Real-Time Sign Language Recognition Using Mediapipe and LSTM Neural Networks

TRAINING/INTERNSHIP/PROJECT REPORT

Submitted in partial fulfillment of the requirements for the award of the degree

Of

-BACHELOR OF TECHNOLOGY-

In

-Department of Electronics and Communication -

By

-Kriti Khurana-
-02501022022-

Guided by

-Dr. Ritu Rani-
-Research Associate-
-IGDTUW -

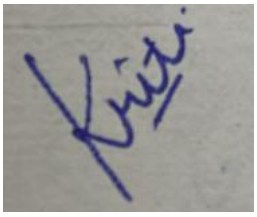
INDIRA GANDHI DELHI TECHNICAL UNIVERSITY FOR WOMEN
NEW DELHI – 110006

CERTIFICATE

I, **Kriti Khurana** , certify that the Internship Project Report entitled "Real-Time Sign Language Recognition Using Mediapipe and LSTM Neural Networks" is done by me and it is authentic work carried out by me at Indira Gandhi Delhi Technical University for Women . For this project, no work has been submitted before for any degree or diploma of the award, to the best of my knowledge and belief.

DR. RITU RANI
(Research Associate)
Centre of Excellence-AI
IGDTUW

Signature of the student:



UNDERTAKING REGARDING ANTI-PLAGIARISM

I, Kriti Khurana hereby, declare that the material/ content presented in the report are free from plagiarism and is properly cited and written in my own words. In case, plagiarism is detected at any stage, I shall be solely responsible for it.

Kriti Khurana

(02501022022)

INDIRA GANDHI DELHI TECHNICAL UNIVERSITY FOR WOMEN

(ESTABLISHED BY GOVT. OF DELHI VIDE ACT 09 OF 2012)
ISO 9001:2015 CERTIFIED UNIVERSITY
KASHMERE GATE, DELHI-110006

WOMEN EDUCATION | WOMEN ENLIGHTENMENT | WOMEN EMPOWERMENT

CENTRE OF EXCELLENCE - ARTIFICIAL INTELLIGENCE

CERTIFICATE OF COMPLETION

This certificate is awarded to

Kriti Khurana

For successfully completing the 7 weeks Summer Internship on "PYTHON & MACHINE LEARNING" from 5th June - 23rd July, 2023 jointly conducted by the COE - AI, AI Club IGDTUW and Anveshan Foundation.

Ishita Saxena
President - AI CLUB
IGDTUW

Dr. Ritu Rani
Research Associate
COE - AI

Prof. Arun Sharma
Coordinator - Centre of Excellence-AI
IGDTUW

UNDERTAKING REGARDING ANTI-PLAGIARISM

I, Kriti Khurana hereby, declare that the material/ content presented in the report are free from plagiarism and is properly cited and written in my own words. In case, plagiarism is detected at any stage, I shall be solely responsible for it.

Kriti Khurana

(02501022022)

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my teammates , Aakriti Tyagi(00101022022) and Aarya Teotia(00201022022), for their invaluable contributions and collaborative efforts in the completion of the project titled “Real-Time Sign Language Recognition Using Mediapipe and LSTM Neural Networks” .Their dedication and teamwork significantly enriched the project. I am also deeply thankful to our mentor, Dr. Ritu Rani(COE-AI, IGDTUW), for her guidance, support, and expertise throughout the project. Her insights and encouragement played a pivotal role in shaping our work. This project has been a collective endeavor, and I extend my appreciation to everyone involved in making it a success.

Kriti Khurana

(02501022022)

DECLARATION

I, **Kriti Khurana**, solemnly declare that the internship project report, Real-Time Sign Language Recognition Using Mediapipe and LSTM Neural Networks **,** is based on my own work carried out under the supervision of **Dr. Ritu Rani**. I assert the statements made and conclusions drawn are an outcome of my work. I further certify that:

I. The work contained in the report is original and has been done by me under the supervision of my supervisor.
II. The work has not been submitted to any other Institution for any other degree/diploma/certificate in this university or any other University of India or abroad.
III. We have followed the guidelines provided by the university in writing the report.
IV. Whenever we have used materials (text, data, theoretical analysis/equations, codes/program, figures, tables, pictures, text etc.) from other sources, we have given due credit to them in the report and have also given their details in the references.

