International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 05 | May - 2022 Impact Factor: 7.185 ISSN: 2582-3930

A Review on Real Time Sign Language Detection Using Tensor Flow Object Detection

Sanket Dhobale,Aniket Kandrikar, Aniket Zullarwar, Ankush Tembhurnikar, Sumeet Manapure,

Prof. Jyotsna Gabhane

1.Department of computer Technology, RTMNU University, India.
2.Department of Computer Technology, Priyadarshini College of Engineering, nagpur, Maharashtra

Abstract

As survey suggest there are around 1.8 million people in India who are dumb and around 7 million people who are dumb. Deaf and Dumb people communicate using the Sign Language and still Sign Language is not considered as the official language. It is difficult for them to communicate with someone who don't understand the sign language, so to solve this problem we are creating a model that can detect sign language using the webcam and we are creating this model using Deep Learning and Tensorflow. This model will help deaf and dumb people to communicate with someone who don't understand the sign language. We are using Convolution Neural Network(CNN) to train the model. We are basically crating our own database using webcam. There are good future scope for this we can work more on it and get the more accurate results and help people.

Keywords - Sign Language Detection, Convolution Neural Network(CNN), Deep Learning, Tensorflow

Introduction

Sign language is an essential tool to bridge the communication gap between normal and hearing-impaired people. However, the diversity of over 7000 present-day sign languages with variability in motion position, hand shape, and position of body parts making automatic sign language recognition (ASLR) a complex system. In order to overcome such complexity, researchers are investigating better systems in sign language recognition over the past two decades. A total of 649 publications related to decision support and intelligent systems on sign language recognition (SLR) are extracted from the Scopus database and analysed. The extracted publications are analysed using bibliometric VOSViewer software to (1) obtain the publications temporal and regional distributions, (2) create the cooperation networks between affiliations and authors and identify productive institutions in this context. Moreover, reviews of techniques for vision-based sign language recognition are presented. Various features extraction and classification techniques used in SLR to achieve good results are discussed. The literature review presented in this paper shows the importance of incorporating intelligent solutions into the sign language recognition systems and reveals that perfect intelligent systems for sign language recognition are still an open problem. Overall, it is expected that this study will facilitate knowledge accumulation and creation of intelligent-based SLR and provide readers, researchers, and practitioners a roadmap to guide future direction.

According to the World Health Organization (WHO) report, the number of people affected by hearing disability in 2005 was approximately 278 million worldwide. Ten (10) years later, this number jumped to 360 million, a roughly 14% increment. Since then, the number has been increasing exponentially. The latest report of WHO revealed that

466 million people were suffering from hearing loss in 2019, which amount to 5% of the world population with 432 million (or 83%) of them being adults, and 34 million (17%) of them are children . The WHO also estimated that the number would double (i.e. 900 million people) by 2050 . In these fast-growing deaf-mute people, there is a need to break the communication barrier that adversely affects the lives and social relationships of deaf-mute people. (1)

Dataset Description

We have created our own dataset using webcam. We basically create datasets for five signs that are use in this project. The sign we used in this model are Hello, Good, Thank You, Yes, No. We divided data into two parts: Testing and Training. We add 70% images into training and 30% images into Testing. Testing folder images will be use to test our model and training folders images will be used to train our model and validation folder images would be used for calculating the efficacy of our model. (2)

Methods

1.Deep Learning

Deep learning is a machine learning technique that teaches computers to do what comes naturally to humans: learn by example. Deep learning is a key technology behind driverless cars, enabling them to recognize a stop sign, or to distinguish a pedestrian from a lamppost. It is the key to voice control in consumer devices like phones, tablets, TVs, and hands-free speakers. Deep learning is getting lots of attention lately and for good reason. It's achieving results that were not possible before. In deep learning, a computer model learns to perform classification tasks directly from images, text, or sound. Deep learning models can achieve state-of-the-art accuracy, sometimes exceeding human-level performance. Models are trained by using a large set of labeled data and neural network architectures that contain many layers. In a word, accuracy. Deep learning achieves recognition accuracy at higher levels than ever before. This helps consumer electronics meet user expectations, and it is crucial for safety critical applications like driverless cars. Recent advances in deep learning have improved to the point where deep learning outperforms humans in some tasks like classifying objects in images. (3)

2.CNN

We Humans are the master at determining objects quickly. When you enter a grocery store, you can separate bananas from other goods such as shoes. Yet to teach those classifications with a computer is very hard. In the past, image classification models used raw pixels to classify the images. You can classify cats by color histogram and edge detection which allows you to classify cats by color and ear shape. This method has been successful but until the method encounters more complex variants. That's where the classical image recognition fails because the model does not account for *other features*. But what are these *other features*? Do you need to tell the model one by one? You will find it a great hassle if not impossible. That's where we can use CNN.

