

Gesture Translator for Hard to Hearing and Tongue-Tied People

Mrs. T. H. Gavhane¹, Ms. S. L. Mortale², Mrs. M. M. Mali³,

Mrs. M. A. Parlikar⁴, Mrs. A. S. Sawalkar⁵

Lecturer, Department of Information Technology^{1,3,4,5}

I/C Head, Department of Information Technology²

Pimpri Chinchwad Polytechnic College, Pune, Maharashtra, India

Abstract: *Vocalize and Language is that the main obstacle for human to speak with the whole world. Thanks to hearing ability we are able to perceive thoughts of every different. Even today we are able to offer commands victimization voice recognition. However, what if one fully cannot hear something and eventually cannot speak. So, The Sign Language is that the main human action tool for hearing impaired and mute folks, associate degreed additionally to make sure a freelance life for them, the automated interpretation of linguistic communication is an in-depth analysis space. With the utilization of image process and artificial intelligence, several techniques and algorithms have been developed during this space. Each linguistic communication recognition system is trained for recognizing the signs and changing them into needed pattern. The projected system aims to produce speech to unarticulated, during this paper the double two-handed Indian linguistic communication is captured as a series of pictures and it's processed with the assistance of Python so it's born-again to speech and text.*

Keywords: Machine Learning, Convolutional Neural Networks (CNN), Sign Language, Communication, etc.

I. INTRODUCTION

Gesture languages square measure authentic on wide and world level. There square measure multiple sign languages in world that square measure regular in use that square measure ASL (American Sign Language) ISL (Indian Sign Language), BSL (Bangla Sign Language), MSL (Malaysian Sign Language). These languages square measure engineered and developed with innumerable work and sensible testing with intention of practicability to the deaf and dumb persons. Any language is made with its word and it's that means. Sign Language is made as \Sign" and \Action of That Sign". Because here we tend to don't seem to be ready to create them perceive that means of sign by writing word. As they're deaf and might not listen from birth therefore we are able to not teach them words.

In linguistic terms, sign languages square measure as made and complicated as any voice communication, despite the common idea that they're not "real languages". Skilled linguists have studied several sign languages and located that they exhibit the elemental properties that exist altogether languages. American language is natural and complete language as it uses hands movements and facial expression for communication purpose. Mainly North American people who are having difficulties for communication due to their hard of hearing fact used this language.

All the more for the most part, both sign and communicated in dialects share the qualities that etymologists have found in all characteristic human dialects, for example, temporariness, semanticist, assertion, profitability, and social transmission. and non-manual articulation. All the more for the most part, both sign and communicated in dialects share the qualities that etymologists have found in all characteristic human dialects, for example, temporariness, semanticist, assertion, profitability, and social transmission.

II. LITERATURE SURVEY

Here we present the literature review of existing techniques:

In Paper [1], Sign language is physical communiqué for contributing the meaning instead of the use of voice to demonstrate communicator's opinion. This paper introduces a simple and green set of rules for function extraction to recognize American Sign Language alphabets from both static and dynamic gestures. The proposed algorithm contains of 4 special techniques: Number of white pixels at the threshold of the picture (NwE), peculiarities. For example, the head margin in Finger period from the centroid point (Fcen), Angles between fingers (AngF) and Differences of angles among hands of the first and remaining frame (delAng). After extracting capabilities from video images, an Artificial Neural Network (ANN) is used to classify the signs.

In paper [2], This paper proposes a novel sign language learning method which employs region of interest (ROI) segmentation pre-processing of input data through an object detection network. As the input, 2D image frames are sampled and concatenated into a wide image. From the image, ROI is segmented by detecting and extracting the area of hands, crucial information in sign language. The hand area detection process is implemented with a well-known object detection network, you only look once (YOLO) and the sign language learning is implemented with a convolutional neural network (CNN). 12 sign gestures are tested through a 2D camera.

