Sign Language to Text

Rajesh Singh¹, Satyam Shekhar², Shashank Shaurya³, Shivang Kumar⁴, Dr. Rekha. K.S⁵

Department of Computer Science and Engineering
The National Institute of Engineering, India

Abstract:
Major development is being done in research for Sign Language Recognition. A lot of techniques has been developed in this field. The Sign Language is used essentially for deaf and dumb people to communicate with each other. This paper shows the sign language recognizing of 28 including backspace and whitespace hand gestures in American sign language. The proposed system contains five modules such as: Setting up the model, caching bottlenecks, adding FC layer to the model, training the model and writing about trained graphs and labels.

Key Words: American Sign Language, Tensorflow, openCV, image recognition, English text.

1. INTRODUCTION

Sign Language is the most natural and expressive way for the hearing impaired people. People, who are not hearing impaired, never try to learn the sign language for interacting with the deaf people. This leads to separation of the deaf people. But we can minimise the difference between the normal people and the deaf community, if we could program the computer in such a way that it can translate sign language to text format. American sign language (ASL) uses one hand gesture to represent each alphabet. The most important aspect is that ASL already has a regular functional database. Many scholars and intellectuals are presently doing work in American Sign Language. Our proposed program is able to understand American Sign Language’s different alphabets. We have also added two extra hand gestures for space and back space. It will help deaf people to express themselves in proper sentence. A real time Sign Language Recognition system is designed and implemented to recognize 26 gestures from the American Sign Language by hand and two extra hand gestures for space and backspace for text generation. The signs are captured at real time by using web cam. The extracted features are compared by using classification algorithm. In order to calculate the sign recognition, the features are compared with testing database. Finally, recognized gesture is converted into text and text get displayed on monitor. Our proposed system provides an opportunity for a deaf and dumb people to communicate with non-signing people without the need of an interpreter. It is very easy to get two completely different signs mixed up which leads to bad miscommunication. So we have used proper dataset which train our model such that miscommunication can be avoided.

2. LITERATURE SURVEY

Every country has its own sign language with a high level of grammatical variations. The sign language exists in India is commonly known as Indian Sign Language (ISL). It has been argued that perhaps the same sign language is used in Nepal, Sri Lanka, Bangladesh, and border regions of Pakistan. Examples of other sign languages are the American Sign Language (ASL), the British Sign Language (BSL), and so on. In general, the semantic meaning of the language components in all sign languages differs, however there are signs with a universal syntax. For instance, a simple gesture by one hand expressing “hi” or “goodbye” has the similar meaning across the world and in all forms of sign languages. ISL is a complete natural language, found in India with its own morphology, phonology, syntax, and grammar. ISL is a visual-spatial language providing linguistic information through hand, arms, face, and head/body gestures. ISL produces both isolated as well as continuous signs. An isolated sign focuses on a single hand gesture, and is an exacting hand configuration and pose represented by a particular image. A continuous sign is a moving gesture, represented by series of images.

3. CONCEPT

A. TENSOR FLOW

TensorFlow is an end-to-end open source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries and community resources that lets researchers push the state-of-the-art in ML and developers easily build and deploy ML powered applications.

B. TRANSFER LEARNING

Transfer learning (TL) is a research problem in machine learning (ML) that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. For example knowledge gained while learning to recognize cars could apply when trying to recognize trucks.

4. IMPLEMENTATION DETAILS

BASIC ALGORITHM
1. Begin
2. Set the model
3. Build a list of training
4. Divide the images into training, testing, validation
5. cache(bottle_necks)
6. Add new softmax and fully-connected layer
7. weights = [ initial default values ]
8. train_model()
9. number_of_epoch = 10000
10. Test(trained_model) == testData
11. End

A. BOTTLE NECKS
The script can take thirty minutes or more to complete, depending on the speed of your machine. The first phase analyzes all the images on disk and calculates and caches the bottleneck values for each of them. This penultimate layer has been trained to output a set of values that’s good enough for the classifier to use to distinguish between all the classes it’s been asked to recognize. That means it has to be a meaningful and compact summary of the images, since it has to contain enough information for the classifier to make a good choice in a very small set of values. The reason our final layer retraining can work on new classes is that it turns out the kind of information needed to distinguish between all the 1,000 classes in ImageNet is often also useful to distinguish between new kinds of objects.

B. TRAINING
Once the bottlenecks are complete, the actual training of the top layer of the network begins. The training accuracy shows what percent of the images used in the current training batch were labelled with the correct class. The validation accuracy is the precision on a randomly-selected group of images from a different set. The key difference is that the training accuracy is based on images that the network has been able to learn from so the network can overfit to the noise in the training data. A true measure of the performance of the network is to measure its performance on a data set not contained in the training data -- this is measured by the validation accuracy. If the train accuracy is high but the validation accuracy remains low, that means the network is over fitting and memorizing particular features in the training images that aren't helpful more generally. Cross entropy is a loss function which gives a glimpse into how well the learning process is progressing. The training’s objective is to make the loss as small as possible, so you can tell if the earning is working by keeping an eye on whether the loss keeps trending downwards, ignoring the short-term noise.

5. RELATED WORKS
A various hand gestures were recognized with different methods by different researchers in which were implemented in different fields. Vision-based approaches, data glove-based approaches, soft computing approaches such as the Artificial Neural Network, Fuzzy logic and others such as PCA, etc have been used to recognize various hand gestures. “Application research on face detection technology uses Open CV technology in mobile augmented reality” introduces the typical technology. Open source computer vision library, Open CV for short is a cross-platform library computer vision based on open source distribution. The Open CV, with C language provides a very rich visual processing algorithm to write it part and combined with the characteristics of its open source. Data gloves and Vision based method are commonly used to interpret gestures for human computer interaction. The sensors attached to a glove that finger flexion into electrical signals for determining the hand posture in the data gloves method. The camera is used to capture the image gestures in the vision based method. The vision based method reduces the difficulties as in the glove based method. “Hand talk-a sign language recognition based on accelerometer and semi data” this paper introduces American Sign Language conventions. It is part of the “deaf culture” and includes its own system of puns, inside jokes, etc. It is very difficult to understand understanding someone speaking Japanese by English speaker. ASL consists of approximately 6000 gestures of common words with spelling using finger used to communicate obscure words or proper nouns. “Hand gesture recognition and voice conversion system for dumb people” proposed lower the communication gap between the mute community and additionally the standard world. The methodology projected interprets language into speech. The system overcomes the deaf people’s needs and improves their way. The proposed system is simple as well as compact compared to the current method and can be carried to any location. This system converts the language in associate text into voice that's well explicable by blind and ancient people. The language interprets into some text kind displayed on the digital
display screen, to facilitate the deaf people likewise. In world applications, this system is helpful for deaf and dumb of us those cannot communicate with ancient person. Conversion of RGB to gray scale and gray scale to binary conversion introduced in the intelligent sign language recognition using image processing. Basically any colour image is a combination of red, green, blue colour. A computer vision system is implemented to select whether to differentiate objects using colour or black and white and, if colour, to decide what colour space to use (red, green, blue or hue, saturation, luminosity).

6. RESULT

Any user can access the program and sit in front of the camera. He/ She can make symbols of American Sign Language (ASL). The system will simultaneously show the corresponding English text.

7. CONCLUSION

Using various concepts of image processing and fundamental properties of image we tried to developed this system. By using recognition of gesture has done successfully.

8. REFERENCES


