ST.TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

AFFILIATED TO MAHATMA GANDHI UNIVERSITY

FingerSpell: An AI-Powered Hand Sign Recognition System for American Sign Language (ASL) Translation

PROJECT REPORT

In partial fulfillment of the requirements for the award of the degree of

BACHELOR OF SCIENCE IN COMPUTER APPLICATIONS [TRIPLE MAIN]

Submitted By:

AKHILA DHANIYA

III BSc Computer Applications [Triple Main]

Register No: SB22CA019

Under the guidance of

Ms. MEGHA GEORGE

Assistant Professor

Department of Computer Applications

2022-2025

ST.TERESA'S COLLEGE (AUTONOMOUS), ERNAKULAM

AFFILIATED TO MAHATMA GANDHI UNIVERSITY

CERTIFICATE

This is to certify that the project report entitled "**FingerSpell: An AI-Powered Hand Sign Recognition System for American Sign Language (ASL) Translation**" is a bona fide record of the work done by **AKHILA DHANIYA (SB22CA004)** during the year 2022-2025 and submitted in partial fulfillment of the requirements for the degree of Bachelor of Science in Computer Applications (Triple Main) under Mahatma Gandhi University, Kottayam.

17/03/2025

Head of the Department

Internal Examiner

ST. TERESA'S COLLEGE DEPT. OF COMPUTERS ERNAKULAM

17/3/'25

External Examiner

Dr. Reseena.mol.N.A

Date: 17/03/2025

DECLARATION

I, AKHILA DHANIYA (Register No: SB22CA004), BSc Computer Applications [Triple Main] student of St. Teresa's College(Autonomous), Ernakulam , hereby declare that the project submitted for Bachelor's Degree in Computer Application is my original work. I further declare that the said work has not previously been submitted to any other university or academic body.

Date: 17/03/2025

Place: ERNAKULAM

AKHILA DHANIYA

ACKNOWLEDGEMENT

I would like to convey my heartfelt gratitude to **Rev. Sr. Nilima (CSST) Manager, Director Rev. Sr. Tessa (CSST) and Principal Dr. Alphonsa Vijaya Joseph** for providing me with this wonderful opportunity to work on a project with the topic FingerSpell: An AI-Powered Hand Sign Recognition System for American Sign Language (ASL) Translation.

I would like to express my profound gratitude to the Head of the Department of Computer Applications and my project guide Ms. Megha George and all other faculty of the department for their contributions to the completion of my project. The completion of the project would not have been possible without their help and insights.

Finally, I take this opportunity to thank all those who have directly or indirectly helped me with my project.

AKHILA DHANIYA

Paper was presented at **The International Conference On Sustainability, Innovation And Social Responsibility** conducted by **Naipunnya Institude Of Management And Information Technology (NIMIT)** on **19th February 2025**.

NAIPUNNYA INSTITUTE OF MANAGEMENT AND INFORMATION TECHNOLOGY (NIMIT)

Naipunnya®
To reach the unreachable

(Affiliated to the University of Calicut, Accredited by NAAC with A++, ISO 9001-2015 Certified)

Pongam, Koratty East, Thrissur District, Kerala State, India, Pincode - 680 308

This is to certify that

Ms. AKHILA DHANIYA

of **St Teresa's College (Autonomous) Ernakulam**

has presented a paper titled

FingerSpell: An AI-Powered Hand Sign Recognition system for American Sign Language (ASL) Translation

in the International Conference on 'Sustainability, Innovation and Social Responsibility' organized by Naipunnya Centre for Research, Naipunnya Institute of Management and Information Technology, Pongam on 19th of February 2025.

Dr. Antony Jose
Director, Naipunnya Centre for Research

Fr. Dr. Paulachan K. J.
Principal & Executive Director

ABSTRACT

Accessibility and communication inclusivity have gained extreme importance in a world that increasingly becomes interconnected. FingerSpell, an AI-based hand sign recognition system, wants to bridge this gap for a population of communication users who utilize American Sign Language (ASL) as their means of primary expression.

