their education, how much they have to do and how much they have already done.

2.4. GAP Analysis for Relevant Work

Table 6 : GAP Analysis for Relevant work

Research	Author(s)	Year	Dataset	Model Used	Metric
Image Processing Based Sinhala Sign Language Recording System	Dissanayake, Maheshi & Herath, H.C.M. & W.A.L.V.Kumari, & Senevirathne, W.A.P.B.	2013	User's own dataset of 15 different gestures	Image processing and Still gesturing techniques (No specific model used)	Accuracy: 93.3%
Convolutional Neural Networks: A Novel Approach for Sinhala	S. Dilakshan and Y. H. P. P. Priyadarshana	2020	Sign Language MNIST	CNN	Accuracy: 89.9%

Sign Recognition System			Kaggle		
EasyTalk: A Translator for Sri Lankan Sign Language using Machine Learning and Artificial Intelligence	Kumar, D. & Kugarajah, Bavanraj & Thavananthan, S. & Bastiansz, G.M.A.S. & Harshanath, S.M.B. & Alosious, J	2020		CNN, RCNN	Accuracy: 97%
Utalk: Sri Lankan Sign Language Converter Mobile App using Image Processing and Machine Learning	Dissanayake, I.S.M., Wickramanayake, P.J., Mudunkotuwa, M.A.S. and Fernando, P.W.N.	2020	User's own dataset, 27 images and video segments	CNN	Accuracy: 90%
Translation of Sri Lankan Sign Language to Sinhala Text: A Leap Motion Technology-based Approach	Rishan, R.M., Jayalal, S. and Wijayasiriwardhane.	2022	Github Groundviews	CNN Leap Motion	Accuracy: 91.82%
Machine Learning Approach for Real Time Translation of Sinhala Sign Language	Hettiarachchi, S. and Meegama, R.	2020	Sliit deskspace, arvix.org	CNN	Accuracy: 91.23%

into Text					
Sri Lankan Sign Language to Sinhala Text using Convolutional Neural Network Combined with Scale Invariant Feature Transform (SIFT)	L.L.D.K.Perera and S.G.V.S.Jayalal.	2021	statistics by United Nations Economic and Social Commission for Asia and the Pacific in 2019	CNN	Accuracy:
Sinhala Sign Language Recognition using Leap Motion and Deep Learning.	Kumarawadu, P. and Izzath, M.	2022	IEEE, International Conference on Cyber-Enabled Distributed	CNN ANN DNN Leap Motion	Accuracy: 90%

2.5. Summary

Our project is an Interactive Sign Language Platform where users can learn some basic Sinhala sign language gestures. By using this technology, we promote interaction between hearing-impaired people and ordinary people. We will construct a system that recognizes hand motions using an image processing system. Additionally, a FAQ system that assists in discovering website information will be developed. In this approach, real-world knowledge has been recorded using a CNN.

CHAPTER 03 : METHODOLOGY

3.1. Chapter Overview

The methods utilised in research, project management, development, and evaluation are covered in this chapter. Here, you can find a number of different processes, appropriate theoretical and methodological management, prioritised designed principles and techniques for solving problems and anticipated project implementation plans.

3.2. Research Methodology

Table 7 : Research Methodology

Research Philosophy	The research philosophy for the system follows pragmatism and an understanding of the linguistic. This methodology involves understanding the nature of sign language, choosing a theoretical framework, data collection, algorithm selection, evaluation metrics and user-centred design.
Research Approach	The study that will be used is deductive because it is founded on established methodologies, also the system needs to be trained using videos, and the performance needs to be evaluated using various metrics.
Research Strategy	The project will follow a grounded theory research strategy, because patterns are derived from the data as a precondition for the study.
Research Choice	The usage of multi methods will be chosen because a wider selection of methods is used and they all fall within the same field of study.
Time zone	The research is cross-sectional, because the time horizon is already established, whereby the data must be collected at a certain point.

3.3. Development Methodology

3.3.1. What is the life cycle model and why?

Our project centres on incorporating the Agile methodology into the project life cycle development process. By utilising the Agile method, we aim to effectively prioritise the crucial elements of the project, resulting in the optimal utilisation of available time for project management. The frequent, short meetings enabled by the Agile approach allow for regular updates to be shared among team members and facilitate the reporting of weekly progress to the supervisor. This enhances the efficiency of the feedback process and enables the seamless integration of any necessary modifications.

3.3.2. Design methodology

Given our adoption of the Agile methodology, we have determined that the Object-Oriented Analysis and Design (OOAD) method is the most appropriate approach for our project's design methodology. The central goal of OOAD is to translate functional requirements into implementable solutions using multiple programming languages, making it an ideal fit for our project needs. While the Structured Systems Analysis and Design Methodology (SSADM) has its advantages, particularly in the application of the waterfall technique in its life cycle development, the rigidity of SSADM can hinder the ability to adapt to changes in requirements, a common occurrence in Agile projects. In contrast, the flexibility of OOAD makes it more suitable for complex projects with many objects and interactions, as opposed to SSADM's suitability for smaller, less complex projects. According to Kendall and Kendall (2005), "Object-oriented analysis and design (OOAD) is more suited to large, complex systems because it emphasises time management through modular design" (Garlan and Shaw, 1993)

3.3.3. Evaluation Methodology => Evaluation metrics and / or benchmarking

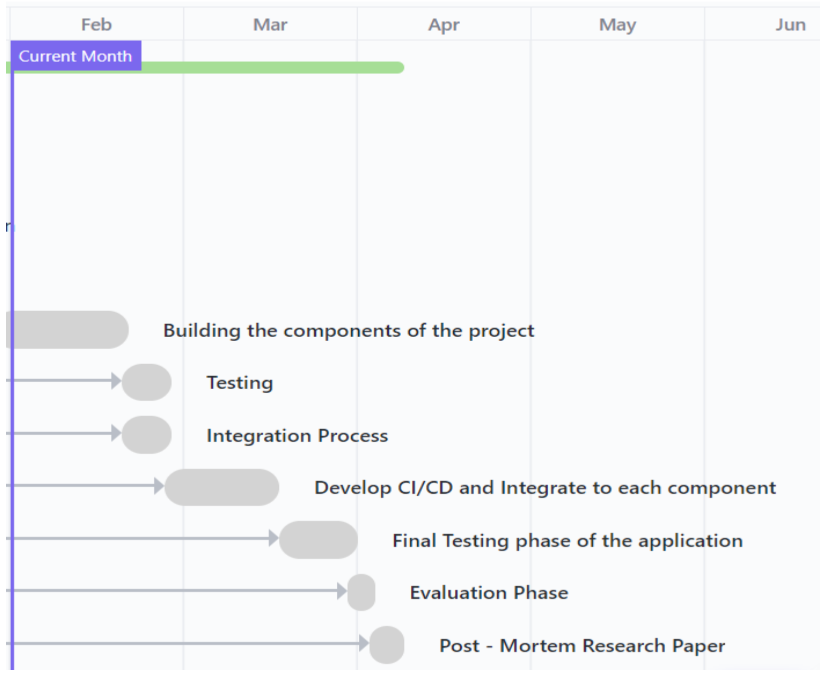
- We intend to evaluate this section using both metrics and benchmarking methods.
- Metrics could offer reliable numerical information on how this project was developed and benchmarking will evaluate the precision of results using

important business KPIs to determine whether this software outperforms the market's rivals.

