iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VII **Month of publication:** July 2021

DOI: https://doi.org/10.22214/ijraset.2021.36437

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

Smart Communication Assistant for Deaf and Dumb People

Ajabe Harshada[1], Baheti Rushikesh[2], Zurange Dhanashree[3], Wakhare Ashwini[4], Prof. Madhuri Gawali[5]
[1, 2, 3, 4, 5] *Jayawantrao Sawant College of Engineering, Pune*

Abstract: ***Communication is the medium by which we can share our thoughts or convey the messages with other person. Nowadays we can give commands using voice recognition. But what if one absolutely cannot hear anything and eventually cannot speak. So the Sign Language is the main communicating tool for hearing impaired and mute people, and also to ensure an independent life for them, the automatic interpretation of sign language is an extensive research area. Sign language recognition (SLR) aims to interpret sign languages automatically by an application in order to help the deaf people to communicate with hearing society conveniently. Our aim is to design a system to help the Deaf and Dumb person to communicate with the rest of the world using sign language. With the use of image processing and artificial intelligence, many 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 aim to provide speech to speechless, in this paper we have introduced Sign Language Recognition using CNN for dynamic gestures to achieve faster results with high accuracy.***
Keywords: Sign Language, Deaf and Dumb, Translator, Image Processing, CNN.

I. INTRODUCTION

Sign languages (also known as signed languages) are languages that use the visual-manual modality to convey meaning [1]. According to WHO 5% of World population is Deaf and dumb. 18 million people are facing hearing loss in India [2]. These people uses Sign language for communication. There are many sign languages. It varies from country to country with its own vocabulary and grammar. Sign language is used by deaf and hard hearing people to exchange information between their own community and with other people. Computer recognition of sign language deals from sign gesture acquisition and continues till text/speech generation. Sign gestures can be classified as static and dynamic [3]. SLR aims to develop algorithms and methods to correctly identify a sequence of produced signs and to understand their meaning. Many approaches to SLR incorrectly treat the problem as Gesture Recognition (GR). So research has thus far focused on identifying optimal features and classification methods to correctly label a given sign from a set of possible signs. However, sign language is far more than just a collection of well specified gestures [4]. The technique used to develop this system focuses on Neural Network based recognition. Sign language recognition module which is gesture to speech/text will capture dynamic gesture using video. Image frames will be created from the video. Then this created images of gestures will be processed by Neural Network Module and provide results in the form of speech or text.

II. RELATED WORK

Various techniques have been used to implement Sign language Recognition using Image Processing with the help of Machine Learning Algorithms. Combinations of preprocessing and feature extraction techniques, like using Scale-Invariant Feature Transform (SIFT), Histogram Of Gradients (HOG), in order to extract feature descriptors from the image and then using Support Vector Machine in order to classify these into various alphabets.

Kamal Preet proposed Video based Sign Language Processing technique where still images from video is created and this image frame then converted to RGB color Model format for processing. Once images created they are converted to grayscale images and then SURF model is used to detect features from the grayscale image [5].

Chaithraj aims to develop system for Two-way communication audio to gestures and gestures to audio. Pattern matching is used for audio to gesture conversion whereas SIFT Algorithm for feature detection from video based images. It produces fairly stable and good results [6].

Omkar Vedak proposed system in which YCbCr color space is used to segment the hand and then processing the image through HOG and then provide it to the model. They trained the SVM classifier using 5000 images [7].

Oriented Fast and Rotated Brief (ORB) which is alterative for SIFT and SURF is used for feature detection and for classification various algorithms such as Random Forest, Naïve Bayes, Support Vector Machine, Logistic Regression, K Nearest Neighbor and Multi-Layer Perceptron. The system is currently only tested against static gesture images [8].

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VII July 2021- Available at www.ijraset.com

III. PROPOSED ALGORITHM

A. CNN Algorithm

1) Step 1: Convolution Operation

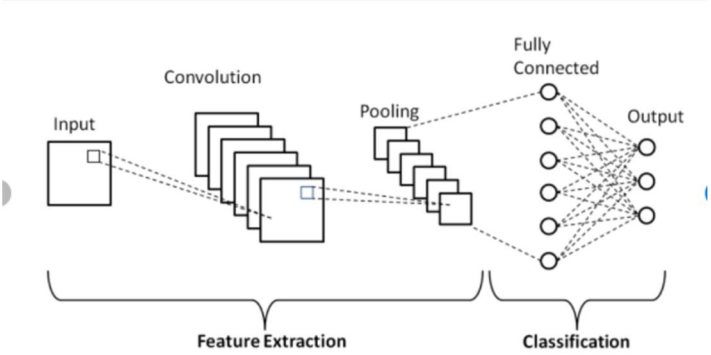
The first building block in our plan of attack is convolution operation. In this step, we will touch on feature detectors, which basically serve as the neural network's filters. We will also discuss feature maps, learning the parameters of such maps, how patterns are detected, the layers of detection, and how the findings are mapped out.

Step 1(b): ReLU Layer: The second part of this step will involve the Rectified Linear Unit or ReLU. We will cover ReLU layers and explore how linearity functions in the context of Convolutional Neural Networks. Not necessary for understanding CNN's, but there's no harm in a quick lesson to improve your skills.