FingerSpell is an innovative project that aims to recognize and interpret American Sign Language (ASL) hand gestures by translating them into text in real time. The system mainly deals with the recognition of individual letters from A to Z by using a webcam to capture hand movements. It does not analyze full hand images but hand landmarks, like finger joints and key points, which makes the process more efficient and accurate in gesture detection.

The core of the project is built using TensorFlow, which helped in training a Fully Connected Neural Network (FCNN) to recognize hand signs based on the tracked key points. The FCNN processes the coordinates collected from hand movements and matches them to the corresponding alphabet letters. This model was trained for hand gesture classification with an accuracy rate of 98%, which is fairly reliable in classifying ASL letters. In order to use it on other devices, a TensorFlow Lite version of the trained model was created for optimal performance with low weight and speed. Hand tracking is carried out by using MediaPipe and patterns are interpreted using the neural network to identify every gesture.

FingerSpell is made available with Streamlit as an interface and, therefore, easily accessible on the web with a user-friendly interface. The present project works with the union of machine learning, computer vision, and neural networks to provide a practical tool that can aid inclusivity. In this context, by overcoming the communication gap for people using ASL-that is, deaf and speech- disabled people—this system reveals how everyday communication can become easier and liberating with modern technology for everyone.

Keyword :ASL Sign Language , Machine Learning, FingerSpell, Tensorflow, Streamlit, MediaPipe.

TABLE OF CONTENT

1. INTRODUCTION

In such a world where technology incessantly redefines interaction between human beings, access to communication tools has never been more necessary than now. For the deaf and hard-of-hearing population, sign language is one of the best ways for linguistic expression. Communication is severely constrained by the lack of understanding and the lack of adoption of sign language by the general non-signers. In order to unite everyone, creative ideas that bridge this divide are desperately needed.

Therefore, we tackle this problem by creating a sign language translator that transforms alphabetic letters (A–Z) into legible text. Recent studies have shown that deep learning techniques, such as CNNs and LSTMs, can achieve high accuracy in real-time ASL recognition, though challenges persist with visually similar hand signs (J. Smith et al., 2021).A Fully Connected Neural Network (FCNN), the architecture selected for our system for its accuracy and efficiency in handling pattern recognition tasks, lies at its core. By utilizing computer vision and artificial intelligence, our technology records hand movements, interprets them in real-time, and converts them into text, enabling smooth communication between sign language users and non-signers.

Actually, the focus on both technical innovation and social effect is what makes this initiative so significant. In addition to being a technological development, its goal is to provide a tool that will enhance comprehension, empathy, and inclusivity in real-world situations. The design, implementation, and evaluation of our FCNN-based system are covered in this paper, showcasing its potential to overcome linguistic and cultural barriers and make the world more accessible to everyone.

2. LITERATURE REVIEW

SI.NO	AUTHOR	FOCUS AREA	KEY FINDINGS	TECHNOLOGY/TOOLS
1	**J.Smithetal. (2021)**	Real-timeASL recognition using deep learning.	Achieved **85% accuracy** using CNN and LSTM; struggled with similar hand signs (M, N).	CNN,LSTM,Data Augmentation
2	**R.Pateletal. (2022)**	Hand gesture recognition using MediaPipe and ML.	Improved real-time recognition, **80%accuracy**, reduced preprocessing time.	MediaPipe Hands, TensorFlowClassifier
3	**L.Wongetal. (2020)**	Transfer learning for sign language recognition.	**90% accuracy**, effectivefeature extraction, but required high computational resources	MobileNetV2,ResNet, Transfer Learning

4	**D.Kimetal. (2019)**	ASL finger spelling using traditional ML algorithms.	SVM performed well with small datasets **(~75%)**;struggled with dynamic hand movements.	SVM,RandomForest, Hand Keypoints
5	**M.Rodriguez et al. (2022)**	Sign language recognition using hybrid deep learning models	CombinedCNN & BiLSTM to recognize ASL gestures with 92% accuracy; improved temporal feature extraction but required high processing power	CNN,BiLSTM, TensorFlow,Data Augmentation

3. METHODOLOGY

The methodology for this hand sign recognition project which uses Fully Connected Neural Networks (FCNN) follows a structured pipeline comprising data collection, preprocessing , model training, and deployment.

