WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Sign Language Detection and Recognization using Deep Learning

[a] Prof. Archana Ghuge, [b] Gaikwad Yash, [c] Shetty Siddhant, [d] Rajbhar Abhishek, [e] Pawar Shantanu

[a,b,c,d,e] Department of Information Technology Sir Visvesvaraya Institute of Technology Chincholi, Nashik-422102, India

A B S T R A C T

Voice and Language is the most thing for human to communicate with each other. Due to hearing ability, we can understand thoughts of each other. Even nowadays we can give commands using voice recognition. But what on the off the chance one cannot hear anything and in the long run cannot talk. So The Sign Language is the main communicating tool for hearing impaired and quite individuals, and additionally to guarantee an free life for them, the automatic interpretation of sign language is an extensive research area. With the use of image processing and artificial intelligence, many ways of techniques and algorithms have been developed in this area. Every sign language recognition system is trained for recognizing the signs and converting them into required pattern. The proposed system point to provide speech to dumbfounded, in this paper the double handed Indian Sign Language is captured as a series of images and it's processed with the help of Python and then it's converted to speech and text.

Keywords: ***Sign Language Detection, Deep Learning, Speech Recognition, CNN.***

1. Introduction

Sign languages are vivid on wide and world level. There are multiple sign languages in world which are regular in use that are ASL (American Sign Language) ISL (Indian Sign Language), BSL (Bangladesh Sign Language), MSL (Malaysian Sign Language). These languages are Built and Developed with lots of work and practical testing with intention of feasibility to the deaf and dumb persons. Any language is created with its word and its meaning. Sign Language is created as "Sign" and "Action of That Sign". Since here we are not able to form them get it meaning of sign by composing word. As they are hard to hearing and can not listen from birth so we will not instruct them word.

We are motivated with aim to use new technologies for way better humanity. We found Deep learning like technologies can be used for conquering the backwardness occurred because of this physical disability.

A irregular person if visited to deaf person and if deaf person is in problem and trying to explain it then it is very difficult to understand what exactly he is trying to say. Delay in detecting his Sign Language can turn into enormous basic issue for that deaf person. These kind of people can not spend normal life. They face communication issues at each point. Also they get boundaries and limitations to their dreams and proficient points. Hence they get demotivated and Inferiority Complex.

Objective is to give them ability to be expressive in ideas and thoughts. They can get helped in increasing their motivation and confidence a nd it will help them to think positively and to conquer that physical inability. To develop system with using latest technologies and tools we are keeping objective to overcome from this worldwide level problem.

This system will definitely can become step into innovation of this worldwide level problem solution. Our system can be Prototype and Proof of Concept for worldwide level solution. This system can be used by Deaf persons and also normal person can have this system with them and deaf person can perform sign in from of camera and sign can be converted to speech.

2. Related Work

[1] Paper Name: A Review on Smart Gloves to Convert Sign to Speech for Mute Community Author Name: Khan Sohelrana, Syed Faiyaz Ahmed ,Shaik Sameer

Description: The mute community all over the globe facing many problems while communicating. The normal and dumb people can communicate only in one way i.e. sign language, but many times communicating with normal persons they noticed difficulty. Therefore, there always exists communication barrier. This communication barrier is seen because a speech impaired person uses gesture to commune with common human being which is not suitable. We are implementing this project to reduce the barrier between dumb and normal person. This device design is based on the embedded system. Flex sensor and NodeMCU are the key components.