Kriti Khurana

(02501022022)

LIST OF ABBREVIATIONS

Abbreviation	Description
LSTM	Long Short Term Memory
RNN	Recurrent Neural Network

TABLE OF FIGURES

Example :

LIST OF TABLES

ABSTRACT/SUMMARY

The research paper focuses on developing a real-time sign language recognition model using the combination of Mediapipe and LSTM neural networks. The primary objective is to address communication challenges faced by deaf individuals and make sign language more accessible. The methodology involves collecting real-time sign language gestures through a webcam, preprocessing the data, and training an LSTM-based neural network for recognition. The model's performance is evaluated, achieving a high accuracy of 92.85% in real-time recognition. Practical applications include integration into digital devices for the hearing-impaired and serving as an educational tool. The study acknowledges limitations, commits to reproducibility, and suggests future research directions for expanding the model's capabilities. Overall, the research demonstrates successful design and potential societal impact in enhancing communication for the hearing-impaired.

INDEX

INTRODUCTION

In our increasingly digitalized world, effective communication is a cornerstone of societal interaction. However, individuals with hearing impairments face unique challenges, relying on sign language—a complex system of hand gestures and expressions. This research addresses the communication gap by introducing a real-time sign language recognition model. Leveraging the capabilities of Mediapipe and LSTM networks, the model aims to provide instant and accurate interpretation of sign language gestures. This introduction sets the stage for exploring existing literature, defining the project's scope, outlining the problem statement, highlighting contributions, and emphasizing the significance of making communication more accessible for the hearing-impaired.

- Significance of Sign Language Communication: Sign language serves as a vital means of communication for individuals who are deaf or hard of hearing. However, a communication gap exists between sign language users and those unfamiliar with this visual language.

- Objective of the Project: The primary goal of this project is to bridge the communication divide by developing a real-time sign language gesture recognition and translation system. Using computer vision and deep learning techniques, the system aims to interpret sign language gestures and convert them into text, fostering effective communication.

- Transition to Computer Vision: Traditional models often faced limitations in capturing the dynamic and nuanced nature of sign language gestures. The introduction of computer vision, exemplified by the Mediapipe Holistic model in this project, allows for the holistic extraction of information from body poses and hand movements in real-time.

- From Interpretation to Translation: While interpreters play a crucial role in bridging language gaps, the proposed system goes a step further by not only interpreting sign language gestures but also translating them into text. This transformative shift enhances communication efficiency and inclusivity.

1.1 SYSTEM REQUIREMENTS

1)Software Requirements

- Operating System: Windows 11 (64 bit)
- Programming Language: Python.

2)Hardware Requirements

- RAM: 16.0 GB
- Graphics: Intel iRIS graphics
- Processor: Intel Core i5

1.2 SYSTEM ANALYSIS

1)FEASIBILITY STUDY:

None

2)ECONOMIC FEASIBILITY:

- Python - free and open source
- Google Colab - Free and Open Source

3)IMPLEMENTATION FEASIBILITY

None

WORK DESCRIPTION

- **Real-time Sign Language Recognition Model Development:**
    - Use the power of Mediapipe and LSTM neural networks to create a model for real-time sign language recognition.
    - Implement a hand gesture tracking system using the MediaPipe framework to capture and analyze sign language gestures accurately.
- **Data Collection and Preprocessing**:
    - Collect real-time sign language gestures through a webcam and use the MediaPipe Holistic model for keypoint extraction.
    - Preprocess the collected keypoints, organizing them into suitable sequences for training the LSTM neural network.
- **Model Training and Optimization:**
    - Develop an LSTM-based neural network using TensorFlow Keras, specifically designed for learning and recognizing sign language gestures.
    - Train and optimize the model using a dataset that includes various sign language gestures, ensuring high accuracy and real-time processing.
- **Evaluation and Testing:**
    - Evaluate the performance of the trained model through comprehensive testing, including accuracy, speed, and adaptability to different sign language gestures.
    - Conduct real-time testing using the MediaPipe Holistic model to detect keypoints from live video frames and predict sign language gestures.
- **Practical Applications and Educational Tool:**
    - Discuss the practical applications of the developed model, such as integration into smartphones, tablets, and digital devices for the hearing-impaired population.
    - Highlight the potential use of the model as an educational tool for teaching sign language to both hearing-impaired individuals and those interested in learning.

WORK OUTCOME

- **High Accuracy and Real-time Recognition**:
  - Achieved a high accuracy of 92.85% in recognizing sign language gestures in real-time video streams.

