

SiNext: The next generation sign language

Shivanshu Bajpai^a, Shubham Sharma^b, Shivanshu Debrani^c, Vikas Srivastava^d

^{a,b,c,d} Bachelor of technology Computer Science and Engineering Galgotias University, Greater Noida, India

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: In today's smart world people are using smart assistants that made their life easier than ever before. The only principle in this technology is the human voice which plays a key role in this. If the voice is the future of computing what about those who cannot hear or speak? The hearing and speaking impairment society aren't able to communicate with it. They can only communicate via sign language.

The aim is to create a model that can speak up on behalf of the user to the smart assistants so that they can also communicate with them and they can also take advantage of modern technology like home automation.

The model uses 'deplearn-knn-image-classifier' to classify the images and 'knn.predictClass' to predict the signs. Model accuracy during the testing was 90.25%.

Keywords: Sign language translation, Gesture recognition, Virtual assistant, Artificial intelligence, Hearing and Speaking Impairment.

1. Introduction

There are 360 million people over the world who are suffering from visual or speaking impairments. This problem makes them feel that they are able to use modern technologies and the world of virtual assistants.[1]

They can only communicate with the help of Sign Language because of their disabilities. Different regions of the world use different types of sign language to communicate with each other. Only India holds 250 languages to communicate with deaf people of 7 million. This idea can reduce the number of languages and make people able to communicate with the virtual world.[2]

Sign language is a language of communication that is completely recognizable by its traditional system, unlike any other language.

Objective and Working:

- The objective is to make a sign language application that can be trained by anyone and people can use it according to their needs.
- The Model can be trained by using any type of symbol according to the user's needs and then after the successful training of the model, it will predict the outcomes according to the symbols and display it on the screen as well as output it from the speaker.
- Pre-processing of the signed input action is performed in the first stage. In the next section, the various regional structures of the previously processed touch image are calculated. In the last section, based on the listed structures of the previous section, the conversion of the signed hand to the text and audio was done.

2. Literature review

The lack of communication can cause the children to learn the basic concept of daily life that can be very essential[3]. This makes them incompetent, unskilled and above all, it makes them dependent on others.

Although they have done a decent job in this paper, their technology is limited only for communicating to each other but not to any smart assistants like ours. We are expanding it to the level so that they can use smart assistants.

In this paper[4] where they are trying to make an electrical circuit based on Arduino. The big problem with Arduino is that it can be a short circuit and deaf people will not be able to find buttons or objects. But in our solution one will not need to touch the system, it will be automatically translated and the assistant will work. As our system does not require any electrical circuit so there will be no risk factors.

With the help of SINEXT the speech-impaired person is able to express his/her thoughts, ask questions, clear doubts, and plenty of other things. Individuals can train the model as per his/her likings which will make it easier to use on a day-to-day basis. Sign language can be hard to remember at times so it is better the individual uses his/her own signs and symbols to express themselves.

3. Methodology

A: Application Summary

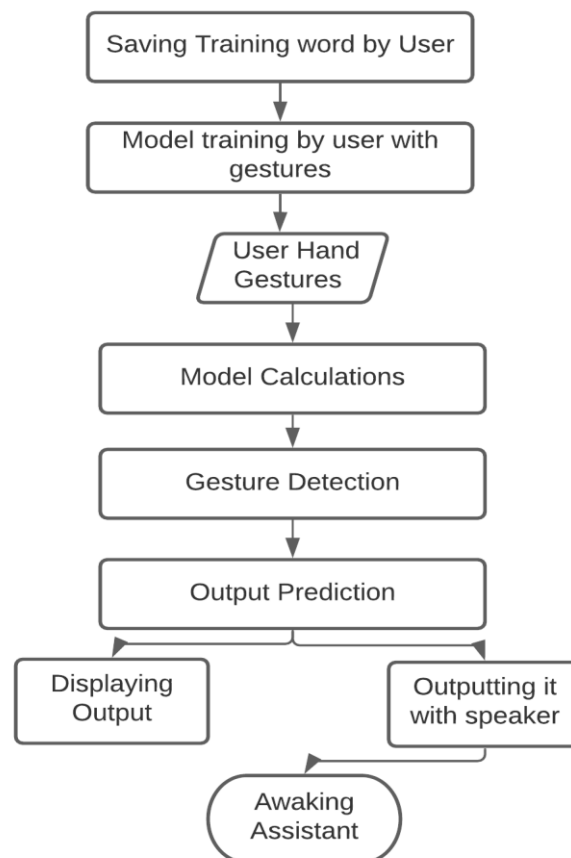


Figure 1. summary of Application working.

