Sign language synthesis from Natural Language Processing

Nijat Mursali
Department of Artificial Intelligence and Robotics
Sapienza University of Rome

Abstract—Sign language is a visual language that individuals with speech and hearing impairments use to communicate in their daily conversations. It is entirely an optical communication language due to its native grammar, which differs fundamentally from that of spoken languages. The fundamental purpose of this paper is to describe the algorithm which takes the spoken language from the user, recognize word(s) with an algorithm and convert words to sign language. As an output, we display the images with the sign language.

1 Introduction

There are more than 7000 known living languages in the world divided in 136 different language groups. Among these 136 language families, Sign language is one and this family contains 136 sign languages all over the world depending upon the region of the world. Sign language is used by hearing impaired people to convey their message.

Sign language is a nonverbal language used by the deaf and hard of hearing to communicate using hand shapes, facial expressions, gestures, and other areas of the body. Because sign languages lack a well-defined structure or grammar, they have no or very little acceptance outside of their limited community. Approximately 72 million of the world's nearly 7 billion inhabitants are deaf or hard of hearing. Out of such a large population, roughly 4.3 million utilize Sign language. The remaining roughly 67 million deaf and hard of hearing persons do not communicate using any sign language. As a result, over 90

In the following paragraphs we will be discussing how we developed the application that is capable of getting the speech from the user and giving the images of the sign language as output.

2 Motivation

A good Sign Language Recognition (SLR) system can break down the barriers that exist between the speech and hearing communities and the speaking culture. The objective of SLR is to create methods and methodologies for correctly recognizing a series of gestures and understanding the meaning of the gestures. SL presents a problem in that it is multi-channel, interpreting meaning in several ways at the same time.

SLR is a difficult and motivating work because to several restrictions and variables. This is not easy task to accomplish because there are not always exact phases we can use to convert to ASL. Thus, we had to think of a way to handle the dataset and accomplish the task.

3 Data collection

Data gathering is an essential component of research in all fields of study, including the sciences, social sciences, technology, humanities, and business. Methods for data collecting may differ per discipline, but the emphasis on ensuring exact and genuine collection leftovers always the same.

We have created a new dataset of ours by gathering the sign language images from different sources. The fundamental source was to use some phases which gave us the opportunity to get images of sign language for different words. In order to parse all the images manually would not be a good idea, thus we have created a new function in order to call the links automatically and parse the images from all of those links. We have used *B*eautifulSoup for this and for all the letters in English language (A-Z) we have checked every single link for the images and added them to local storage.

```
url = f'https://www.babysignlanguage.com/dictionary-letter/?letter={letter}'
response = requests.get(url)
soup = bs4.BeautifulSoup(response.text, 'html.parser')
content = str(soup.find_all('div', {"class": "single-letter-card"}))
soup1 = bs4.BeautifulSoup(content, 'html.parser')
```

Fig. 1. Getting the links for all letters

As shown in the image, we have first added all the alphabet letters in English language and for those letters, we have checked the URL response and for specific element we

get all the links. Then we have added those links to the list for further processing.

Then, for all the links, we are making the GET request for specific elements where images are added by their source and name. Then, within the loop, we are making the request to get the images and add them to the local storage.

4 Development

4.1 Interface

Before moving to sign language, we need to discuss how we built the structure of the application. We have used the *T*kinker which is a Python binding to the Tk GUI toolkit. It is the standard Python interface to the Tk GUI toolkit, and is Python's de facto standard GUI. We have created a simple interface where the user can click the button to get the speech from the user and then the algorithm will check which words the user has said. The following image shows the interface of the application.

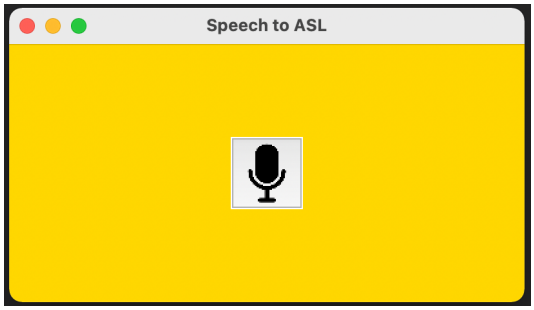

Fig. 2. Getting the links for all letters

In order to get the input from the microphone, you need to press the microphone button. After pressing the button, the input is recorded and added to a variable called "text".

In the following process, we use NLTK which is a suite of libraries and programs for symbolic and statistical natural language processing for English written in the Python programming language. We use "word tokenize" from NLTK library to get the tokenized words. For this process, we also check if sentences have some punctuations. As we know, in English language ' is used so often, for instance "don't". Thus, we just removed the punctuations.

4.2 Algorithm

We have also used Google for getting the alphabet of English in Sign language. So, if there are not any images for the word we are getting the directory of the alphabet and adding the images to the list.

Then, we have used the *w*ebbrowser library which comes with Python to export the images to GIF file in order to display them in the new browser.

5 Results

To test the proposed work, we have taken some simple sentences from different users. For the moment, we take simple sentences with 2-5 words not to have problems with the algorithm, thus, we have taken mostly used English words and sentences into consideration.

As an example, we can say the sentence is "I wanted to go to America." In this case, if we tokenize the words, we will get the list of words which are "I", "want", "to", "go", "america". Thus, our application checks every words in the dataset, if it exists or not, so if yes, it adds the image to the list of images. If the image does not exist in the dataset for that specific word, it divides the words into characters and adds the alphabetical character of each word into the list of images. For instance, in our sentence we had "to", so we can show it as an alphabetical characters. The following figure shows the output of the sentence we have mentioned above.

Fig. 3. Output of spoken language into ASL

6 Conclusions

This research paper exhibits an optimal approach, to accomplish the algorithm which makes it possible to get the speech input from the user and apply the Natural Language Processing (NLP) to convert the words to initials forms and display the images of the words in sign language. In order to accomplish this task, we have gathered around more than 1000 images and 26 alphabet letters in English language. For the NLP, we have used specific NLTK library and algorithms to convert the words into the initial form. We also took the some specific punctuations in English language. Thus, if no word is in the database, it checks the characters of the unknown word and displays the images of letters in English alphabet.

7 Future work

This research may be expanded to identify more complex sentences by adding our dataset of images which could explain the phases in English language. This research work can also be extended to recognize the video from the camera for English alphabet, words and sentences.

Acknowledgements

I would like to thank my professor Christian Napoli for his guidance throughout this project. I would also give my huge thanks to my colleagues for testing the application and specifying the accuracyy the have got for this project.

References

[1] American Sign Language, https://www.nidcd.nih.gov/health/american-sign-language

[2] Ss, Shivashankara S, Dr.Srinath. *American Sign Language Recognition System: An Optimal Approach.* . International Journal of Image, Graphics and Signal Processing. 10.