



Making Alexa Responding To Sign Language

Chahat Sheikh¹, Imbesat Asad², Lubna Ali³, Sinela Anwar⁴, Sofiya Khanam⁵

Student, Department of Computer Science and Engineering, ACET, Nagpur, India¹⁻⁵

Abstract: Alexa usually responds to voice commands but what about the people who cannot hear and speak. Everyone should adapt new technologies which will ease their life. So to overcome this problem we are Making Amazon Echo to Respond to sign language. This device will help the deaf mute to interact with Alexa by showing gestures and actions. Here camera will be used to interpret signs from the user then the signs will be converted to text and speech and then Alexa's response will be transcribed. For developing this device there is need of A neural network to interpret the signs, A text to speech system to speak the interpreted sign to Alexa by using Web Speech API for speech synthesis to speak out the detected label, A speech to text system to transcribe the response from Alexa for the user again by using Web Speech API for transcribing the Echo's response which responds to the query clueless to the fact that it came from another machine. And then an interface for which we have to design a browser which connects the network and web Speech API which can be opened in laptop with webcam this interface will combine all together and make Alexa to respond sign language.

Keywords: Alexa, Amazon Echo , CNN, RNN, LRCN.

I. INTRODUCTION

Alexa usually responds to voice commands but this device will help the deaf mute to interact with Alexa. Here camera will be used to interpret signs from the user then the signs will be converted to text and speech and then Alexa's response will be transcribed. If voice is the future of computing interfaces, what about those who cannot hear or speak. So to resolve a part of this problem we are designing a project on Making Amazon Echo to Respond to sign language. The things we are needed are a neural network to interpret the signs (i.e. convert video of signing into text). A text to speech system to speak the interpreted sign to Alexa, a speech to text system to transcribe the response from Alexa for the user, A device (laptop/tablet) to run this system and an Echo to interact with and an interface that ties this all together.

There is both a visual and temporal aspect to a sign, so to combine a CNN with an RNN where the output of the last convolutional layer (before classification) is fed into an RNN as sequences. The technical term for these is Long-Term Recurrent Convolutional Networks (LRCN).

By using a 3D convolutional network where the convolutions would be applied in three dimensions, with the first two dimensions being the image and the third dimension being time. However these networks require a lot of memory. Instead of training a CNN on individual frames from a video stream, instead train it only on optical flow representations which would represent the pattern of apparent motion between two consecutive frames. With this was that it would encode the motion which would lead to a more general model for sign language. Use a Two Stream CNN where the spatial stream would be a single frame (RGB) and the temporal stream would use the optical flow representation.

II. PROPOSED SYSTEM

Alexa usually responds to voice commands but this device will help the deaf-mute to interact with Alexa. Here camera will be used to interpret signs from the user then the signs will be converted to text and speech and then Alexa's response will be transcribed.

There is an existing system which is an artificial intelligence app that acts as an interpreter between deaf people and Alexa. For it to work, users must sign at a webcam that's connected to a computer. The app translates the ASL signs from the webcam into text and reads them aloud for Alexa to hear. When Alexa talks back, the app generates a text version of the response for the user to read.

III. IMPLEMENTATION AND DISCUSSION

The Proposed System will be implemented according to the given flow:

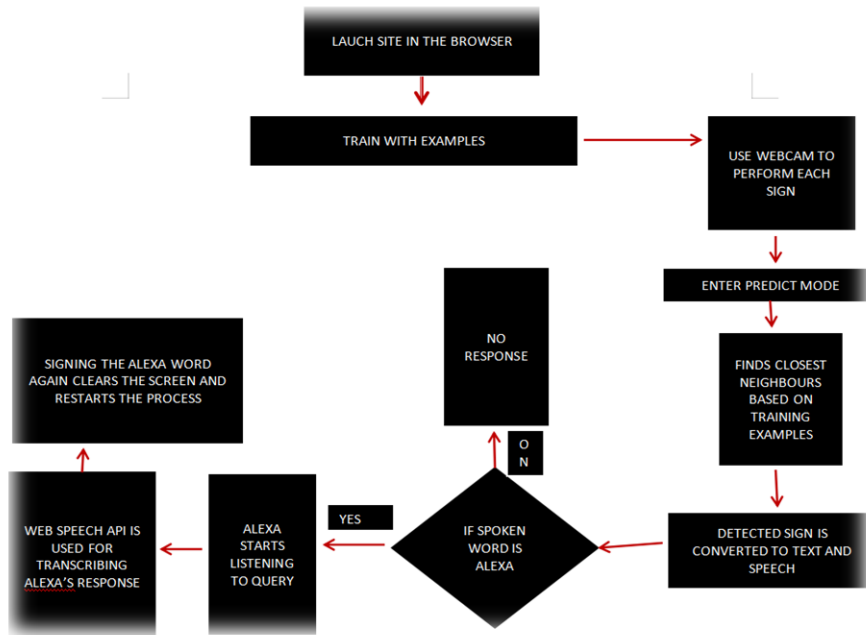


Fig. 1 System Flow

IV. CONCLUSION

Alexa cannot respond to the people who cannot hear and speak, so with this system we can overcome this problem. And the people who cannot hear and speak will be able to adopt this technology, which will ease their life. In the future this system can be converted into an app and can be customized according to the preferred language of the user.

REFERENCES

1. React js:<https://reactjs.org/>
2. Blog on Dialogue Management:<https://www.amazon.science/blog/science-innovations-power-alexas-conversations-dialogue-management>
3. <https://www.slideshare.net/AmazonWebServices/fsi300-bringing-the-brains-behind-alexas-to-financial-services>
4. <https://smarternutter.com/how-does-alexas-work-technically/>
5. Blog on Amazon Lex: <https://techcrunch.com/2017/04/20/amazon-lex-the-technology-behind-alexas-opens-up-to-developers/>
6. <https://towardsdatascience.com/how-amazon-alexas-works-your-guide-to-natural-language-processing-ai-7506004709d3>