VISION QUEST

PROJECT SYNOPSIS

OF MAJOR PROJECT

(SESSION 2023-24)

BACHELOR OF TECHNOLOGY

C.S.E

KIET Group of Institutions, Delhi-NCR, Ghaziabad (UP)

Department of Computer Science and Engineering

SUBMITTED BY:-

NAME- HARSHVARDHAN GUPTA

CLASS- VIIA

ROLL NO- 2000290100067/66

NAME- JASPREET SINGH

CLASS- VIIB

ROLL NO- 2000290100075/05

BATCH- (2020-2024)

OCTOBER,2023

CERTIFICATE

This is to certify that the work contained in the project entitled “Vision Quest”, submitted by Harshvardhan Gupta (2000290100067) and Jaspreet Singh (2000290100075) is a record of bonafide works carried out by them under my supervision and guidance. The contents embodied in the project are being submitted as a part of the Final Year/Major project for the undergraduate curriculum and have not been submitted for the award of any other degree or diploma in this or any other university.

Date : 30-10-2023

Place: Ghaziabad

_____

MR. RAHUL KUMAR SHARMA

ASSISTANT PROFESSOR

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

DECLARATION

We hereby certify that:-

a. The work contained in the project is original and has been done by us under the supervision of our supervisor.

b. The work has not been submitted to any other Institute for any degree or diploma.

c. We have conformed to the norms and guidelines given to us by the Project Review Committee of our department.

d. Whenever we have used materials (data, theoretical analysis, and text) from other sources, we have given due credit to them by citing them in the text of the project and giving their details in the references.

Date : 30-10-2023

Place: Ghaziabad

HARSHVARDHAN GUPTA (2000290100067) _____

JASPREET SINGH (2000290100075) _____

ACKNOWLEDGEMENT

We would like to express our deep gratitude to our project guide Mr. Rahul Kumar Sharma, Assistant Professor, Department of Computer Science & Engineering, for his guidance with immense knowledge and encouragement. We are grateful to Dr. Vineet Sharma, Head of the Department, Computer Science & Engineering, for providing us with the required facilities for the completion of the project work.

We express our thanks to all teaching faculty of Department of Computer Science & Engineering, whose suggestions during reviews helped us in the accomplishment of our project. We would like to thank all non-teaching staff of the Department of Computer Science & Engineering, for providing great assistance in accomplishment of our project.

We would like to thank our parents, friends, and classmates for their encouragement throughout our project period. At last, but not the least, we thank everyone for supporting us directly or indirectly in completing this project successfully.

HARSHVARDHAN GUPTA (2000290100067) _____

JASPREET SINGH (2000290100075) _____

ABSTRACT

Deaf and mute people use sign language to communicate. Unlike acoustically conveyed sound patterns, sign language uses hand gestures, facial expressions, body language and manual communication to convey thoughts. Due to the considerable time required in learning Sign Language, people find it difficult to communicate with specially abled people, creating a communication gap. Hence conventionally, people face problems in recognizing sign language. Moreover, different countries have their respective form of sign gesture communication which results in non-uniformity. The ISL (Indian Sign Language) used in India is largely different from the American Sign Language used in the US, mostly because of the difference in culture, geographical and historical context. Somewhere between 138 and 300 different types of sign language are currently being used throughout the world. Sign language structure varies spatially and temporally. We have identified these as a major barrier in communication with a significant part of society. And hence, we propose to design a system that recognizes different signs and conveys the information to people.

The component of any sign language consists of hand shape, motion, and place of articulation. When combined, these three components (together with palm orientation) uniquely determine the meaning of the manual sign. For sign language identification, sensor-based and vision-based methods are used. In vision-based gesture recognition technology, a camera reads the movements of the human body, typically hand movements and uses these gestures to interpret sign language, whereas in sensor-based methods, Realtime hand and finger movements can be monitored using the leap motion sensor. We aim at developing a scalable project where we will be considering different hand gestures to recognize the letters and words. We plan to use different deep learning models to predict the sign. This may be developed as a desktop or mobile application to enable especially abled people to communicate easily and effectively with others. However, this project can later be extended to capture the whole vocabulary of ASL (American Sign Language) through manual and non-manual signs.

