IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, July 2022
Impact Factor: 6.252

Convolutional Neural Network Based Bidirectional Sign Language Translation System

Anil Sonawane[1], Vinit Shahane[2], Onkar Devkar[3], Prof. Ghadage R. A[4]
Students, *Department of* Computer Engineering[1, 2, 3]
Assistant Professor, Department of Computer Engineering[4]
Vishwabharati Academy's College of Engineering, Ahmednagar, India

Abstract: *This paper focuses on experimenting with different segmentation approaches and unsupervised learn- ing algorithms to create an accurate sign language recognition model. To more easily approach the problem and obtain reasonable results, we experimented with just up to 10 different classes/letters in our self- made dataset instead of all 26 possible letters. We collected 12000 RGB images and their corresponding depth data using a Microsoft Kinect. Up to half of the data was fed into the auto encoder to extract features while the other half was used for testing. We achieved a classification accuracy of 98selected set of test data using our trained model. In addition to the work, we did on static images, we also created a live demo version of the project which can be run at a little less than 2 seconds per frame to classify signed hand gestures from any person. The problem we are investigating is sign language recognition through unsupervised feature learning. Being able to recognize sign language is an interesting computer vision problem while simultaneously being extremely useful for deaf people to interact with people who don't know how to understand American Sign Language (ASL). Hand gesture is one of the methods used in sign language for non-verbal communication. Various sign language systems have been developed by many makers around the world but they are neither flexible nor cost-effective for the end users.*

Keywords: Image processing, Noise removal, Feature extraction and matching, Static and dynamic gesture

I. INTRODUCTION

The goal of this project was to build a neural network able to classify which letter of the American Sign Language (ASL) alphabet is being signed, given an image of a signing hand. This project is a first step towards building a possible sign language translator, which can take communications in sign language and translate them into written and oral language. Such a translator would greatly lower the barrier for many deaf and mute individuals to be able to better communicate with others in day to day interactions. This goal is further motivated by the isolation that is felt within the deaf community. Loneliness and depression exists in higher rates among the deaf population, especially when they are immersed in a hearing world. Large barriers that profoundly affect life quality stem from the communication disconnect between the deaf and the hearing. Some examples are information deprivation, limitation of social connections, and difficulty integrating in society.

Most research implementations for this task have used depth maps generated by depth camera and high resolution images. The objective of this project was to see if neural networks are able to classify signed ASL letters using simple images of hands taken with a personal device such as a laptop webcam. This is in alignment with the motivation as this would make a future implementation of a real time ASL- to-oral/written language translator practical in an everyday situation.

- The sign language is a very important way of communication for deaf-dumb people. In sign language each gesture has a specific meaning. So therefore complex meanings can be explain by the help of combination of various basic elements.
- Typically Sign recognition is related as image understanding. So here we are going to use Tensor- flow Object Detection API Model using Python.
- The system will use a webcam for the capturing images and pre-processing of the signs will be done by using Microsoft Visual Studio as an IDE and OpenCv library.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, July 2022
Impact Factor: 6.252

II. MOTIVATION AND PROBLEM DEFINITION

2.1 Objectives

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

2.2 Scope of Project

This system will definitely can become step into innovation of this global level problem solution. Our system can be Prototype and Proof of Concept for global level solution. This system can be used by Deaf and 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 text or speech.

2.3 Motivation of Project

The 2011 Indian census cites roughly 1.3 million people with “hearing impairment”. In contrast to that numbers from India’s National Association of the Deaf estimates that 18 million people –roughly 1 per cent of Indian population are deaf. These statistics formed the motivation for our project. As these speeches impairment and deaf people need a proper channel to communicate with normal people there is a need for a system. Not all normal people can understand sign language of impaired people. Our 5 project hence is aimed at converting the sign language gestures into text that is readable for normal people.

2.4 Problem definition

Speech impaired people use hand signs and gestures to communicate. Normal people face difficulty in understanding their language. Hence there is a need of a system which recognizes the different signs, gestures and conveys the information to the normal people. It bridges the gap between physically challenged people and normal people

III. LITERATURE REVIEW

Convolutional Neural Network based Bidirectional Sign Language Translation System

This paper focuses on experimenting with different seg-mentation approaches and unsupervised learning algorithms to create an accurate sign language recognition model. To more easily approach the problem and obtain reasonable results, we experimented with just up to 10 different classes/letters in the our self-made dataset instead of all 26 possible letters. We collected 12000 RGB images and their corresponding depth data using a Microsoft Kinect.