2) Step 2: Pooling In this part, we'll cover pooling and will get to understand exactly how it generally works. Our nexus here, however, will be a specific type of pooling; max pooling. We'll cover various approaches, though, including mean (or sum) pooling. This part will end with a demonstration made using a visual interactive tool that will definitely sort the whole concept out for you.

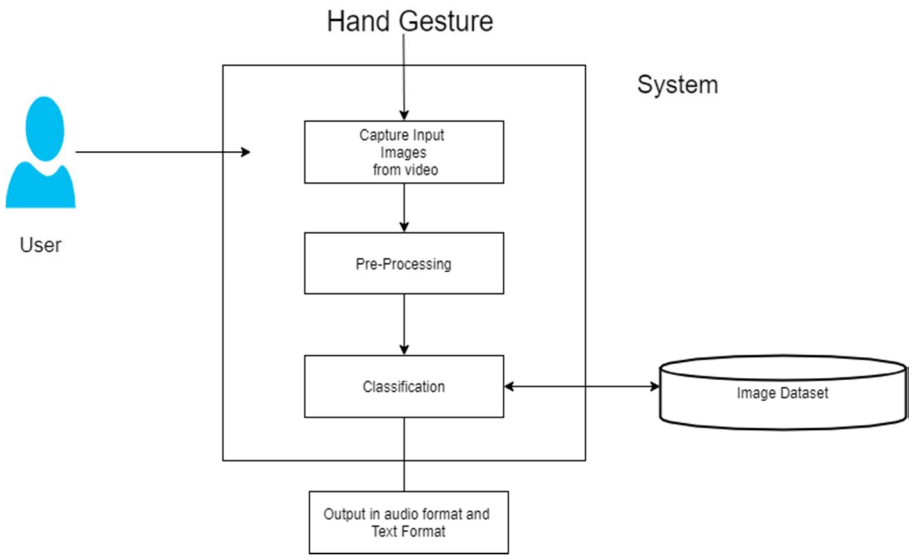
3) Step 3: Flattening this will be a brief breakdown of the flattening process and how we move from pooled to flattened layers when working with Convolutional Neural Networks.

4) Step 4: Full Connection In this part, everything that we covered throughout the section will be merged together. By learning this, you'll get to envision a fuller picture of how Convolutional Neural Networks operate and how the "neurons" that are finally produced learn the classification of images.



CNN Architecture

IV. SYSTEM ARCHITECTURE



©IJRASET: All Rights are Reserved

V. WORKING OF SYSTEM

We are proposing a system that will use Deep learning algorithm i.e. Convolutional Neural Network. Our planned model will be trained with around 100-500 images of and with increasing epoch in order to increase accuracy. The reason we are using CNN is it has multiple layers hence it will help into training model with easy manner. We will sue Open Computer Vision Technology simultaneously to interact with camera, to take live input from camera. We will set and define different signs with images and that images will be trained with algorithm. Person will have to perform sign in front of camera. After taking live input from camera the sign will be recognized. Recognized sign will give text output and it will be translated to audio sound. So system will work as Sign to Speech.