In paper [3], Sign language is a type of language that uses manual communication to convey meaningful messages to other people. They proposed a system called Dynamic tool for American Sign Language (ASL) finger spelling interpreter which can consistently classify the letters a-z. The dataset consists of a set of American Sign Language videos.

In paper [4], author introduce a hand sign language recognition framework is proposed for various Bangla alphabets using Artificial Neural Network (ANN). For that, initially the input image is normalized and the skin area is extracted on the basis of the YCb Cr values corresponding to human skin color. The extracted area i.e., hand sign area is converted into a binary image and the gaps in the binary hand sign area are filled through the morphological operations. After that, the boundary edge of the hand sign area is extracted through the canny edge detector and extracts the hand sign region of interest (ROI). Finally, features are extracted from the hand sign ROI using Freeman Chain Code (FCC). The ANN is used for training and classifies the hand sign images. The proposed method is tested using various hand sign images and results are presented to demonstrate the efficiency and effectiveness.

In paper [5], Any hand gesture can be represented by sets of feature vectors that change over time. Recurrent Neural (RNNs) are suited to analyse this type of sets thanks to their ability to model the long-term contextual information of temporal sequences. In this paper, RNN is trained by using as features the angles formed by the finger bones of the human hands. The selected features, acquired by a Leap Motion Controller (LMC) sensor, are chosen because the majority of human hand gestures produce joint movements that generate truly characteristic corners. The proposed method, including the effectiveness of the selected angles, was initially tested by creating a very challenging dataset composed by a large number of gestures defined by the American Sign Language (ASL).

In paper [6], author has shown 38 Bangla Sign detection using SIFT and CNN, and used classification techniques like SVM/ANN. They have taken all training and testing images in fixed illumination, but cannot detect two handed gestures.

In paper [7], author proposed accuracy of optimum distance of Kinect Sensor to recognize the Thai finger-spelling and there are 16 gestures were done. This system is strong and small amount of calculation to detect the signer's hand in the complex background.

In paper [8], they searched for the edge detection pre-processing and Skin detection pre-processing. And Edge detection gives more accuracy than skin detection.

In paper [9], author proposed a sentence-based sign language recognition system based on motion data. System uses KNN solution and HMMs.

In paper [10], author proposed a tool which will provide communication opportunity for common people to

understand mode of communication. They used CNN classifier and proposed a tool that will capture the ASL gestures made by deaf and dumb people in real time and classify those gestures into text and voice.

III. PROPOSED METHODOLOGY

We are offering a gadget with a view to use gadget learning set of rules i.e., CNN Convolutional Neural Network. Our planned version can be educated with round 100-500 snap shots of and with growing epoch as a way to increase accuracy. The reason we are the use of CNN is it has a couple of layers as a result it will help into schooling model with smooth manner. We will use Open Computer Vis ion Technology concurrently to interact with digicam, to take stay input from camera. We will set and dene different symptoms with photos and those pictures could be trained with set of rules. Person can have to carry out register in front of camera. After taking live enter from camera the sign might be recognized. Recognized sign will supply textual content output and it will be translated to audio sound. So, system will work as Sign to Speech.

A random man or woman if visited to deaf person and if deaf character is in problem and looking to explain it then it's far very difficult to recognize what exactly he is trying to s ay. Delay in detecting his Sign language can turn into big critical trouble for that deaf character. This kind of people can't spend everyday life. They face communication issues at every point. Also, they get limitations and limitations to their desires and expert aims. Hence, they get demotivated and Inferiority Complex. This machine will definitely can grow to be step into innovation of this global degree hassle solution.

Our device can be prototype and proof of concept for global stage solution. This machine can be used by Deaf and Dumb persons and also regular person can have this machine with them and deaf character can carry out register the front of camera and signal can be transformed to txt or speech. Objective is to provide them capability to be expressive in thoughts and thoughts. They can get helped in increasing their motivation and condense and it will help them to suppose definitely and to overcome that bodily disability. To develop gadget with the usage of trendy technologies and tools we're keeping objective to conquer from this worldwide level trouble.