A Novel Feature Extraction for American Sign Language Recognition

Sign language is physical communication for contributing the meaning instead of using voice to demonstrate communicator’s opinion. This paper introduces a simple and efficient algorithm for feature extraction to recognize American Sign Language alphabets from both static and dynamic gestures. The proposed algorithm comprises of four different techniques: Number of white pixels at the edge of the image (NwE), Finger length from the centroid point (Fcen), Angles between fingers (AngF) and Differences of angles between fingers of the first and last frame (delAng). After extracting features from video images, an Artificial Neural Network (ANN) is used to classify the signs. The result of these experiments is achieved up to 95.

An effective sign language learning with object classification using ROI

This paper proposes a novel sign language learning method which employs region of interest (ROI) segmentation preprocessing 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

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, July 2022
Impact Factor: 6.252

neural network sign gestures are tested through a 2D camera. The results show that, compared to the method without ROI segmentation, the accuracy is increased by 12

Bangla Sign using SIFT CNN
Very few researches were done in detecting Bangla sign language. Most of the researchers in this field used SVM, ANN or KNN as the classifier. In this paper, we try to implement a Bangla sign language system that uses SIFT feature extraction and Convolutional Neural Network (CNN) for classification. We also show that using SIFT feature increases the accuracy of CNN for detecting Bangla Sign language.

"Crawling and Classification Strategies for Generating Multi lan- gauge corpus for sign language"
Although there is considerable sign language content available online, it can be hard to locate content in a specific sign language on a particular topic. The Sign Language Digital Library (SLaDL) aims to improve access through the generation of a multi-language corpus of sign language video. SLaDL uses a combination of crawling to collect potential sign language content and applying multimodal sign language detection and identification classifiers to winnow the collected videos to those believed to be in a particular sign language. Here we compare the quantity and variety of sign language videos located via breadth-first, depth-first, and focused crawling strategies. Then we examine the accuracy of different approaches to combining textual metadata and video features for the 3-way classification task of identifying videos in American Sign Language (ASL), British Sign Language (BSL), and without- sign language. Finally, due to the high computational cost of generating the video features used for classification, we explore the tradeoffs when using a cascading classifier and when generating features based on motion in sampled frames on classifier accuracy.

IV. WORKING OF PROPOSED SYSTEM

In this project, we have implemented an automatic sign language gesture recognition system in real- time, using tools learnt in computer vision and machine learning. We learned about how sometimes basic approaches work better than complicated approaches. Despite trying to use a smart segmentation algorithm, the relatively basic skin segmentation model turned out to extract the best skin masks. We also realized the time constraints and difficulties of creating a dataset from scratch. Looking back, it would have been nice to have had a dataset already to work off of. Some letters were harder to classify in our live demo such as "a" vs "i" since they only differ by a very small edge (the "i" has the pinky pointing up). Although our classification system works quite well as has been demonstrated through tables and images, there's still a lot of scope for possible future work.

