Project Report

for

SIGN LANGUAGE RECOGNITION WITH MACHINE LEARNING (SLRWML)

Version 1.0

Prepared By:

Group ID: 202004

Batuhan BAYRAKTAR 201611007
Bersu OĞUZ 201611042
Numan SÜME 201615047
Öznur ÜSTÜNDAĞ 201611061

Due Date: 03 January 2021

Table Of Contents

Abstract

After the 2010s, technological developments focused on artificial intelligence and machine learning. Despite this, technologies that help people with disabilities have not improved as much as other technologies. They should have the same opportunity as other people. This project aims to bring deaf people closer to other people. Sign language is very important for communication. According to the project, Deaf people use that application to translate what they want to say to Turkish. In addition to that, people who need to communicate with deaf people also want that translation. The other usage of that technology is easy writing like what voice assistants do. Using sign language, users can easily send messages. In order to help programmers, videos explaining how the project works, open-source code, and libraries used are shared in the literature. This article aims to summarize the literature and research of previous projects that were the subject of this project. We also examine machine learning and artificial intelligence science literature to evaluate how to work in this field.
Keywords: Sign language, deaf people, machine learning, artificial intelligence

Öz

2010'lu yıllardan sonra teknolojik gelişmeler yapay zeka ve makina öğrenmesi alanlarında yoğunlaştı. Buna rağmen engelli insanlar için yapılan gelişmeler diğer gelişmelerin yanında eksik kaldı. Engelli insanlar da diğer insanlarla aynı fırsatlara eşit olmalıdır. Bu proje duyma engelli insanlarla diğer insanları daha çok yakınlaştırmayı amaçlar. Bu iletişim için de işaret dili büyük önem arz etmektedir. Bu projeye göre duyma engelli insanlar, uygulamayı kullanarak anlatmak istedikleri şeyleri işaret dilinden Türkçe'ye çevirebilirler. Buna ek olarak duyma engelli birisiyle iletişim kurmak istediğinde diğer insanlar da bu uygulamayı kullanarak onlara işaret dili tercümesini gösterebilirler. Bu uygulamanın diğer bir kullanım alanı da sesli asistanlar gibi çalışır. Nasıl sesli asistan kullanarak hızlı mesaj yazabiliyorsak, aynı şekilde işaret dili kullanarak da aynı işlemi

gerçekleştirebiliriz. Programcılara yardımcı olmak amacıyla literatürde projenin nasıl çalıştığını anlatan videolar, açık kaynak kodu ve kullanılan kütüphaneler paylaşılmıştır. Bu makale, bu projenin konusu olan önceki projelerin literatürünü ve araştırmalarını özetlemeyi amaçlamaktadır. Bu alanda nasıl çalışılacağını değerlendirmek için makine öğrenimi ve yapay zeka bilimi literatürünü de inceliyoruz.

Anahtar Kelimeler: İşaret dili, sağır insanlar, makine öğrenimi, yapay zeka

1. Introduction

1.1 Motivation

The motivation of the Sign Language Convertor project has been explained below:

1 There are lots of people who suffer from hear loss and related diseases. Thus, increasing usage of digital world. The sign language convention has been becoming a significant factor for those people.

2 The demand of the sign language translator has been increasing because of the wide usage and the translator requires a translator person because there are not enough people who are able to translate so developing and computer programs will figure out the problem.

3 Using machine learning algorithms will decrease the workload of the computer memory and give more true results day by day related to the using program. It will increase reliability.

1.2 Problem Statement

It is obvious that the biggest advantage of people is communication with each other. Nevertheless some of us born with some disabilities. On the other hand, some of us lose our abilities due to trauma, accident, or illness. One of these problems is deafness. There are 466 million people in the world with disabling hearing loss. This is over 5% of the world's population. In addition, nearly 1 million people have hearing problems in Turkey as well. This problem makes

people's life more difficult. They express their thoughts and feelings thanks to sign language. However, the rest of society doesn't know how to communicate with sign language. In this project, we aim that figure out this disconnection.

1.3 Solution Statement

In this part, the Solution that we planned to handle this disconnection will be explained. There is some example of that project but there is not Turkish support. According to the problem statement and previous information, we will implement a mobile application and a web site to translate Turkish to sign language. Thanks to that application, deaf people can express their thoughts and feelings in sign language, and the application translates it to Turkish as a text. If they want, they can make it as voice via google assistant or Siri. On the other hand, people who want to communicate with deaf people can also use that application on the opposite side. People also use it for texting when they can not use their hands for example while cooking or walking under rain etc.

2. Literature Search

2.1 Introduction

According to the data of the World Health Organization (WHO), 466 million deaf people live worldwide and more than 3 million of them living in Turkey. Unfortunately, only 10 percent of people who are deaf today have hearing aids. Also, some of the people can not hear totally and hearing aids can not help them. Therefore, sign language is of great importance for deaf people. Using this language, they can communicate with the people they meet. We designed this project to make life easier for those with a high percentage (those without hearing aids).

