Real-Time Sign Language Translator

Prina Kamakotti

13171780

Aim

The main objective of this project is to create an real-time sign language interpreter in the hopes of promoting communication between people with hearing and speech disability and the world of hearing. To facilitate automated translation, we use Machine learning models. Machine Learning is an application of artificial intelligence which is used to analyse data and apply the results in other situations automatically. Machine Learning enables computers to learn and grow from the data provided, which mimics a human beings ability to grow from past experiences. The main goals of this project are as follow:

1. Image processing to recognise hand gestures.
2. Select appropriate machine learning model and train the model with various sign languages.
3. Validate and deploy the project.

Background

Translating is a tough endeavour even amongst spoken languages, let alone translation of sign language. Sign language is the main form of communication of people with hearing or speech disabilities. According to the World Health Organization, any type of hearing or speech impairment affects about 5 per cent of the world's population. This is a vast number of people affected, nearly 450 million human beings. So sign language to them is the choice of communication. Like every spoken language, which has variations and dialects according to the location of origin, the world currently has more than a hundred different sign languages. To people with no hearing or speech disabilities, learning sign language is usually a neglected cause, hence communicating with people who have such disabilities is extremely arduous.

Machine translation of sign language has been available from 1977, albeit with limited capabilities. Sign Language translation began through the use of hand gesture recognition equipment such as gloves with motion sensors etc. This method to interpret hand gestures has a severe limitation with its efficiency and the movement restrictions. So, with the current digital age development, cameras and digital recordings replaced the old hand glove hardware, making it much easier to interpret sign language. As seen in the objectives, the first hurdle in the translation process is to recognise the hand gestures, and a literature review of this

problem shows that there have been several methods to address this issue. Cohen et al. (2003) proposed an ensemble method which used Hidden Markov Models, which is a statistical machine learning method along with traditional machine learning models. Boulay et al. (2003) developed a method to recognise human postures in video sequences with the help of a three-dimensional human model and to compare the sequence with it. From the literature review, It can be assumed that the many proposed methods can all be used to identify the hand gesture, and it is advised to choose the most suitable method for the data.

The main point of differentiation between the traditional and the proposed method is the fact that the proposed method can perform translations in real-time, contrary to the traditional methods where the translation should be a pre-recorded digital video or photographs. Also, the points of similarity with the traditional methods lie in the recognition stage, where similar methods are used to extract the hand gestures from the person signing.

Real-Time Sign Language Translator

Introduction

Any real-time application of machine learning requires two things, a front-end that captures the data, in this case, it is the video or pictures from a user's webcam and a back-end which should return the predicted translation to be displayed. The front end of the project is responsible for collecting the hand gestures, and a function should be provided to split the video in frames to capture all the hand signs. These gestures are then fed to the back-end, which consists of the recognition model and the prediction model, both of which make use of machine learning algorithms. As the video captured might not only consist of the hand but also the whole human body, it is essential to include a bounding box to isolate the hand signs from the environment. The boundary box can be interfaced on the screen, and the user will have to perform the sign language within the bounding box present in the display. After this is developed the first phase of the project would be complete. Since the back-end is responsible for the recognition and prediction, there should be a significant emphasis on the specifications in this phase of the project, few of the key specifications are as follow:

1. **Data Gathering**: As mentioned in the above sections, there are many sign languages present in the world, hence gathering data about them is crucial. Since this is an unconventional method of sign language translation, data sets for most of the sign languages are not present, hence it is advised to create a dataset for the words required. And this can be achieved by downloading videos of people performing sign language and enclosing a bounding box at each sign with the meaning behind them. This part of the project will have the most manual human work.

2. **Selecting a Machine learning Model for Real-Time recognition:** The second phase of the back-end development consists of selecting and training a machine learning model. Since this project deals with identifying an object, algorithms which deal with object identifying like YOLO(You Only Look Once) algorithms or algorithms like Long Short Term Memory which is suitable for time-series predictions.

3. **Designing the User Interface:** In this stage of the project, it is crucial to make the application user-friendly, as it should be easy to use for people with little or no technical background. Hence the application can be widely used and can bring about benefits to all.