[2] Paper Name:A Gesture-to-Emotional Speech Conversion by Combining Gesture Recognition and Facial Expression Recognition Author Name:Nan Song,Hongwu Yang*

Description:—This paper proposes a facial expression integrated sign language to emotional speech conversion method to solve the communication problems between healthy people and speech disorders. Firstly, the characteristics of sign language and the features of facial expression are obtained by a deep neural network (DNN) model. Secondly, a support vector machine (SVM) are trained to classify the sign language and facial expression for recognizing the text of sign language and emotional tags of facial expression. At the same time, a hidden Markov model-based Mandarin-Tibetan bilingual emotional speech synthesizer is trained by speaker adaptive training with a Mandarin emotional speech corpus. Finally, the Mandarin or Tibetan emotional speech is synthesized from the recognized text of sign language and emotional tags. The objective tests show that the recognition rate for static sign language respectively. Subjective evaluation demonstrates that synthesized emotional speech can get 4.0 of the emotional mean opinion score. The pleasure-arousal-dominance (PAD) tree dimensional emotion model is employed to evaluate the PAD values for both facial expression and synthesized emotional speech. Results show that the PAD values of facial expression are close to the PAD values of synthesized emotional speech. This means that the synthesized emotional speech can express the emotions of facial expression.

[3] Paper Name:Hidden Markov model-based Sign Language to Speech Conversion System in TAMIL Author Name:Aiswarya V, Naren Raju N, Johanan Joy Singh S, Nagarajan T, Vijayalakshmi P

Description:Quick-eared and articulately speaking people convey their ideas, thoughts, and experiences by vocally interacting with the people around them. The difficulty in achieving the same level of communication is high in the case of the deaf and mute population as they express their emotions through sign language. An ease of communication between the former and the latter is necessary to make the latter an integral part of the society.

The aim of this work is to develop a system for recognizing the sign language, which will aid in making this necessity a reality. In the proposed work an accelerometer-gyroscope sensor-based hand gesture recognition module is developed to recognize different hand gestures that are converted to Tamil phrases and an HMMbased text-to-speech synthesizer is built to convert the corresponding text to synthetic speech.

[4] Paper Name:A Translator for American Sign Language to Text and Speech Author Name:Vi N.T. Truong,Chuan-Kai Yang,Quoc-Viet Tran

Description:In the year 2001, Viola and Jones's study is a milestone in developing an algorithm capable of detecting human faces in real time. The original technique was only used for the face detection, but many researchers have applied it for the detection of many other objects such as eyes, mouths, car's number plates and traffic signs. Amongst them, the hand signs are also detected successfully. This paper proposed a system that can automatically detect static hand signs of alphabets in American Sign Language (ASL). To do that, we adopted the two combined concepts AdaBoost and Haar-like classifiers. In this work, to increase the accuracy of the system, we use a huge database for training process, and it generates impressive results. The translator was implemented and trained using a data set of 28000 samples of hand sign images, 1000 images for each hand sign of Positive training images in different scales, illumination, and the data set of 11100 samples of Negative images. All the Positive images were taken by the Logitech Webcam and the frames size were set on the VGA standard 640x480 resolution. Experiments show that our system can recognize all signs with a precision of 98.7. Input of this system is live video and output is the text and speech.

[5] Paper Name:Design and Implementation of A Sign-toSpeech/Text System for Deaf and Dumb People Author Name:DalalAbdulla, Shahrazad Abdulla and Rameesa Manaf ,Anwar H. Jarndal.

Description: This paper presents an approach for designing and implementing a smart glove for deaf and dumb people. There have been several researches done in order to find an easier way for non-vocal people to communicate with vocal people and express themselves to the hearing world. Developments have been made in sign language but mainly in American Sign Language. This research aims to develop a sign to Arabic language translator based on smart glove interfaced wirelessly with microcontroller and text/voice presenting devices. An approach has been developed and programmed to display Arabic text. The whole system has been implemented, programmed, cased and tested with very good results.