3.4. Project Management Methodology

3.4.1. Schedule using the Gantt chart after doing a WBS

Figure 2 : Gantt Chart



3.4.2. Deliverables, milestones and dates of deliverables

Table 8 : Deliverables

Deliverables / Milestone	Due Date
Literature Review	Week 03
Project Proposal	Week 05
Software Requirement Specification (SRS)	Week 10
Learning the Tech stack	Week 10

Building the components of the project	Week 14
Testing	Week 17
Integration Process	Week 17
Develop CI /CD Pipelines and Integrate to each component	Week 19
Final Testing Phase of the application	Week 22
Evaluation Phase	Week 23
Post - Mortem Research Paper	Week 24

3.5. Chapter Summary

Making conversations easier between non-hearing impaired and hearing impaired people in Sri Lanka is still at the research level and is currently being developed. Therefore, we have designed and started developing a Sinhala sign language learning platform to make communication easier using modern technologies, concepts, and an integrated methodology. This is a very important milestone that we have reached, and the above-stated methodology proves the product's success in a pre-production view.

CHAPTER 04 : SOFTWARE REQUIREMENT SPECIFICATION

4.1. Chapter Overview

This chapter provides a detailed analysis on the system requirements for SignIt. The stakeholders of the system are first identified and their responsibilities are outlined. The benefits and drawbacks of various techniques for requirement gathering are discussed and reviewed. The use case diagram and its definitions are included during the requirement analysis stage. Finally, a scope definition is used to specify the system's functional and non-functional requirements, which are then prioritised in relation to the function.

4.2. Rich Picture

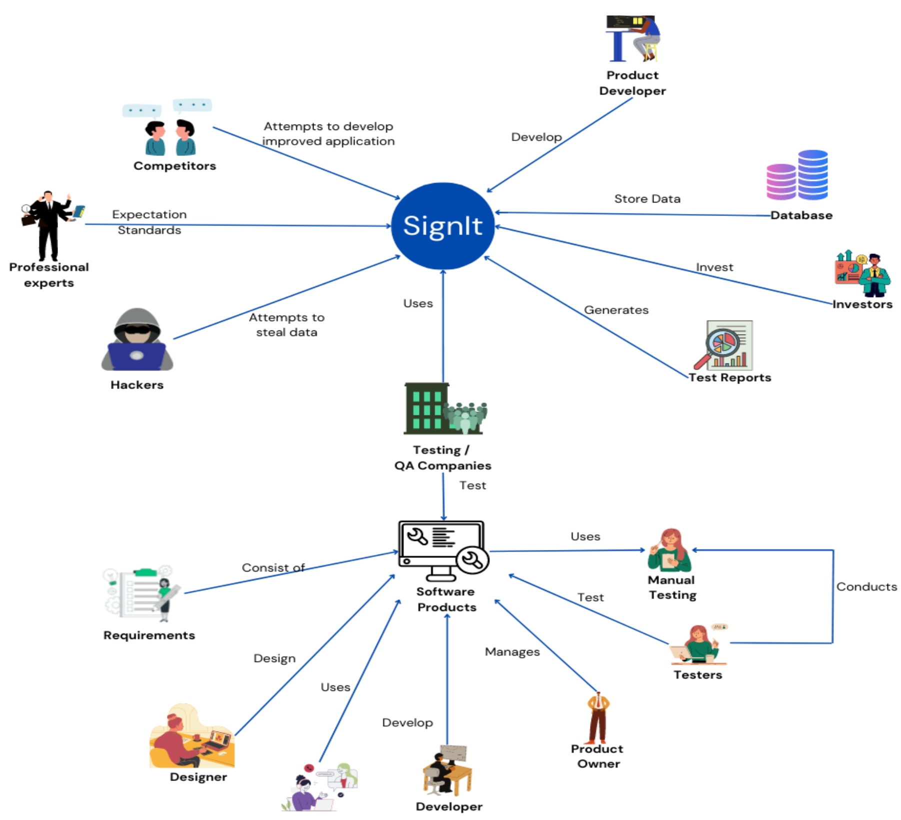


Figure 3 : Rich picture

4.3. Stakeholder Analysis

The system's established stakeholders are shown in the onion diagrams, along with a system overview and each stakeholder's role within it.

