We Talk

An initiative towards hearing more – We Talk is an innovation towards making the world better for the deaf and dumb people around us. Our objective is to bring together a culmination of AI driven touch and gesture systems along with smart recognition of hand gestures to convert these to speech, and drive speech recognition to go beyond textual conversion by linking these through a library of speaking systems followed by the hearing impaired. This mammoth initiative hopes to enhance gesture recognition processes as a touchless way of advanced Human-Machine Interaction which can be utilized by everybody.

The team

Our team is nothing special but are just a bunch of laymen with a dream of achieving greatness in life. United together in our sophomore year of college, we went from nobody to being roommates to winning the SIH (Smart India Hackathon) together. Now we dream of achieving something more and imagine cup provides us the perfect opportunity to showcase who we are.

Sreyansh Baranwal, BNM Institute of Technology, Computer Science Enginering, Grad. Year – 2024, Lazy but productive, an embodiment of patience and enthusiasm

Ajitesh Kumar Soni, BNM Institute of Technology, Computer Science Engineering, Grad. Year - 2024, The Tech savvy individual who has a tech point with every statement he makes

Soham K Chattopadhyay, BNM Institute of Technology, Computer Science Engineering, Grad. Year - 2024, The guy who has his way with his words and his tech, the guy who has it when it counts

Abhishek Ranjan, BNM Institute of Technology, Computer Science Engineering,
Grad. Year – 2024, The compiler and the data master who bring it all together.

Mentors:

Self-mentored

The Concept

Communication through visible cues becomes critical for achieving business, social and personal outcomes for the differently abled. Technology can be a powerful partner in this regard with our persona of choice, 'hearing impaired' people who are aware of the digital realm and ore familiar with smartphones or personal computers.

Indeed, a lot of mechanisms help the deaf community understand and be understood by hearing people thanks to artificial intelligence delivered subtly through mediums like meeting software like Zoom, Teams etc. which require multiple switches between video, audio and textual consumption. AI can truly become an asset to enhance accessibility and inclusion, in this regard.

Our objective is to bring together a culmination of AI driven touch and gesture systems along with smart recognition of hand gestures to convert these to speech, and drive speech recognition to go beyond textual conversion by linking these through a library of speaking systems followed by the hearing impaired. This mammoth initiative hopes to enhance gesture recognition processes as a touchless way of advanced Human-Machine Interaction which can be utilized by everybody.

Target Audience or Market:

As of December 2021, according to the World Federation of the Deaf and dumb, there are more than 70 million deaf and dumb people worldwide.

A World Health Organization report says around 63 million people (about twice the population of California) in India suffer from either complete or partial deafness, and of these, at least 50 lakhs are children.

To ensure our technology is implemented with collaboration with tech giants such as Microsoft, we aim to provide a better world for such especially abled people especially when they face grievances during online video calls

Personas

People who are innovative and are unable to communicate due to their disability often feel discarded and rejected in our world. In the world of the internet, it makes it near impossible for such people to have a livelihood. To prevent such scenarios from occurring and to make sure they feel normal about their condition and can communicate, our technology comes to a rescue here.

Small case implementations are the start, and the end is when such people no longer feel left behind both in the real and the virtual world.

Feedback

The project built by the team is the first under its kind. Since it has been under development, the opportunity of getting feedbacks haven shown up yet.

The concept around which our idea is built upon has received a lot of praise worthy compliments and comments that has kept the team motivated to continue working.

How it works:

Phase 1: How a sign language translation platform that enables virtual interaction between hearing and speech impaired individuals might work:

The platform is a sign language translation platform that allows hearing and speech impaired individuals to communicate virtually.

When you first use the platform, you are prompted with a "train gestures" page where the website asks you to upload and train two gestures: one for starting a sign, and one for stopping a sign. These gestures are used to mark the beginning and end of a sign language word or phrase.