This page is designed for the application summary and it describes how our application is going to work and how you can use the application.

The summary of the application is that it will save the desired keywords for the gestures that he/she wants to say to the smart assistant to do or to ask the things. Then the model will train according to the gestures saved by the user. When it comes to detecting the gestures it will simply take the picture of the user doing the gesture with the help of a camera and convert it into pixels and calculate its threshold and matches with the trained data's threshold and predicts the output for that particular gesture. The output will get converted into text and speech so that smart assistants can listen to it as well as deaf people can also listen to it.

B: Registration

This page mainly focuses on the keyword that the user is going to say and based on those keywords the model will train.

1. // const image = dl.fromPixels(this.video);

This constant will save the pixels taken from the image.

dl is a deeplearn instance imported as dl.

this.video refers to the video instance captured during the training.

2. // this.knn.addImage(image, this.training)

It will add the current image to the KNN classifier for classification.

this.training refers to the training phrase.

3. //const exampleCount = this.knn.getClassExampleCount()

This will count the examples for the particular image and save that same image those times so that our model predicts fluently.

C: Model Training

This page is used to train the model. Users can train their model by using the appropriate sign for the particular word and there will be a default sign too for error catching.

```
// this.knn.load()
    .then(() => this.startTraining());
}
```

This particular line will train the model for us based on the saved images into the KNN addImages matrix using the algorithm K - nearest neighbor [Wikipedia].

Let $C_n^{w_{ni}}$ denote the weighted nearest classifier with weights $\{w_{ni}\}_{i=1}^n$. Subject to regularity conditions ^[further explanation needed] on the class distributions the excess risk has the following asymptotic expansion^[12]

$$\mathcal{R}_{\mathcal{R}}(C_n^{w_{ni}}) - \mathcal{R}_{\mathcal{R}}(C^{Bayes}) = (B_1 s_n^2 + B_2 t_n^2) \{1 + o(1)\},$$

for constants B_1 and B_2 where $s_n^2 = \sum_{i=1}^n w_{ni}^2$ and $t_n = n^{-2/d} \sum_{i=1}^n w_{ni} \{i^{1+2/d} - (i-1)^{1+2/d}\}$.

The optimal weighting scheme $\{w_{ni}^*\}_{i=1}^n$, that balances the two terms in the display above, is given as follows: set $k^* = \lfloor Bn^{\frac{4}{d+4}} \rfloor$,

$$w_{ni}^* = \frac{1}{k^*} \left[1 + \frac{d}{2} - \frac{d}{2k^{*2/d}} \{i^{1+2/d} - (i-1)^{1+2/d}\} \right] \text{ for } i = 1, 2, \dots, k^* \text{ and}$$

$$w_{ni}^* = 0 \text{ for } i = k^* + 1, \dots, n.$$

With optimal weights the dominant term in the asymptotic expansion of the excess risk is $\mathcal{O}(n^{-\frac{4}{d+4}})$. Similar results are true when using a [bagged nearest neighbour classifier](#).

Figure 1.1 how K - nearest neighbor weighted the nearest classifier.

The software is used to help visually and speaking impaired people to meet up the virtual world of virtual assistants without speak. Inability to communicate between deaf-mutes and smart systems. Accessibility issues to virtual assistants that arise for people with speech disabilities.

Sign language is the only obvious mode of communication well known and recognizable to them.

The data will only be stored in the user's machine with the help of browser cookies and it will get deleted once the cookies are deleted. So we are maintaining user data confidentiality too.

4. Application Workflow

The working of the application describes in the following steps:

1. The user will train the model by saving the text that he/she wants to say on behalf of their voice to the smart assistant.
2. The user will train the model with the gestures respectives to the text that he/she wants to say to the smart assistant according to their convenience.
3. Once the models get trained with enough examples it will start predicting the examples using the predefined library from tensorflow.js named KNN Image classifier.
4. When the user will show the symbols to the machine it will capture the picture with the help of a camera and convert it into pixels and check if it will match any trained threshold before or near that threshold. If it is so then it will display the output accordingly.
5. Then it will convert the text into speech for the people who are deaf so that they can hear that output.
6. And the last thing it will start listening to the output from the smart assistant for the query response so that deaf users will get the output from the smart assistant.



