

SIGN LANGUAGE TO TEXT CONVERTER

PRESENTED BY:

- AHD MAMDOUH
- OMAR MOHAMED
- ALI SALAH
- EYAD MOHAMED

UNDER SUPERVISION OF: DR. NOHA EL-SAYAD

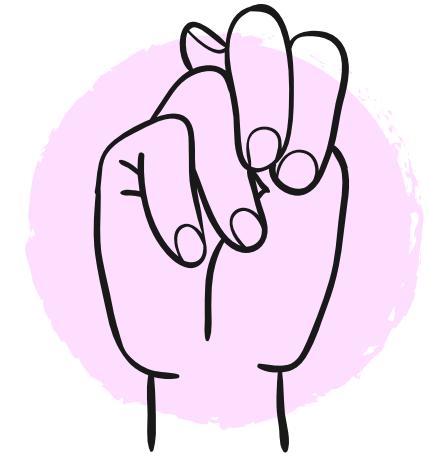
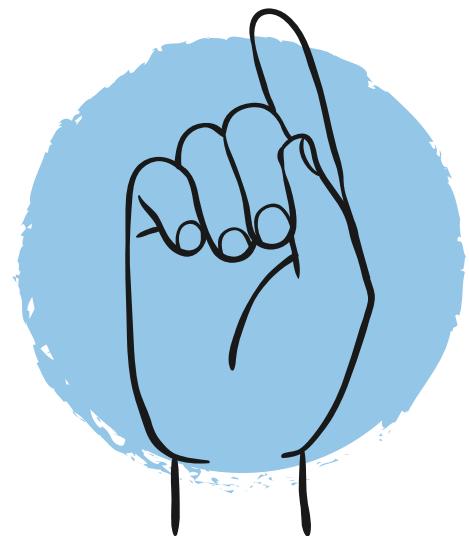


TABLE OF CONTENTS

01

INTRODUCTION

02

PROBLEM STATEMENT &
OBJECTIVE

03

DIAGRAMS

04

PROJECT MODULES

05

AI MODEL

06

DATA SET
&
CNN

07

FUTURE WORK

01

INTRODUCTION

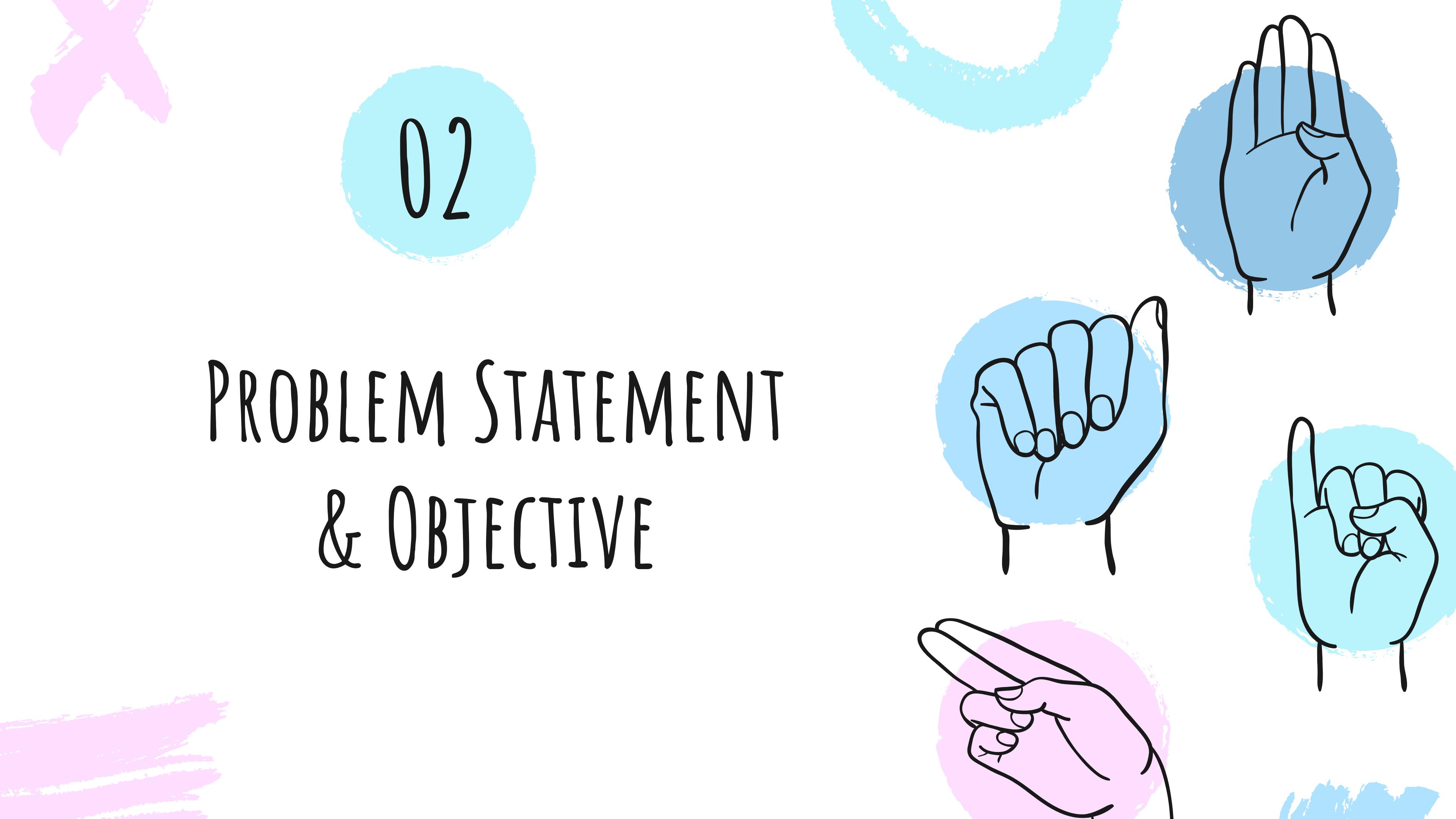
INTRODUCTION

Over 70 million deaf individuals worldwide rely on sign language to communicate – to learn, to work, and to belong.