A static hand gesture recognition model can recognize fingerspelling-based hand gestures to form a complete word by combining each gesture. But such models are probable to not recognize moving signs, such as the letters “J” and “Z”. Adding on, fingerspelling for big words and sentences is not a feasible task and temporal properties do not get captured.

Thus, we aim at performing dynamic hand gesture recognition by applying video classification. We wish to develop a system that may take in sign language as input from the webcam, convert it into text and display the relevant translation as output. That way, we aim at smoothing the overall process of communication between a specially able person and the generalized world. We strongly believe, being born with a certain limitation should never act as a barrier to communication.

LIST OF ABBREVIATIONS

Abbreviation Description:

API Application Programming Interface

ASL American Sign Language

BSL British Sign Language

HCI Human Computer Interface

ISL Indian Sign Language

OpenCV Open-Source Computer Vision Library

SLR Sign Language Recognition

Table of contents:-

INTRODUCTION

Communication is very crucial to human beings, as it enables us to express ourselves. We communicate through speech, gestures, body language, reading, writing or through visual aids, speech being one of the most used among them. The study by the World Health Organization (WHO) reports that more than 7% of the world's population has hearing impairment. It is estimated that about 900 million people will experience hearing loss in 2050. According to WHO 2018 report in India, about 63 million people suffer from hearing impairment. Those suffering from speech and hearing loss find it impossible to communicate with the conventional world and vice versa. Use of sign language is the only way to communicate. But not all people understand sign language, and hence, visual aids or an interpreter are used for communication. However, these methods are rather cumbersome and expensive, and can't be used in an emergency. Sign Language mainly uses manual communication to convey meaning. This involves simultaneously combining hand shapes, orientations, gestures and movement of the hands, arms, or body to express the speaker's thoughts.

Sign Language consists of fingerspelling, which spells out words character by character, and word level association which involves hand gestures that convey the word meaning. Fingerspelling is a vital tool in sign language, as it enables the communication of names, addresses and other words that do not carry a meaning in word level association. Despite this, fingerspelling is not widely used as it is challenging to understand and difficult to use. Moreover, there is no universal sign language and very few people know it, which makes it an inadequate Alternative for communication. A system for sign language recognition that classifies hand gestures and finger spelling can solve this problem. Various techniques are used, and their accuracies are recorded and compared in this report.

Hence, The Hand Sign Language Recognition System aims to create a technological solution for the deaf and hearing-impaired community, addressing the communication gap between sign language users and those who do not understand sign language. Sign language is a vital mode of communication for many individuals, and this project seeks to develop a system that can recognize and interpret hand signs, converting them into text or spoken language, thereby facilitating effective communication.

Technologies used in Developing Android Applications-

Languages:

Python is chosen for their flexibility, performance, and extensive libraries for computer vision and machine learning.

Software Required:

OpenCV, TensorFlow, and PyTorch are essential software tools for image processing, deep learning, and neural network development.

FIELD OF PROJECT- The project primarily falls within the domains of Computer Vision, Machine Learning, and Human-Computer Interaction. Computer vision is essential for detecting and tracking hand gestures, machine learning is used to train the recognition models, and human-computer interaction focuses on creating an intuitive user interface.

STATEMENT ABOUT THE PROBLEM:-

The problem addressed by this project is the limited accessibility to sign language interpreters and the challenges faced by deaf individuals when communicating with non-deaf individuals who do not understand sign language. The Hand Sign Language Recognition System aims to provide a real-time solution for interpreting sign language, reducing the dependence on human interpreters, and enhancing accessibility in various aspects of life, such as education, healthcare, and public services.