4.3.1. Onion Model

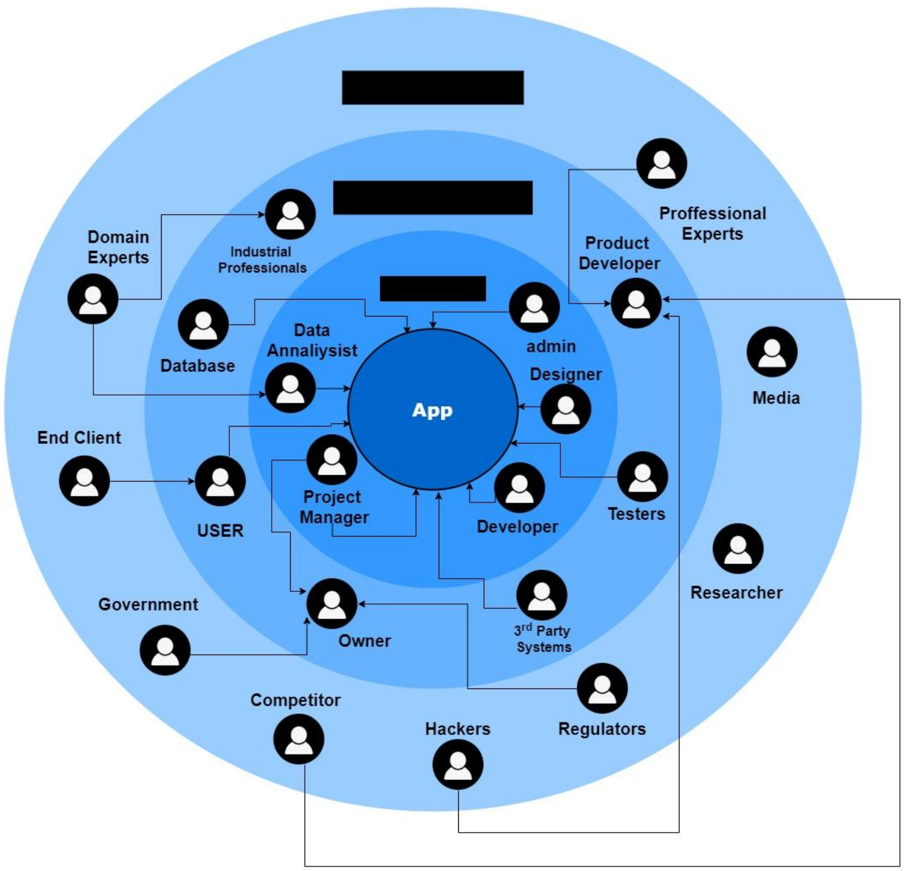


Figure 4 : Onion model

4.3.2. Stakeholder Viewpoints

Table 9 : Stakeholder Analysis

Stakeholder	**Role**	**Benefits**
System Admin	Functional Administration	Utilise the application and set up the essential environment.
Data scientist, ML Engineers, Software Engineers	Functional Maintenance	Deploy the task and design, develop the model.
Users, 3rd party systems	Functional Beneficiary	Deploy the created application and integrate it with other systems, programs.
Product owner	Functional Beneficiary	Owner of the proposed system or program.
Product manager	Management Support	Managing and developing the process of application to ensure the effectiveness of the project.
Regulator	Quality regulator	Confirm that the application fits into its own individuality considering all the private policies.
Product developers	Functional Maintenance	They are in control of designing, maintaining and updating the system regularly.
NLP/NL experts	Expert	Gives an idea of the technology and methodologies to be used.

Domain experts	Expert	Give a domain perspective about the technology and methodologies used.
Technical writer	Functional Support	Support in the documentation process.
Researcher	Educational Support	Explore the current domain and provide ideas to improve the proposed process and techniques.
Competitor	Negative Stakeholder	Competes directly with the proposed system which could lead us to come up with the most innovative and fresh ideas.
Hacker	Negative Stakeholder	Unauthorised access can lead to problems which are sorted by difficulty and numbered by variety.

4.4. Selection of Requirement Elicitation Techniques / Methods

In order to elicit and analyse needs, three basic approaches were used to collect them. As far as we self brainstorm, these techniques helped us identify the elements essential for our learning platform and uncover new features that weren't included in other systems.

4.4.1. Literature Review and Gap Analysis of Existing Systems

We began the information-gathering process for this phase by examining the characteristics of existing systems and the implementations that are presently in place. This demonstrated the platform's suggestions for improvement and provided a way to contrast the new features of our system with those of other systems.

Advantages	Disadvantages
• Main features and functions that should be present in the learning platform were identified. • GAP analysis in the literature review showed areas in which this learning platform could be improved from other models.	• As there were no systems discovered that were designed for the instruction of Sinhala sign language, not all current systems that were studied were learning platforms, just Sign languages to text translators. This indicates how the objectives of the two systems that were compared are different.

4.4.2. Questionnaire

To represent a wider audience of the general public, this was done using a google sheet questionnaire. In this manner, we could also learn about their requirements and demands. The sample population was asked questions and given the opportunity to offer their own suggestions for this learning platform. The sample composition considered for the questionnaire was the general population. Anyone of eligible age was taken as a candidate for a person who'd be interested in a sign language learning system. More details on the general sample composition is discussed in the discussion of results.

Advantages	Disadvantages
• This reaches a larger population of individuals. • Google forms provides tools to analyse results in bar charts and pie charts and	• Some questions would have been misunderstood • Some would have answered in a manner that could negatively affect

other useful visualisation tools. ● Anyone who has knowledge of using a smart device could fill out the form, with no extra effort needed. ● Only takes a few minutes to complete the survey.	our requirements-gathering process. ● The older population may or may not have known how to respond to a Google Form, so that reduces the target population for gathering information.

4.4.3. Interviews

Interviews can be performed with individuals who have a connection to this area of study or who have dealt with the subject topic in question professionally. We conducted an interview with a person who had experience dealing with sign language users and who thought our approach may be useful in the corporate sector. Therefore the sample population considered for this interview was a domain expert in the field of business and sign language.

Advantages	**Disadvantages**
● Questions were asked directly and with interviews, we receive a more direct and personal answer, than in surveys. ● The person being interviewed has more knowledge than us and can therefore give more professional suggestions for the platform. ● Interviews allow the interviewee to speak freely and provide an answer they wish to give rather	● Interviews are time confusing and we have to respect the times when they are free to conduct the interviews, as opposed to our own schedules. ● We may receive complicated answers which weren't the answers we were hoping to receive. ● The questions might puzzle the interviewee and would have had

than short-set and rigid answers.	to move on to another question instead. ● Only a limited number of interviews can take place as they are time-consuming, as mentioned earlier.

The primary functions required for this learning platform were defined in the initial research done for present systems. The GAP analysis revealed additional properties that have not been included in the current systems, indicating innovative functionalities to be explored in the learning platform.

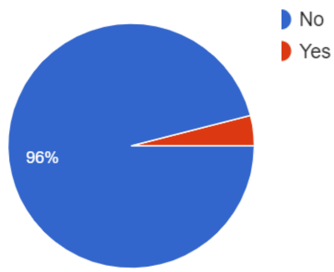
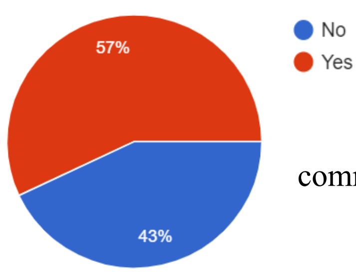
We found that both findings of the survey and interview are beneficial to us, therefore we used the survey results to evaluate the requirements of the general public and the interview questions to recognize the professional components of the system.

4.5. Discussion of Results

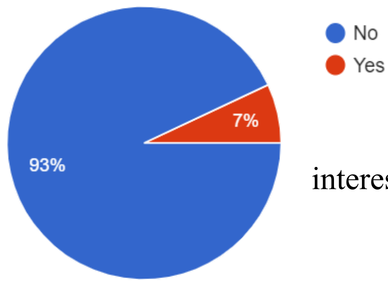
4.5.1. Questionnaire Discussion

The questionnaire's first four questions were designed to elicit basic information about the respondents. Such as the email address, name, age and the industry they work in. Later, the questions were tailored to our needs.