We examined how existing projects work and what they benefit from. In addition, our literature research enabled us to identify gaps in the current literature. These shortcomings will contribute to the later stages of the project.
In the next section, the literature of previous projects is discussed. In the fourth section, we compare our project and previous projects. We mention the advantages of our project according to previous work. In the fifth part, we explain and conclude the contribution of the literature of the projects to us.

2.2 Related Work

The traditional way to prevent this issue is to use a mechanical tool for ears to improve the sound of the environment for deaf people. However, these devices only increase the sound, for totally deaf people this sort of device is not useful. There have been many studies on sign language in the past. We reviewed these studies, made improvements, and tried to fix the shortcomings we found in our project. We will examine a few examples made in the past below.

2.2.1 Real-Time Malaysian Sign Language Translation using Colour Segmentation and Neural Network[1]

Recommended automatic sign language translators provide a real-time English translation of Malaysian SL. A wearable glove is used in combination with colors to help identify fingertips in the color spectrum of the camera.
This project focuses solely on finger typing, which takes different approaches respectively. The working process takes place in 4 stages; First, read the frame, dissect the color, find the center of the image, and finally determine the result.
Finally, trained neural networks are used to describe signs to be translated into English.

2.2.2 A real-time continuous gesture recognition system for sign language[2]

First, a motion input is detected and analyzed, then motion is analyzed according to four parameters: stance, position, orientation, and motion.

The developers implemented a prototype system with a 250-word dictionary and collected 196 educational phrases in Taiwanese Sign Language (TWL). This system uses hidden Markov models (HMMs) for 51 basic postures, 6 orientations, and 8 motion principles.

The accuracy rate obtained in this study is approximately 80.4%.

2.2.3 A vision-based sign language recognition system using tied-mixture density HMM[3]

In this project, a vision-based vocabulary Chinese sign language recognition (SLR) system is presented.The system consists of two parts.

In the first part, hand sensing, background subtraction and pupil detection techniques are largely combined to show the movements in different environments with the help of appropriate colored gloves.Principal component analysis (PCA) is used to interpret finger features in more detail.

In the last section, the Bound Mix Density Hidden Markov Models (TMDHMM) framework for SLR is proposed.

The accuracy of this project reaches up to 92.5%.

2.2.4 A Video-Based Indian Sign Language Recognition System (INSLR) Using Wavelet Transform and Fuzzy Logic[4]

In this project, a simplified Video-Based Indian Sign Language Recognition System (INSLR) has been designed using a combination of various image processing techniques and computational intelligence techniques. At the same time, a wavelet-based video segmentation technique is used that detects many hand signals and head movements. The properties of hand gestures are understood using the feature vectors of the image using elliptical Fourier annotations.
Finally, the INSLR system uses a sound system to play recognized movements along with text output. The system was tested by 10 different programmers using a data set of 80 words and sentences.

It shows that it is 96% correct according to the results of the tests.

2.3 Sign Language Translator Using Machine Learning

According to research which was done in 2017, There is currently more than 3 million people have been suffering from deafness. Previous works work in English and translate word by word to text. Also, one of the project use gloves to recognize. These are not useful and efficient in real life. Translating word by word does not make sentences and it is hard to understand for people. Besides, people can not carry gloves all the time. People need a real-time translator to sign language to human languages. In addition to that, there is not any study in Turkey. In this project we want that solve this problem. Deaf people can use applications, which in their mobile phone, to communicate with other people. In addition to that, their relatives also want to translate human languages to sign language. To do this, people can use it with video or using the front camera of the mobile phone.

Application translate it to writing and using voice assistants, other people can hear what deaf people sign. Another application area is similar to the voice assistant. While preparing some food or walking in the rain, writing text messages is very difficult. People use voice assistants at these times. Deaf people haven't that opportunity. To give them that chance, this application can be used as an assistant. Lastly, for online videos, movies, series, or tv shows, there exist subtitles but some people may want to sign language translation.

2.4 Conclusion

As a result, some projects have been carried out and designed to make life easy for 466 million deaf people around the world. There are not many applications for a sign language translator. The ones that exist could not use efficiently in real life. Based on these projects, we decided to identify the deficiencies and make an easier and more understandable project. Our project has been approaching issues in a different way. In this way, it will be ensured that deaf people without hearing aids can easily communicate with anyone they want to use this project.

3. Software Requirements Specification

3.1 Introduction

The following subsections are an overview of the entire “Software Requirements Specification” (SRS) document.

3.1.1 Purpose

This document has been prepared to explain briefly what is the “Sign Language Recognition with Machine Learning” and why it will be implemented. The required information will be investigated in subsections.

3.1.2 Scope

The sign language has been being used by the people who suffer from hearing loss. These people express themselves with a specific type of language which is called “Sign Language”.[1] The sign language is like the other languages, each region has its own specific sign language like normal language. The common thing in sign language is that; each expression is stated with fingers and some arm movement.