A. Architecture

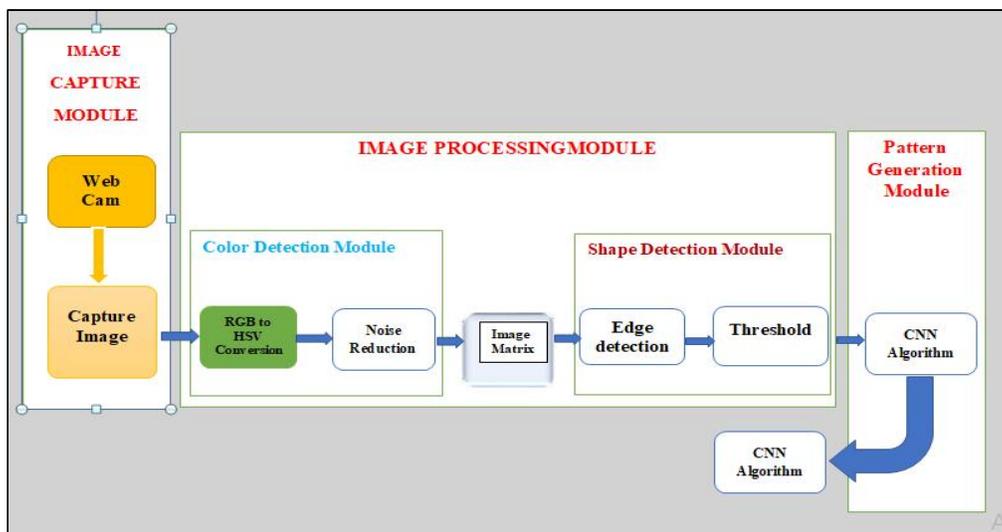


Figure 1: Proposed Architecture

B. Algorithms

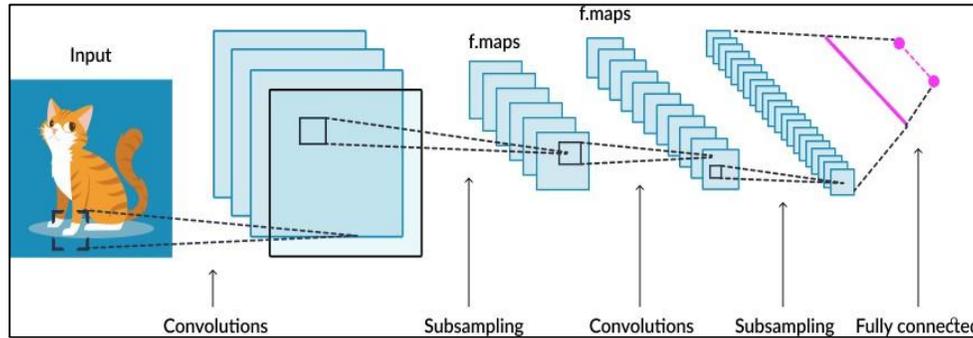


Figure 2: CNN Model

Therefore, applying features from a specific layer only to classification seems to be a process that does not utilize learned CNN's potential discriminate power to its full extent. This inherent property leads to the need for fusion of features from multiple layers. To address this information to predict an image correctly. In some cases, features from the lower layer carry more discriminative power than those problem, we propose a method of combining features from multiple layers in given CNN models.

Moreover, already learned CNN models with from the top layer of the CNN are utilized for classification; however, those features may not contain enough useful training images are reused to extract features from multiple layers. Convolutional neural networks (CNN) have been widely used in automatic image classification systems. In most cases, features the proposed fusion method is evaluated according to image classification benchmark data sets, CIFAR-10, NORB, and SVHN. In all cases, we show that the proposed method improves the reported performances of the existing models by 0.38,3.22 and 0.13, respectively.

