Project Synopsis

on

Sign Language Recognition

Submitted as a part of course curriculum for

Bachelor of Technology

in

Computer Science
& Engineering

Submitted by

Ashish kumar sharma (2000290100032)

Ayush chauhan (2000290100040)

Divyansh Sheoran (2000290100059)

Under the Supervision of

Prof. Shalini Kapoor

Assistant Professor

KIET Group of Institutions, Ghaziabad
Department of Computer Science
Dr. A.P.J. Abdul Kalam Technical University

2022-2023

DECLARATION

We hereby declare that this submission is our work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

Signature of Students

Name: Ashish Kumar
Sharma
Roll No.: 2000290100032
Date: 11-11-2022

Name: Ayush Chauhan
Roll No.: 2000290100040
Date: 11-11-2022

Name: Divyansh Sheoran
Roll No.: 2000290100059
Date: 11-11-2022

CERTIFICATE

This is to certify that Project Report entitled “YouTube Transcript Summarizer” which is submitted by **Ashish kumar sharma, Ayush Chauhan, Divyansh Sheoran,** in partial fulfilment of the requirement for the award of degree B. Tech. in Department of Computer Science & Engineering of Dr A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

Date:

Supervisor Signature
Prof. Shalini Kapoor
(Assistant Professor)

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the synopsis of the B.Tech Mini Project undertaken during B.Tech. Third Year. We owe a special debt of gratitude to Prof. Shalini Kapoor (Assistant Professor), Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his/her constant support and guidance throughout the course of our work. His/Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his/her cognizant efforts that our endeavours have seen the light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Vineet Sharma, Head of the Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his full support and assistance during the development of theproject. We also do not like to miss the opportunity to acknowledge the contribution of allthe faculty members of the department for their kind assistance and cooperation during the development of our project.

Last but not the least, we acknowledge our friends for their contribution to the completion of the project.

Signature:
Date :
Name :
Roll No:

ABSTRACT

The goal of the Sign Language Recognition project is to create a system that can understand and recognize sign language motions by utilizing the powerful JavaScript hand tracking framework Handfree.js. The aim is to enable smooth contact with technology by filling in the gaps in communication for people who have hearing loss. This paper gives a thorough summary of the project, covering its history, goals, methods, specifics of implementation, outcomes, and suggestions for the future.

In order to achieve accurate hand tracking, the project integrates Handfree.js with a highly developed sign language recognition module. A methodical strategy is utilized, involving the gathering of datasets, preprocessing, choosing models, and training procedures. The design and development of the sign language recognition module, the integration of Handfree.js into the system, and the creation of an intuitive user interface are all included in the implementation phase.

The results of the assessment provide the system's performance metrics and accuracy along with a comparison to other sign language recognition systems currently in use. The study explores how the results should be interpreted, pointing out the shortcomings of the system and suggesting areas for improvement in the future.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

1.1 Introduction

2 The deaf and dumb community uses sign language as their primary form of communication. A variety of hand gestures, finger movements, facial, head, and eye movements are combined to form their language. Sign language has a precise grammar of its own. The project's goal is to replace the traditional human- to- human method of communication with a relatively novel one—human- computer communication. In the past, sign language has been disregarded. As a result, our study focuses on deaf-mute individuals whose problems are frequently ignored by society. They're considered to be the "other" variety. They can only be communicated with through sign language.

3 With the help of this initiative, those with exceptional abilities will be able to communicate not only with those who are considered "normal" by society, but also with those who are blind because the text recognized by the sign will be transformed into speech. Therefore, a method combining computer vision and deep learning with machine learning will be employed to accomplish this goal. In summary, this method takes the key elements of gestures and motions from a video and recognizes the sign in real-time. Numerous sign language systems have been developed, however they have proven to be exceedingly expensive and difficult to use. This project is the foundation for clear documentation and a user- friendly GUI for sign language recognition.

3.1 Problem Statement

The static gesture consists of hand gestures, whereas the latter includes motion of hands, head, or both. Sign language is a visual language and consists of 3 major components, such as finger-spelling, word-level sign vocabulary, and non-manual features. Finger-spelling is used to spell words letter by letter and convey the message whereas the latter is keyword-based. But the design of a sign language translator is quite challenging despite many research efforts during the last few decades. Also, even the same signs have significantly different appearances for different signers and different viewpoints. This work focuses on the creation of a static sign language translator.

3.2 Objective

The primary objectives of this project are to advance the field of automatic sign language recognition by leveraging the capabilities of handfree.js, a JavaScript library that facilitates hands-free interactions through hand tracking. This project aims to develop an innovative and accessible solution for real-time sign language interpretation, addressing the communication barriers faced by the deaf and hard-of-hearing communities.