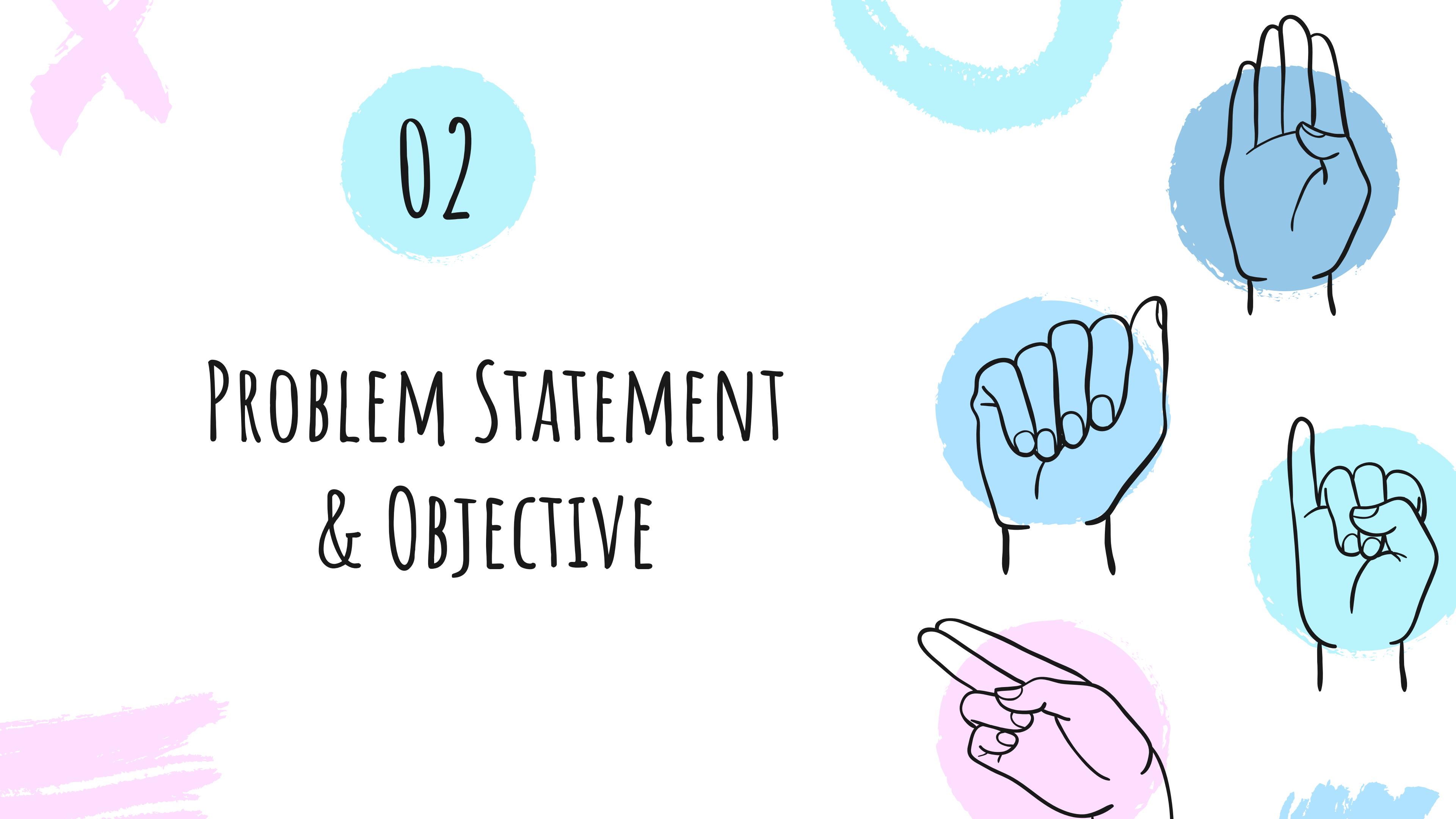
Yet, millions of interactions are lost every day due to a communication gap between signers and non-signers.

With AI and computer vision, we aim to bridge this gap – turning hand gestures into real-time, readable language to enable equal access and connection for all.