Question	Do you suffer from hearing loss?
Aim of the question	to count the number of people who suffer from hearing loss
Observation	

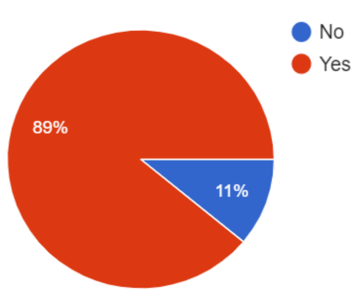
	It has been discovered that 4% of the participants suffer from hearing loss. While remaining 96% of the participants do not have any hearing losses.
Conclusion	
According the responses given by the participants that the number of people suffering from hearing loss appears to be low in Sri Lanka	
Question	Have you ever faced difficulties when communicating with a hearing-impaired individual?
Aim of the question	Finding out if people have any difficulties communicating with hearing impaired people.
Observation	
	We can see that 57% of respondents have had no difficulty communicating with hearing impaired people. While 43% responded that they were having troubles when communicating with people who are deaf.
Conclusion	
We can conclude that more than half of the population has no difficulty communicating with deaf people. It could also be because they have not come into interaction with any impaired individuals. We could also assume that hearing-impaired people do not expose themselves to society a lot.	

Question	If yes, share your experience
Aim of the question	To understand the problems they face when communicating with hearing-individuals
Observation	
The majority of those who responded said they couldn't understand what they were saying. Some of them stated that there was some misunderstanding as a result of improper communication with hearing-impaired people. The majority of the participants expressed regret over their inability to communicate effectively with them.	
Conclusion	
It is clear that communication between hearing-impaired and non-hearing-impaired people is ineffective. When deaf people use sign language to communicate and non hearing-impaired people use normal language, both parties are unable to understand what the other person is saying and some misunderstandings may occur. Because of this type of misunderstanding, both parties rarely communicate with one another.	
Question	Would you like to reduce the communication gap between people using spoken language and hearing-impaired individuals?
Aim of the question	To know whether both normal and deaf people like to communicate with each other
Observation	



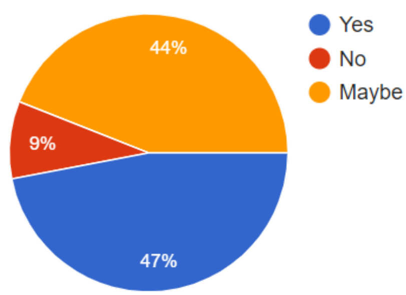
93% of the participants were willing to reduce the gap between people using spoken language and hearing-impaired individuals while only 7% were interested in reducing that gap.

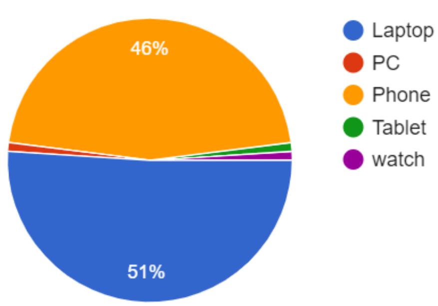
Conclusion	
Even though a small number of people oppose the idea of reducing communication, the majority of people clearly prefer to reduce the communication gap between hearing-impaired and normal language speakers. It is clear that language is the primary barrier for both parties.	
Question	Have you ever dealt with Sinhala Sign Language?
Aim of the question	Determining the number of people who have been exposed to Sinhala sign language
Observation	



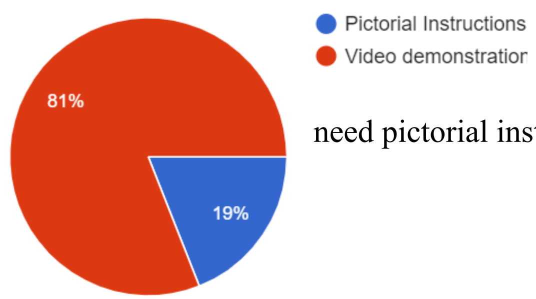
Only 11% of the respondents have not dealt with Sinhala sign language. But we can see 89% of them have dealt with SSL.

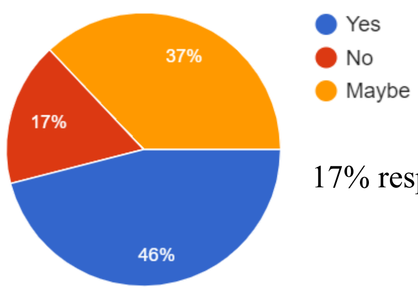
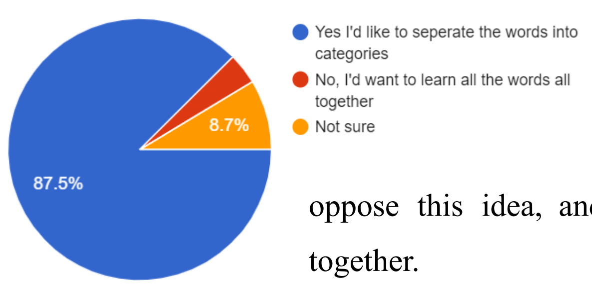
Conclusion	
According to this survey, more than three-quarters of people have dealt with SSL. As a result, we can say that SSL is not a foreign language to Sri Lankans. In some way Sri Lankans are aware about the use of SSL and exposed to the use of it.	

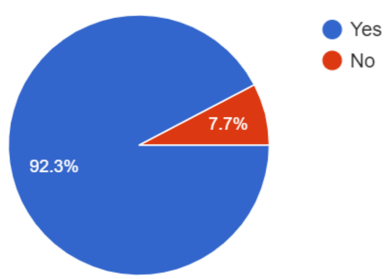
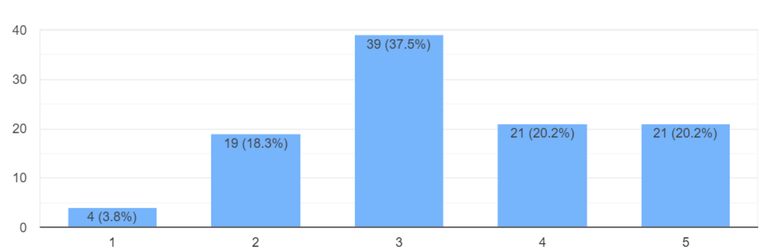
<table>
<tr><td colspan="2"></td></tr>
<tr><td>Question</td><td>Would you like to learn Sinhala Sign Language using a website?</td></tr>
<tr><td>Aim of the question</td><td>To know whether people are interested in learning SSL</td></tr>
<tr><td colspan="2">Observation</td></tr>
<tr><td>


