SIGN LANGUAGE INTERPRETER

Bhupendra Singh Chaudhary - 23200179
Sahil Chalkhure - 23206052
ACM40960 Projects in Maths Modelling, MSc Data and Computational Science,
University College Dublin

Abstract

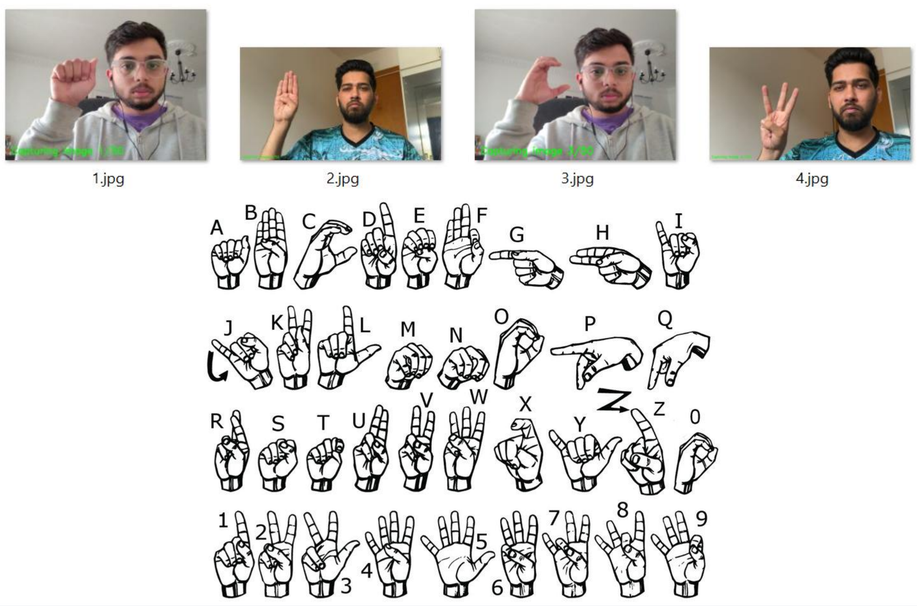
Sign language is an important communication tool for people who are deaf. This project presents a **real-time sign language interpreter** for recognising **American Sign Language** letters and numbers using a **webcam**. It uses **MediaPipe** to extract hand landmarks and a **Random Forest classifier** for gesture recognition. The **self-collected dataset** was processed, resulting in a highly **accurate model** with perfect evaluation metrics, including an **AUC** and **precision-recall score** of **1.00**. The system effectively bridges communication gaps by **translating gestures into text or speech**, with potential future improvements focused on expanding gesture vocabulary and adapting to more complex scenarios.

Motivation

This project is motivated by the **communication challenges** faced by **deaf people**. Although **sign language** is commonly used in the deaf community, it often creates barriers when communicating with those who don't understand it. This can cause problems in everyday situations like at school or in healthcare. Our project is designed to **bridge this gap** by translating sign language into **text or speech**, making it easier for everyone to communicate. That's what inspired us to develop this idea.