Figure 2. Actions For Performing Some Operations[7]

5. Results

Comparative Study

A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
Y	Z				
0	1	2	3	4	
5	6	7	8	9	

Figure 3. for helping to use signs during the model training

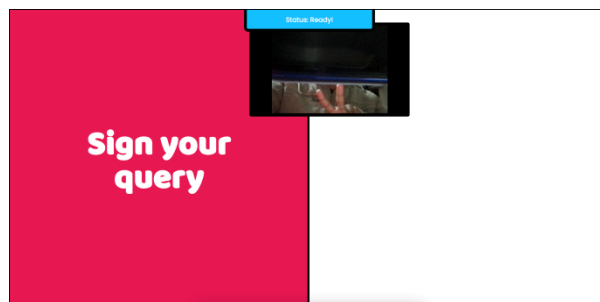


Figure 4. Response according to the symbol

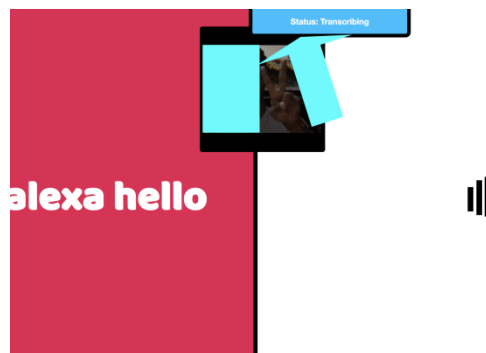


Figure 5. Response of victory symbol to 'alexa hello'

6. Conclusion

1. As there are approximately 360 million speech-impaired people on Earth[1] and most of them find it very difficult to communicate. So with SINEXT, this very problem can be addressed efficiently.

2. The biggest advantage of this model is that it can be used with home automation so that people who are not able to walk, can also take it into use.
3. Moreover, the model is not pre-trained, so the user can mold it or train it as per his requirements.
4. There will be no barriers or translators. It will be easy to communicate with each other as a platform makes it easy to use.
5. With SINEXT we are giving speech impaired people to upgrade the meetup skills that can be life-changing for them.

7. Some aggregated data sources

Only India holds 250 sign language translators for over 7 million visually impaired people[5]. About 5% population in the world is struggling to express their ideas and represent to the world by not saying anything or lack of connection between the world.

8. Future Optimization

This can be also used in the creation of home automation so that each and every single unit in the house can be controlled by the signs.

References

1. Mundial, Banco. "Informe mundial sobre la discapacidad 2011." (2011). [Google Scholar]
2. L. Boppana, R. Ahamed, H. Rane and R. K. Kodali, "Assistive Sign Language Converter for Deaf and Dumb," 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), Atlanta, GA, USA, 2019, pp. 302-307, doi: 10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00071.
3. Kumar, Vinay K., R. H. Goudar, and V. T. Desai. "Sign language unification: The need for next generation deaf education." *Procedia Computer Science* 48 (2015): 673-678.
4. Das, Abhinandan, Lavish Yadav, Mayank Singhal, Raman Sachan, Hemang Goyal, Keshav Taparia, Raghav Gulati, Ankit Singh, and Gaurav Trivedi. "Smart glove for Sign Language communications." In 2016 International Conference on Accessibility to Digital World (ICADW), pp. 27-31. IEEE, 2016.
5. Hai, Pham The, Huynh Chau Thinh, Bui Van Phuc, and Ha Hoang Kha. "Automatic feature extraction for Vietnamese sign language recognition using support vector machine." In 2018 2nd International Conference on Recent Advances in Signal Processing, Telecommunications & Computing (SigTelCom), pp. 146-151. IEEE, 2018.
6. Harini, R., R. Janani, S. Keerthana, S. Madhubala, and S. Venkatasubramanian. "Sign Language Translation." In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 883-886. IEEE, 2020.
7. He, Siming. "Research of a sign language translation system based on deep learning." In 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM), pp. 392-396. IEEE, 2019.