A key goal is to create a user-friendly application that accurately recognizes and translates sign language gestures into written or spoken language. By integrating handfree.js, which provides robust hand tracking and gesture recognition functionalities, the project seeks to achieve high accuracy and responsiveness in recognizing a wide array of sign language gestures. This includes developing a comprehensive dataset of signs, training machine learning models to interpret these gestures, and optimizing the system for real-time performance.

Furthermore, the project aims to enhance the accessibility and inclusivity of digital communication platforms. By providing a reliable tool for sign language recognition, it facilitates smoother interaction between sign language users and those unfamiliar with it, promoting greater social inclusion and equality. This aligns with the broader objective of leveraging technology to bridge communication gaps and support diverse modes of expression.

Additionally, the project focuses on the scalability and adaptability of the solution. This involves ensuring that the system can be easily extended to accommodate different sign languages and dialects, as well as being compatible with various devices and platforms. By doing so, the project aspires to create a versatile tool that can be used in various contexts, such as education, customer service, and personal communication.

In summary, this project aims to make significant contributions to the field of automatic sign language recognition by developing a real-time, accurate, and inclusive solution using handfree.js. By focusing on user-friendliness, accessibility, and scalability, the project endeavours to enhance communication and inclusivity for sign language users worldwide.

CHAPTER 2 LITERATURE REVIEW

Christopher Lee and Yangsheng Xu developed a glove-based gesture recognition system that was able to recognize 14 of the letters from the hand alphabet, learn new gestures and able to update the model of each gesture in the system in online mode, with a ate of 10Hz. Over the years advanced glove devices have been designed such as the Sayre Glove, Dexterous Hand Master and Power Glove.

The most successful commercially available glove is by far the VPL Data Glove . It was developed by Zimmerman during the 1970's. It is based upon patented optical fiber sensors along the back of the fingers. Star-ner and Pentland developed a glove-environment system capable of recognizing 40 signs from the American Sign Language (ASL) with a rate of 5Hz. Another research is by Hyeon-Kyu Lee and Jin H. Kim presented work on real-time hand-gesture recognition using HMM (Hidden Markov Model). Kjeldsen and Kendersi devised a technique for doing skin-tone segmentation in HSV space, based on the premise that skin tone in images occupies a connected volume in HSV space. They further developed a system which used a back-propagation neural network to recognize gestures from the segmented hand image.

Etsuko Ueda and Yoshio Matsumoto presented a novel technique a hand-pose estimation that can be used for vision-based human interfaces, in this method, the hand regions are extracted from multiple images obtained by a multi viewpoint camera system, and constructing the "voxel Model"[6] . Hand pose is estimated. Chan Wah Ng, Surendra Ranganath presented a hand gesture recognition system, they used image furrier descriptor as their prime feature and classified with the help of RBF network. Their system's overall performance was 90.9%. Claudia Nolker and Helge Ritter presented a hand gesture recognition modal based on recognition of finger tips, in their approach they find full identification of all finger joint angles and based on that a 3D modal of hand is prepared and using neural network

CHAPTER 3 PROPOSED METHODOLOGY

The proposed system is designed to develop a real-time sign language detector using the TensorFlow Object Detection API, leveraging transfer learning for enhanced model performance. The initial step involves data acquisition, where images of sign language gestures are captured via a webcam using Python and OpenCV. These images form the dataset required for training the model.

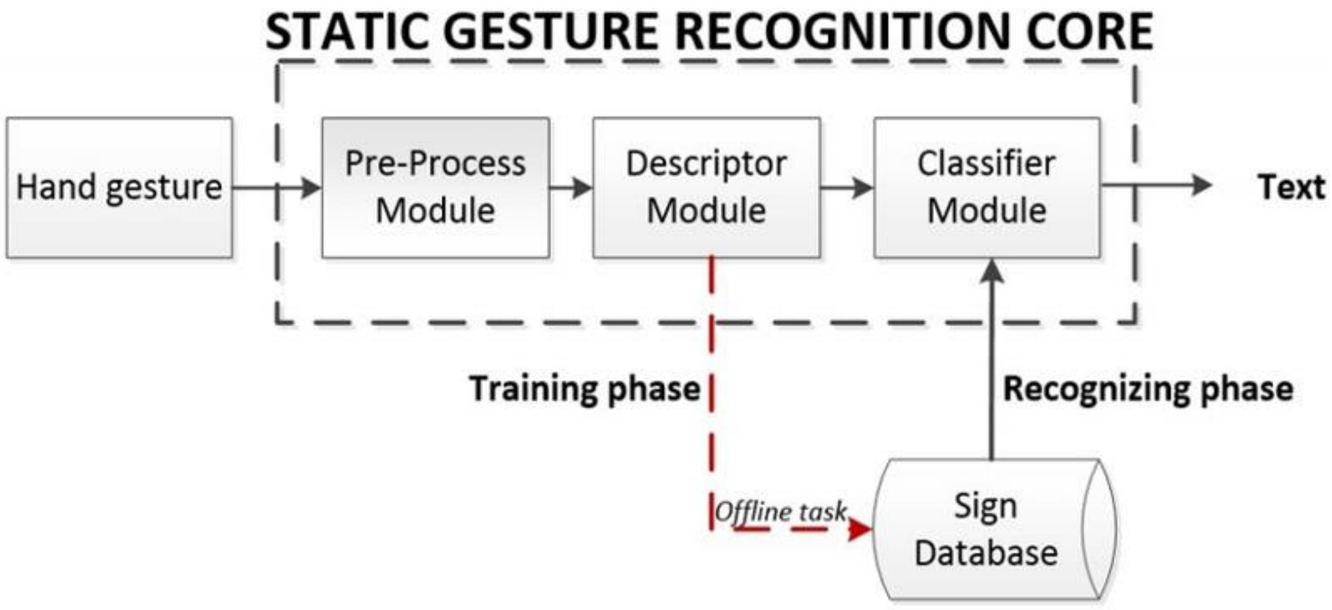
Following data acquisition, a labeled map is created, representing all the objects within the model. This map contains the label for each sign (alphabet) along with a unique ID assigned to each one, ranging from 1 to 26. These labels correspond to the 26 letters of the alphabet, providing a reference to identify the class name during model training and inference.

