**PROJECT BY:**

UMAIMA KHAN

AYESHA SALEEM

SIGN TO ALPHABET CONVERTOR

**Contents**

[**ABSTRACT** 2](#_Toc189872432)

[**INTRODUCTION** 2](#_Toc189872433)

[**DATASET DESCRIPTION** 2](#_Toc189872434)

[**CHALLENGES:** 3](#_Toc189872435)

[**METHODOLOGY:** 3](#_Toc189872436)

[**RESULT** 4](#_Toc189872437)

[**ANALYSIS / DISCUSSIONS:** 4](#_Toc189872438)

[**CONCLUSION:** 5](#_Toc189872439)

**ABSTRACT**

The Alphabet to Sign Language Converter is an innovative real-time application designed to bridge the communication gap between hearing individuals and the deaf or hard-of-hearing community. The system uses computer vision and machine learning to recognize hand gestures corresponding to the English alphabet and translates them into their respective letters. Built using Python, Flask, OpenCV, and Mediapipe, the application provides an interactive platform for users to learn and practice the American Sign Language (ASL) alphabet. The project incorporates multiple modes (Easy, Medium, Hard, and Freestyle) to cater to users of varying skill levels, making it a versatile tool for both beginners and advanced learners**.**

# **INTRODUCTION**

Sign language is a critical mode of communication for millions of people worldwide. However, learning sign language can be daunting for beginners due to the complexity of hand gestures and the lack of accessible learning tools. This project aims to address this challenge by developing a real-time alphabet-to-sign language converter that leverages cutting-edge technologies like computer vision and machine learning.

The application captures hand gestures via a webcam, processes them using the Mediapipe library, and predicts the corresponding letter using a pre-trained Random Forest classifier. The system is designed to be intuitive and user-friendly, with a web interface built using Flask. The project not only serves as a learning tool but also demonstrates the potential of technology to enhance accessibility and inclusivity.

# **DATASET DESCRIPTION**

The dataset used for training the Random Forest classifier is a critical component of the project. It consists of hand gesture images corresponding to each letter of the English alphabet. Below is a detailed breakdown of the dataset:

***Data Collection:***

Hand gestures for each letter of the alphabet were captured using a webcam. The Mediapipe library was used to extract 21 hand landmarks for each gesture, representing key points on the hand (e.g., fingertips, knuckles, wrist).

***Preprocessing:***

The hand landmarks were normalized relative to the wrist position to ensure consistency across different hand sizes and orientations. Additional features, such as the distance between specific landmarks, were computed to enhance the model's accuracy.

***Dataset Composition:***

The dataset includes multiple samples for each letter to ensure robust model performance. The dataset was split into training and testing sets, with 80% of the data used for training and 20% for testing.

# **CHALLENGES:**

Some letters with similar hand shapes (e.g., 'P' and 'R') posed challenges during data collection and preprocessing. These letters required additional samples and feature engineering to improve classification accuracy.

# **METHODOLOGY:**

The project follows a structured and systematic approach to achieve real-time hand gesture recognition and letter prediction. Below is a detailed explanation of the methodology:

***Hand Detection and Landmark Extraction:***

The Mediapipe library is used to detect hands in the video feed and extract 21 hand landmarks for each detected hand. The landmarks are normalized relative to the wrist position to ensure consistency across different hand sizes and orientations.

***Feature Extraction:***

The normalized landmarks are used as features for the machine learning model. Additional features, such as the distance between specific landmarks, are computed to enhance the model's accuracy.

***Machine Learning Model:***

A Random Forest classifier is trained on the preprocessed dataset to predict the corresponding letter based on the hand landmarks. The Random Forest algorithm was chosen due to its ability to handle high-dimensional data and its robustness to overfitting. The model is saved as a joblib file (random\_forest.joblib) for real-time inference.

***Real-Time Prediction:***

The webcam captures the user's hand gestures in real-time. The hand landmarks are extracted and passed to the trained model for prediction. The predicted letter is displayed on the screen along with visual feedback.

***Web Interface:***

The Flask framework is used to create a web interface for the application. The interface provides a live video feed of the hand gesture recognition process. The interface is designed to be intuitive and user-friendly, with clear instructions and visual cues.

# **RESULT**

The application successfully recognizes hand gestures corresponding to the English alphabet and displays the predicted letter in real-time. Below are the key results:

***Accuracy:***

The Random Forest classifier achieves high accuracy in predicting letters, particularly for gestures with distinct hand shapes (e.g., 'A', 'B', 'C'). The model performs well in the Easy and Medium modes, where users are guided through the process of forming each letter.

***Real-Time Performance:***

The system performs well in real-time, with minimal latency between gesture recognition and letter prediction. The application is able to process video frames at a rate of 20-30 frames per second, ensuring a smooth user experience.

***User Feedback:***

Users found the application to be intuitive and easy to use. The inclusion of different modes was particularly appreciated, as it allowed users to progress at their own pace.

# **ANALYSIS / DISCUSSIONS:**

The project demonstrates the potential of computer vision and machine learning to enhance accessibility and inclusivity. Below is a detailed analysis of the system's performance and areas for improvement:

***Strengths:***

The system is able to recognize a wide range of hand gestures with high accuracy. The inclusion of multiple modes makes the application accessible to users of varying skill levels. The real-time performance of the system is impressive, with minimal latency between gesture recognition and letter prediction.

***Challenges:***

Some letters with similar hand shapes (e.g., 'P' and 'R') pose challenges for the model, leading to occasional misclassifications. The system relies on a webcam for input, which may limit its usability in low-light conditions or with low-quality cameras.

***Future Work:***

* Expand the dataset to include more samples for challenging letters, improving the model's accuracy.
* Incorporate additional sign languages (e.g., British Sign Language, Indian Sign Language) to make the application more versatile.
* Develop a mobile version of the application to increase accessibility and usability.

**CONCLUSION:**

The Alphabet to Sign Language Converter is a successful implementation of a real-time hand gesture recognition system. By leveraging computer vision and machine learning, the application provides an effective tool for learning and practicing the ASL alphabet. The project demonstrates the potential of technology to enhance communication and accessibility for the deaf and hard-of-hearing community.

The system's ability to recognize hand gestures in real-time, combined with its user-friendly interface and multiple modes of operation, makes it a valuable learning tool for both beginners and advanced learners. Future work could focus on expanding the system to include words, phrases, and other sign languages, further increasing its utility and impact.