Significance and Innovation

As mentioned in the above sections, the importance of such an application cannot be stressed enough, an application which can produce a real-time translation of sign language benefits the above mentioned 450 million people with speaking or hearing disabilities in the world. Imagine what a person with such disability as that has to go through, a small task like ordering coffee might be immeasurable hard since the other person might not be aware of sign language. Lots of opportunities are lost to such people due to the negligence of the importance of sign language. One such common example would be in the recruitment field; at this time and age, most companies use video conferencing tools to do interviews for jobs, for people who have said disabilities, this is a lost opportunity. Another aspect is the fact that this gives the people with said disabilities a way to fit in; they can be understood; they can get their words across without an interpreter.

The innovation in this project is that it provides real-time sign language translation, conventional sign language communication approaches use digital recordings of signing and perform recognition on it. As machine learning is a robust technology, it is possible to train many sign languages and all of them can be translated.

Project Timeline
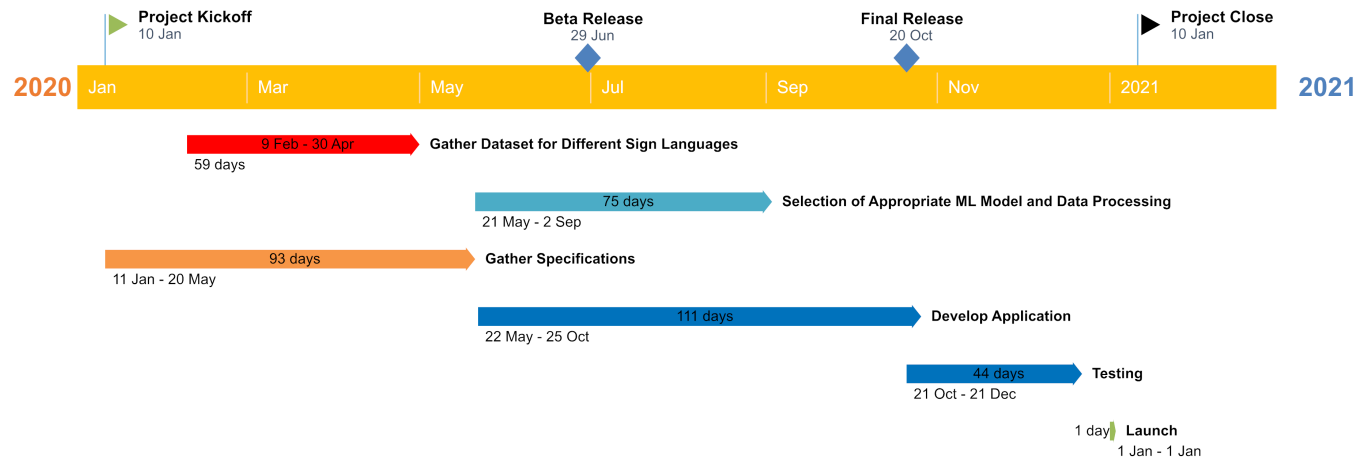
Project Timeline

The figure above depicts the Gnatt Chart for the project timeline, and the project is divided into five subprojects which are as follow:

1. **Data Gathering:** Since this task requires a lot of manual work, a period of two months has been allocated to this sub-project, in this task it is essential to gather data on the different types of sign languages found all over the world.

2. **Specification or Project Scope :** One of the crucial elements of any project is to define the scope of the project, i.e. it is essential to know the boundaries of the project and what we set to achieve before we start, and since there are usually changes made to the specifications made midway, this is an ongoing and lengthy process. In this task, the specifications and scope are modified continuously and checked to see if the project does not deviate from the scope.

3. **Model Selection And Data Pre-Processing:** This task in the project requires the help of data experts who are adept in machine learning models and processing data, selecting the appropriate model and processing is a crucial element and determines the accuracy of the whole project.

4. **Application Development:** In this stage of the project, the interface of the application along with all the underlying features are developed, this stage requires the expertise of software developers who are proficient in front-end.

5. **Testing:** In this stage of the project, the application is thoroughly tested among a select few participants and feedback, and usability is tested. Any bugs or corrections in the applications is fixed during this stage. Once the application passes this stage, it can be launched

Outcome of Project

Once all the scope elements of the project are satisfied, the project is expected to provide real-time sign language translation in various sign languages, from the previous sections we've come to know the extent of people who have hearing and speaking disabilities, Since this application seeks to serve this niche in the market. And since this is a narrow market, having a competitive advantage is easy. Rather than monetary gain, this project seeks to improve the lives of the people living with such disabilities and hopes to provide them with better opportunities in life.

