Real Time Sign Language Recognisation

Submitted by:

Nikita (102317245)

Kashika Jindal(102317154)

BE Third Year

COPC

Submitted to:

Dr. Anjula Mehto

Assistant Professor

THAPAR INSTITUTE
OF ENGINEERING & TECHNOLOGY
(Deemed to be University)

Computer Science and Engineering Department

Thapar Institute of Engineering and Technology, Patiala

November 2025

TABLE OF CONTENTS

1.INTRODUCTION

Sign Language is a structured, expressive, and visual form of communication primarily used by individuals with hearing or speech impairments. However, despite being a complete and natural language, most people who are not hearing-impaired do not understand sign language. This creates a significant communication barrier between the Deaf community and the rest of society.

To address this challenge, this project aims to develop an **automated Sign Language Recognition (SLR) system** that can interpret hand gestures in real-time and convert them into meaningful text or speech. Such systems play an essential role in assistive technology, enabling smoother interaction in education, public services, workplaces, hospitals, and daily communication.

The system integrates **computer vision**, **machine learning**, and **deep learning** to detect, analyze, and classify hand gestures from live video input. Instead of using images, this project uses **hand keypoint landmarks** generated using the MediaPipe Hands model—making the system lightweight, fast, and more accurate even with limited data.

The sign language recognition system consits of several components:

1.Data collection: Hand gesture data is collected using live webcam input that includes english aphpabets from (a-z) excluding j and z which requires motion.

2.Preprocessing: The collected data is preprocessed to enhance the quality,remove noise and extract relevant features using opencv and mediapipe.

3.Model Training: Deep learning model i,e.Convolutional neural networks(CNNs) is trained on the preprocessed data to learn the mapping between input gestures and their corresponding meanings.

4.Real-time recognition: The trained model converts sign language alphabets to text when captured using camera.

2. PROBLEM STATEMENT

2.1 The challenge

Communication between individuals who use sign language and those who rely on spoken language often becomes difficult due to the absence of a common medium. In many real-life situations, especially in places like hospitals, schools, public offices, and workplaces, immediate communication is crucial, yet trained sign-language interpreters are rarely available. Even when they are available, interpreter services can be costly, inconvenient, or impractical for everyday conversations. This communication barrier leads to misunderstandings, delays, and reduced accessibility for people who depend on sign language.

2.2 The technical need

To address this communication gap, there is a need for an intelligent, automated system capable of recognizing sign language gestures accurately in real time. However, traditional image-based gesture recognition systems are computationally heavy and require large datasets and powerful hardware. This project instead focuses on a lightweight pipeline that uses hand landmarks rather than full images, allowing the model to perform efficiently even on normal devices without GPUs.

2.3 The solution

The Sign Language Recognition System developed in this project translates static hand gestures into alphabets using machine learning, enabling seamless interaction between sign-language users and non-users. The use of hand keypoints significantly reduces computational complexity while maintaining accuracy, making real-time prediction possible. With a lightweight TFLite model, the system becomes accessible, scalable, and affordable for everyday communication needs. Overall, this project offers a practical solution that improves accessibility and inclusion by bridging the communication gap through technology.

3. OVERVIEW OF THE DATASET USED

The dataset used in this project consists of hand landmark coordinates extracted from images of American Sign Language (ASL) gestures for the alphabets A to Z, excluding J and Z. The remaining 24 alphabets are represented by static hand signs,Each sample in the dataset contains the x and y coordinates of 21 hand landmarks, giving a total of 42 numerical features per record.The dataset contains a total of 24,000 rows, with 1,000 samples for each of the 24 gesture classes. Every row consists of:

42 numerical values (x_1, y_1, x_2, y_2, … x_{21}, y_{21}) representing hand keypoints.

1 label indicating the corresponding alphabet (A, B, C, …, excluding J, Z).