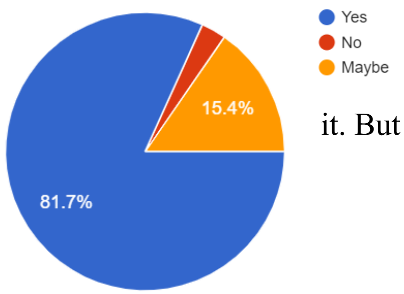
</td><td>47% of respondents said they would like to learn SSL, while 47% said they were unsure. Only 9% of respondents said they disliked learning SSL.</td></tr>
<tr><td colspan="2">Conclusion</td></tr>
<tr><td colspan="2">It can be seen that the majority of people want to learn SSL, which will aid in closing the communication gap between hearing-impaired and non-hearing-impaired people. Only a small number of people are uncertain whether they like to learn SSL or not.</td></tr>
<tr><td colspan="2"></td></tr>
<tr><td>Question</td><td>Which device would you most like to use for the learning platform website?</td></tr>
<tr><td>Aim of the question</td><td>Understanding the user's needs so that we can design our website accordingly.</td></tr>
<tr><td colspan="2">Observation</td></tr>
<tr><td>


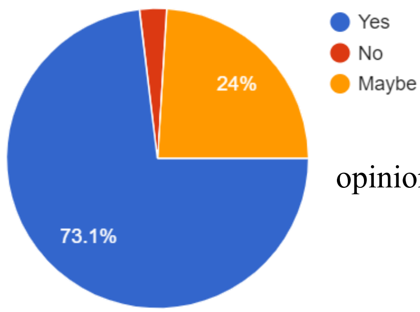
</td><td>It is observed that 46% of the participants like to learn using phones and 51% like to learn using the laptop. Only 1% of them like to learn using PC, Tablet and watch.</td></tr>
</table>

Conclusion	
We can see that the majority of users prefer to learn using their phones because they are very convenient to use. Using a laptop also appears to be more effective than using a PC, tablet, or watch.	
Question	Would you rather learn how to sign words from pictorial instructions or from a short video clip of the signed word?
Aim of the question	This is also questioned in order to gain a better understanding of the user's requirements.
Observation	
	81% of the participants are willing to learn using a video demonstrator while only 19% need pictorial instructions to learn SSL.
Conclusion	
Most of the users prefer video demonstrations rather than pictorial instructions. Since learning using video is more reliable, the majority of the users are likely to learn SSL using video.	
Question	Would you be comfortable with using your device's camera to test out your knowledge? (No information would be stored from the camera for the user's privacy.)
Aim of the question	By asking this question we can reassure that users are comfortable

	with using their cameras.
Observation	
	46% of the participants are willing to use the cameras during the learning process. And 37% of the participants are not sure about using the cameras while 17% responded saying they are not willing.
Conclusion	
We can observe that most of the users are comfortable with using their cameras but only a small number of people oppose this idea.	
Question	Would you prefer dividing your lessons into categories such as "Greetings", "Emotions"..etc?
Aim of the question	To determine whether users are comfortable learning in categories
Observation	
	87.5% of the participants are willing to study using categories, while 8.7% are not sure about their choice. But 3.8% oppose this idea, and they are willing to learn all the words together.
Conclusion	
Many of the participants are interested in learning words through categorization. It is	

simple to learn by categorising the words.	
Question	Would you like to keep track of your learning progress on how far you've come?
Aim of the question	To determine whether users prefer to track their progress.
Observation	
	Only 7.7% do not want to track their progress. Remaining 92.3% of the participants prefer to track their progress.
Conclusion	
The majority of people like to keep track of their progress, which helps them learn better.	
Question	How useful will this website be in your day-to-day communication?
Aim of the question	To determine whether or not this website will be useful.
Observation	
	37.5% of the respondents have responded neutrally. Where only 3.8% have said it is not useful. While 21.2% of the participants have said it is useful in their day-to-day life.

Conclusion	
As per the survey, the majority of the participants agreed that it may or may not be useful. There is no certainty about the usefulness of this website in day-to-day life.	
Question	Do you believe that this system would be useful in the corporate sector when interacting with hearing-impaired employees?
Aim of the question	Determining whether this website will be useful in corporate sectors
Observation	
	81.7% of the participants say that the website is useful in the corporate sectors. While only 2.9% have opposed it. But 15.4% of the participants are not sure about this idea.
Conclusion	
It is very likely to be observed that most of the participants have agreed that the website will be useful in corporate sectors.	
Question	Would it be useful to customise this website according to the user/ organisation using this website? (Customizable words/phrases)
Aim of the question	To know whether the idea of customising the website according to the user preference is acceptable.
Observation	



Of the respondents, 73.1% of them agreed that customising the website according to user preference is useful. Only 2.9% of the participants oppose this opinion. While 24% say it may or may not be useful.

Conclusion
It is clearly observable that the majority of the people like to customise the website according to their preferences.

4.5.2. Interview Discussion

Mr. Tharusha Wikramasinghe, Manager of a large-scale cooperation, was interviewed as a part of the requirement elicitation technique. We learned that they use sign languages to interact with clients at their branch. During the interview, he mentioned that sign languages are utilised in two branches, as well as personnel who are deaf. We discovered that there will be 300 to 400 hearing-impaired employees by 2025.

As such, our interviewer inquired as, how people react when they visit these outlets. He stated that the majority of the time, the first impression is not as optimistic as expected. Customers eventually understood the concept of their business and began to visit on a regular basis, appreciating their efforts. We also learned that Sinhala Sign Language is not widely used in Sri Lanka. He also stated that the sign languages used in their two branches are not the same, and that some gestures are made up. He also suggested that we could develop a system that could be used in all corporate sectors.

One of our interviewers asked him about using a sign language platform. He responded that it would be very useful and inclusive if we created a learning platform that would assist all corporate sectors in learning which is customizable.

4.6. Summary of Findings

Model	Description	Sample data	F1 Score	Complexity
3D Hand Pose with MediaPipe and TensorFlow (Google, 2021)	This updated version of the hand position detection model has enhanced 2D accuracy, novel 3D support, and the ability to forecast critical spots on both hands at once. One of the most frequent requests from the developer community was for support for multi-hand tracking, which was included in this release.	Our dataset was used	59	Machine time was high
Sign-language-gesture-recognition (Hthuwal, n.d.)	The goal of the research is to recognize sign language using deep neural networks that combine spatial and temporal features. The dataset contains 46 Argentine Sign Language gesture categories. The goal is to improve communication for those who primarily communicate through sign language.	Their own data set was used. (Argentinian Sign Language Gestures.)	95.2	Machine time was high
Sign Language Detection using ACTION RECOGNIT	A key point detection model was used to construct key points for an action detection model in order to decode sign language. Tensorflow and Keras were used to create a deep neural network with LSTM layers. The model	Their own data set was used	100	Machine time was high

ION with Python | LSTM Deep Learning Model (Nochnack, 2021)	used MediaPipe Holistic Key Points to detect actions and predict sign language from video sequences in real time.			