The interest to sign language has been increasing with growing media. Today each TV series, news and movies have been using sign language to reach the deaf people or people who communicate with sign language. Thus, sign language stuff and similar applications have been being used widely nowadays.

The Sign Language Recognition has been providing to convert sign language to written language faster, safely and without human beings. Since it is not feasible technically to convert all media on the internet to written language from sign language, it will be useful to use an algorithm that does the same thing.

The software that will be prepared may use the dataset that is available on the internet but also it will be needed to generate some dataset. The sign language is used by many people, so it contains small differences for every person, the software should understand the difference and ignore it by machine learning algorithms.
The objectives of the project are:

- To convert the sign language that is perceived by camera to written language.
- To do conversion as much as faster.

- To do conversion truly,namely, it has to ignore expression not related with sign language-gestures-
- To generate meaningful sentences rather than putting all words side by side.
- To store user's type of speaking and understand words easily from personel speaking

3.1.3 Definitions, Abbreviations, Acronyms

AI	Artificial Intelligence
ML	Machine Learning
SLRWML	Sign Language Recognition with Machine Learning

3.1.4 References

[1] “https://tr.wikipedia.org/wiki/İşaret_dili (Last Access : 30 November 2020)”

3.1.5 Overview

This document had been prepared to give clear information about the software that converts sign language to written language. This document includes all technical detail and required specification for this purpose.

3.2 Overall Description

This section describes the requirements for SLRWML. To be easy to understand, this part of the SRS provides a basis for requirements. The detail definitions are mentioned in chapter 3.

3.2.1 Product Perspective

SLRWML is an application that has both a web interface and an app interface.

3.2.1.1 External System Interfaces

Since SLRWML is an independent system, there is no system interface with any other system.

3.2.1.2 User Interfaces

SLRWML has 1 user, but there are 2 different interfaces according to the users' request.

Translation From Image To Sound And Text Interface

The user interface is designed for everyone. Therefore, the interface should be simple and convenient. In this interface, there is a screen at the bottom of the screen that detects and shows the user's movements, and there is a text box that converts these movements into a written text. Next to the text box there is a loudspeaker sign. This sign will also enable you to hear the text aloud.

Translation from Text to Image Interface

The user interface is designed for everyone. There is a text box in this interface where you can enter text. At the top is a screen that translates the entry in this text box into sign language.

3.2.1.3 Hardware Interfaces

Server Side

The PC will have a 64 bit architecture. The computer will have a hard disk large enough to hold close to 100 sign language data.

Client Side

Any phone or personal computer that can support any Windows environment with mouse support or run on android devices is acceptable.

3.2.1.4 Software Interfaces

Server Side

The developed application will be accessed using a web browser or by downloading it from the google play store. Approximately 100 sign language information will be uploaded to a database.

Client Side

The required software product on the client side is an internet browser that supports at least HTML version 3.2, java enabled, camera enabled and any operating system that can run browsers.

3.2.1.5 Communication Interfaces

The default communication protocol for data transmission between server and the client is Transmission Control Protocol/ Internet Protocol (TCP/IP). At the upper level Hyper Text Transfer Protocol (HTTP, default port=80) will be used for communication between the web server and client.

3.2.1.6 Memory Constraints

SLWML will not use any additional resources. User's computer running the web browser or android phone must have enough physical memory to run this program.

3.2.1.7 Operations

The SLRWML operations that users need are explained below.

• Users will first see 2 different translation options. These are image-to-sound / text translation and image-to-text translation. By choosing one of these options, they will be directed to the relevant page.

• On the image-to-audio and text-to-text translation page, the user shows the words they want to tell in sign language and the computer detects it, then

translates the detected word into text and displays it in the text box. Optionally, the texts in the text box can be translated into sound and listened.

• On the text-to-image translation page, users type the word they want to tell in the text box. Then, the sign language equivalent of this word is projected as an animation on the display screen.

3.2.2 Product Functions

The system has 1 different function for 1 different user group:

User: Chooses what kind of translate he wants to do, then writes or displays the inputs requested by the system.

3.2.3 User Characteristics

There will be 1 types of users:

User: The user must have a smartphone or web browser with an internet connection in order to use the SLRWML website. It also needs an android phone with internet connection in order to use SLRWML's application.

3.2.4 Constraints

The size of the hard disk should be close to 100 data that the system wants to keep in memory.

3.2.5 Assumptions and Dependencies

Each user will have the appropriate hardware and software configuration specified in sections 2.1.3, 2.1.4, 2.1.5.

3.2.6 Apportioning of Requirements

Language usage requirements such as server type, communication type and secure connection should be determined. The problems that the user will experience in the application should be determined and appropriate solutions should be produced for them.

3.3 Specific Requirements

3.3.1 External Interface Requirements

There are no external interfaces for this system.