02



PROBLEM STATEMENT & OBJECTIVE

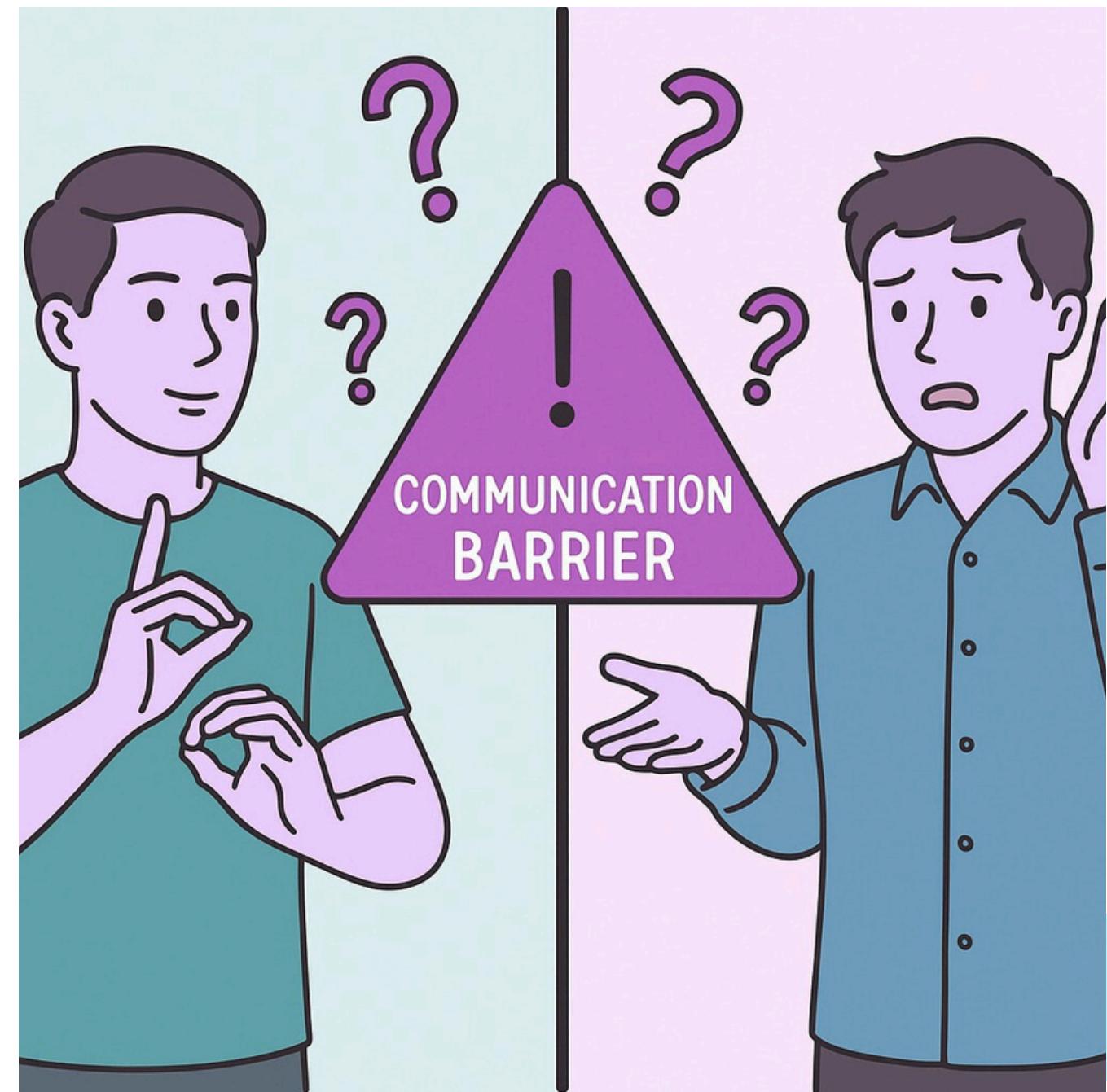
PROBLEM STATEMENT

Sign language is vital – but not universally understood.”

Despite the widespread use of sign language among the deaf and hard-of-hearing community, most people don't understand it.

This creates a communication barrier in:

- **Education**
- **Employment**
- **Healthcare**
- **Everyday interactions**

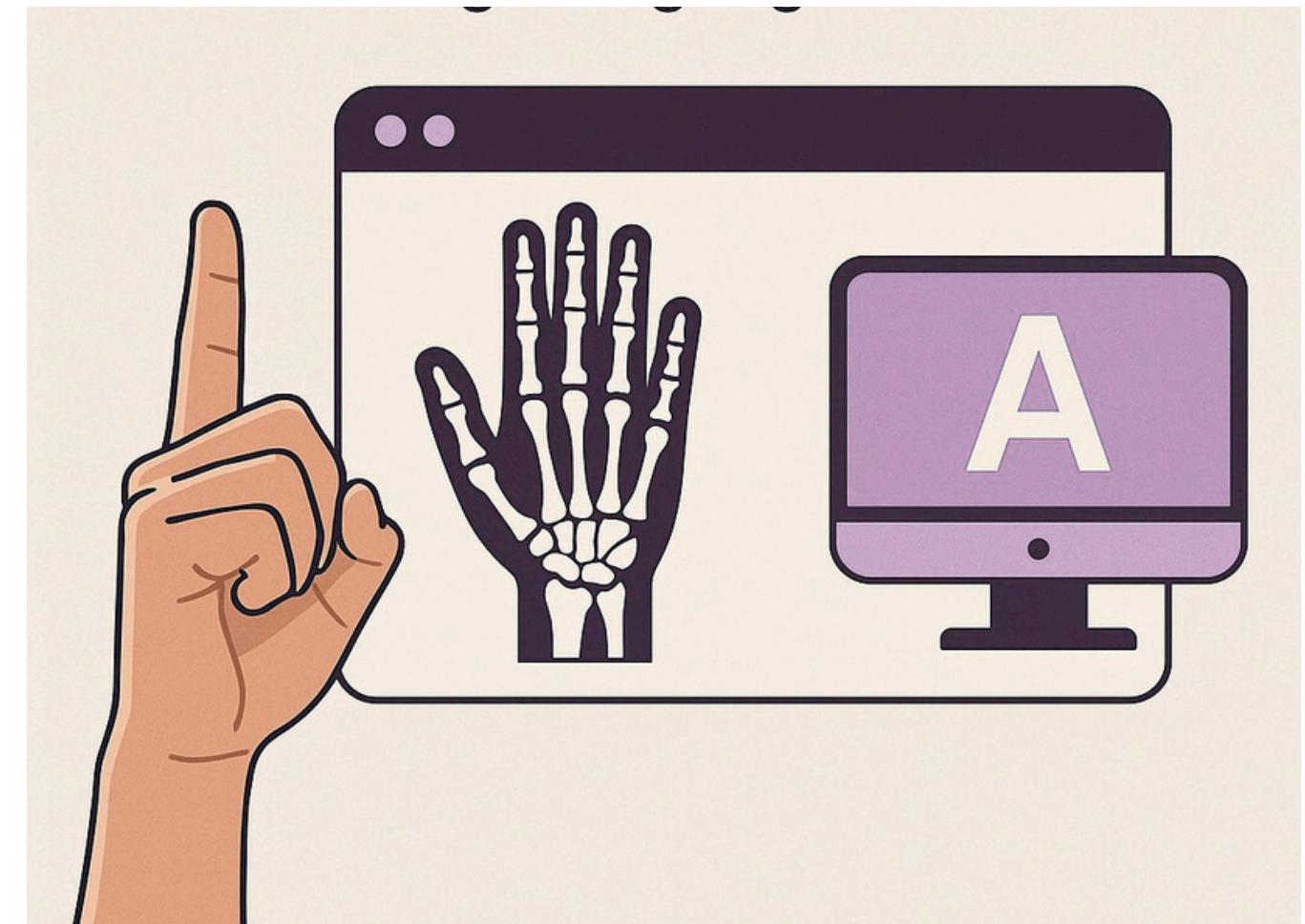


OBJECTIVE

To bridge the communication gap using computer vision and AI.”

Design and implement a vision-based system that:

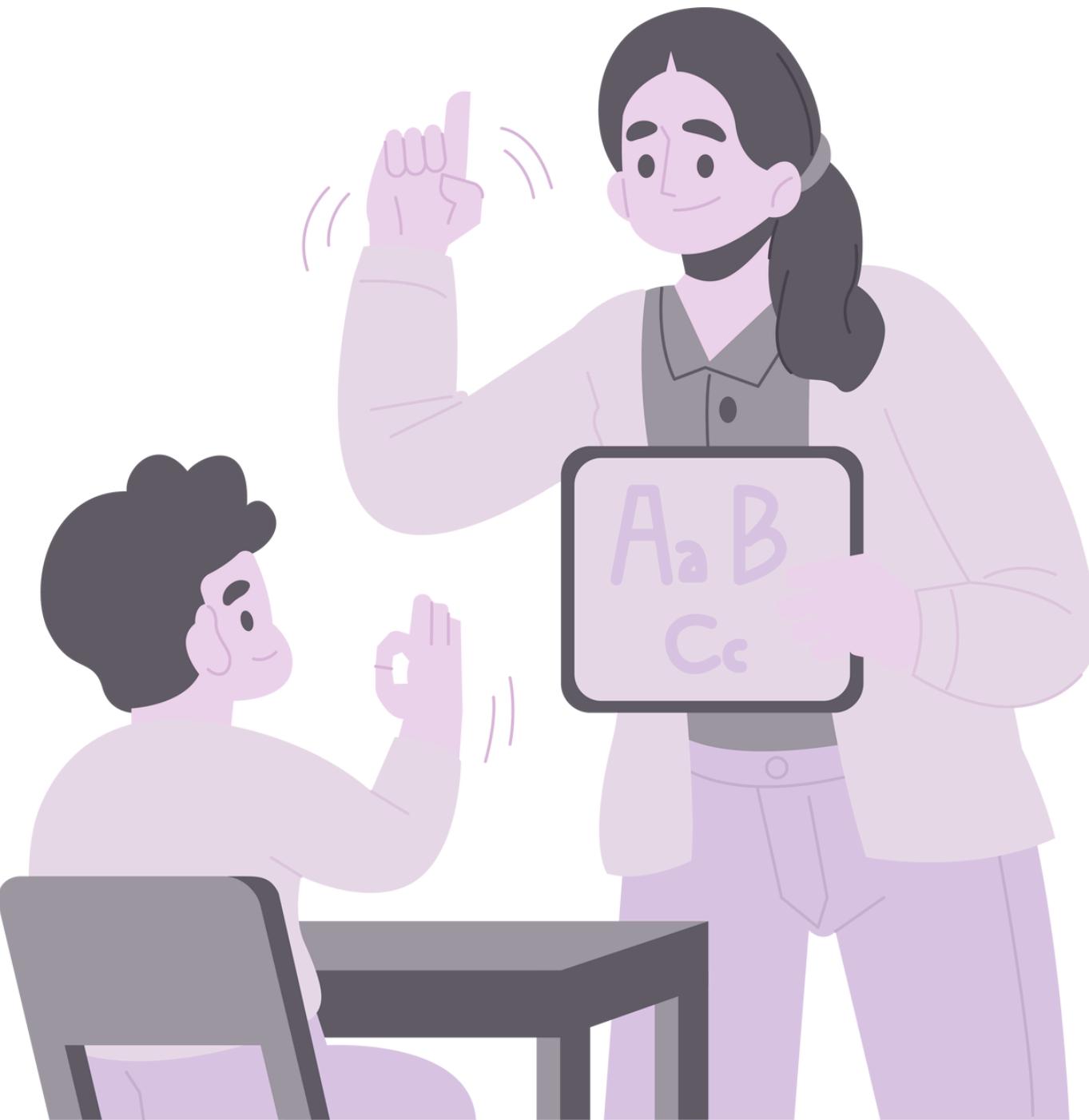
- 🎥 **Captures hand gestures using a live camera feed**
- 👉 **Extracts precise hand landmarks using MediaPipe technology**
- 🧠 **Recognizes American Sign Language (ASL) letters with a CNN model**
- 💬 **Converts the output into text**



The goal is a low-cost, real-time, and scalable solution to make communication more inclusive.