Table 10 : Summary of Findings

Findings	**Literature Review**	**Questionnaire and Interview**
There are no online systems as of now to learn SSL	X	X
Accessibility to camera to test user's knowledge is novel	X	X
Systems should support left and right handed users and users with different skin tones		X
This area of sign language lacks data and research.	X	X
Environment disturbances and background interruptions may disrupt the final outcome from the system.	–	–
Most existing systems only focuses on	X	

ASL or BSL		
This system should fully support hearing impaired users	X	X

4.7. Context Diagram

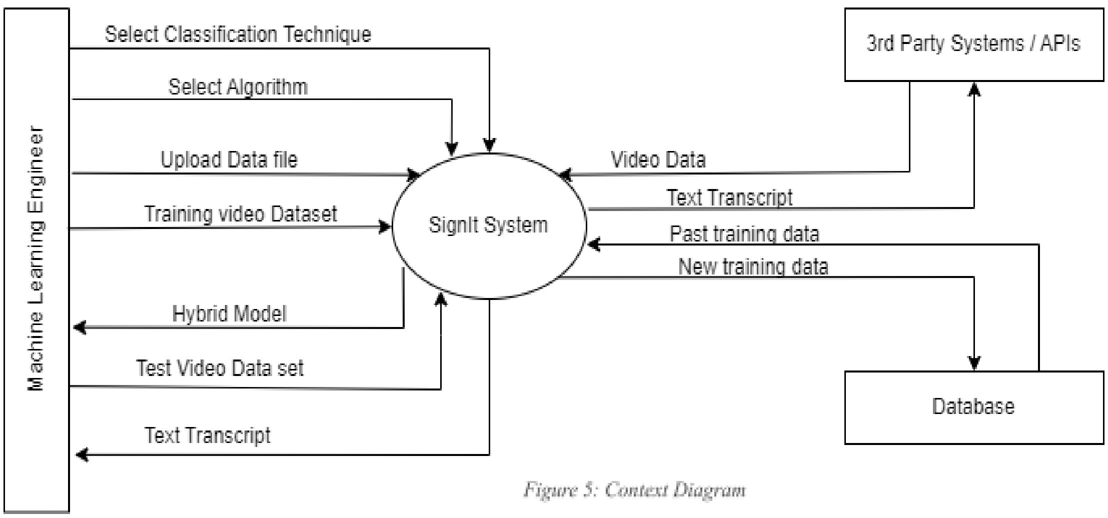


Figure 5: Context Diagram

4.8. Use case Diagram

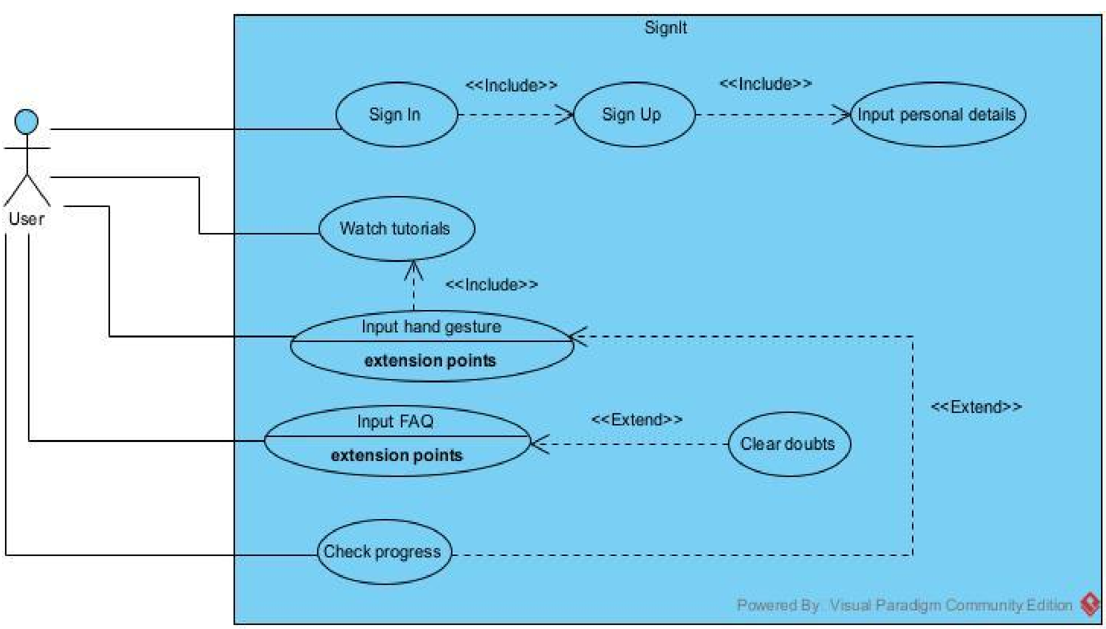


Figure 6 : Use case Diagram

4.8.1. Use case Descriptions

4.8.1.1. Use Case - Sign up

SignIt		
Use Case	Sign up	
Section	Main	
Actor	User	
Purpose	Signing up to the website	
Overview	If the user is interested in learning SSL he/ she has to sign up to our website in order to proceed with the lessons	
Precondition	The user must not have an existing account with our website	
Typical cause of events	**Actor Action**	**System Response**
	Opens the website browser	Shows sign up/ sign in block
	Select sign up	Redirects to the sign up page
	Provide required credentials	Approve sign up
Include	• Include personal information such as, name, email address • User creates username and password	

	• System creates a new account for the user
Exclude	• User entering an existing username • User entering an already existing password • Password requirements are not satisfied

4.8.1.2. Use case - Sign in

SignIt		
Use Case	Sign In	
Section	Main	
Actor	User	
Purpose	Signing in to the website	
Overview	The Sign In process allows users to access their account by providing their login credentials.	
Precondition	Users must have signed up to our website.	
Typical cause of events	**Actor Action**	**System Response**
	Opens the website browser	Pops up sign up / sign in block
	Select sign in	Redirects the site to sign in page
	Provide sign in requirements	Approve sign in and directs the site to home page

Include	● System validate user credentials ● System redirects the user to the home page upon successful sign in
Exclude	● User enters incorrect credentials