3.2 Data Collection

The dataset has hand gesture coordinates of American Sign Language (ASL) alphabets from A to Z stored as CSV file. Instead of using raw images, the dataset includes hand landmark coordinates extracted using MediaPipe Hand Landmark Model. Each alphabet has around 1500 samples, providing a diverse and well-balanced dataset for training and evaluation.

3.3 Data Preprocessing

To improve the accuracy and efficiency of the model, different preprocessing steps are applied to the dataset. Initially we converted the 21 handkey points (x, y) for each gesture into relative coordinates so model learns gesture patterns independent of the positioning of hands within the frame. Then the coordinates are flattened into a single vector to input into a neural network. Additionally, to avoid being affected by the large coordinate variations Normalization is performed by scaling the values within a fixed range. Preprocessing helps the model pay attention to only the structure and shape of hand gestures and not their absolute positions, which in turn improves generalization and reduces noise.

3.4 Model Building & Training

In the Model Building and Training stage of FingerSpell, we designed a Sequential Fully Connected Neural Network (FCNN) using TensorFlow Keras to recognize American Sign Language (ASL) letters A-Z. TensorFlow, along with Keras as a high-level API, has become a standard framework for training deep learning models due to its scalability and ease of use.The model architecture consisted of multiple Dense (Fully Connected) Layers with ReLU activations to learn complex patterns from hand landmarks. We used Dropout layers to prevent overfitting by randomly deactivating neurons during training. The output layer utilized a Softmax activation to convert model predictions into probabilities for multi-class classification. We compiled the model with the Adam optimizer for adaptive learning and categorical crossentropy loss suitable for multi-class classification, while monitoring accuracy as the performance metric.

The model was trained on a preprocessed dataset with 1000 epochs and a batch size of 128**,** ensuring a balanced training speed and memory usage. We used validation data to monitor the model's performance and avoid overfitting.

3.5 Model Testing and Evaluation

We perform Inference Test and display the Confusion matrix as well as the Classification Report. Inference testing is the process of using a trained machine learning model to make prediction**s** on new, unseen data. It is the final stage of the model lifecycle where the model's performance is evaluated in real-world scenarios. Tough there is minor difficulties in recognizing similar gestures like M,N,T and S, the model achieved an accuracy of 98% .

3.6 Deployment

The Model is deployed using Streamlit. Streamlit is a widely-used Python library that simplifies the process of building websites with just a few lines of code. Specifically designed for machine learning engineers, Streamlit provides a powerful framework with helpful tools for seamlessly integrating ML models and dataset files into web applications (Shukla et al., 2021).Through the Streamlit-based web interface users can interact with the model using webcam input. The interface captures the hand gestures, processes the landmarks, and maps them onto their corresponding alphabets, allowing seamless and interactive sign language recognition. The letters are then appended to words ,implementing the sign to text conversion(Kemka et al.,2023).