Budget

Product	Quantity	Cost per product (AUD $)	Total Cost
Desktop/Laptops	10	1,500	15,000
Camera	5	700	3,500
Sensors	10	500	5,000
Supporting Softwares	As required	1,000	1,000
Miscellaneous Necessities	-	15,000	15,000
Servers	2	2000	4,000
Office Space (Optional)	-	1200$ per week	57,600
Bills (Optional)	-	150$ per week	7,200
Marketing	As required	10,000	10,000

Staff Position	Base Salary per hour (AUD $)	Total Cost
Project Manager	70	134,400
Release Co-ordinator	60	115,200
Data Analysts	50	96,000
Software Developers	50	96,000
Product Testers	30	57,600
Marketing Analysts	45	86,400
Human Resources	40	76,800
System Architects	60	115,200

Contingency Allowance (15%)	134,385
Total Budget:	1,030,285 AUD

Personnel Required

The personnel required for the successful completion of this project are as follows:

Project Manager: 1

Release Coordinator:1

Software Developers: 5

Data Analyst: 5

Product Testers: 30

Marketing Analysts: 2

System Architects: 5

References

Abe, M., Sakou, H. & Sagawa, H. 1994, 'Sign language translation based on syntactic and semantic analysis', Systems and computers in Japan, vol. 25, no. 6, pp. 91-103.

Akmeliawati, R., Ooi, M.P.-L. & Kuang, Y.C. 2007, 'Real-time Malaysian sign language translation using colour segmentation and neural network', IEEE, pp. 1-6.

Bantupalli, K. & Xie, Y. 2018, 'American Sign Language Recognition using Deep Learning and Computer Vision', IEEE, pp. 4896-9.

Boulay, B., Bremond, F. & Thonnat, M. 2003, 'Human posture recognition in video sequence'.

Bragg, D., Koller, O., Bellard, M., Berke, L., Boudrealt, P., Braffort, A., Caselli, N., Huenerfauth, M., Kacorri, H. & Verhoef, T. 2019, 'Sign Language Recognition, Generation, and Translation: An Interdisciplinary Perspective', arXiv preprint arXiv:1908.08597.

Cohen, I., Sebe, N., Garg, A., Chen, L.S. & Huang, T.S. 2003, 'Facial expression recognition from video sequences: temporal and static modeling', Computer Vision and image understanding, vol. 91, no. 1-2, pp. 160-87.

Deza, A. & Hasan, D., 'MIE324 Final Report: Sign Language Recognition'. Dreuw, P., Stein, D., Deselaers, T., Rybach, D., Zahedi, M., Bungeroth, J. & Ney, H. 2008, 'Spoken language processing techniques for sign language recognition and translation', Technology and Disability, vol. 20, no. 2, pp. 121-33.

Gilbert, J.A. & YU, S.-y. 2009, 'Sign language translation system', Series Sign language translation system Google Patents.

Kau, L.-J., Su, W.-L., Yu, P.-J. & Wei, S.-J. 2015, 'A real-time portable sign language translation system', IEEE, pp. 1-4.

Madhuri, Y., Anitha, G. & Anburajan, M. 2013, 'Vision-based sign language translation device', IEEE, pp. 565-8.

Parton, B.S. 2005, 'Sign language recognition and translation: A multidisciplined approach from the field of artificial intelligence', Journal of deaf studies and deaf education, vol. 11, no. 1, pp. 94-101.

Sakiyama, T., Oohira, E., Sagawa, H., Ohki, M., Sagara, K., Inoue, K., Obuchi, Y., Toda, Y. & Abe, M. 1997, 'Sign language generation apparatus and sign language translation apparatus', Series Sign language generation apparatus and sign language translation apparatus Google Patents.

Starner, T., Weaver, J. & Pentland, A. 1998, 'Real-time american sign language recognition using desk and wearable computer based video', IEEE Transactions on pattern analysis and machine intelligence, vol. 20, no. 12, pp. 1371-5.

Tolba, M. & Elons, A. 2013, 'Recent developments in sign language recognition systems', IEEE, pp. Xxxvi-xlii.

Zhao, L., Kipper, K., Schuler, W., Vogler, C., Badler, N. & Palmer, M. 2000, 'A machine translation system from English to American Sign Language', Springer, pp. 54-67.

Video Pitch Link

https://youtu.be/E9jNFtC7BGM

GitHub Link

https://github.com/prinakk/A3_ML_UTS2019