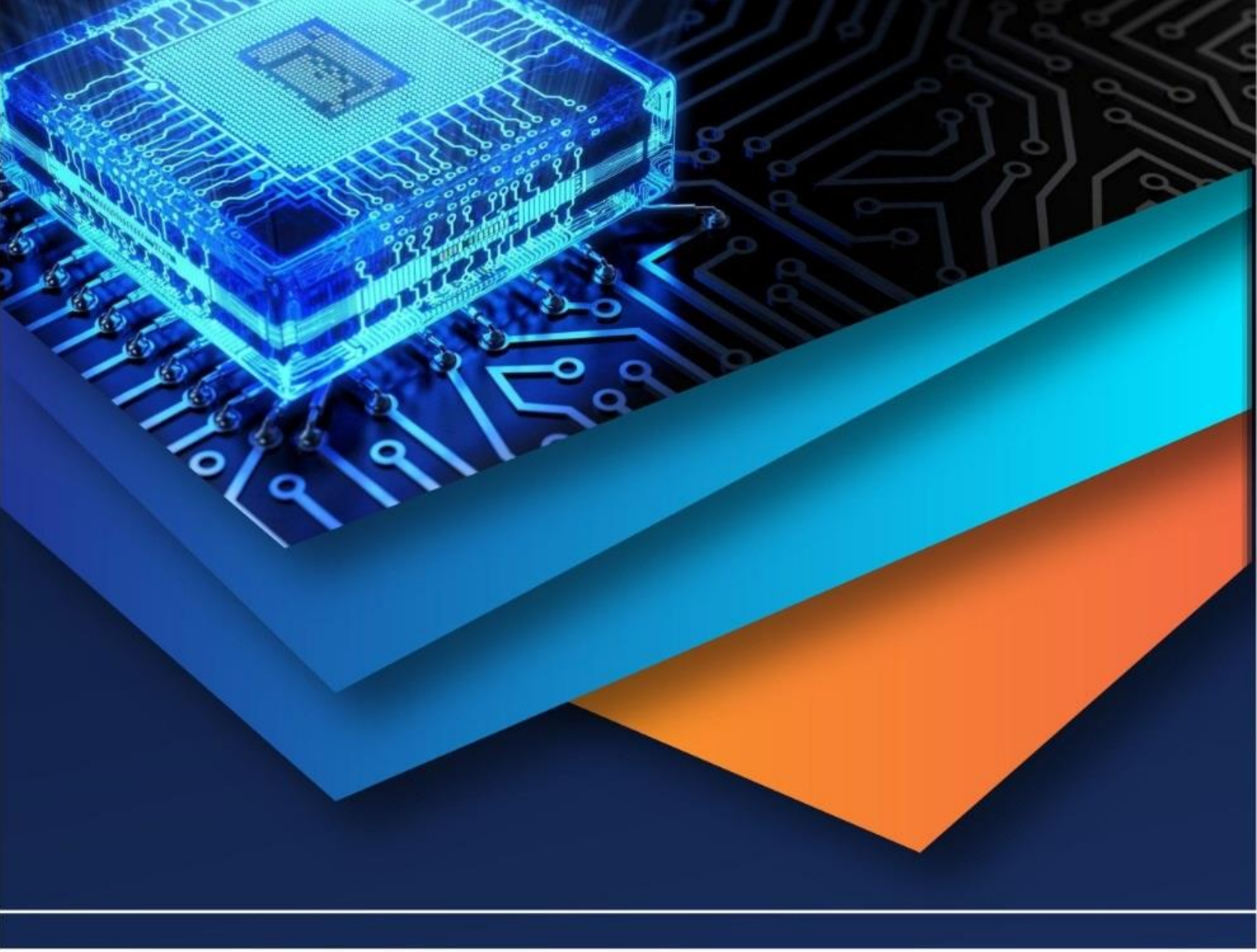
VI. CONCLUSION AND FUTURE WORK

Sign Language recognition can help bridge gap between people with hearing or speech impairment or who use sign language with the rest of society who haven't known to sign language.The Convolution Neural Network gives substantially high accuracy for recognition of dynamic gestures. Our application system can be enhanced in future to give much functionality. Image Processing part and pattern matching algorithm can be combined to form Two-Way communication system i.e. it should be capable of converting normal language to sign language and vice versa. We try to recognize signs which include motion. Moreover it can focus on converting the sequence of gestures into text i.e. word and sentences and then converting it into the speech which can be heard.

REFERENCES

[1] https://en.wikipedia.org/wiki/Sign_language
[2] https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss
[3] Ashok Kumar Sahoo , Gouri Sankar Mishra , Kiran Kumar Ravulakollu "Sign language recognition: State of the art"
[4] http://personal.ee.surrey.ac.uk/Personal/H.Cooper/research/papers/SLR-LAP.pdf
[5] Kamal Preet Kour, Dr. Lini Mathew "Sign Language Recognition using Image processing" by International Journal of Advanced Research in computer Science and software Engineering in August-2017
[6] Chaithraj, Prathima S, Kavana H R, Priyanaka S, Girish C "Sign Language Converter for Deaf and Dumb People in Two-Way Communication for Regional Languages" April-2018NCESC – 2018 (Volume 6 – Issue 13),
[7] Omkar Vedak, Prasad Zavre, Abhijeet Todkar, Manoj Pati "Sign language Interpreter using Image processing and Machine Learning" April-2019
[8] Anmol Mittal, SavitojSingh, VasudevAwatramani "Hand gesture recognition using Image processing and Feature Extraction Techniques" 2020
[9] https://www.researchgate.net/figure/Schematic-diagram-of-a-basic-convolutional-neural-network-CNN-architecture-26_fig1_336805909
[10] Athitsos, V., Sclaroff, S.: Estimating 3D hand pose from a cluttered image. In: Procs. of CVPR, vol. 2. Madison WI, USA (2003)
[11] https://en.wikipedia.org/wiki/Convolutional_neural_network
[12] https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-
[13] Bin Xie, Xiaoyu He & Yi Li. (2018). RGB-D static gesture recognition based on convolutional neural network. The Journal of Engineering, 1515-1520.
[14] Jalal, M. A. (2018). American Sign Posture Understanding with Deep Neural Networks. International Conference on Information Fusion (FUSION) (p. 7). IEEE.
[15] Ma, M. X. (2018). Design and Analyze the Structure based on Deep Belief Network for Gesture Recognition. International Conference on Advanced Computational Intelligence (ICACI) (p. 5). IEEE.
[16] Shanta, S. S. (2018). Bangla Sign Language Detection Using SIFT and CNN. International Conference on Computing, Communication and Networking Technologies (ICCCNT) (p. 6). IEEE.
[17] Ye, Y., Tian, Y., Huenerfauth, M., & Liu, J. (2018). Recognizing American Sign Language Gestures from Within Continuous Videos. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2145-214509, IEEE.
[18] Daniel Kelly, John Mc Donald & Charles Markham. (2011). Weakly Supervised Training of a Sign Language Recognition System Using Multiple Instance Learning Density Matrices. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 526-541.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)