

American Sign Language Interpreter

Scope:

The scope of our project entails creating a system to assist with the use and learning of American Sign Language through image processing, machine learning and web-app database querying. The main objective of this project is twofold. The first goal is to be able to track hand gestures with a camera, recognize various ASL words or letters and translate these words or letters into text. The second goal is to be able to assist with learning sign language. This would be done by providing the program with text, and actively tracking a user's hand gestures while providing feedback on the accuracy of the gestures compared to the input text.

Requirements:

The key technical challenges will include successfully recognizing the asl hand signs as words, in noisy environments that vary in lighting, angle, and interfering objects in the background. Since a language includes many words and phrases, we will likely need to limit the capacity of our project to recognize the letters and a subset of the words in ASL so that we can train the program within a reasonable time frame. Additionally, there are words in ASL that are motions as opposed to single frame pictures. The inputs could be formatted as multi-frame videos instead of pictures to recognize these moving signs. Since we want our product to translate in real time for communication, the processing and evaluation of the signals should ideally happen within a few seconds at most, with accuracy above 95%.

Implementation:

The implementation of our American Sign Language interpreter will utilize a high frame rate and resolution camera for accurate gesture recognition of the signs. We will process the image feed from this camera to train our neural network to better recognize the gestures that make up ASL. To this end we will require Amazon Web Service credits (or an equivalent) to facilitate accelerated neural network training. We will also require a tablet/screen for text and speech feedback generated from the interpretation of ASL. The main bulk of the implementation will require us to develop a hand tracking algorithm that utilizes the high frame rate and resolution camera to accurately determine the position and orientation of the hands and possibly by extension arms. This data will then be used to train our neural network to recognize the ASL signs which we will then convert to text/speech on our output tablet. Encompassing all this will be a web app which we will develop to simplify the experience of gesture recognition for normal users and enable us to also create an ASL teaching tool.

Minimum Viable Product:

Our MVP will be a system running primarily through a web-app that will recognize ASL words and phrases with a 95% and be able to translate them into English text/speech. Alongside that, it will be able to receive text input and track a user's attempt at the ASL translation of the input text. This system should also be able to provide a response to gestures and hand signs within 750ms so that the user experience is fairly seamless.