[6] Paper Name:On Design and Implementation of A Sign-toSpeech/Text Syst Author Name:Anwar Jarndal ,Ahmed Al-Maflehi

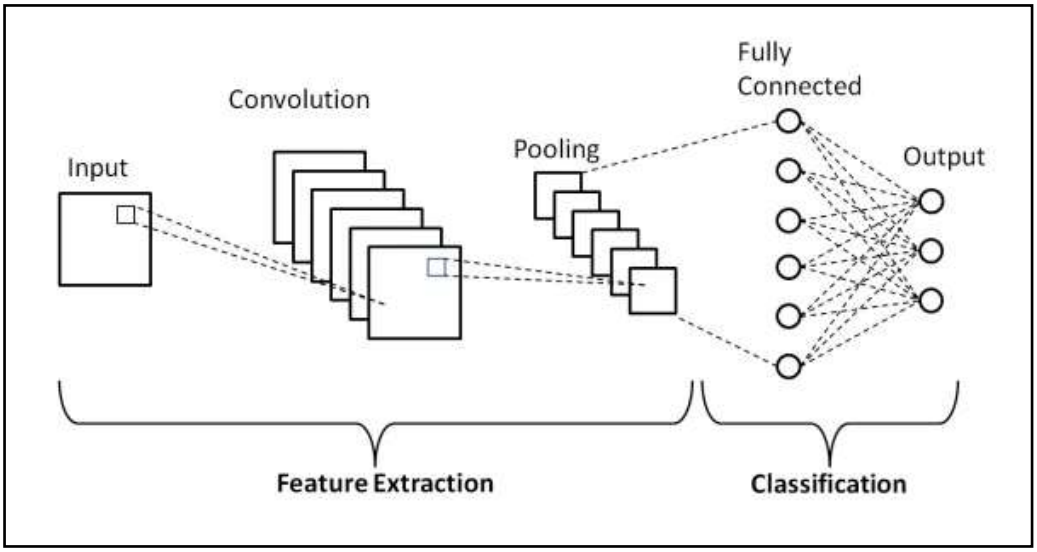
Description: In this paper two approaches have been investigated for designing and implementing a sign to speech/text translator. The approaches have been developed and implemented to display a dual language (Arabic and English) text and voice. In the first part of this paper, a vision-based system is developed and demonstrated. In the second part, a glove-based system is designed and implemented. The second system is based on wireless interfaced glove, microcontroller and presenting devices to translate the sign to Arabic/English language.

[7] Paper Name:Talking Hands – An Indian Sign Language to Speech Translating Gloves Author Name:S Yarisha Heera1, Madhuri K Murthy1,Sravanti V S1

Description:According to the recent statistics about 7.5population is hearing impaired and Indian Sign Language is the only mode of communication used by them. In this paper we have presented an approach that gives a technique for improving Sign Language Recognition system. In the proposed method; we will be using sensors which are incorporated on a glove to detect the gestures and convert it to speech with the help of a Bluetooth module and an Android Smart Phone. The gloves will help in producing artificial speech which provides an environment similar to daily communication which is hard to achieve for speech impaired people.

3. Methodology

The model uses Convolutional Neural Networks (CNN) a deep learning technique to recognize the hand gestures and identify sign language according to provided speech. The CNN model is a neural network that contains layers which takes exactly one input and gives one output. These layers are grouped together and forms a network. The CNN model is mostly used for analysing the images and videos because it gives more accuracy in classification. The below shown figure is the architecture of CNN model.



Architecture of CNN Model

There are two main parts in the architecture of CNN model. They are Feature Extraction and Classification.

- CNNs rely on feature extraction to analyse and understand the underlying patterns and structures within images. This extraction separates, identifies, and analyses relevant features that are crucial for the specific task at hand, such as image classification.
- CNNs classify images using a linked process: feature extraction identifies patterns, and the fully connected "classification layer" analyzes these features to assign one of multiple possible classes. All neurons in this final layer connect, allowing complex analysis for accurate image categorization.