4.8.1.3. Input personal details

SignIt		
Use Case	Input personal detail	
Section	Main	
Actor	User	
Purpose	While signing up user must provide required personal details	
Overview	When the system directs the user to the sign up page, user must provide the required details to complete the signup process	
Precondition	User must have clicked the relevant button to direct into the sign up page User must not have an existing account User must have a valid email address	
Typical cause of events	**Actor Action**	**System Response**
	Enters the sign up page	Asks for required personal details
	Provide the required	Check for detail

	details	validation
		Create an account after validating
Include	• User provides required and valid details • System confirms the details and create a new account for the user	
Exclude	• User does not provide the required details and detailed being invalid • Failing to provide the details in relevant fields • System encounters a technical issue	

4.8.1.4. Watch tutorials

SignIt	
Use Case	Watch tutorials
Section	Main
Actor	User
Purpose	Provide users with instructional videos that offer visual guidance and examples to help them better understand the features and functionality of our platform and to enhance the user experience.
Overview	After successfully signed into the website, user must watch the tutorials
Precondition	User must have signed into the website

Typical cause of events	Actor Action	System Response
	Sign Up to the page	Give access to the lessons
	Select tutorial option	Displays the list of available tutorial videos
	Select the relevant tutorial video	Displays the relevant video to the user
Include	• System displays the available tutorial videos • User clicks the relevant tutorial • System plays the video, the user can control the video by pasusing, rewinding, fast forward etc.	
Exclude	• System does not have any relevant videos • Poor quality videos or error in playing the video	

4.8.1.5. Input hand gesture

SignIt	
Use Case	Input hand gesture
Section	Main
Actor	User
Purpose	In order to proceed with the lesson, user must input the hand gesture they need to learn

Overview	After users watch the relevant tutorials, they must proceed with the learning. In order to complete the learning process user must gesture the relevant sign through their web camera or mobile phone camera	
Precondition	User must have watched the relevant tutorials	
Typical cause of events	**Actor Action**	**System Response**
	Proceed with the learning process	Provides the learning materials
		Displays the input screen
	Gestures the word	Authenticate the input
Include	• System displays the input screen • User enters the hand gesture they want to learn • System displays appropriate results	
Exclude	• User enters an invalid gesture • System being unable to recognize the hand gesture	

4.8.1.6. Input FAQ

SignIt	
Use Case	Input FAQ
Section	Main
Actor	User

<table>
<tr><td>Purpose</td><td colspan="2">Provide answers to the most commonly asked questions from the user</td></tr>
<tr><td>Overview</td><td colspan="2">If users have any questions or concerns about any aspect of this website, they can use the FAQ section as a resource to resolve their problems.</td></tr>
<tr><td>Precondition</td><td colspan="2">User must have access to interact with the FAQ section of our website</td></tr>
<tr><td rowspan="4">Typical cause of events</td><td>Actor Action</td><td>System Response</td></tr>
<tr><td>Presented with a question</td><td>Search for relevant questions and answers</td></tr>
<tr><td></td><td>Provide answers</td></tr>
<tr><td>Will clarify the questions with the provided answers</td><td></td></tr>
<tr><td>Include</td><td colspan="2">User must be able to search for keywords related to their question
FAW must be in a clear and understandable format</td></tr>
<tr><td>Exclude</td><td colspan="2">FAQ does not assist with technical issues</td></tr>
</table>

4.8.1.7. Clarify the doubts

SignIt	
Use Case	Clarify the doubts

Section	Main	
Actor	User	
Purpose	Clear the doubts that users encounter while using our website	
Overview	The user can receive support and clarification on any topic related to the learning material by accessing the FAQ section	
Precondition	User must have access to interact with the FAQ section of our website	
Typical cause of events	**Actor Action**	**System Response**
	Visit the FAQ section	
	Search for relevant answer	Provide with relevant questions and answers
	Clarify the doubts using the results provided by FAQ section	
Include	● Website must be able to provide with the relevant answers	
Exclude	● User's question violate the website's terms and conditions ● Website failing to provide the expected answers	

4.8.1.8. Check Progress

SignIt		
Use Case	Check progress	
Section	Main	
Actor	User	
Purpose	To have a track of the work completed by the user	
Overview	After correctly performing the hand gesture, the user will receive feedback indicating the result. The feedback will be saved to the progress bar, which the user can access at any time to view their progress.	
Precondition	User must have completed at least one task.	
Typical cause of events	**Actor Action**	**System Response**
	Makes the gesture	Validate the input and provide feedback
	Receives a feedback	Record the feedback to the progress section
Include	● User must have completed at least one lesson ● User has access to the progress bar	
Exclude	● User has not completed any lesson ● User is not logged into their account ● ‘User does not have access to the progress tracking section	

4.9. Functional Requirements

4.9.1. Priority level indications

This is the reference used to map the priority to each functional requirement.

Table 11 : Priority Indication

Priority Level	Description
Critical	The system's main features and functionalities
Important	Not mandatory, but is thought to be needed
Non-important	Out of scope requirements

4.9.2. Functional Requirements

Table 12 : Functional Requirements

	Requirement and Description	Priority
FR01	Accepting video input from device	Critical
	The device must be able to accept videos as input.	
FR02	Removing unnecessary background disruption	Critical
	Because background features can interfere with gesture recognition, unnecessary background disruptions are avoided.	
FR03	Enhancing video input	Critical

	Improving video input and applying it to building endeavours	
FR04	Breaking down video input into frames	Critical
	Using the trained model, break the videos into several set of frames	
FR05	Adding the image filters for broken down frames	Important
	After breaking the videos into set of frames, adding filters according to the trained model	
FR06	Extracting the gestures done by user	Critical
	Convert the videos to a series of frames.	
FR07	Accept the signed gesture	Critical
	Comparing the input gestures with the pre-existing datasets	
FR08	Determining the word or phrase signed	Critical
	Identifying the gestured word or phrase signed by the user	
FR09	Generating the results from the model	Critical
	After identifying the gestures generate results and give pop up messages	

FR10	Displaying the results and suggestions	Important
	Giving pop up messages including the results and suggestions	
FR11	Support user in all lighting background	Important
	The system must be able to recognize user inputs from various backgrounds.	
FR12	Support user with different skin tones	Important
	System should accommodate users with various skin tones	
FR13	Should support users signing from either hand	Important
	System must be able to recognize even user inputs the gesture in different hands (left or right)	
FR14	GUI support for users	Important
	The GUI of the system must be user-friendly.	
FR15	FAQ system	Not-Impo rtant
	A FAQ section has been added to the website to help users clear up any questions they may have.	

FR16	Progress checker for users	Important
	A progress checker has been added to keep track of the user's progress.	

4.10. Non - Functional Requirements

Accuracy:The system's accuracy is crucial, and it should translate speech data into text data. More transcription errors imply that the model or system is unusable.

Performance: The train and test sets' combined data quantity (from speech data to lexicon) is enormous and will continue to grow over time. such that when using additional data, the model training time will be longer.