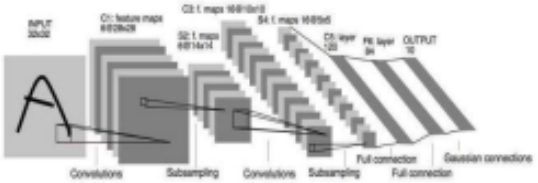

Fig. CNN Architecture

CNN is a type of neural network model which allows us to extract higher representations for the image content. Unlike the classical image recognition where you define the image features yourself, CNN takes the image's raw pixel data, trains the model, then extracts the features automatically for better classification. Because the human brain is designed to capture patterns in order to classify an object, changing the points where you focus your observation also changes your interpretation of the overall image. Similar to how the human brain works, CNN distinguishes meaningful features in an image in order to classify the the test set to train the model, the model could adjust the loss function with the test dataset. This will base the training with the test dataset and is a common cause of overfitting. Therefore during the training we need to use validation datasets then ultimately test the model with the unseen test set. (4)

3. Keras and Tensorflow

TensorFlow 2 is an end-to-end, open-source machine learning platform. You can think of it as an infrastructure layer for differentiable programming. It combines four key abilities:

- Efficiently executing low-level tensor operations on CPU, GPU, or TPU.
- Computing the gradient of arbitrary

differentiable expressions.

- Scaling computation to many devices
- Exporting programs ("graphs") to external runtimes such as servers, browsers, mobile and embedded devices.

Keras is the high-level API of TensorFlow 2: an approachable, highly-productive interface for solving machine learning problems, with a focus on modern deep learning. It provides essential abstractions and building blocks for developing and shipping machine learning solutions with high iteration velocity. Keras empowers engineers and researchers to take full advantage of the scalability and cross platform capabilities of TensorFlow 2: you can run Keras on TPU or on large clusters of GPUs, and you can export your Keras models to run in the browser or on a mobile device. (5)

Conclusion

Hand gestures are a powerful way for human communicaton, with lots of potential applications in the area of human computer interaction. Vision-based hand gesture recognition techniques have many proven advantages compared with traditional devices. However, hand gesture recognition is a difficult problem and the current work is only a small contribution towards achieving the results needed in the field of sign language recognition. This paper presented a vision-based system able to interpret hand gestures from the American Sign Language and convert them to text or speech. After that we also did the opposite. We have been able to convert text to sign language. The proposed solution was tested in real time situations, were it was possible to prove that obtained classification models were able to recognize all the trained gestures being at the same time user independent, important requirements for this type of systems. The selected hand features, in conjunction with machine learning algorithms, proved to be very efficient, allowing their application in any real-time sign language recognition systems. As future work it is intended to keep improving the system and make experiments with complete language datasets. As a final conclusion one can say that although there is still much to do in the area, the proposed solution is a solid foundation for the development of any vision-based sign language recognition user interface system. The sign language grammar can be easily changed and the system configured to train the new language gestures.

References

[1] Joyeeta Singh Indian sign language recognition using eigen value weighted euclidean distance based classification technique.International Journal of Advanced Computer Science and Applications 4, 2 (2013). NEHA V. TAVARI, P. A. V. D. Indian sign language recognition based on histograms of oriented gradient. International Journal of Computer Science and Information Technologies 5, 3 (2017), 3657–3660.

[2] NMANIVAS. Gesturerecognition system. https://github.com/nmanivas/Gesture RecognitionSystem.

[3] PADMAVATHI. S, SAIPREETHY.M.S, V. Indian sign language character recognition using neural networks. IJCA Special Issue on Recent Trends in Pattern Recognition and Image Analysis, RTPRIA (2018).

[4] SAKSHI GOYAL, ISHITA SHARMA, S. S. Sign language recognition system for deaf and dumb people. International Journal of Engineering Research Technology 2, 4 (April 2019).

[5] SCHAKENBERG Schakenberg code for bag of visual words on github.https://github.com/shackenberg/Minima lBag-of-Visual-Words-Image- Classifier. [6] XIAOLONG TENG, BIANI WU, W. Y., AND LIU, C. A hand gesture recognition system based on local linear embedding, April 2005..

[6] Pansare, Jayashree R., Shravan H. Gawande, and Maya Ingle. "Real-time static hand gesture recognition for American Sign Language (ASL) in complex background." (2012).

[7] Mohandes, Mohamed, Junzhao Liu, and Mohamed Deriche. "A survey of image-based arabic sign language recognition." 2014 IEEE 11th International Multi-Conference on Systems, Signals & Devices (SSD14). IEEE, 2014.

[8] Mohandes, Mohamed, Mohamed Deriche, and Junzhao Liu. "Image-based and sensor-based approaches to Arabic sign language recognition." IEEE transactions on human-machine systems 44.4 (2014): 551-557.

[9] Mohandes, Mohamed, Mohamed Deriche, and Junzhao Liu. "Image-based and sensor-based approaches to Arabic sign language recognition." IEEE transactions on human-machine systems 44.4 (2014): 551-557.

[10] Wu, Shuqiong, and Hiroshi Nagahashi. "Real-time 2D hands detection and tracking for sign language recognition." 2013 8th International Conference on System of Systems Engineering. IEEE, 2013.