3.3.2 Functional Requirements

Assuming there is a pre-existing database of Turkish word and sign language gestures as animation. Users don’t have to sign up or sign in. Applications will be available for everyone.

3.3.2.1 Image-to-Sound / Text Conversion Functions

Devices must have a camera for detection. The gestures which are detected via camera, will translate to text messages. For voice message, Devices must have a voice assistant such as Google Assistant or Siri. Text messages will be converted to voice via assistants.

3.3.2.2 Text to Image Conversion Functions

There will be two options for input. Users can write down as a text. Otherwise, the user can use the phone's voice assistant. After translation, output will be shown as animation on the screen.

3.3.2.3 Common Functions

There is just one type of user. So, all functions are common

3.3.3 Performance Requirements

We aim that people can use it in daily life. Because of that, translation should take less than 5 second.

3.3.4 Security Requirements

The security is provided using HTTPS over SSL protocol for secure communication between the client and the server. Also, secure network protocol TCP/IP protocol in SQL Server will be used.

3.3.5 Design Constraints

No planning was made about that section.

3.3.6 Software System Attributes

Functionality: Application is designed to translate sign language to human language and human language to sign language.

Reliability: Apllication will work properly for every OS system on mobile phones.

Security: Users are not allowed to reach the database directly, so data will be saved in this way. User information will not share with 3th party applications.

Safety: Safety will be covered by using TCP/IP on the network side.

Performance: Translation should take less than 5 second.

Flexibility: Responsive web design (RWD) will be used in GUI design. RWD is an approach that makes web pages render well on a variety of devices and window or screen sizes.

Scalability: Application is designed to cover the needs of all users.

Portability: The system works on all operating systems.

Availability: As long as there is access to the internet, there will be access to the application without any interrupts.

Usability: GUI will be designed to make users comfortable and familiar with the application. Every type of human can easily use applications. GUI will be mapped according to that.

Maintainability: Application will be updated according to coming feedback from users and improving the word database of application. Database will be backed up monthly in case of data loss.

3.3.7 Other Requirements

Users must have a mobile device for using the application. Applications must be installed on mobile phones. For database connection, Mobile devices have to have internet connection. Users should know basic phone usage.

4. Software Design Description

4.1 Introduction

Sign Language has been used to express feelings, thoughts without speaking because of some diseases related to the speaking ability or lack of hearing ability which is called deafness. The Sign Language Recognition with Machine Learning provides a translation between sign language and normal language. The translated sign language can be easily converted to other languages or vice versa.

4.1.1 Purpose

This project has been aiming to generate a useful, fast translation among sign language and normal language. The program will use machine learning to improve correctness of the translation.

4.1.2 Scope

The Software Design Description has been explaining what will be used in the project, what the system will look like. The required functions, definitions, basics of the project are explained in the report. The main structure of this project will be explained briefly in the following sections.

4.1.3 Definitions, Abbreviations, Acronyms

● CSV - A file format that contains strings, numbers separated with comma. - Comma(C)-Separated(S) Variables(V).

● GIF- This means Graphics Interchange Format. It is a bitmap image format.

● Machine Learning - A branch of science that deals with perception of computers, datasets or data types to design algorithms that allow learning based on data types.

● Python- Python is a programming language that allows users to build programs, it is preferred by data scientists, AI programmers or similar people because of its usability.

4.1.4 References

There are no references that were used for section 1.

4.1.5 Overview

The following sections will be investigated through this report.

- In section 2, used tools in the design of the system and the purpose of usage will be investigated.
- In section 3, the architecture of system, system existence and physical location will be mentioned.
- In section 4, the system interface will be explained.
- In section 5, the user interface of the system will be mentioned for design.
- In section 6, the process design will be mentioned as sketches.
- In section 7, the dataset design and relations of data will be mentioned.

4.2 Design Considerations

4.2.1 Approach

- Software architecture diagram-It was used to show how users of a typical software system can interact with external systems, data sources and services.

- Hardware architecture diagram-This diagram presents the hardware (servers, workstations) that are interconnected by a network, as well as the technical and application components that are deployed on this hardware.

- We made the dataset to keep the information of the users who reported the error with the dataset.

4.2.2 Tools Used

- For ER diagrams no specific software will be used, instead of software, “draw.io” will be used.
- The Phyton will be used as a programming language on the “Anaconda”.
- For analyzing the pictures, a sort of webcam will be used.
- “OpenCV” library, “PyTorch” and “onnx” will be used also.
- Keras
- Tensorflow
- For the documentation, MS office documentation will be used.

4.2.3 Constraints

The perceiver of the computer will perceive the movement of the hand instead of perceiving face expression.

4.2.4 Assumptions and Dependencies

The Sign Language will be expected as in American. Thus, other languages will not be expected from the user.