The Deaf and mute community can only communicate using sign language. Sign language involves simultaneously combining hand shapes, orientations, gestures and movement of the hands, arms, or body to express the speaker's thoughts. Because of cultural, geographic and historical differences, there exists over 300 different types of sign languages around the world.

The ISL (Indian Sign Language) used in India is very different from the American Sign Language used in the United States. This causes inconsistency of sign languages around the world. Moreover, learning sign language requires significant amount of time and effort. This makes it difficult for the conventional world to learn and hence interact with the deaf and mute community. According to a recent study, out of every thousand kids born, 2 to 3 of them are deaf or hard-of-hearing, and, as degrees of hearing loss go, there are 16 to 30 times more children who are identified as Deaf (having a Profound 91+dB hearing loss) than hard-of-hearing. For those deaf or hard of hearing children, only 10% of parents & family learn sign language to communicate with them.

We identify this as a major barrier in communicating with a significant part of the society.

NEED FOR HAND SIGN LANGUAGE RECOGNITION SYSTEM:-

1. Facilitate Communication: The primary need is to facilitate effective communication for the deaf and hearing-impaired community, enabling them to express themselves and understand others.
2. Reduce Reliance on Interpreters: By automating sign language recognition, the system reduces the need for human interpreters, making communication more independent and accessible.
3. Improve Accessibility and Inclusivity: The system contributes to making education, healthcare, and public services more accessible and inclusive for the deaf and hearing-impaired population.

OBJECTIVES:-

The main objectives of this project are to contribute to the field of automatic sign language recognition and translation to text or speech. In our project, we focus on static sign language hand gestures. This work focused on recognizing the hand gestures which includes 26 English alphabets (A-Z) and 10 digits (0-9) using Deep Neural Networks (DNN). We created a convolution neural networks classifier that can classify the hand gestures into English alphabets and digits. We have trained the neural network under different configurations and architectures like LeNet-5 [2], MobileNet V2 [3], and our own architecture. We used the horizontal voting ensemble technique to achieve the maximum accuracy of the model. We have also created a web application using Django Rest Frameworks to test our results from a live camera.

Some of the points are:

1. Develop Recognition System: Create a real-time hand sign detection and recognition system that can accurately identify and interpret signs.
2. Translation Capability: Enable translation of sign language into text or spoken language for non-sign language users.
3. Intuitive User Interface: Implement an intuitive and user-friendly interface that allows easy interaction with the system.
4. High Accuracy and Speed: Strive to achieve high accuracy and speed in recognizing signs, ensuring effective and timely communication.

LITERATURE REVIEW

INTRODUCTION

1. Chuan CH, Regina E, Guardino C (2014) American Sign Language recognition using leap motion sensor. In: 13th IEEE international conference on machine learning and applications (ICMLA), pp 541–544 Chuan et al. developed an American Sign Language recognition system using leap motion sensor. The system was classified using K-Nearest Neighbour and Support Vector machine and the accuracy of 72.78% and 79.83% was achieved respectively.
2. "Deep Convolutional Neural Networks for Sign Language Recognition" G.Anantha Rao, Guntur (DT) Extraction of complex head and hand movements along with their constantly changing shapes for recognition of sign language is considered a difficult problem in computer vision.
3. K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. 2D CNN are used to extract spatial features of input images while RNN are employed to capture the long-term temporal dependencies among input video frames. VGG16 pretrained on ImageNet to extract spatial features and then feed the extracted features to a stacked GRU.
4. J. Carreira and A. Zisserman. Quo vadis, action recognition? a new model and the kinetics dataset. CVPR, 2017. 3D convolutional networks are used which are able to establish not only the holistic representation of each frame but also the temporal relationship between frames in a hierarchical fashion. Inflate 2D filters of the Inception network trained on ImageNet, thus obtaining well-initialized 3D filters.
5. Recognizing American Sign Language Gestures from within Continuous Videos One of the main challenges is that in actions in continuous videos, the temporal boundaries of a specific movement are not very clear. This paper detects their temporal locations from within continuous videos, by collecting an ASL dataset that has been annotated with the time-intervals for each ASL word.