Figure 1. 21 landmarks of hand captured by mediapipe

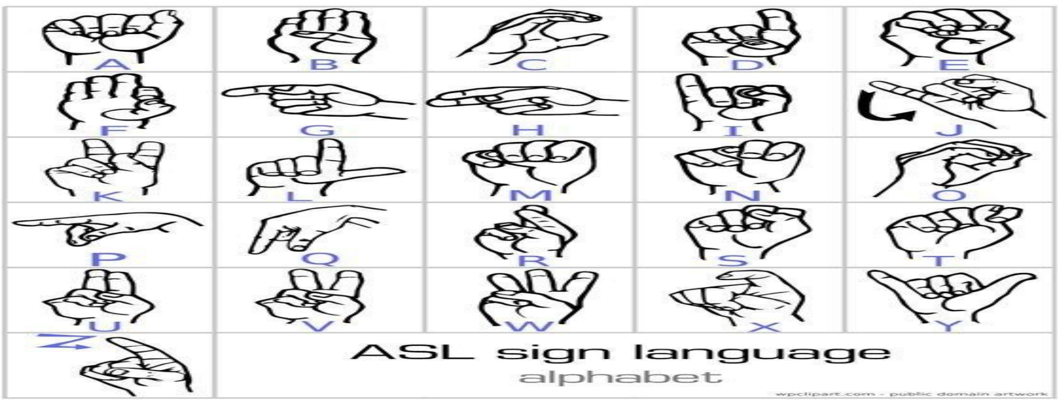

Figure 2. American sign language alphabets

4. PROJECT WORKFLOW

The project involves capturing dataset using opencv and mediapipe and includes preprocessing to train the model using convolutional neural network . Below is the flowchart depicting the workflow:

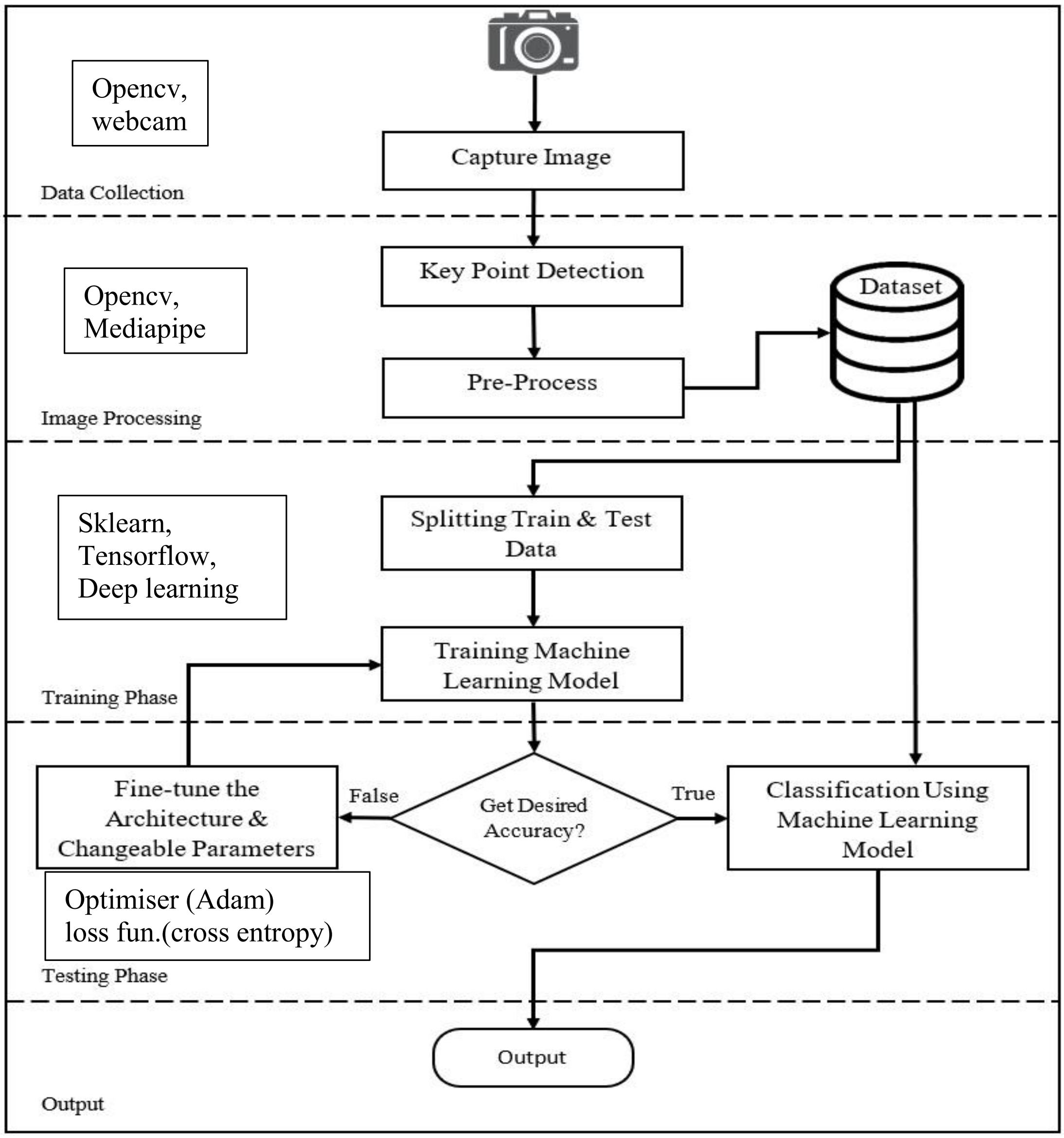

Figure 3. Flowchart

5. RESULTS

The data was split into training and test sets using a train-test ratio of 75:25, i.e., 75% for training and 25% for validation/testing. This allows the model to learn patterns on the training set and validate generalization on unseen data (test set).

The model is a fully connected (Dense) neural network built using TensorFlow/Keras.

1. Input layer: 42 neurons (for 42 keypoint coordinates)
2. Dropout: 0.2 (helps prevent overfitting)
3. Dense: 20 neurons, ReLU activation
4. Dropout: 0.4
5. Dense: 10 neurons, ReLU activation
6. Output layer: 24 neurons, Softmax activation

Optimiser and loss function ;

Adam is used as an optimiser which uses adaptive learning rate optimization algorithm and It works well for small datasets and non-linear data, which is suitable for hand keypoint recognition. sparse_categorical_crossentropy is used as loss function because it is a multi-class classification problem, and the labels are integers (0-23) rather than one-hot encoded vectors. The model was trained for **100 epochs** and that the results are given below:

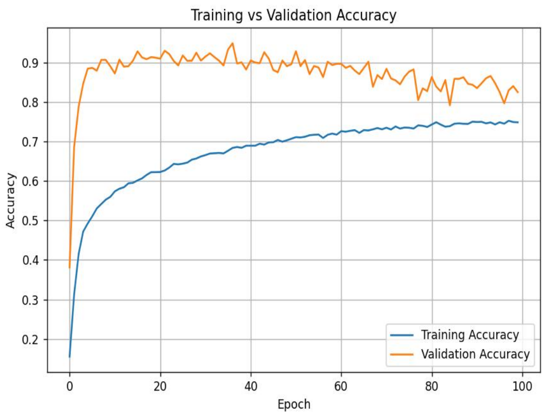

Figure 4. Graph for accuracy

	Training	Validation
Accuracy	75.5%	82.4%
Loss	0.748	0.437

Figure 5. Table for accuracy and loss

6. CONCLUSION

The Sign Language Recognition system developed in this project successfully demonstrates how technology can bridge the communication gap between the Deaf community and the hearing population. By using hand landmark coordinates instead of full images, the model remains lightweight, fast, and efficient, making it suitable even for devices without advanced hardware. The system achieves reliable accuracy in predicting static ASL alphabets and proves that hand keypoint–based recognition can be an effective alternative to traditional image-based methods.

Through this project, we were able to implement the complete pipeline—data collection, preprocessing, model training, and real-time prediction—showing how computer vision and machine learning can be combined to solve meaningful real-world problems. Although the model is currently limited to static gestures and excludes letters like J and Z, it provides a strong foundation for future enhancements such as dynamic gesture recognition, sentence formation, integration with mobile apps, and real-time speech output.

Overall, this project highlights the potential of AI-driven assistive technology to promote accessibility, inclusivity, and independence for individuals with hearing impairments, making communication smoother and more natural in everyday life.