To facilitate efficient model training, TFRecords of the training and testing data are generated using the generate_tfrecord utility. TFRecord is a binary storage format used by TensorFlow, which significantly enhances the performance of the data import pipeline. The use of binary files ensures that data occupies less disk space, copies faster, and can be read more efficiently from the disk, thus reducing the overall training time of the model.

The TensorFlow Object Detection API, combined with transfer learning, allows the model to leverage pre-trained networks, thereby improving the detection accuracy and reducing the amount of training data required. This approach is particularly beneficial for recognizing complex sign language gestures with high precision in real-time applications.

In summary, the system aims to create an effective and efficient real-time sign language detection tool by capturing and processing image data, creating a labeled map, and utilizing TFRecords for optimized training of the TensorFlow Object Detection API. This results in a robust model capable of accurately interpreting sign language gestures, contributing to improved communication for the deaf and hard-of-hearing communities.

3.1 Flowchart



CHAPTER 5 DIAGRAMS

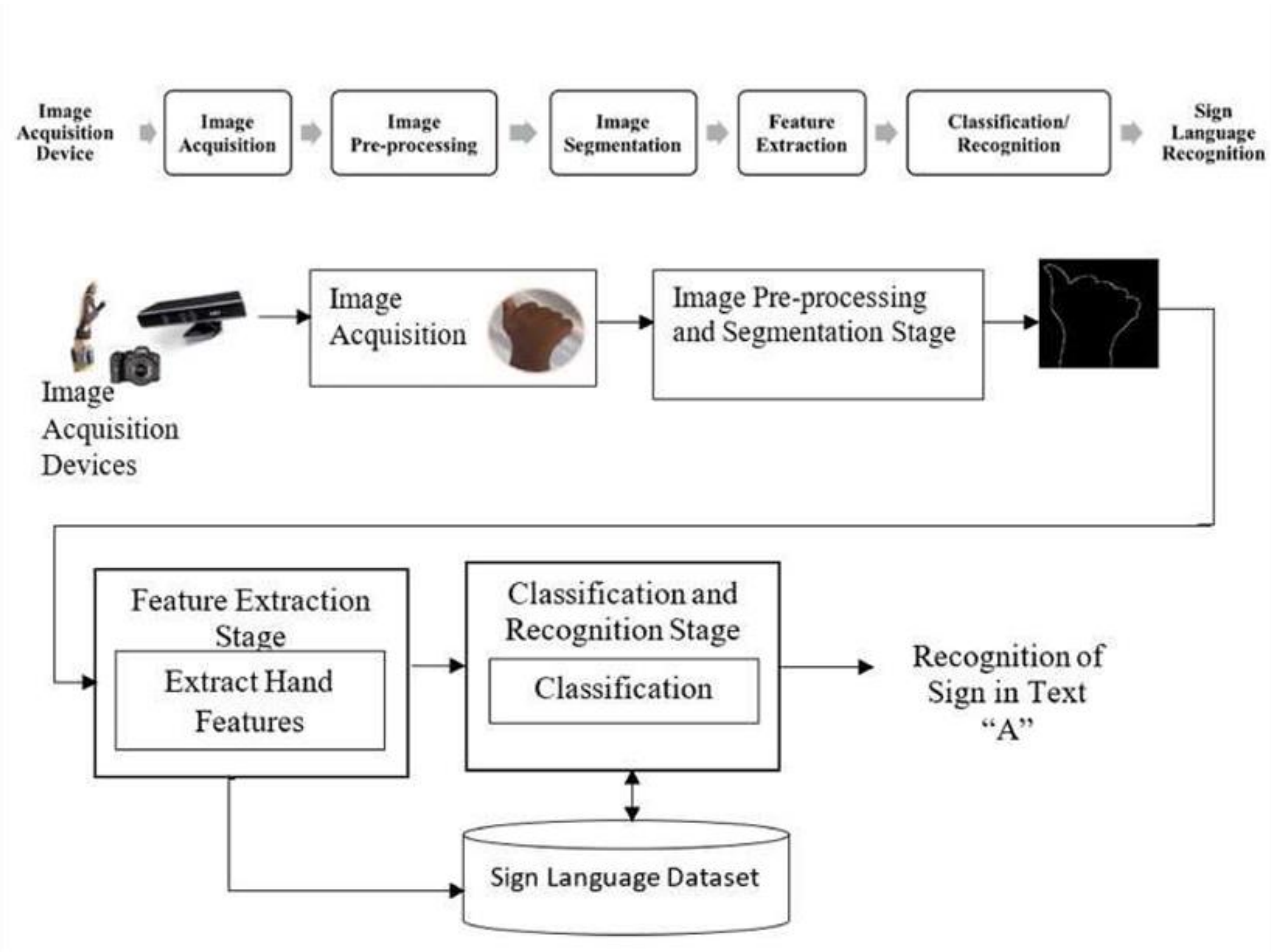


fig.1 Project Stages

CHAPTER 6 CONCLUSION

This project, exploits the power of handsfree.js and machine learning toward the real-time interpretation of sign language gestures, setting a new standard in sign language recognition. Designed for those highly proficient in sign language, it offers an easy and accessible means of communicating through the use of handsfree.js for seamless gesture tracking, along with machine learning for exact interpretation. With its potential applications in online communication platforms, educational resources, and accessibility aids, it promises great inclusivity and access for persons with hearing impairment. Its flexibility and scalability provide a platform for further development and enhancements in the assistive technologyarea.