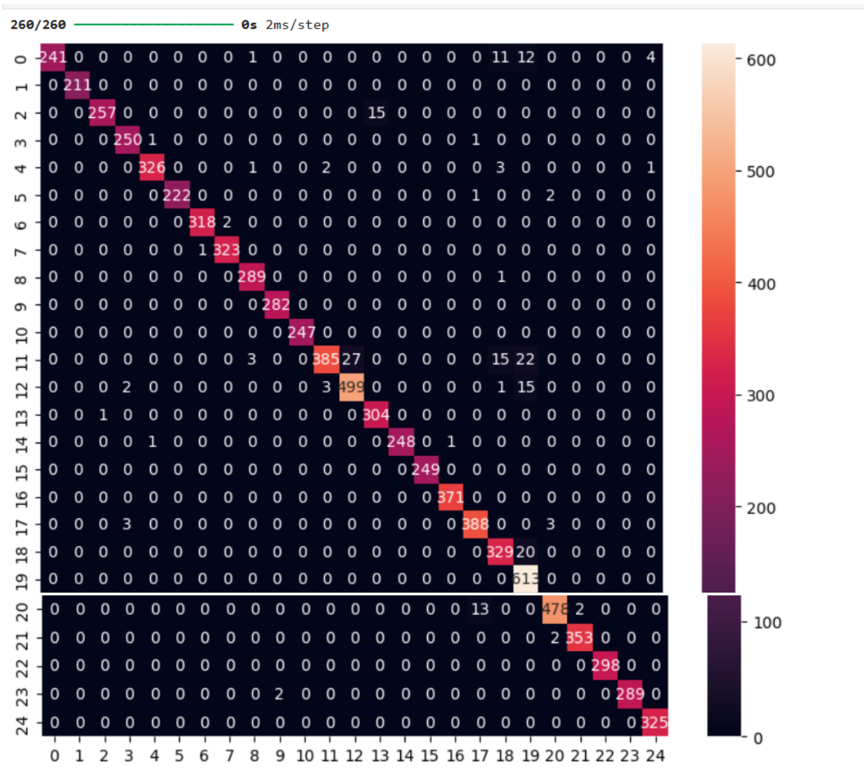
3.7 Model Architecture

The project uses Fully Connected Neural Network (FCNN) ,a type of deep learning framework for stacking layers in a linear fashion. FCNNs have been widely used for pattern recognition tasks, demonstrating their effectiveness in image and gesture classification (LeCun et al., 2015). The input layer takes a 42-dimensional vector: 21 landmarks × 2 coordinates. The first hidden layer consists of 64 neurons with ReLU (Rectified Linear Unit) activation, which helps introduce non-linearity and improves learning. Then a dropout layer follows to reduce overfitting. The second hidden layer has 32 neurons, also using ReLU activation, followed by another dropout layer for further regularization. Finally, the output layer applies a softmax activation function, enable multi-class classification across 26 classes.

4. RESULT

Here we evaluates the performance of FCNN models for Sign language Recognition. The results were assessed based on key metrics such as accuracy, precision, recall,and F1-score, derived from the classification reports and confusion matrices.

CONFUSION MATRIX & CLASSIFICATION REPORT



The confusion matrix is mostly diagonal and denoting a good classification of most of the test samples into the respective categories. The non-presence of sizable off-diagonal values signifies the minimal number of misclassifications, thereby demonstrating the model's efficiency in classifying separate classes. The diagonal values reflect the balance among the classes, showing that the model is not biased toward any specific class. This contributes to the overall reliability and accuracy of the model.

Classification Report

	precision	recall	f1-score	support
0	1.00	0.90	0.95	269
1	1.00	1.00	1.00	211
2	1.00	0.94	0.97	272
3	0.98	0.99	0.99	252
4	0.99	0.98	0.99	333
5	1.00	0.99	0.99	225
6	1.00	0.99	1.00	320
7	0.99	1.00	1.00	324
8	0.98	1.00	0.99	290
9	0.99	1.00	1.00	282
10	1.00	1.00	1.00	247
11	0.99	0.85	0.91	452
12	0.95	0.96	0.95	520
13	0.95	1.00	0.97	305
14	1.00	0.99	1.00	250
15	1.00	1.00	1.00	249
16	1.00	1.00	1.00	371
17	0.96	0.98	0.97	394
18	0.91	0.94	0.93	349
19	0.90	1.00	0.95	613
20	0.99	0.97	0.98	493
21	0.99	0.99	0.99	355
22	1.00	1.00	1.00	298
23	1.00	0.99	1.00	291
24	0.98	1.00	0.99	325
accuracy			0.98	8290
macro avg	0.98	0.98	0.98	8290
weighted avg	0.98	0.98	0.98	8290

- The model has overall **Accuracy** of **98**% .
- **Macro Average and Weighted average is also 98%** suggests that, on average, each class is performing well, meaning the model is not biased toward a particular class.
- Most classes have precision, recall, and F1-scores close to 1.00 (100%), meaning very few misclassifications. A few classes have slightly lower recall values (e.g., class 11), indicating some missed predictions for these classes.
- The support values indicate the number of test samples per class, ensuring no class is underrepresented.