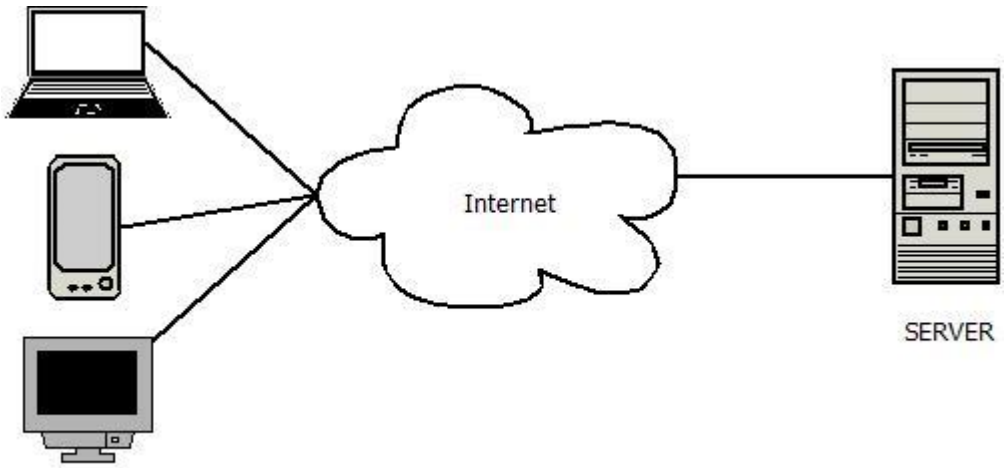
4.3 Architecture

4.3.1 Software Architecture

We do not use database design in this project. We will keep the data in the text file with the extension 'csv' instead of database. There will be pictures other than this text file. The format of the pictures will be GIF.

4.3.2 Hardware Architecture

The hardware architecture to be used in this project is client-server architecture. The client-server model describes how a server provides resources and services to one or more clients. When a client requests a connection to a server, the server can accept or deny the connection. If the connection is accepted, the server establishes and maintains a connection with the client over a specific protocol.



In this project, the client will make a request to the SLRWML application server and the connection will be made using the TCP / IP protocol.

4.4 System Interfaces

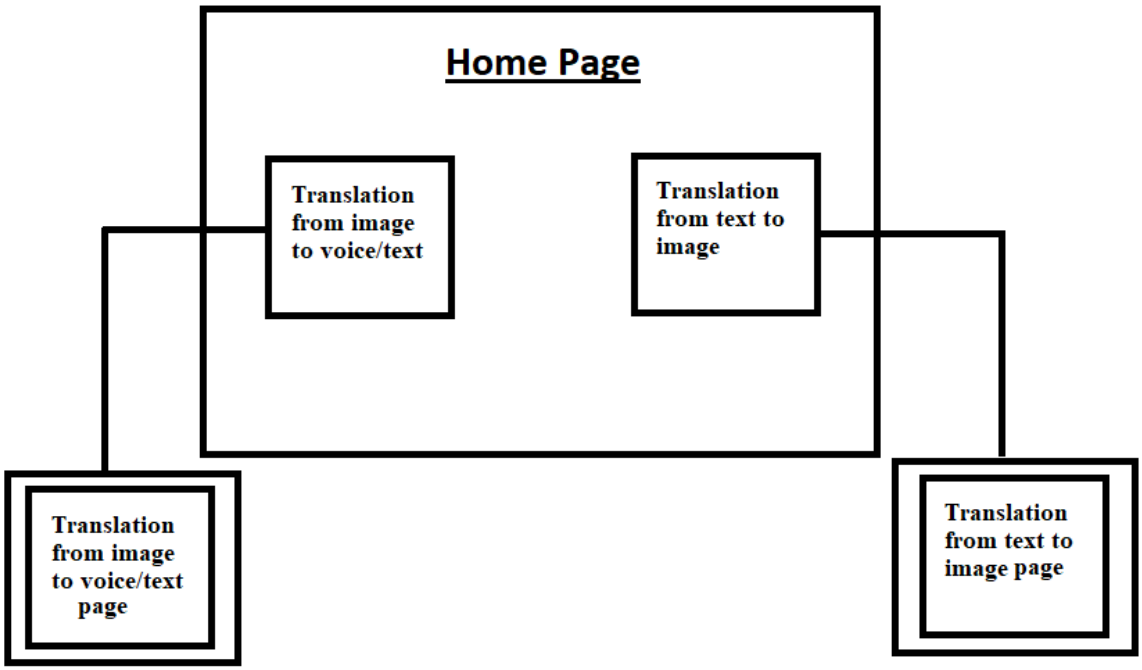
4.4.1 External System Interfaces

There is no external interface in this system.

4.5 User Interface Design

The general view of the site is shown in this section.

4.5.1 Navigation



- Users will be directed to "Home Page" by opening the application they downloaded on their android phone or by using the URL link.

- There are 2 different buttons on the homepage. The first is "Translation from image to voice/text" and the other is "Translation from text to image" button.