Looking ahead, the future scope for this project lies in many exciting avenues for exploration and improvement. First, continuous refinement of the machine learning model may further increase accuracy and robustness to enable precise recognition of complex sign language gestures. Second, expansion of the dataset and training of algorithms on diverse sign language variations can enhance the versatility and inclusivity of the system for a wider base of users. Third, addition of more functionalities, such as natural language processing that would translate sign language into text or speech, could make the system more useful in many scenarios. Last, use of feedback from the users and collaboration with experts of sign language linguistics could ultimately make sure that the system evolves to meet the changing needs of its users and promote accessibility and inclusivity in digital communication and interaction

REFERENCES

[1] Lorena P Vargas et al "Sign Language Recognition System using Neural Network for Digital Hardware Implementation" 2011 J. Phys.: Conf. Ser. 274 012051

[2] Brandon Garcia, Sigberto Alarcon Viesca, "Real-time American Sign Language Recognition with Convolutional Neural Networks", Stanford University, 2016, pp. 225-232

[3] Pigou L., Dieleman S., Kindermans PJ., Schrauwen B. (2015) Sign Language Recognition Using Convolutional Neural Networks. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol 8925. Springer, Cham.

[4] Muskan Dhiman, Dr. G.N Rathna, "Sign Language Recognition", 2017, Summer Research Fellowship Programme of India's Science Academies, NIT Hamirpur

[5] Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9,** 611–629 (2018).

[6] Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh, Eero P. Simoncelli, "Image Quality Assessment:From Error Visibility to Structural Similarity", 2004, IEEE Transactions on Image Processing, Vol. 13, No. 4, April 2004

[7] Muja, Marius & Lowe, David. (2009). Fast Approximate Nearest Neighbors with Automatic Algorithm Configuration.. VISAPP 2009 - Proceedings of the 4th International Conference on Computer Vision Theory and Applications. 1. 331-340.

[8] Jiuxiang Gua,, Zhenhua Wangb,, Jason Kuenb, Lianyang Mab, Amir Shahroudyb, Bing Shuaib, Ting Liub, Xingxing Wangb, Li Wangb, Gang Wangb, Jianfei Caic, Tsuhan Chenc, " Recent Advances in Convolutional Neural Network ", Oct 2017, arXiv:1512.0710v6

[9] T. Mitchell, "Machine Learning", McGraw Hill 1997, pp. 11-13

[10] React Documentation, ReactJS Facebook, Silicon Valley, California [Online] https://reactjs.org/docs/getting-started.html

[11] Tensorflow Documentation , Google, [Online] https://www.tensorflow.org/api_docs