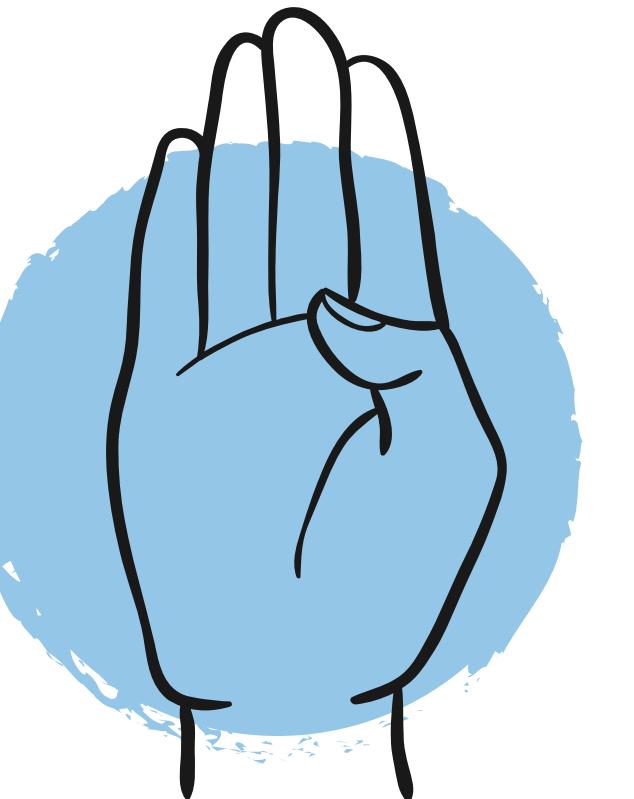
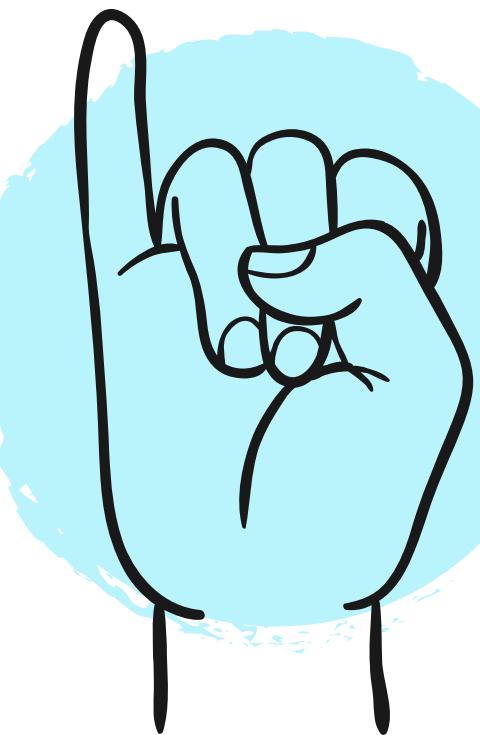
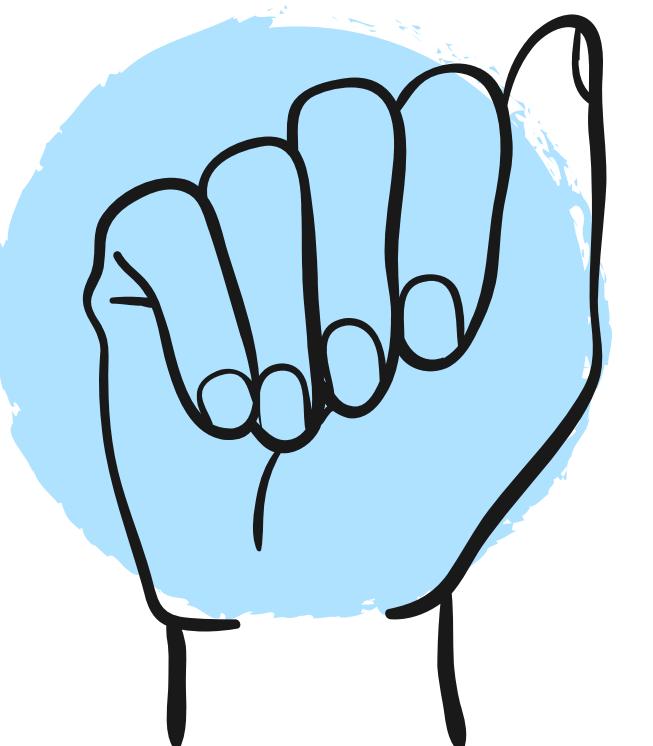
TARGET USERS

- **Deaf and Hard-of-Hearing individuals – to communicate with people who don't understand sign language.**
- **Non-signers (general public) – to interpret ASL letters from deaf users.**
- **Educators and Students – as an assistive tool in inclusive classrooms.**
- **Customer service and public service sectors – to bridge communication gaps in banks, hospitals, and airports.**