5. IMPLEMENTATION

The system is being implemented with Streamlit, providing an interactive web interface for real-time hand sign recognition. Integrated with a webcam to capture live hand movements, it uses MediaPipe for landmark detection. The model predicts the corresponding letter and displays it on the interface by bounding boxes to represent detection confidence. Users can save the predicted letters to create words, adding a space when the space key is pressed, as the options are provided to clear and restart to enhance the usability.

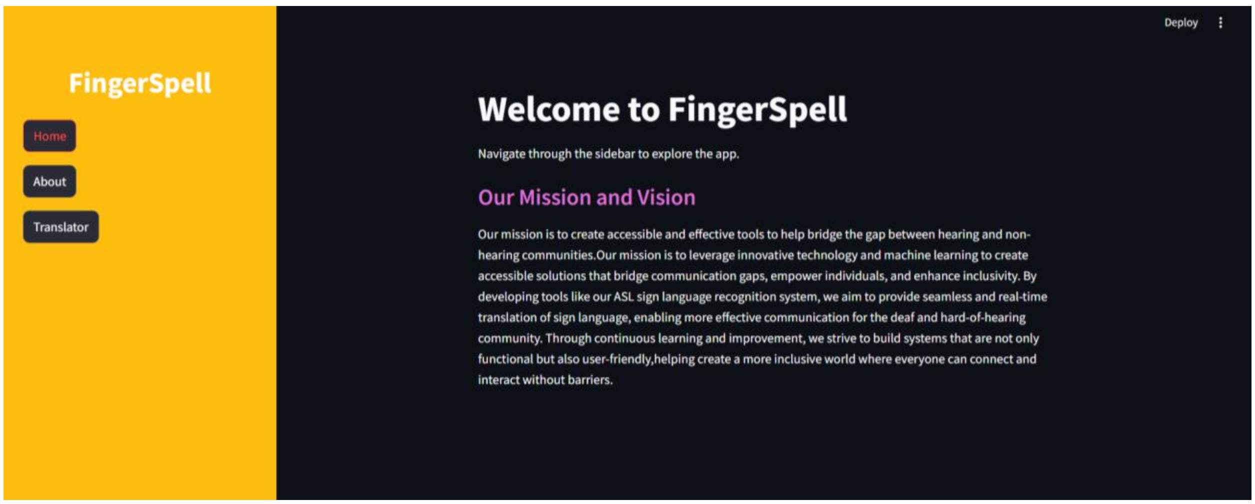


Figure 5.1

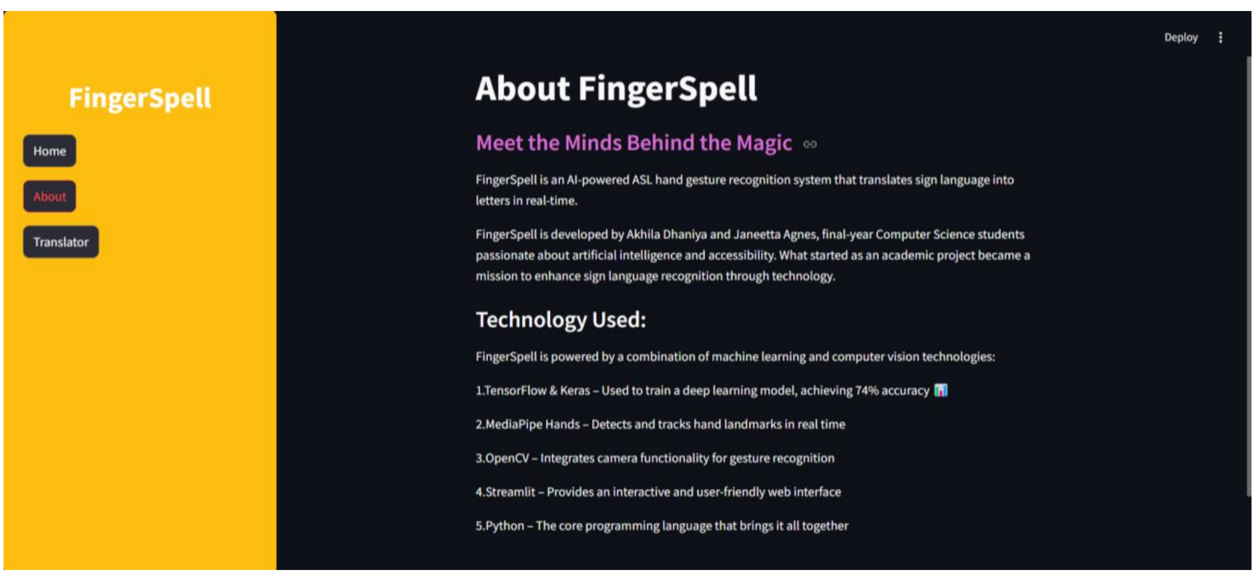


Figure 5.2

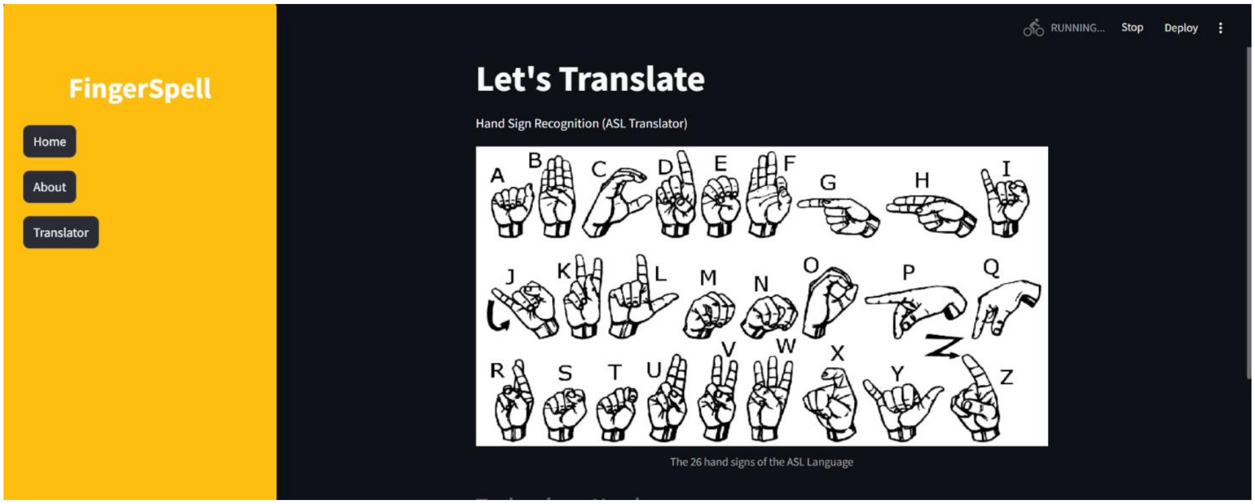


Figure 5.3

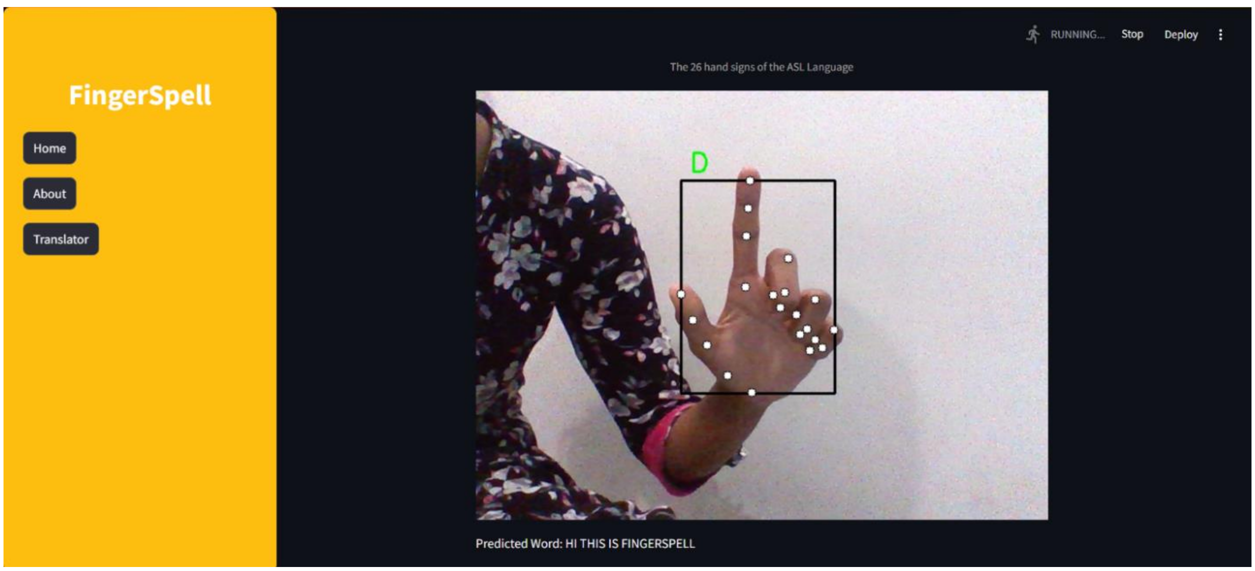


Figure 5.4

6. CONCLUSION

ASL sign language recognition system developed in this project demonstrates a big step forward to make communication access more enhanced for the deaf and hard-of-hearing community.It successfully applies advanced machine learning techniques through MediaPipe hand landmark Model and TensorFlow model training in hand gesture recognition of the whole alphabet.