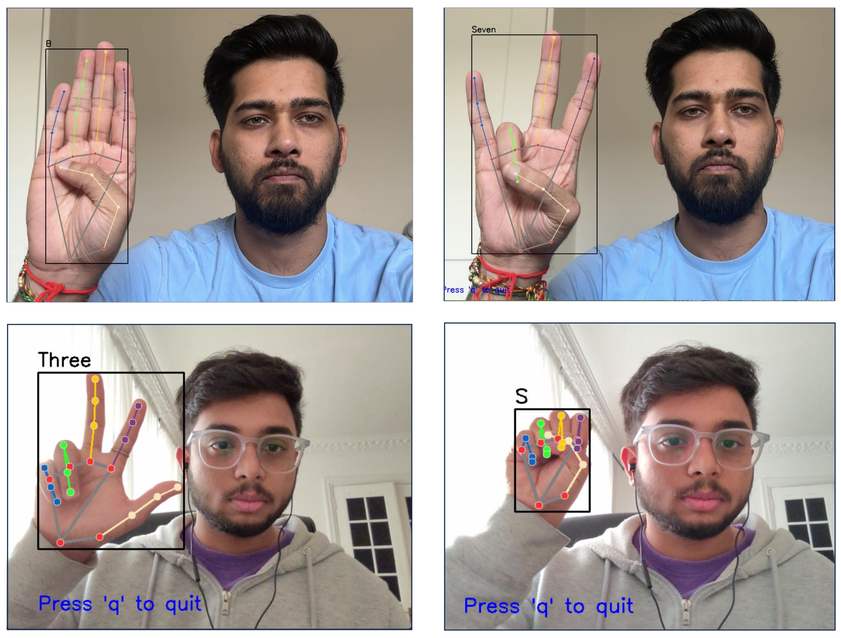
Dataset Used

The **dataset** we used for this project was collected using a **laptop webcam**. It includes **images of a hand showing gestures** for **letters (A-Z)** and **numbers (0-9)**. The images were captured in **real-time** and labelled. **200 samples** were collected for each gesture to improve **model training**. The dataset was organized into labelled folders for easy processing. The **gestures** we trained are given in the image below.



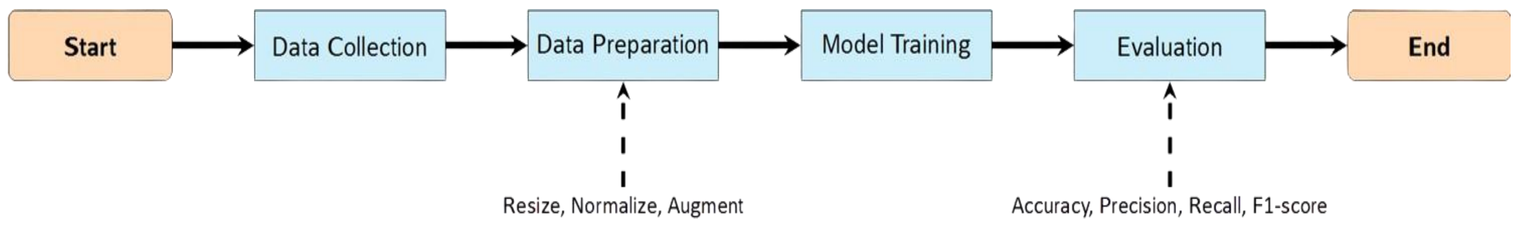
Results

- The model showed **high accuracy** in recognising **American Sign Language gestures**, performing consistently across different **letters and numbers**.
- The **confusion matrix** showed that most gestures were correctly classified with very few errors.
- The **ROC** and **precision-recall curves** confirmed the model's **reliability**.
- The real-time application allows users to **create words** using the predicted **letters and numbers** giving them more **control** over the output.
- The system effectively translates hand gestures into **audio and visual outputs**, proving its **practical usefulness**.



Methodology

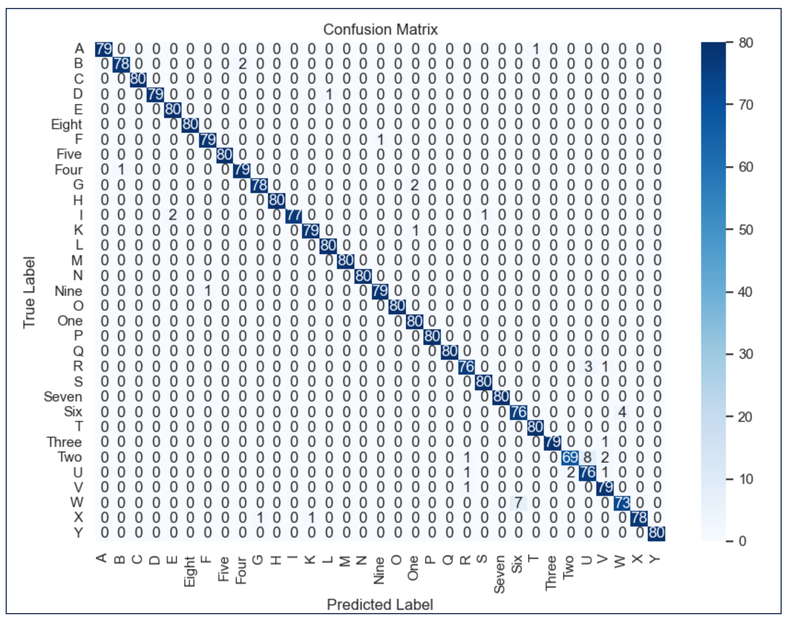
- **Data Collection**: The data for the hand gestures are captured using a **webcam** and stored in **image format** in labeled folders, each representing a **letter or number**.
- **Data Preprocessing**: Images are then processed using **MediaPipe** to extract key **hand landmarks**, which are essential for recognising gestures.
- **Model Training**: A **RandomForest classifier** is trained on the extracted features and evaluated using metrics like **accuracy, confusion matrices and ROC curves**.
- **Real-time Inference**: A **real-time app** predicts gestures from live **webcam feed** and plays an **audio file** matching the recognised gesture for smoother interaction.