PERFORMANCE PARAMETERS	RESULTS
Accuracy	92.85%
Precision	96.42%
F1-score	93.53%

TABLE 1. RESULT TABLE SHOWING ACHIEVED VALUES OF ALL THE PERFORMANCE METRICES

- **Integration into Digital Devices:**
  - Demonstrated the model's integration into smartphones and other digital devices, enhancing accessibility for the hearing-impaired population.
- **Improved Communication for Hearing-Impaired Individuals:**
  - Enabled more natural and expressive communication for hearing-impaired individuals, potentially improving their quality of life and participation in various activities.
- **Educational Tool for Sign Language Learning:**
  - Provided a tool for real-time feedback and guidance in learning sign language, benefiting both hearing-impaired individuals and those interested in acquiring sign language skills.
- **Hybrid Approach with MediaPipe and LSTM Networks**:
  - Introduced a novel hybrid approach by combining the capabilities of MediaPipe and LSTM networks for sign language recognition.

CONCLUSION

In conclusion, this research endeavors to break down barriers in communication for individuals with hearing impairments by introducing a real-time sign language recognition model. The fusion of Mediapipe and LSTM neural networks has resulted in a robust system capable of understanding and interpreting sign language gestures with a high degree of accuracy.

The research journey began with a comprehensive literature review, unveiling the societal impact of communication challenges faced by the deaf and hard-of-hearing communities. The scope was meticulously defined, focusing on the integration of key technologies and their application in real-time sign language recognition.

Through the phases of data collection, preprocessing, model development, and evaluation, the project demonstrated commendable success. The model's accuracy of 92.85% signifies its potential to provide instantaneous and accurate sign language interpretation, fostering a more inclusive digital landscape.

FUTURE SCOPE.

Multimodal Integration:

- Future developments may explore the integration of multiple modalities, such as facial expressions and body movements, to enhance the accuracy and context-awareness of sign language recognition. This could lead to a more holistic interpretation of sign language communication.

Adaptive Learning Models:

- Implementing adaptive learning models that can personalize the recognition system based on individual signing styles and preferences would enhance user-specific accuracy. Incorporating machine learning techniques for continuous adaptation will be crucial for ongoing improvement.

Real-Time Translation Capabilities:

- Extending the system to include real-time translation features could facilitate communication between sign language users and non-signers. Developing language translation capabilities for various spoken languages would broaden the system's applicability.

Collaboration with Augmented Reality (AR):

- Integrating the system with AR technologies could provide an immersive sign language learning experience. AR applications can overlay sign language interpretations onto the real world, creating interactive and engaging learning environments.

Global Standardization Efforts:

- Participation in global standardization efforts for sign language gestures can contribute to the development of universally accepted recognition systems. Collaboration with international organizations and communities will aid in establishing common benchmarks and practices.

By addressing these future avenues, the real-time sign language recognition system can evolve to meet the dynamic needs of users, furthering its impact on communication accessibility for individuals with hearing impairments.

BIBLIOGRAPHY

[1] https://github.com/rabBit64/Sign-language-recognition-with-RNNand-Mediapipe

[2] https://www.google.com/search?client=safari&rls=en&q=https%3A%2F%2Fgithub.com%2FTachionstrahl%2FSign&ie=UTF-8&oe=UTF8

[3] https://academic.oup.com/jdsde/article/23/4/341/5025968

[4] https://www.google.com/search?client=safari&rls=en&q=http%3A%2F%2Fwww.dep.go.th&ie=UTF-8&oe=UTF-8

[5] https://arxiv.org/abs/1406.1078

RESUME

Kriti Khurana

Delhi,India -110085
8178567131
kriti025btece22@igdtuw.ac.in

Passionate and driven B.Tech student with a focus on machine learning and a commitment to innovation.

EXPERIENCE

COE-AI Department Of IGDTUW— *Machine Learning Intern*
Completed an intensive training program in Python and machine learning techniques.
Developed and implemented machine learning models for various projects.

Projects:

- **Real Time Sign Language Recognition using Mediapipe LSTM Deep Learning Model:**

Designed and trained a Long Short-Term Memory (LSTM) deep learning model for real time sign language recognition .
Conducted extensive research on the same as a major project during the internship

- **Emotion Recognition using CNN:**

Developed a Convolutional Neural Network (CNN) model for emotion recognition.

- **Exploratory Data Analysis (EDA) on Body Fat Prediction:**

Conducted comprehensive EDA to analyze and prepare data for accurate body fat prediction.

Skills:

- Programming Languages: Python, SQL, C++, C
- Machine Learning and AI: Completed internship with AI Club IGDTUW, participated in and contributed to multiple machine learning projects, authored one research paper.
- Project Management: Experience in event management as a member of IEEE and Innerve societies at IGDTUW.
- Ongoing Project: Developing a DCFM system using NLP model

EDUCATION

Indira Gandhi Delhi Technical University , New Delhi—
B.tech, Electronics And Communication
Graduation year -2026