4. Working

1. **Real-time Input Capture:**
    - Camera: Captures sign language gestures as a sequence of video frames.
    - Audio: Records spoken language input via a microphone.
2. **Preprocessing:**
    - Noise removal: Filtering techniques are applied to enhance the quality of both video and audio signals.
3. **Feature Extraction (Sign Language to Audio):**
    - Convolutional Layers: Extract spatial features from video frames, identifying patterns in hand shapes, movements, and facial expressions.
    - Pooling Layers: Down sample feature maps, reducing dimensionality and focusing on salient information.
4. **Intermediate Representation:**
    - Vector Format: Extracted features are condensed into a vector representation, capturing essential information for classification or translation.
5. **Classification (Sign Language to Audio):**

- CNN Classifier: Uses the vector representation to classify video frames into corresponding words or phrases, leveraging a pre-trained dataset of sign language examples.

6. **Speech Synthesis:**
   - Text Conversion: Classified words or phrases are converted into text format.
   - Audio Generation: Text is transformed into spoken audio using text-to-speech (TTS) techniques.
7. **Output:**
   - Sign Language to Audio: Produces spoken audio output corresponding to the signed input.
   - Speech to Sign Language (not yet implemented): Will generate a sequence of images representing sign language gestures for the spoken input.

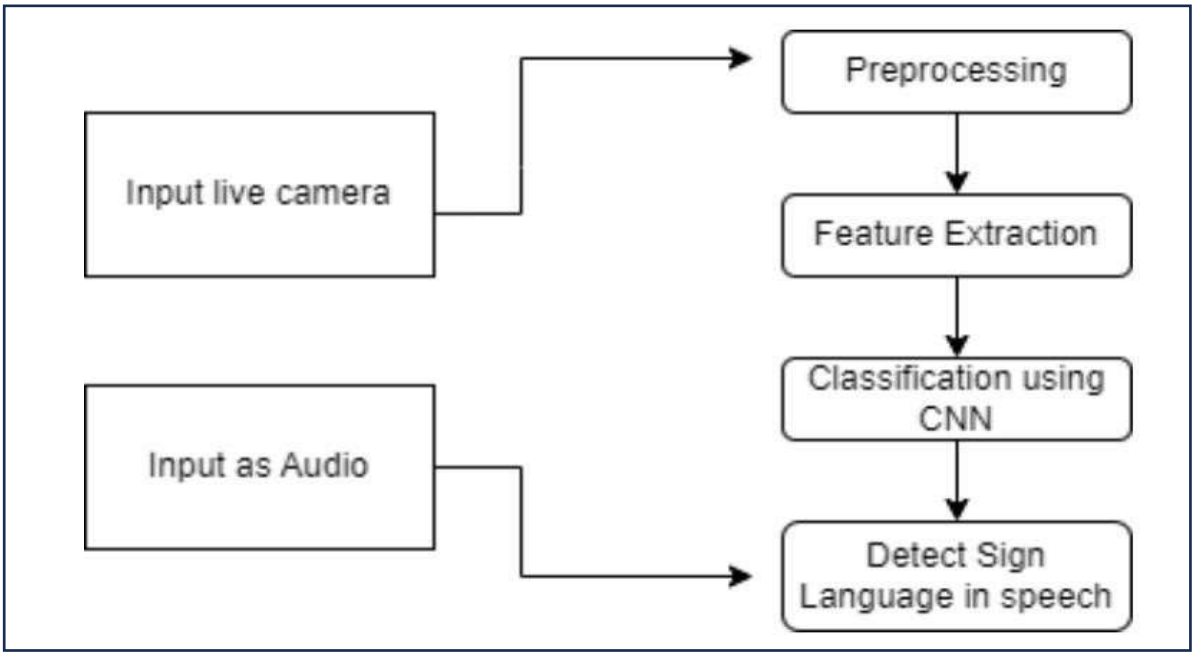


Fig. System Architecture

4. Conclusion

Sign Language is a tool to reduce the communication gap between deaf-mute people and normal person. This system which is proposed above gives the methodology which aims to do the same as the two-way communication is possible. This method proposed here facilitates the conversion on the sign into speech. This overcomes the requirement of a translator since real time conversion is used. The system acts a voice of the per son who is deaf-mute. This project is a step towards helping a specially challenged people. This can be further enhanced by making it more user friendly, efficient, portable, compatible for more signs and as well as dynamic signs. This can be further improvised so as to making it compatible for the mobile phones using the built-in camera of the phone. We can increase the distance at which it can be used by using a longer trans-receiver module or over Wi-Fi.