Exploratory Data Analysis

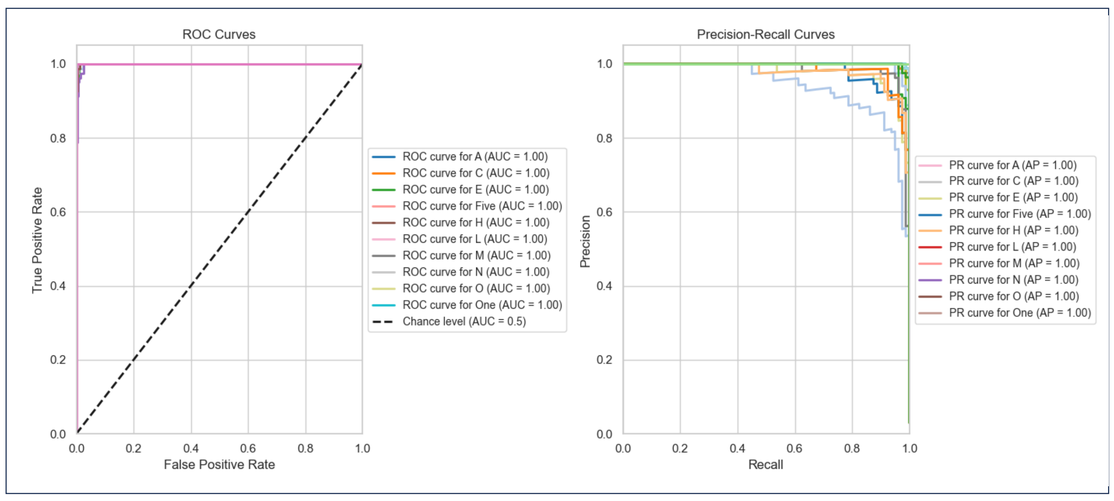
CONFUSION MATRIX

The confusion matrix shows that the model is mostly **accurate**, but it does make some **small mistakes**. While most gestures are **correctly recognized**, there are a few mix-ups, especially between similar signs like **'W' and 'Six'**. The diagonal pattern still indicates that the model gets most predictions right, showing that it's generally **reliable**. However, some errors suggest that the model could be improved for even better accuracy. Overall, it performs well in **real-time**, but some adjustments could help reduce these minor mistakes.



ROC AND PRECISION RECALL CURVES

The **ROC curves** show that the model does a great job, with most gesture classes having **AUC scores close to 1.00**, meaning it can easily tell the difference between signs. The **Precision-Recall curves** are also strong, with lines mostly near the top, showing high **precision and recall**. Although there are slight dips from perfect scores, these are small, meaning the model is still very reliable for real-time use, with only a few chances of getting confused between similar gestures. Overall, it's a **strong and effective tool** for **accurate sign language recognition**.



Discussion

The project does a good job of recognising common hand signs accurately and provides **real-time visual & audio feedback**. It works well even on basic hardware and its **clear visual output** makes it easy for users to see how their signs are being recognised but it struggles with complex signs and finds it **difficult to handle continuous hand movements**. The system's **performance also drops in low light** or when hands are partially hidden, yet the project offers a **strong foundation** for building an **effective sign language recognition tool**.

Conclusion and Future Scope

This project shows how **computer vision** and **machine learning** can create a sign language interpreter for **letters and numbers.** The **option to create sentences** using predicted letters and numbers gives **users more control** and makes the system **easier to use** in real-time. The system works well but can be improved by **expanding the gesture set** to include **more words**, **improving performance** in various lighting and backgrounds and **adding support for two-handed gestures**. Additionally, **increasing the number of samples** collected could better fit the model and improve overall accuracy.

Scan to view

GitHub Repositories

bhupendrachaudhary08

sahilchalkhure26

References

1. "Hand Gesture Recognition using MediaPipe" by Zhang, S.
2. "Sign Language Recognition Using Convolutional Neural Networks." by Kazemi, V.
3. Google MediaPipe Documentation: https://github.com/google-ai-edge/mediapipe