Usability: The command prompt is used for everything, from setting up the system to testing it with new speech data. The system will be more user-friendly if it has a user interface for uploading voice data and displaying written transcripts.

Security: The system must save the client speech data for testing and training purposes. So that the system may be protected against unauthorised access and data misuse.

Table 13 : Non- Functional Requirements

	Requirements and Description	Specification	Priority
NFR01	Model must have high accuracy	Accuracy	Important
NFR02	All type of users must be able to use the system without much effort	Usability	Important
NFR03	System and data access should be restricted based on role.	Security	Important
NFR04	User friendly interface	Usability	Important

NFR05	The system should generate error messages.	Performance	Important
NFR06	The website must be responsive	Performance	Non-important

4.11. Chapter Summary

In this chapter, the stakeholder's views and demand was analysed for the SSL system using various methods of data collection. Main use cases and definitions defined the functional and non-functional requirements along with their priorities for implementing the System.

CHAPTER 05 : SOCIAL, LEGAL, ETHICAL AND PROFESSIONAL ISSUES

5.1. Chapter Overview

The SLEP analysis aids in the assessment of all types of external influences on the product being built. This allows you to see the big picture of the entire/macro environment. All legal, social, ethical, and professional aspects of our project SignIt were thoroughly examined and addressed.

5.2. SLEP Issues and Mitigation

5.2.1. Social Issues

- To address privacy concerns, the platform could be built with data protection and security in mind. Only with their explicit consent could user data be collected and used for the sole purpose of providing the service. To protect user data from misuse or unauthorised access, strong data protection measures could be implemented.
- Since the platform is designed for SSL, careful consideration is given to the cultural norms, values, and practices of the corresponding community to ensure that the platform respects their cultural values.

5.2.2. Legal Issues

- The datasets employed for building the model were acquired with due permission, and we have received video consent letters from each of the individuals involved.
- The collected data is not shared with any external parties.
- The research papers that we have referred to and related works have been properly cited, and credit has been given to the relevant authors.

5.2.3. Ethical Issues

- The Google forms used for review do not reveal the user's identity.
- The datasets for the AI models were chosen by limiting the search to that follows proper ethic framework.

- User data is anonymised when appropriate, throughout the preparation processes before use in model training, and the data collection procedure was ethically sound.

5.2.4. Professional Issues

- Thorough testing and quality assurance measures will be implemented to identify and address technical issues prior to the platform's public release, thereby ensuring its smooth functionality and optimal performance for the target audience.
- The product is continuously developed and deadlines are scheduled and fulfilled according to the plan.
- Regular weekly supervisor meetings are held, and group meetings are scheduled at least once a week. Additionally, communication is maintained between meetings through phone calls and text messages.
- We used email for official communication with the supervisor, while Whatsapp was utilised for efficient day-to-day communication.

5.3. Chapter Summary

This chapter describes the Social, Legal, Ethical, and Professional issues related to our project and how they are alleviated. Developers can gain a comprehensive view of the project's macro-environment and make informed decisions to minimise potential negative impact on their project by analysing these four factors.

References

[1]World Health Organization. (2022). *Hearing loss*. Retrieved February 18, 2023, from https://www.who.int/health-topics/hearing-loss#tab=tab_1

[2] Dissanayake, M., Herath, H. C. M., W. A. L. V. Kumari, & Senevirathne, W. A. P. B. (2013). Image Processing Based Sinhala Sign Language Recording System. In The Institution Of Engineers Sri Lanka (IESL) (pp. 1-7).

[3] Dilakshan, S. and Priyadarshana, Y.H.P.P., 2020. Convolutional Neural Networks: A Novel Approach for Sinhala Sign Recognition System. In 11th IEEE Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON) (pp. 0141-0146).

[4] ResearchGate. (n.d.). Sign Language - Science Topic. Retrieved February 26, 2023, from https://www.researchgate.net/topic/Sign-Language

[5]Perera, L.L.D.K., & Jayalal, S.G.V.S. (2021). Sri Lankan Sign Language to Sinhala Text using Convolutional Neural Network Combined with Scale Invariant Feature Transform (SIFT). In 2021 International Conference on Advanced Research in Computing (pp. 1-4).

[6] Kumarawadu, P., & Izzath, M. (2022). Sinhala Sign Language Recognition using Leap Motion and Deep Learning. Journal of Information and Communication Technologies, 4(1), 54-68.

[7] McCombes, S. (2019). Developing Strong Research Questions | Criteria and Examples. Scribbr. [online] Available at: https://www.scribbr.com/research-process/research-questions/ [Accessed 25 February 2023]

[8] Hettiarachchi, S., & Meegama, R. (2020). Machine Learning Approach for Real-Time Translation of Sinhala Sign Language into Text. In Proceedings of the International Conference on Advances in Computing and Technology (ICACT) (pp. 187-191).

[9] Rishan, R.M., Jayalal, S., & Wijayasiriwardhane, T.K. (2022, February). Translation of Sri Lankan Sign Language to Sinhala Text: A Leap Motion Technology-based Approach. In Proceedings of the 2nd International Conference on Advanced Research in Computing (ICARC) (pp. 218-223). IEEE.

[10] Kumar, D., Kugarajah, B., Thavananthan, S., Bastiansz, G.M.A.S., Harshanath, S.M.B., & Alosious, J. (2020). EasyTalk: A Translator for Sri Lankan Sign Language using Machine Learning and Artificial Intelligence. In Proceedings of the 2020 7th International Conference on Information Technology for Social Good (pp. 506-511).

[11] Dissanayake, I.S.M., Wickramanayake, P.J., Mudunkotuwa, M.A.S., & Fernando, P.W.N. (2020). Utalk: Sri Lankan Sign Language Converter Mobile App using Image Processing and Machine Learning. In 2nd International Conference on Advancements in Computing (ICAC).

[13] Garlan, D. & Shaw, M. (1993). A Comparison of Object-Oriented and Structured Analysis and Design. IEEE Transactions on Software Engineering, 19(12), pp. 1321-1336.

[14] Hthuwal, A. n.d., Sign Language Gesture Recognition, GitHub repository, viewed 26 February 2023, https://github.com/hthuwal/sign-language-gesture-recognition.

[15] Nochnack, N. 2021, Action Detection for Sign Language, GitHub repository, viewed 26 February 2023, https://github.com/nicknochnack/ActionDetectionforSignLanguage.

[16] Google 2021, 'Real-time 3D Handpose Estimation with MediaPipe on TensorFlow', TensorFlow blog, viewed 26 February 2023, https://blog.tensorflow.org/2021/11/3D-handpose.html.