Scope:

1. A Survey of Hand Gesture Recognition Methods in Sign Language Recognition Sign Language Recognition (SLR) system, which is required to recognize sign languages, has been widely studied for years. The studies are based on various input sensors, gesture segmentation, extraction of features and classification methods. This paper aims to analyse and compare the methods employed in the SLR systems, classification methods that have been used, and suggests the most promising method for future research. Due to recent advancement in classification methods, many of the recent proposed works mainly contribute on the classification methods, such as hybrid method and Deep Learning. This paper focuses on the classification methods used in prior Sign Language Recognition system. Based on our review, HMM-based approaches have been explored extensively in prior research, including its modifications.

This study is based on various input sensors, gesture segmentation, extraction of features and classification methods. This paper aims to analyse and compare the methods employed in the SLR systems, classifications methods that have been used, and suggests the most reliable method for future research. Due to recent advancement in classification methods, many of the recently proposed works mainly contribute to the classification methods, such as hybrid method and Deep Learning. Based on our review, HMM-based approaches have been explored extensively in prior research, including its modifications. Hybrid CNN-HMM and fully Deep Learning approaches have shown promising results and offer opportunities for further exploration.

2. Communication between Deaf and Mute People and Normal People

   Chat applications have become a powerful media that assist people to communicate in different languages with each other. There are lots of chat applications that are used different people in different languages but there is not such a chat application that has facilitated to communicate with sign languages. The developed system is based on Sinhala Sign language. The system has included four main components as text messages are converted to sign messages, voice messages are converted to sign messages, sign messages are converted to text messages and sign messages are converted to voice messages. Google voice recognition API has used to develop speech character recognition for voice messages. The system has been trained for the speech and text patterns by using some text parameters and signs of Sinhala Sign language is displayed by emojis. Those emojis and signs that are included in this system will bring the normal people closer to the disabled people. This is a 2-way communication system, but it uses pattern of gesture recognition which is not very reliable in getting appropriate output.

3. Intelligent Sign Language Recognition Using Image Processing

   Computer recognition of sign language is an important research problem for enabling communication with hearing impaired people. This project introduces an efficient and fast algorithm for identification of the number of fingers opened in a gesture representing an alphabet of the Binary Sign Language. The system does not require the hand to be perfectly aligned to the camera. The project uses image processing system to identify, especially English alphabetic sign language used by the deaf people to communicate. The basic objective of this project is to develop a computer based intelligent system that will enable dumb people significantly to communicate with all other people using their natural hand gestures. The idea consisted of designing and building up an intelligent system using image processing, machine learning and artificial intelligence concepts to take visual inputs of sign language's hand gestures and generate easily recognizable form of outputs. Hence the objective of this project is to develop an intelligent system which can act as a translator between the sign language and the spoken language dynamically and can make the communication between people with hearing impairment and normal people both effective and efficient. The system is we are implementing for Binary sign language, but it can detect any sign language with prior image processing.

4. Sign Language Recognition Using Image Processing

One of the major drawbacks of our society is the barrier that is created between disabled or handicapped persons and the normal person. Communication is the only medium by which we can share our thoughts or convey the message but for a person with disability (deaf and mute) faces difficulty in communication with normal person. For many deaf and dumb people, sign language is the basic means of communication. Sign language recognition (SLR) aims to interpret sign languages automatically by a computer in order to help the deaf communicate with hearing society conveniently. Our aim is to design a system to help the person who trained the hearing impaired to communicate with the rest of the world using sign language or hand gesture recognition techniques. In this system, feature detection and feature extraction of hand 23 gesture is done with the help of SURF algorithm using image processing. All this work is done using MATLAB software. With the help of this algorithm, a person can easily train a deaf and mute.

