Bilkent University

Department of Computer Engineering

Senior Design Project

Project short-name: YouTalkWeSign

Project Specifications Report

Abdurrezak Efe, Yasin Erdoğdu, Enes Kavak, Cihangir Mercan

Supervisor: Hamdi Dibeklioğlu
Jury Members: Varol Akman, Mustafa Özdal
Innovation Expert: Mustafa Sakalsız

Progress Report
October 9, 2017

This report is submitted to the Department of Computer Engineering of Bilkent University in partial fulfillment of the requirements of the Senior Design Project course CS491/2.

Contents

1 Introduction

1.1 Description

YouTalkWeSign will be a web application which is designed to serve people with hearing impairments or deafness. Through an on-screen avatar, the app will translate the spoken words in YouTube videos into sign language. The app can be run by replacing the "youtube" word in the video address with "youtalkwesign". Therefore, our website is http://www.youtalkwesign.com/.

For most of the videos, YouTube has speech-to-text subtitles functionality. For these videos, we will simply get the speech text from the YouTube. Then, we will translate the text to the sign language. For some of the videos, there is no speech-to-text subtitles support. For example, live videos do not have it. For these videos, we will first convert speech to text. Then, we will translate the text to the sign language.

While we are doing this conversion, we will be using the state of the art solutions of deep learning like Convolutional Neural Networks, Recurrent Neural Networks and maybe even GANs. The training data will be collected from a TV channel's morning news that also includes sign language explanation along itself. The model is to recognize speech and generate corresponding text, then from the text train the sign language correspondence.

Therefore, for example, when user makes this request to our website: https://www.youtalkwesign.com/watch?v=oxi7HcgWga0,

we will generate the sign language from the text transcript of the video. Then, we will display it on the right side of the screen near the video. Hence, the video and the sign language avatar will be played simultaneously.

1.2 Constraints

1.2.1 Implementational Constraints

- For a collaborative contribution to this project, Git software revision system will be used and implementation will be controlled and conducted through GitHub platform.
- The model will be designed in Python with Deep Learning frameworks such as TensorFlow and Keras.
- When speech-to-text conversion is needed, Google Speech API will be used.
- For back-end work of our website, we will use Java Spring Boot with Thymeleaf support on the front-end.

1.2.2 Economical Constraints

- We bought one-year hosting and domain for our website.
- The training data, frameworks and libraries to be used will be acquired with no fee except Google Speech API.
- For users, there will be no need for signing up for watching the videos with our sign language avatar.

1.2.3 Language Constraints

- The app only works with the videos that consist of spoken English language.
- If the popularity obliges, we will add other languages of demand.

1.3 Professional and Ethical Issues

Our application will be storing the data of the viewed videos from our website to show the list of the most viewed videos. We will not share this data with third parties.

2 Requirements

2.1 Functional Requirements

- YouTalkWeSign does not require anything more than a web browser and internet connection.
- The users are going to have to write "youtalkwesign" instead of "youtube" in Youtube links to get the sign language equivalent along the video itself.
- The video and the sign language avatar can be paused and continued simultaneously.
- The users can watch the video in a theatre mode. That is, the video takes %80 of the screen and the avatar will take up to other %20.

- The most viewed videos link will be available at our website navigation bar and the users will be able to see what is the trend.

2.2 Non-functional Requirements

- YouTalkWeSign will have a good-looking and attractive user interface. It will be user friendly and easy to use.
- YouTalkWeSign will be suitable for extensions. For example, we can increase the number of languages supported. Also, the other platforms like Twitch and Dailymotion can also be integrated.
- A user focused system may be implemented to enhance the quality of the usage- via recommendations etc.
- We are aiming to make the conversion to the sign language as much as fast because we do not want our users to wait a lot. For live videos, they may wait longer than the normal videos. However, we will try to decrease the delay as much as possible.
- The system will be reliable such that we will retrieve the video from the YouTube without making any changes on the original version.