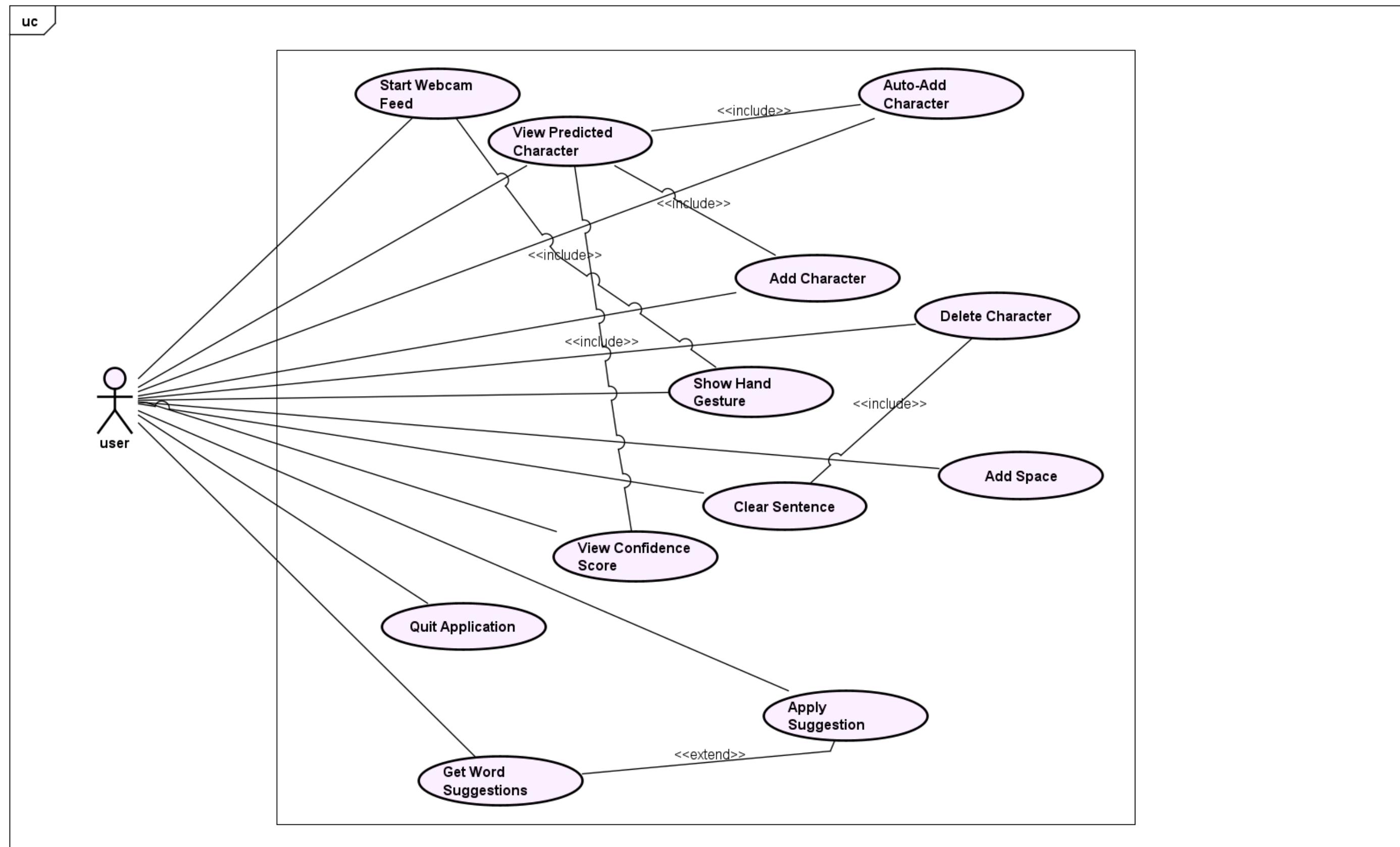


03

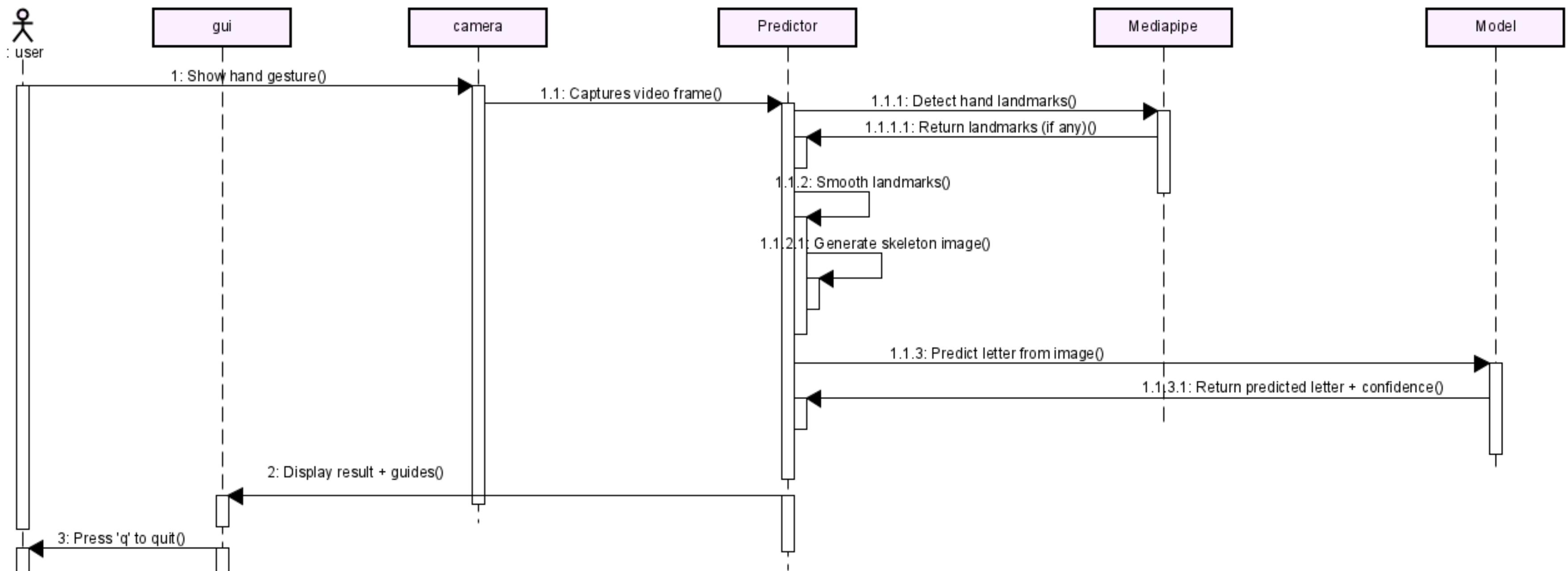
DIAGRAMS



USE CASE DIAGRAM



SEQUENCE DIAGRAM



04

PROJECT MODULES

PROJECT MODULES

- The project is composed of key modules working together.
- Each module plays a role in converting sign gestures into readable text.

1. Data Management

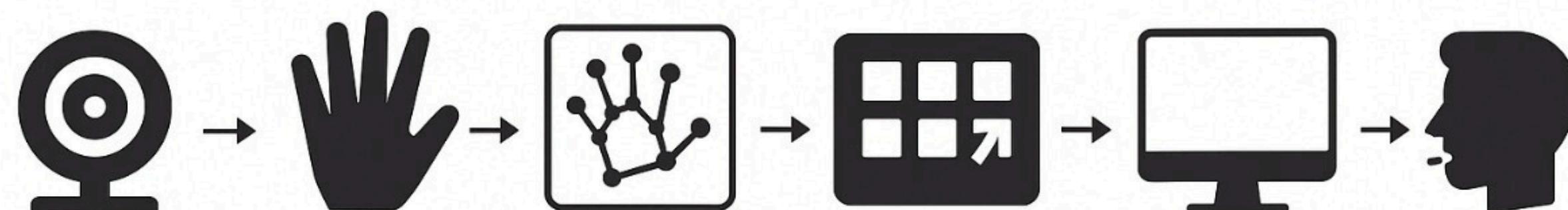
- Organize dataset into training, validation, and test sets

2. Real-Time Video Capture

