Sign Language Translation System for Dumb and Deaf

S. S. Kadam¹, Bhagyashree Sarode², Chinmayee Modak³, Khushboo Singh⁴

¹Professor, Dept. of Information Technology, Bharati Vidyapeeth College of Engineering, Navi Mumbai, India
²,³,⁴Student, Dept. of Information Technology, Bharati Vidyapeeth College of Engineering, Navi Mumbai, India

Abstract: Communication is a very basic and important necessity for human beings. People convey their feelings, needs, ideas through communication. But there are lots of people who are born dumb and deaf or become dumb or deaf for years due to some medical problems. These people face a lot of difficulties in communication. Sign language is used by these people for communication. These sign languages are made using human gestures mostly involving hand gestures. A translator is required for communication between a person who knows sign language and a person who does not know the language. But many situations occur where the translator is not available. This creates a communication gap. This gap can be reduced with the help of deep learning. The main motto of this work is to provide the sign to text and audio translation. The proposed system captures the image using OpenCV and passes it through different layers of the Convolutional neural network. The sign is then converted to the text. And this text is converted to audio. This will help non-sign speakers to communicate more efficiently.

Keywords: Sign language, Deep learning, OpenCV, Convolution Neural Network, Text to speech conversion.

1. Introduction

Having strong communication skills aids in all aspects of life – from professional life to personal life and everything that falls in between but unfortunately people with hearing impairment and speaking inability find it difficult to communicate. Sign language is a ubiquitous solution to overcome this barrier. But non-sign speakers find it very difficult to understand this non-verbal means of communication as a result Sign language interpreter is required to fill the communication gap.

The presented work tends to build a system that can be used by a non-sign speaker to understand the sign language. The proposed system will be trained on sign language alphabets and common English phrases. The system will be the Convolutional neural network.

The extracted features will pass through various layers of CNN such as Convolutional Layer, Pooling Layer, and Fully Connected Layer. The output layer will use Softmax to predict the results.

2. Motivation

We got motivated to build such a product by observing our society on the daily basis. While travelling in the local transport the dumb and deaf people face many difficulties in communicating their feelings to the person who is having no knowledge of sign languages. There come times when a person wants to communicate with the dumb and deaf person but due to the lack of knowledge of the sign language the communication becomes difficult. Such a product will help to decrease this communication gap and help in the social development of the society.

3. Methodology

The input is captured through the camera using OpenCV and passed through the CNN classifier. In the classifier, the captured input is initially taken and passed through the convolution layer. This layer extracts features from the images. The next layer is the Pooling layer, which reduces the number of parameters when the images are too large. The number of convolution layer and pooling layer can be increased for more accuracy.

After passing through several convolution and pooling layers the output is flattened into a vector and sent into a fully connected layer. The last layer is the Softmax layer. This layer is used to show the output. The output is generated in text format. Further, the text is converted into an audio format using Google text to speech API.

Fig. 1. Methodology

4. Requirements

A. Dataset

The dataset is built for all the alphabets and numerical using hand gestures. The dataset contains a large number of samples of images in different positions and conditions for different scenarios. This helps in enhancing the overall result.
CNN model takes an input image, process it and classify it under certain categories. In CNN the input images are passed through several layers like the convolution layer, Pooling layer, fully connected layers (FC) and apply Softmax function.

B. **CNN Model**

<table>
<thead>
<tr>
<th>Sign</th>
<th>Alphabet</th>
<th>Sign</th>
<th>Alphabet</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>R</td>
<td>B</td>
<td>M</td>
</tr>
<tr>
<td>C</td>
<td>O</td>
<td>D</td>
<td>Y</td>
</tr>
<tr>
<td>E</td>
<td>P</td>
<td>F</td>
<td>V</td>
</tr>
<tr>
<td>G</td>
<td>K</td>
<td>H</td>
<td>U</td>
</tr>
<tr>
<td>I</td>
<td>X</td>
<td>L</td>
<td>Q</td>
</tr>
</tbody>
</table>

![Sign language dataset](image)

Fig. 2. Sign language dataset

![Layers in CNN](image)

Fig. 3. Layers in CNN

C. **OpenCV**

OpenCV stands for Open Computer Vision which is a library in Python used for Image Processing. This will be required to perform various actions and processing of the captured images.

5. **Proposed outcomes**

- The application will translate sign language into the English alphabets.
- The text generated as the output will be converted into an audio format.
- By using this application dumb and deaf people can communicate with others smoothly without any barriers.
- The communication gap between the dumb and deaf people and the people not knowing sign language will decrease.
- This application will be used as a tool for giving knowledge of sign language to the parents of deaf so that they can teach their deaf children.

6. **Conclusion**

The paper proposed will help in the development of a product that will decrease the communication gap between the dumb and deaf and the other people. This system can be beneficial economically as compared to the other existing systems. The accuracy of the system can be enhanced by increasing the size of the dataset and enhancing the model used for classification.

References