- When the "Translation from image to sound / text" button is pressed, it is directed to the desired page. This page asks for camera permission when

first opened.After this permission is given, words are explained using sign language through the camera, and then the words corresponding to this sign language appear in the text box below. The user can press the loudspeaker button if they want and listen to the text box out loud.

- When the "Translate from text to image" button is clicked, it is directed to the desired page. The text is written in the text box. Then, when the 'translate' button is pressed, an animation appears at the top, this animation shows the equivalence of those words in sign language.

4.5.2 Screen Definitions

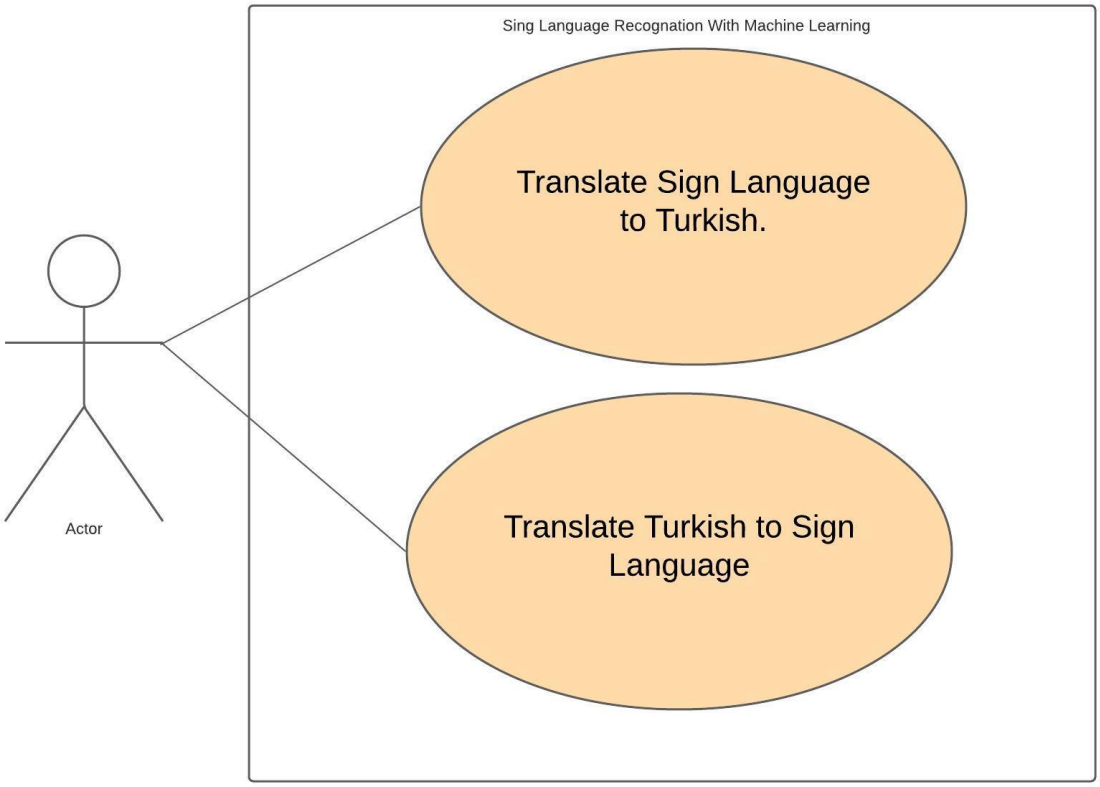
4.5.2.1 Home Page

4.5.2.2 Translation from text to image page

4.5.2.3 Translation from image to sound / text

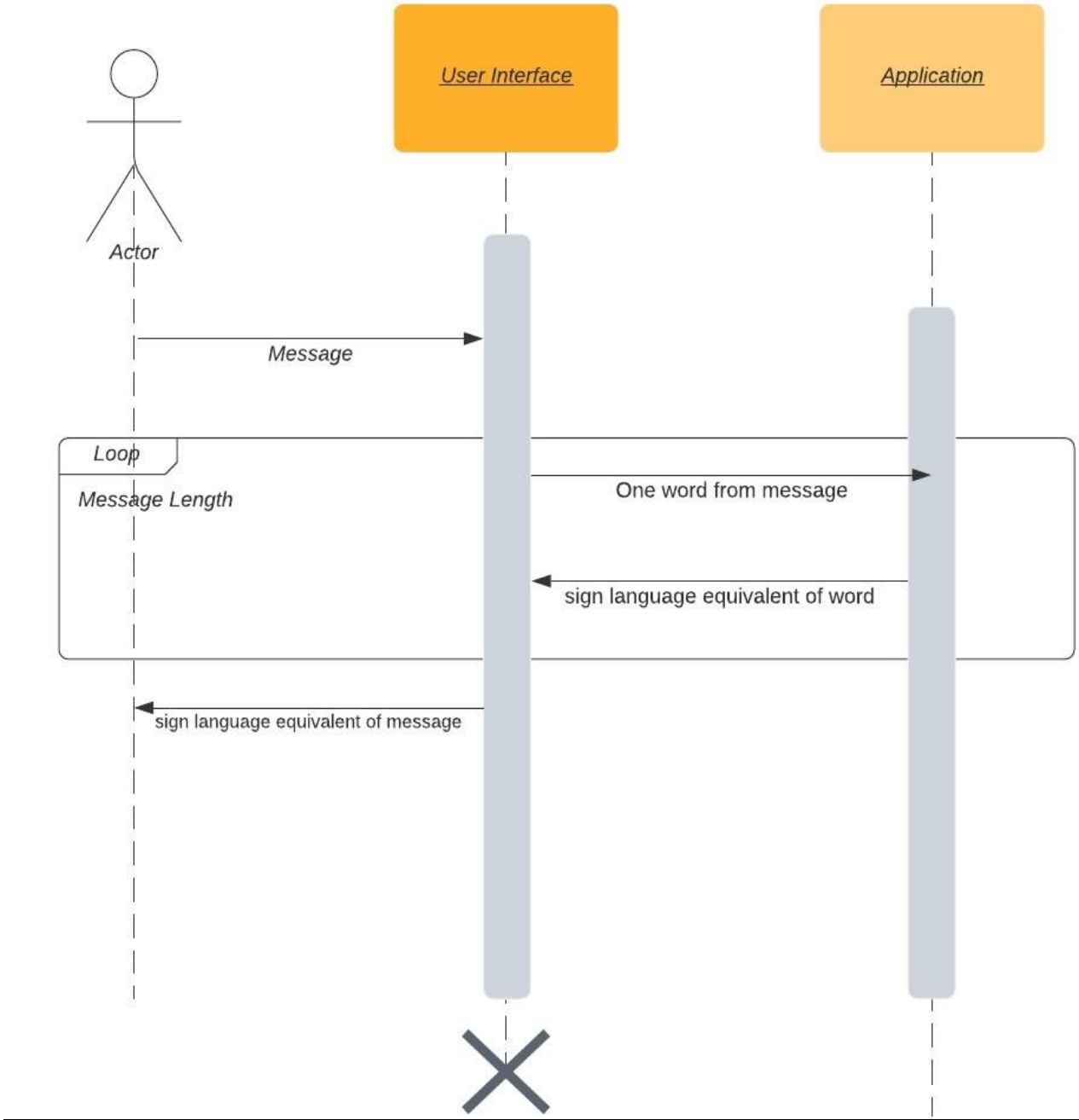
4.6 Process Design

4.6.1 Use Case

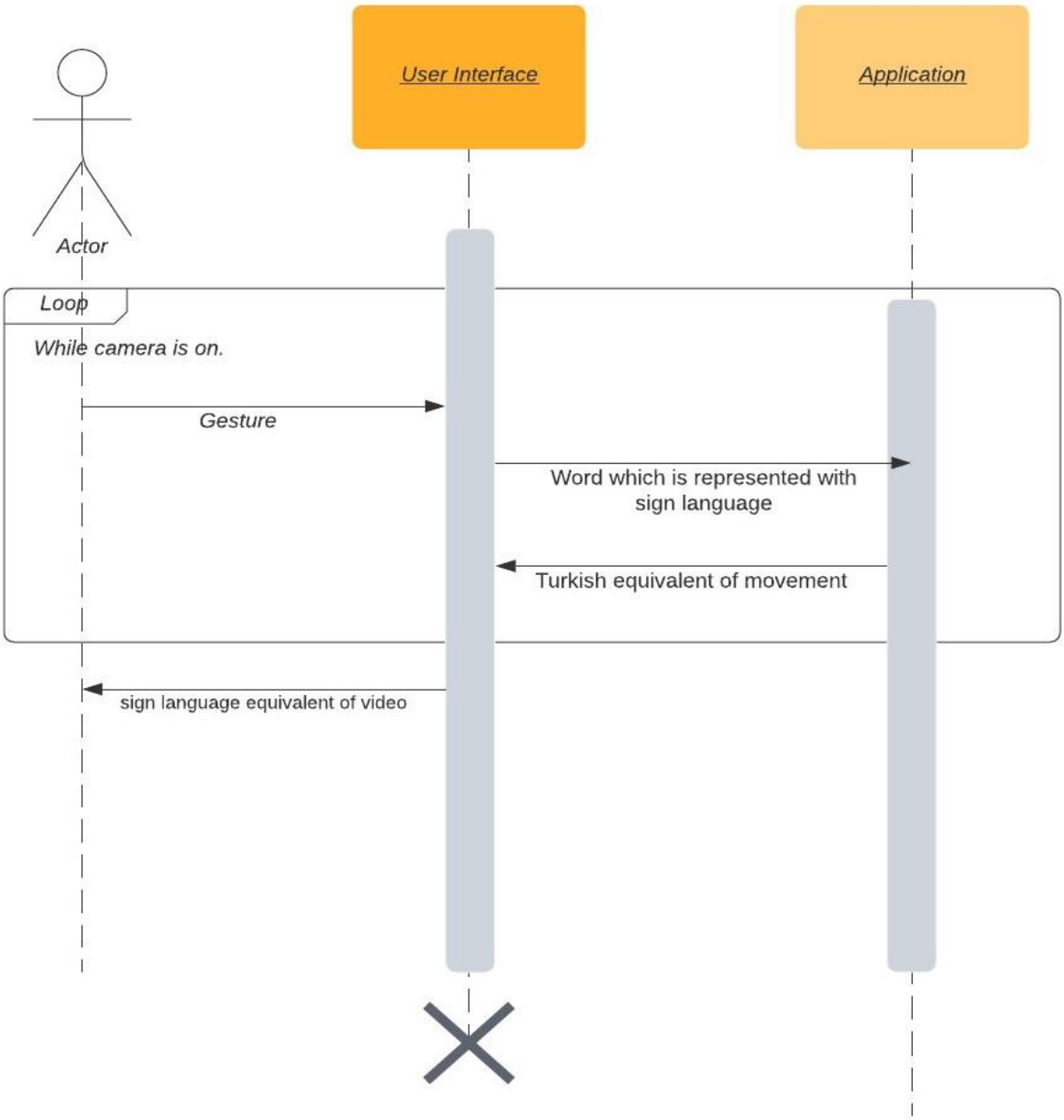


4.6.2 Sequence Diagrams

4.6.2.1 Text to Sign Language:



4.6.2.2 Sign Language to text:



4.7 Dataset Design

This section describes the how SLRWML handle with data. In this project, we prefer that hold data in csv files instead of dataset management systems.

4.7.1 CSV File Definitions

4.7.1.1 Word.csv

This csv file builded for storing words and signs which are used in daily life constantly.

- WordID: Field that holds unique ID for each word.
- Word: Field that holds Turkish words.
- Sign: Field that holds path of gif which represents sign language equivalent of word