- Uses webcam stream
- Crops fixed region for hand detection

3. CNN Model Architecture

- Achieves high accuracy for A-Z prediction



Webcam
Input

MediaPipe
Landmark
Detection

Skeleton
Generation

CNN
Prediction

UI
Feedback

Output
Text/
Speech

PROJECT MODULES

4. MediaPipe Hand Landmark Extraction

- Detects 21 hand landmarks

5. Skeleton Image Generation

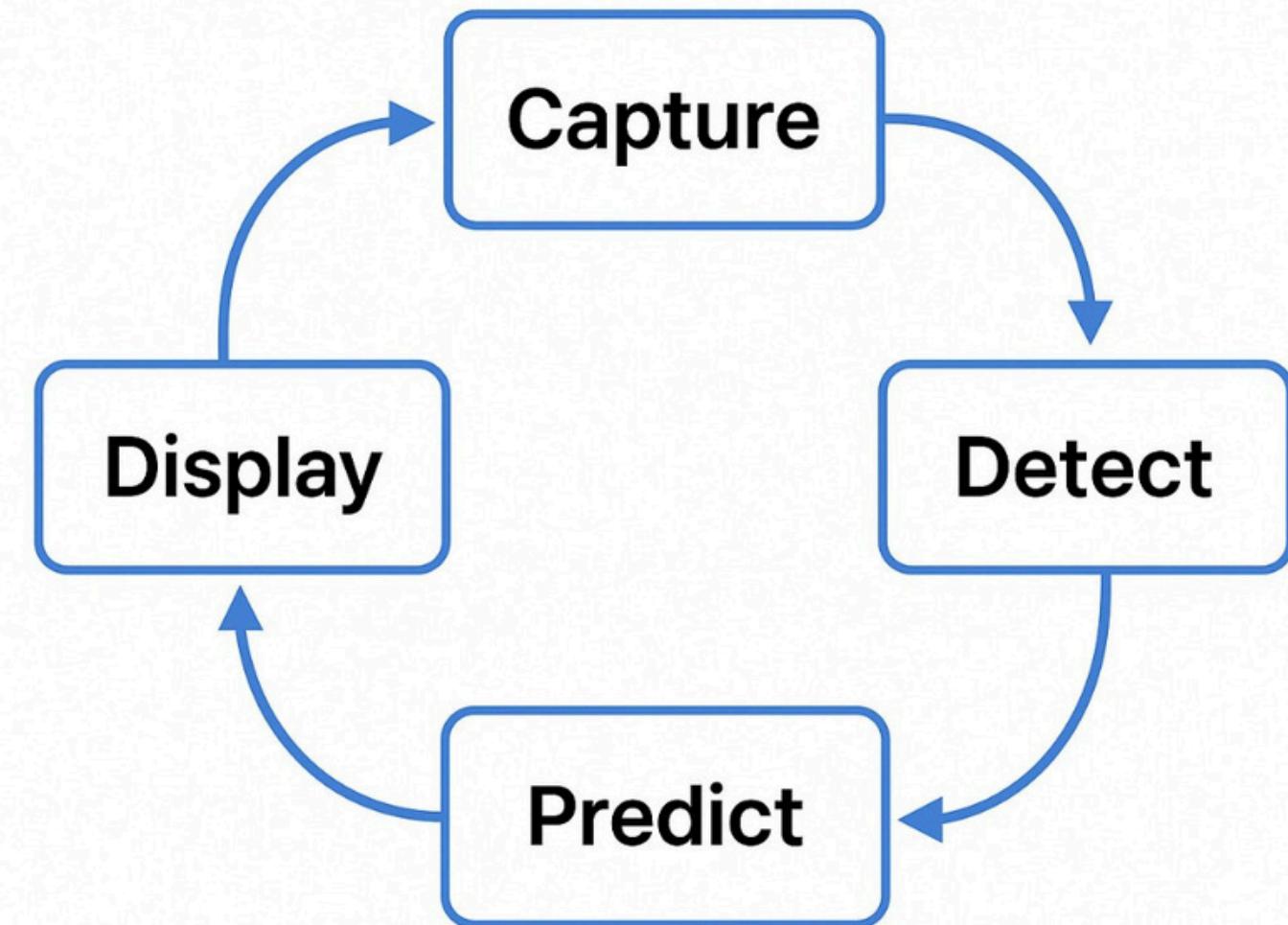
- Converts landmark data to skeleton images
- input format for CNN