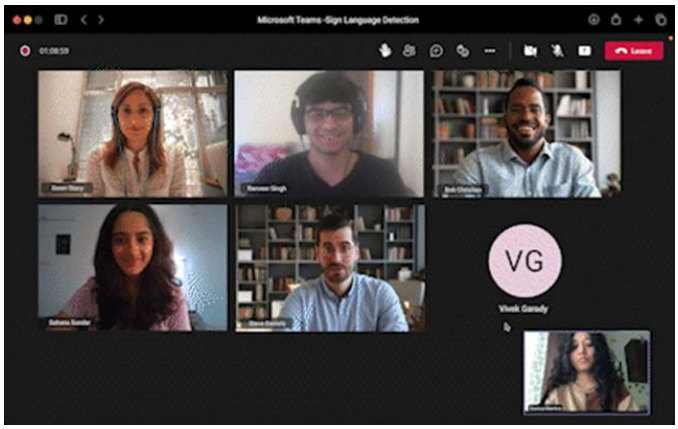
After training the start and stop gestures, you can then upload and train new words and their corresponding gestures on the platform. These words and gestures can then be used in the next feature of the platform, which is video calling.

To use the video calling feature, you simply select a contact from your list of friends or colleagues and initiate a call. During the call, you can use the trained gestures and words to communicate with your contact. The platform will automatically translate your gestures and words into text or audio in real-time, allowing your contact to understand what you are saying.

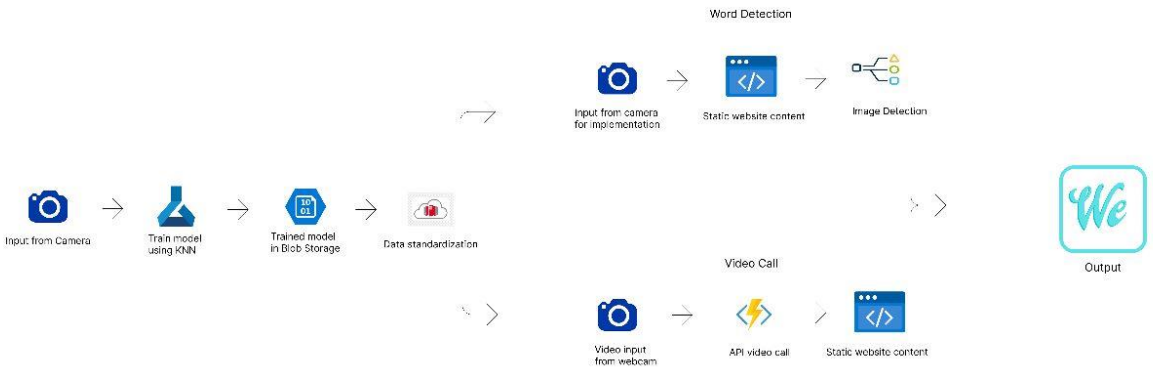
The platform also includes features like instant messaging and group chat, which allow you to communicate with multiple contacts at once using the trained gestures and words. This makes it easy for hearing and speech impaired individuals to stay connected and communicate with others, even if they are not in the same location.

Overall, this sign language translation platform provides a convenient and accessible way for hearing and speech impaired individuals to communicate with others virtually, using sign language gestures and words.

Phase-2: Integrate phase 1 into online video calling platform like Microsoft teams for better user experience



How it works:



This project takes the input from the web-camera from the device through which the user needs to access the data.

After the access is granted, the model is trained by various hand gestures and it is implemented using the KNN algorithm. A detailed study about the algorithm is given in the below section.

The data is then absorbed and stored in the blob storage of the azure storage. This is the primary piece of azure technology that is used in our project.

The data undergoes scrutinization and standardization and gives valid input for the machine to consider.

The data taken from the camera the first time is for the start and the stop gesture, now the input is taken again for the actual word that has to be translated. Image detection occurs and the model is further trained.