4.7.1.2 Character.csv

This csv file builded for storing letters and numbers. The words which are not in word.csv will represent thanks to letters.

- CharID: Field that holds unique ID for each character.
- Character: Field that holds numbers and letters.
- Sign: Field that holds path of gif which represents sign language equivalent of characters.

5. Conclusions

This report contains all the information and documents about our senior project "Sign Language Recognition With Machine Learning". The aim of this project is to facilitate communication between deaf people and other people. In the project, those who want to communicate with deaf or deaf people use it to translate what they want to say into Turkish and also translate Turkish into sign

language. For our senior project, we did in-depth research on related studies, applications, available tools, data sets about our project, available source code and programming language. After completing the literature section, we decided on the project requirements with our consultant and created an SRS report covering these requirements. After determining the needs, we decided on the necessary design and methods to develop this application in our SDD report. As a result of our research, there are many translation applications in foreign languages that work towards this purpose, but there is no detailed application in Turkish. At the same time, translations are not made word by word in studies.

We are grateful to Assist. Dr. Sibel TARIYAN ÖZYER for helping us at every stage of our project.

References

[1] "Akmeliawati, R., Ooi, M. P. L., & Kuang, Y. C. (2007, May). Real-time Malaysian sign language translation using color segmentation and neural network. In Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE (pp. 1-6).
IEEE" https://ieeexplore.ieee.org/document/4258110

[2] "Liang, R. H., & Ouhyoung, M. (1998, April). A real-time continuous gesture recognition system for sign language. In Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on (pp. 558-567). IEEE." https://ieeexplore.ieee.org/document/671007/authors#authors

[3] "Zhang, L. G., Chen, Y., Fang, G., Chen, X., & Gao, W. (2004, October). A vision-based sign language recognition system using tied-mixture density HMM. In Proceedings of the 6th international conference on Multimodal interfaces (pp. 198- 204).
ACM" https://www.researchgate.net/publication/221052241_A_vision-based_sign_language_recognition_system_using_tied-mixture_density_HMM

[4] "Kishore, P. V. V., & Kumar, P. R. (2012). A video-based Indian Sign Language Recognition System (INSLR) using wavelet transform and fuzzy logic. International Journal of Engineering and Technology, 4(5),537." https://www.researchgate.net/publication/269839692_A_Video_Based_Indian_Sign_Language_Recognition_System_INSLR_Using_Wavelet_Transform_and_Fuzzy_Logic

[5] “https://standards.ieee.org/standard/830-1998.html“

[6] “https://www.cse.msu.edu/~cse870/IEEEXplore-SRS-template.pdf”

[7] Sketching Graph Tool – “https://www.lucidchart.com/”

[8] IEEE SDD Description – “http://cengproject.cankaya.edu.tr/wp-content/uploads/sites/10/2017/12/SDD-ieee-1016-2009.pdf”