5. Sign Language Interpreter using Image Processing and Machine Learning

   Speech impairment is a disability which affects one's ability to speak and hear. Such individuals use sign language to communicate with other people. Although it is an effective form of communication, there remains a challenge for people who do not understand sign language to communicate with speech impaired people. The aim of this paper is to develop an application which will translate sign language to English in the form of text and audio, thus aiding communication with sign language. The application acquires image data using the webcam of the computer, then it is preprocessed using a combinational algorithm and recognition is done using template matching. The translation in the form of text is then converted to audio. The database used for this system includes 6000 images of English alphabets. We used 4800 images for training and 1200 images for testing. The system produces 88% accuracy.

6. Hand Gesture Recognition based on Digital Image Processing using MATLAB

   This research work presents a prototype system that helps to recognize hand gesture to normal people in order to communicate more effectively with the special people. Aforesaid research work focuses on the problem of gesture recognition in real time that sign language used by the community of deaf people. The problem addressed is based on Digital Image Processing using Color Segmentation, Skin Detection, Image Segmentation, Image Filtering, and Template Matching techniques. This system recognizes gestures of ASL (American Sign Language) including the alphabet and a subset of its words.

7. Gesture Recognition System

   Communication plays a crucial part in human life. It encourages a man to pass on his sentiments, feelings and messages by talking, composing or by utilizing some other medium. Gesture based communication is the main method for Communication 24 for the discourse and hearing weakened individuals. Communication via gestures is a dialect that utilizations outwardly transmitted motions that consolidates hand signs and development of the hands, arms, lip designs, body developments and outward appearances, rather than utilizing discourse or content, to express the individual's musings. Gestures are the requirement for hearing and discourse hindered, they pass

on their message to others just with the assistance of motions. Gesture Recognition System is the capacity of the computer interface to catch, track and perceive the motions and deliver the yield in light of the caught signals. It enables the clients to interface with machines (HMI) without the any need of mechanical gadgets. There are two sorts of sign recognition methods: image- based and sensor-based strategies. Image based approach is utilized as a part of this project that manages communication via gestures motions to distinguish and track the signs and change over them into the relating discourse and content.

Feasibility study

1. Technical Feasibility:

Objective: Assess whether the required technology and resources are available and can be effectively implemented.

Considerations:

- Availability of hardware (webcam or depth-sensing camera).
- Software tools (OpenCV, TensorFlow, PyTorch) and their compatibility.
- Technical expertise and human resources required for development.

Outcome: If technical feasibility is confirmed, it suggests that the project can be implemented from a technological perspective.

2. Economic Feasibility:

Objective: Evaluate whether the project is financially viable and economically sustainable.

Considerations:

- Project costs, including hardware, software, and personnel.
- Potential cost savings, such as reduced reliance on human sign language interpreters.
- Revenue generation potential (if applicable, e.g., through commercialization).

Outcome: Economic feasibility assesses whether the project's benefits outweigh its costs and whether it's financially sustainable.

3. Operational Feasibility:

Objective: Determine whether the project can be integrated into existing operations and systems.

Considerations:

- Compatibility with existing hardware and software infrastructure.
- User acceptance and ease of integration into various applications (e.g., education, healthcare).

Outcome: Operational feasibility examines how smoothly the project can be integrated into real-world applications without major disruptions.

METHODOLOGY

1. Data Collection: Collect a diverse dataset of sign language gestures, including different sign languages and variations.
2. Preprocessing: Prepare the data through image and data preprocessing to make it suitable for model training.
3. Model Development: Train deep learning models (e.g., convolutional neural networks) for hand sign recognition using the prepared dataset.
4. User Interface: Develop an intuitive user interface, which may include touchscreens or other input devices for users to communicate with the system.
5. Real-time Processing: Implement real-time sign language recognition to ensure immediate communication.
6. Testing and Validation: Conduct rigorous testing and validation to evaluate the system's accuracy and usability, involving real users and experts in sign language.
7. Refinement: Continuously improve the system based on user feedback and identified issues.
8. Documentation: Create comprehensive project documentation to ensure future reference, maintenance, and development.