Proposed System Algorithm having following steps –

1. To create image data set, sign images are captured in threshold form to get exact sign shape.
2. 2000 images per sign hence 10,000 images for 5 signs.
3. Threshold converts image RGB to black and white color only.
4. Then we get exact shape of hand.
5. Then we define CNN model and add layers in it.
6. Split data into testing as 20% images and training 80% images, then train model.
7. Trained model is then saved.
8. The model loaded and through webcam we send input to model and model predicts it's sign on camera screen.
9. To hear the predicted text, we "C" button.
10. On pressing "C" we can hear the sign text.

IV. RESULT AND DISCUSSIONS

Table 1: Comparison of Accuracy

Comparison of Convolutional Neural Network Algorithm with Artificial Neural Network algorithm for Accuracy	
Algorithm	Accuracy
Convolutional Neural Network algorithm (CNN)	98%
Artificial Neural Network algorithm (ANN)	88%

V. CONCLUSIONS AND FUTURE SCOPE

The proposed device intention to offer speech to speechless, on this paper the double surpassed Indian Sign Language is captured as a series of pictures and it is processed with the help of Python and then it is converted to speech and text. With the help of photo processing and Artificial Intelligence, many techniques and algorithms were developed in this area. Every sign language recognition device is skilled for spotting the signs and converting them into required pattern.

This proposed device is only able to detect hand gesture and not whole-body gesture. There are many improvements to be made into this device. Currently it can't be able to detect whole body gesture so in future we can extend this project. Already CNN algorithm is popular algorithm in image processing system and it will definitely make huge impact on Indian Sign Language.

REFERENCES

- [1] "A Novel Feature Extraction for American Sign Language Recognition" Ariya Thongtawee, Onamon Pinsanoh, Yuttana Kitjaidure 2018 Bio medical Engineering International Conference (BMEiCON-2018).
- [2] "An effective sign language learning with object classification using ROI" Sun mok Kim, Yangho Ji, and Ki-Baek Lee 2018 Second IEEE International Conference on Robotic Computing.
- [3] "Dynamic Tool for A merican Sign Language" Prateek SG Jagadeesh J International Conference on Advances in Co mputing, Co mmun ication Control and Networking (ICACCCN2018).
- [4] "Hand Sign Language Recognition for Bangla A lphabet Based and ANN" Proceedings of the 2017 4th International Conference on Advances in Electrical Engineering 28-30 September, 2017, Dhaka, Bangladesh Mohammad Mahadi Hasan, Md. Khaliluzzaman, Shabiba Akhtar Himel and Rukhsat Tasneem Chowdhury.
- [5] Exploiting Recurrent Neural Networks and Leap Motion Controller for the Recognition of Sign Language and Semaphoric Hand Gestures" Danilo Avola, Member, IEEE, Marco Bernard i, Student Member, IEEE, Luigi Cinque, Senior Member, IEEE, Gian Luca Foresti, Senior Member, IEEE, and Cristiano Massaroni, Student Member, IEEE, 2018.
- [6] Bangla Sign Detection using SIFT and CNN", Shirin Shanta, Saif Taifur Anwar, Md. Rayhanul Kab ir, IEEE, 9th ICCCNT 2018.
- [7] Reliability and Accuracy of Thai Sign Language Recognition with Kinect Sensor", Chana Chansri, Jakkree Srinonchat 2016 IEEE.
- [8] The Comparison of Some Hidden Markov Models for Sign Language Recognition" Suharjito, Herman Gunawan, Narada Thiracitta, Gunawan Witjaksono, The 1st 2018 INAPR International Conference.
- [9] "User-dependent Sign Language Recognition Using Motion Detection", Mohamed Hassan, Khaled Assaleh, Tamer Shanableh, 2016 International Conference on Computational Science and Computational Intelligence.
- [10] Dynamic Toll for American Sign Language Finger Spelling Interpreter, Prateek S G, Siddarth R, P. G. Sunitha Hiremath, Jagdeesh J, Smitha Y, Neha Tarannum, International Conference on Advances in Computing, Communication Control and Networking (ICACCCN2018), 2018 IEEE.