Team Name : Pravachaka

(Automatic Closed Caption generation using AVSR)

	Name	Branch and Semester	Contact Number	Email- ID
Team Leader	Suhas R	EC – 6th Sem	8197413187	suhas7thfeb2000@gmail.com
Member 1	Prashantha T P	EC – 6th Sem	7899571984	tpprashantha@gmail.com
Member 2	Karthikeya Moleyar	EC – 6th Sem	9844459828	moleyarcentral@gmail.com
Transaction ID (anju.marina.lobo@oksbi)	P200223174020496729958O			

Note:

1. One can participate either as a part of a team or an individual basis. Switching teams is not allowed.
2. The uploaded ideas will be screened to go to the second round.
3. Judging : competition entries shall be judged, or winners selected based on the following criteria
   - Is the problem worth solving
   - How innovative or novel is the idea
   - Scientific accuracy
   - Social impact
   - Scalability
4. Decisions of IIC JSSSTU in respect of all matters to do with the competition will be final and no correspondence will be entertained.
5. In second round, the selected teams will have to present their idea in front of the jury panel.
6. Payment of INR 50 should be made to the UPI ID **anju.marina.lobo@oksbi** and submit the transaction ID above.
7. Idea should be submitted in **.pdf** format.

Abstract:

The closed caption creation is coming to existence as a new industry with boom in the video data on the internet. This process is manual and many of the times it is not accurate and is also not in synchronous with the visuals and speech of the video. This process can be made faster and also efficient if machine learning is incorporated to it. Instead of merely listening to the audio and typing it, we can train the computer to train a model to automatically use both audio and video signals to predict the word being said in the video. This idea can be transformed into a product to generate closed captions for the video viewers or also open caption generator for the video makers.

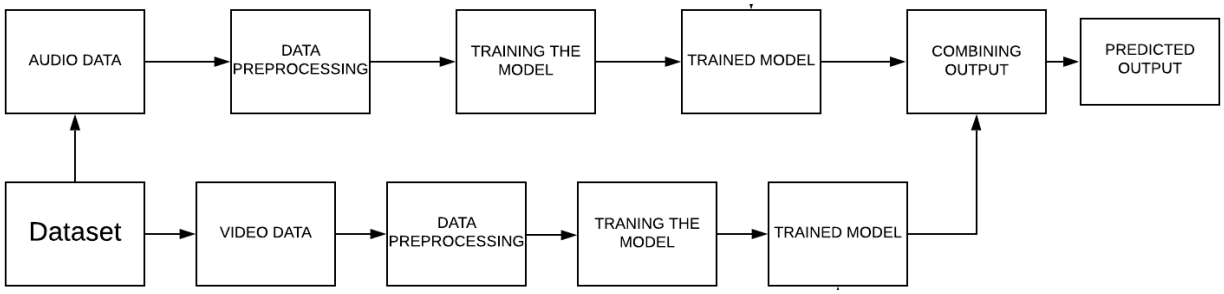
Introduction:

We all enjoy watching videos and movies. These videos have become an integral part of our life today. Everybody watches at least one video daily. It can be a standup comedy, news, movies, speeches or even educational videos. All these things have something in common the audio information. It is common observation that the most of the times, people opt for the closed captions, in order understand what the person in the video is saying even if the person has knowledge of the language of the speaker, especially while watching movies or watching some educational videos. This can be because of the noisy environment and the difficulty due to speaker's accent. The closed captions are truly a boon when this is the case. How are these closed captions prepared? They are either done by a stenographer, who listens to audio and types the closed captions or by looking into the scripts. In other words it is done manually. So this idea of ours tries to reduce the work of manual typing of the closed captions and also increase the accuracy of the closed captions.

Motivation:

To prepare the closed captions for a movie it takes about 3 to 4 weeks. In the same way, each and every video released requires a considerable amount of time to prepare closed captions. This makes many videos to be released without any closed captions. Having closed captions also allows the people who are hard of hearing to read the conversation or dialogue of the video. Implementing this idea will also be a contribution to the rapidly developing the artificial intelligence field. Having the closed captions of any video also makes it easy to translate the video into different languages.

Methodology :



As shown in the block diagram, we will train a model to predict the word that is being said by the person in the video based on the visual and audio features of the video. Implementing this project will surely require a lot of computational power as the number of words being used in the world is enormous and the pronunciation differs from a place to place.

Social Impact:

For small scale video maker, it is not feasible to make closed captions for all his video. He may either not have enough to time to spend time to make closed captions for his video and may not even be able to afford someone to create closed captions for him. So this idea of automated closed captions will surly help him to create one for his video.

The closed captions improve the level of understanding. The closed captions in video helps to enhance viewers the vocabulary and also to make digital notes for an educational videos.

The automated closed captions also will significantly reduce the language barrier. It is easier to translate a text into any desired language compared to the direct audio signal. Thus contributing a lot to increase in viewership.

Market Survey :

According to last updated facts by WHO (World Health Organization), about 466 million people around the world are deaf and also estimates by 2050, the count will reach to 900 million people who are hard of hearing. These are the people who cannot enjoy the videos as the other customers do. So, this increases the viewership of the videos and in turn increases the income of the video creator.

There will also be a rise in the usage and traffic of internet, which will benefit the internet service providers and also has potential for establishment of new service providers.

This idea can be converted into a software. This software can be marketed to potential customers like film makers, video creators and even TV channels who can convert it into the open captions.

The data is like the fuel which drives the 21st century. This idea to become a reality requires a huge amount of data. The dataset prepared itself will be highly valuable.