Tools Used:

- LabelImg
- TensorFlow Object Detection API
- Transfer Learning
- OpenCV
- Google Collab

FUTURE SCOPE AND CONCLUSION

Future Scope-

In future works, the model can be improved upon the following points:

i) Upgrading the dataset and training to recognize combination of static gestures: Some words like “Bye” use multiple hand gestures to signify one word. Also, when speaking a sentence, a lot of gestures are used. The model can be efficiently trained to implement these features.

ii) Interpretation of facial expression by real-time system: A few words like “Congratulations” “Sorry”, etc. use facial expressions to convey the meaning. Our system can only recognize hand-gestures for now. In future, inclusion of facial expressions will ensure addition of more words and better accuracy of result.

iii) Inclusion of different types of sign language: As previously stated, there are over 300 different types of sign languages used around the world. Our model recognizes ASL and ISL types only. But with training more datasets, the model can be advanced for use around the world.

iv) Integrating more words into the vocabulary: Currently, our model can accurately recognize only 7 sign language - Hello, I Love You, yes, No, Sorry, Thank You and Please. With inclusion of more words into the vocabulary, the model can be widely used.

v) Use of advanced technologies: With time, the technologies advance, and in future, the different tools can be upgraded to include a more efficient tool and hence more accurate results. TensorFlow model that has been used can be interchanged with another model as well.

Conclusion

Sign languages are kinds of visual languages that employ movements of hands, body, and facial expression as a means of communication. Sign languages are important for especially abled people to have a means of communication. Through it, they can communicate and express and share their feelings with others. The drawback is that not everyone possesses the knowledge of sign languages which limits communication. This limitation can be overcome using automated.

Sign Language Recognition systems which will be able to easily translate the sign language gestures into commonly spoken language. In this paper, it has been done by TensorFlow object detection API. The system has been trained on the American Sign Language alphabet dataset.

The system detects sign language in real-time. For data acquisition, images have been captured by a webcam using OpenCV which makes the cost cheaper. Though the system has achieved a high average confidence rate, the dataset it has been trained on is small and limited.

REFERENCES

[1] Daiyi Peng, Xuanyi Dong, Esteban Real, Mingxing Tan, Yifeng Lu Hanxiao Liu, Gabriel Bender, Adam Kraft, Chen Liang, Quoc V. Le, "PyGlove: Symbolic programming for automated machine learning", Advances in Neural Information Processing Systems, Vol. 33, pp. 96-108, 2020.

[2] Zhang Shujun and Qun Zhang, "Sign language recognition based on global-local attention", Journal of Visual Communication and Image Representation, Vol. 80, Article id: 103280, 2021.

[3] I. A. Adeyanju, O. O. Bello, and M. A. Adegboye, "Machine learning methods for sign language recognition: A critical review and analysis", Intelligent Systems with Applications, Vol. 12, Article id: 200056, 2021.

[4] Brandon Garcia, Sigberto Alarcon Viesca, “Real-time American Sign Language Recognition with Convolutional Neural Networks”, Stanford University Stanford, CA, 2016

[5] Sawant Pramada, Deshpande Saylee, Nale Pranita, Nerkar Samiksha, Mrs.Archana S. Vaidya, “Intelligent Sign Language Recognition Using Image Processing”, IOSR Journal of Engineering (IOSRJEN), Volume 3, Issue 2, PP 45-51, 2013

[6] Pratibha Pandey, Vinay Jain, “An Efficient Algorithm for Sign Language Recognition”, (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 6, 2015.

[7] Anup Kumar, Karun Thankachan, Mevin, M. Dominic, “Sign Language Recognition”, IEEEIndia, 2016.