REFERENCES

[1] Neha Poddar, Shrushti Rao, Shruti Sawant, Vrushali Somavanshi, Prof. Sumita Chandak (Feb 2015), “Study of Sign Language Translation using Gesture Recognition”, International Journal of Advanced Research in Computer and Communication Engineering.

[2] Jayshree R. Pansare, Maya Ingle (2016),” Vision-Based Approach for American Sign Language Recognition Using Edge Orientation Histogram”, at 2016 International Conference on Image, Vision and Computing.

[3] Arslan Arif, Syed Tahir Hussain Rizvi, Iqra Jawaid, Muhammad Adam Waleed, Techno-Talk: An American Sign Language (ASL) Translator, at CoDIT’16 - April 6-8, 2016, Malta.

[4] Justin K. Chen, Debabrata Sengupta, Rukmani Ravi Sundaram,” Sign Language Gesture Recognition with Unsupervised Feature Learning”.

[5] Matheesha Fernando, Janaka Wijayanayaka,"Low cost approach for Real Time Sign Language Recognition" at 2013 IEEE 8th International Conference on Industrial and Information Systems, ICIIS 2013, Aug. 18-20, 2013, Sri.

[6] Aarthi M Vijayalakshmi, "SIGN LANGUAGE TO SPEECH CONVERSION", at 2016 FIFTH INTERNATIONAL CONFERENCE ON RECENT TRENDS IN INFORMATION.

[7] Caixia Wu, Chong Pan,YufengJin, Shengli Sun, and Guangyi Shi Shaoxing, "Improvement of Chinese Sign Language Translation System based on Collaboration of Arm and Finger Sensing Nodes", At The 6th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems June 19-22, 2016.

[8] Poonam Chavan1, Prof. Tushar Ghorpade2, Prof. Puja Padiya, "Indian Sign Language to Forecast Text using Leap Motion Sensor and RF Classifier" at 2016 Symposium on Colossal Data Analysis and Networking (CDAN).

[9] Ms.Manisha D. Raut, Ms. Pallavi Dhok, Mr. Ketan Machhale, Ms. Jaspreet Manjeet Hora, "A System for Recognition of Indian Sign Language for Deaf People using Otsu's Algorithm", International Research Journal of Engineering and Technology (IRJET).

[10] C. Swapna, Shabnam Shaikh "Literature survey on hand gesture recognition for video processing" International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE).

[11] Rini Akmeliawati, Melanie Po-Leen Ooi and Ye Chow Kuang, "Real-Time Malaysian Sign Language Translation Using Colour Segmentation and Neural Network", IEEE on Instrumentation and Measurement Technology Conference Proceeding, Warsaw, Poland 2006, pp. 1-6.

[12] Azadeh Kiani Sarkalehl, Fereshteh Poorahangaryan, Bahman Zan, Ali Karami, "A Neural Network Based System for Persian Sign Language Recognition" IEEE International Conference on Signal and Image Processing Applications 2009.

[13] Incertis, J. Bermejo, and E. Casanova, "Hand Gesture Recognition for Deaf People Interfacing," The 18th International Conference on Pattern Recognition, 2006 IEEE.

[14] R. Feris, M. Turk, R. Raskar, K. Tan, and G. Ohashi. "Exploiting depth discontinuities for vision-based fingerspelling recognition". In IEEE Workshop on Real-time Vision for Human-Computer Interaction, 2004..

[15] T. Starner and A. Pentland, "Real-time American sign language recognition from video using hidden markov models", Technical Report, M.I.T Media Laboratory Perceptual Computing Section, Technical Report No. 375, 1995.