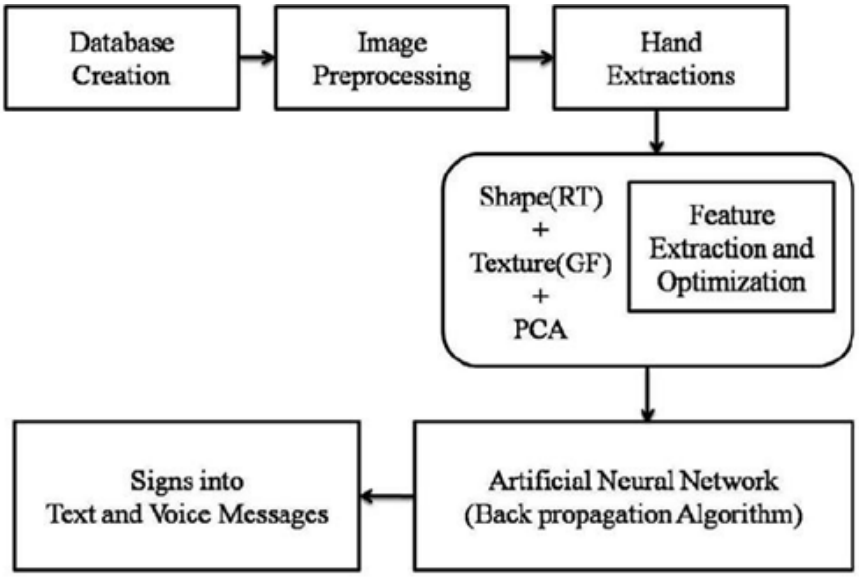


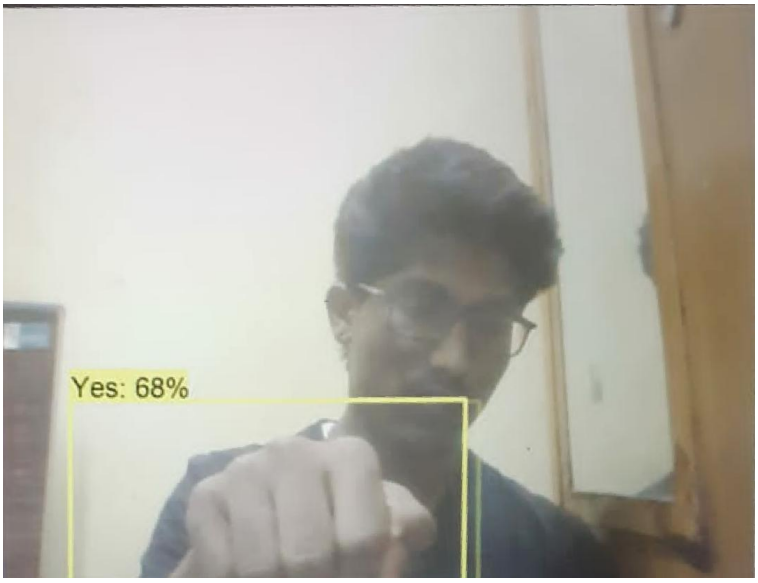
Fig.: System Architecture

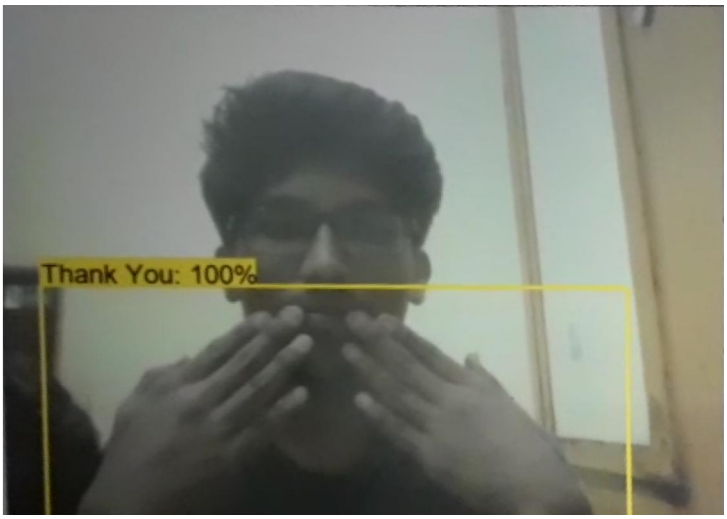
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, July 2022
Impact Factor: 6.252

V. SOFTWARE INTERFACE

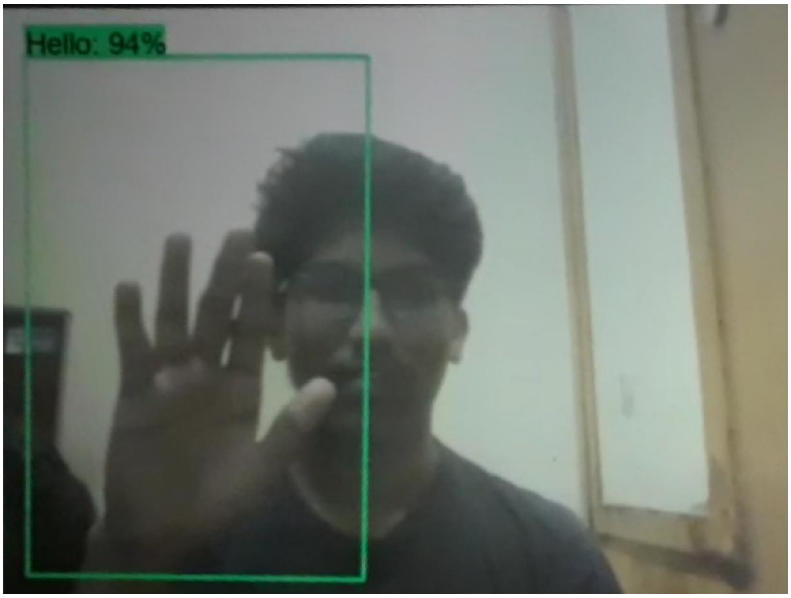
- **Jupyter Notebook :** Project Jupyter is a project and community whose goal is to develop open-source software, open-standards, and services for interactive computing across dozens of programming languages. JupyterLab is the latest web-based interactive development environment for notebooks, code, and data. Its flexible interface allows users to configure and arrange workflows in data science, scientific computing, computational journalism, and machine learning. A modular design invites extensions to expand and enrich functionality.
- **OpenCv :**OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today's systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