Although the model shows minor difficulties in distinguishing between similar gestures, such as 'M' ,'N' and ‘T’, it is still promising to achieve an accuracy of 98%. The real-time gesture recognition and user-friendly deployment on Streamlit show how the system may be used practically for translating American Sign Language into text.The project not only showcases the power of deep learning but also works towards making more inclusive communication technologies. Further improvements in the model's accuracy and including more features such as word prediction and recognition of multi-hands would push the system into further utility and impact.

Potential future improvements for this project include scaling up the dataset by including a wider variety of hand shapes, lighting setups, and backgrounds to enhance model generalization. Implementing dynamic gesture recognition for complete words and phrases, instead of single letters would greatly boost the practical usability of the system. Further improvement in real-time detection accuracy and latency optimization would make the model more responsive and friendly. Adding deep learning-powered hand tracking, multi-angle gesture analysis, and multilingual sign language support to further extend its reach and efficiency. All in all, this project demonstrates the power of AI-powered sign language recognition and lays the groundwork for future developments in gesture-based human-computer interaction and assistive communication technologies.

7. REFERENCES

- S. M. Mohid, M. N. K. Chowdhury, and M. S. Islam, "Real-Time American Sign Language Recognition Using Mediapipe and TensorFlow," *International Journal of Computer Applications*, vol. 183, no. 47, pp. 1–6, Mar. 2022.
- R. Kadam, A. Suryawanshi, and P. Padiya, "Hand Gesture Recognition System Using Machine Learning and Mediapipe," *Procedia Computer Science*, vol. 192, pp. 1793–1802, 2021.
- S. Gupta, R. Singh, and A. Sharma, "Comparative analysis of Fully Connected Neural Networks (FCNN) and Convolutional Neural Networks (CNN) for image-based applications," in *Proceedings of the 2022 IEEE Conference on Artificial Intelligence and Machine Learning (AIML)*, 2022, pp. 112–118.
- J. Huang, W. Zhou, and H. Li, "Sign Language Recognition Using Deep Learning: A Review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 5, pp. 1766–1782, May 2022.
- R. V. Lopes et al., "A Comparative Study of Machine Learning Algorithms for Hand Gesture Recognition," *International Journal of Human-Computer Interaction*, vol. 39, no. 3, pp. 457–472, 2022.
- OpenCV Developers, "Real-Time Hand Detection and Tracking," *OpenCV Documentation*, 2023.