Further as a feature, we have a video call option which shows the real time implementation of our project and shows also the future opportunity for us to work on our project

These services from azure of data storage and we page hosting are very helpful as they make the process very easy, efficient, time-conserving and resource-conserving

Core Technologies

K Nearest Neighbour: -

KNN (K-nearest neighbours) is a classification algorithm that works by finding the K-closest examples in the training data and using them to predict the class of a new data point. It can be used to classify sign language gestures by finding the K-nearest examples in the training data and using their classes to predict the class of the new gesture.

To use KNN for sign language recognition, you would first need to collect a large dataset of sign language gestures and their corresponding labels. Then, you would need to extract features from each gesture, such as the position and orientation of the hands and fingers, the movement of the arms and body, and so on. These features would be used to represent each gesture as a vector in a feature space, which can be used to find the K-nearest neighbours.

Once you have trained a KNN model on your sign language dataset, you can use it to classify new gestures by finding the K-nearest neighbours in the feature space and using their labels to predict the class of the new gesture. For example, if you are trying to classify a new gesture as "hello," "goodbye," or "thank you," the KNN model would find the K-nearest gestures in the training data and use their labels to predict the class of the new gesture.

Azure Storage:

Azure Storage is a Microsoft-managed service that provides storage for structured and unstructured data in the cloud. It includes a variety of storage options, including Azure Blob storage, Azure File storage, Azure Table storage, and Azure Queue storage.

Azure Blob storage is a service for storing large amounts of unstructured object data, such as text or binary data. It is typically used to store data that is frequently accessed, such as images, videos, and documents.

Azure File storage is a service that provides managed file shares in the cloud. It is commonly used to store files that are shared among multiple applications or between users.

Azure Table storage is a service that provides a NoSQL key-value store for semi-structured data. It is used to store large amounts of structured data that does not require complex relationships between data items.

Azure Queue storage is a service that provides a message queue for storing and retrieving messages. It is used to decouple applications and enable communication between them.

Azure Storage offers high availability, durability, and scalability, making it suitable for storing and managing large amounts of data.

Azure web hosting:

Azure is a cloud computing platform and infrastructure created by Microsoft for building, deploying, and managing applications and services through a global network of Microsoft-managed data centres. Azure offers a variety of services for hosting web applications, including Azure App Service and Azure Virtual Machines.

Azure App Service is a fully managed platform-as-a-service (PaaS) that enables you to build, deploy, and scale web, mobile, and API apps. It includes support for multiple languages and frameworks, such as .NET, Java, Node.js, PHP, and Python. Azure App Service also provides features like automatic scaling, built-in

monitoring, and integration with Azure Active Directory for identity management.

Azure Virtual Machines is a service that enables you to create and manage virtual machines in the cloud. It allows you to deploy a custom image or choose from a range of preconfigured images, including Windows and Linux operating systems. With Azure Virtual Machines, you have full control over the configuration of your virtual machine and can install and run any software you need.

The Business Plan:

The plan is to incorporate our services with Microsoft Teams and make it viable software for all the especially abled people, namely the deaf and dumb. With the noble cause in mind, the plan involves either selling the API to a client for a substantial amount of loyalty or by establishing a subscription-based mechanism for users who are willing to avail further services.

Competition:

Mainly, we have 2 competitors in this field SiMAX a German based company and hand talk which targets to sell on the text to sign translation to another website, however, no one is considering speech to text conversion this is where our solution come to play which is to provide real-time two-way sign to text and speech to sign translations using our model in online meeting platform like Microsoft teams etc. and to provide translation services through our project.

Business Model

We aim to provide sign translations services to different websites through our application which will enable us to reach more audience at the same time make revenue which help our idea to sustain in the market.

As we proceed towards the future, online education is becoming a boon and thus for our project to cope up and generate profit with the transition of modes

from offline to online learning, a subscription-based mechanism is offered in which case, the school willing to teach sign languages via this method generally pays an amount to avail the services of our project.

This not only is a milestone for the project but also for human evolution for adaptation towards such technology which unites the world and gets rid of stigma.