VI. PROJECT IMPLEMENTATION





IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, July 2022
Impact Factor: 6.252



VII. ADVANTAGES AND APPLICATIONS

7.1 Advantages

1. It is useful for deaf and dumb people.
2. PaHMMs can improve the robustness of ASL recognition even on a small scale.

7.2 Applications

1. The system is use for deaf and dumb people.
2. The approach for extracting signer position independent features is very powerful for sign language recognition system in practical applications sessions online.

VIII. CONCLUSION AND FUTURE WORK

In this project, we have implemented an automatic sign language gesture recognition system in real-time, using tools learnt in computer vision and machine learning. We learned about how sometimes basic approaches work better than complicated approaches. Despite trying to use a smart segmentation algorithm, the relatively basic skin segmentation

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 2, July 2022
Impact Factor: 6.252

model turned out to extract the best skin masks. We also realized the time constraints and difficulties of creating a dataset from scratch. Looking back, it would have been nice to have had a dataset already to work off of. Some letters were harder to classify in our live demo such as" a" vs" i" since they only differ by a very small edge (the" i" has the pinky pointing up). Although our classification system works quite well as has been demonstrated through tables and images, theres still a lot of scope for possible future work.

8.1 Future Work

The future scope for this project holds a broad scope across varied domains. Initially, it can be to build an application using 'TensorFlow lite' software and then integrate the Bluetooth application that has built for text to speech conversion. The data-set can be improved for more classification of more gestures. Several gestures include the movements of hands and hence to identify these gestures, a video classification algorithm can be incorporated to make the system, a complete ecosystem.

IX. ACKNOWLEDGEMENTS

The completion of our project brings with it a sense of satisfaction, but it is never complete without those people who made it possible and whose constant support has crowned our efforts with success. One cannot even imagine our completion of the project without guidance and neither can we succeed without acknowledging it. It is a great pleasure that we acknowledge the enormous assistance and excellent co-operation to us by the respected personalities.

REFERENCES

[1]. Lance Fernandes, Prathamesh Dalvi, Manisha Bansode. Convolutional Neural Network based Bidirectional Sign Language Translation System. Proceedings of the Third International Con- ference on Smart Systems and Inventive Technology (ICSSIT 2020) IEEE Xplore Part Number: CFP20P17-ART; ISBN: 978-1-7281-5821-1. IEEE, 2020.

[2]. P. Y. Simard. Best practices for convolutional neural networks applied to visual document anal- ysis, August 2003. Seventh International Conference on Document Analysis and Recognition.

[3]. X. Teng. A hand gesture recognition system based on local linear embedding, April 2005. Journal of Visual Languages and Computing.

[4]. M. Ranzato. Efcient learning of sparse representations with an energy-based model, 2006. Courant Institute of Mathematical Sciences.

[5]. D. Metaxas. Sign language and human activity recognition, June 2011. CVPR Workshop on Gesture Recognition.

[6]. S. Sarkar. Segmentation-robust representations, matching, and modeling for sign language recog- nition, June 2011. CVPR Workshop on Gesture Recognition, Co-authors: Barbara Loeding, Ruiduo Yang, Sunita Nayak, Ayush Parashar.

[7]. P Vijayalakshmi and M Aarthi. Sign language to speech conversion. In 2016 International Conference on Recent Trends in Information Technology (ICRTIT), pages 1–6. IEEE, 2016.

[8]. Tecperson. Sign language mnist. . [Online]. Available: https://www.kaggle.com/datamunge/ sign- language-mnist, 2017.

[9]. Shaheer Bin Rizwan, Muhammad Saad Zahid Khan, and Muhammad Imran. American sign language translation via smart wearable glove technology. In 2019 International Symposium on Recent Advances in Electrical Engineering (RAEE), volume 4, pages 1–6. IEEE, 2019.