6. Gesture Prediction Engine

- Predicts sign language letters
- Calculates model confidence

7. User Interface & Visual Feedback

- Displays bounding boxes and labels
- Real-time feedback to users



All modules work together in a continuous loop. The system keeps detecting and predicting gestures until the user exits the app

AI MODEL

05

AI MODEL

Model Overview

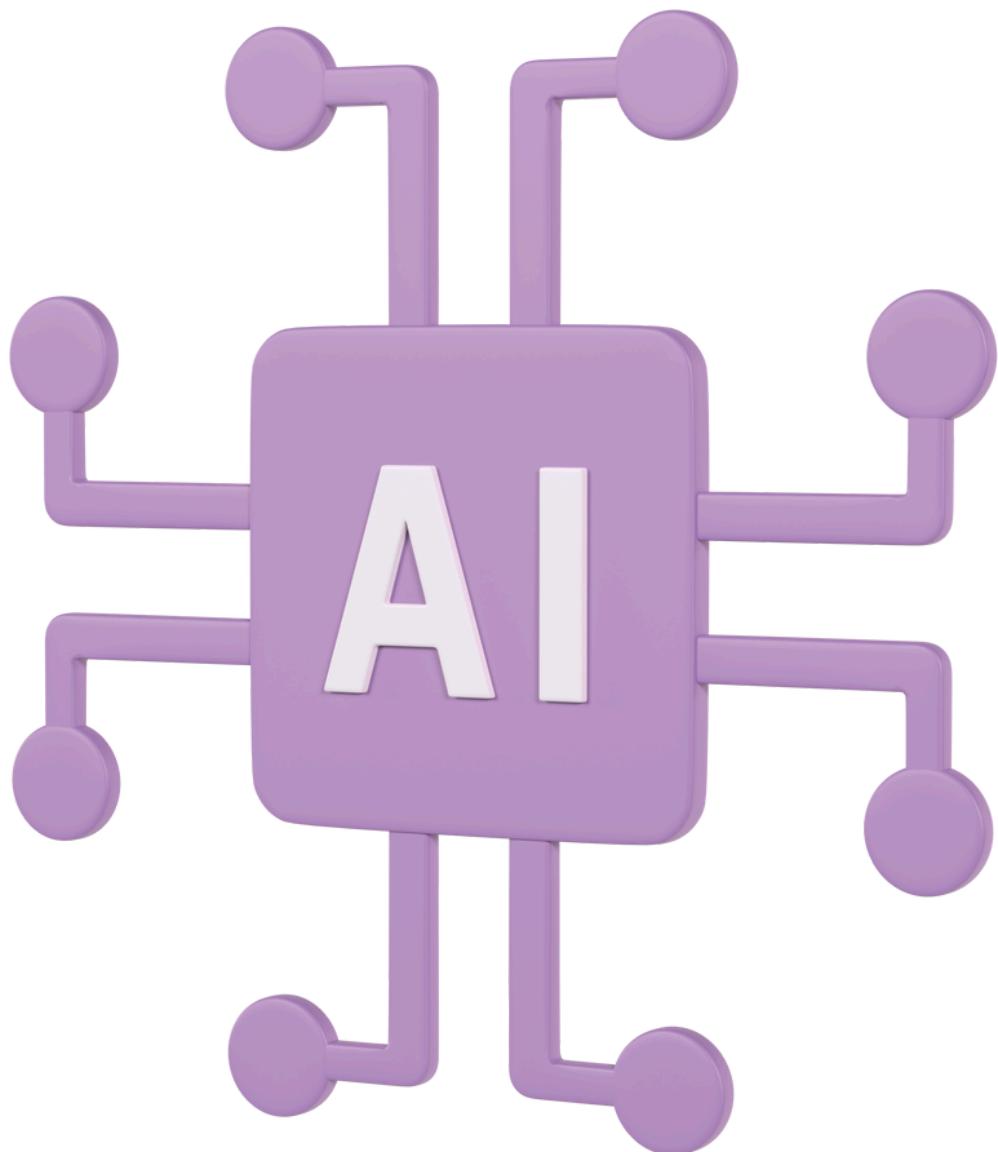
- The model takes in **400x400 RGB images of hand landmarks** and outputs one of **26 letters (A-Z)**.

Training Platform:

- Training was done on Google Colab using TensorFlow/Keras.

Iterations:

- The model was trained more than one times to fine-tune performance.
- Final training used 50 epochs, which led to the highest accuracy.



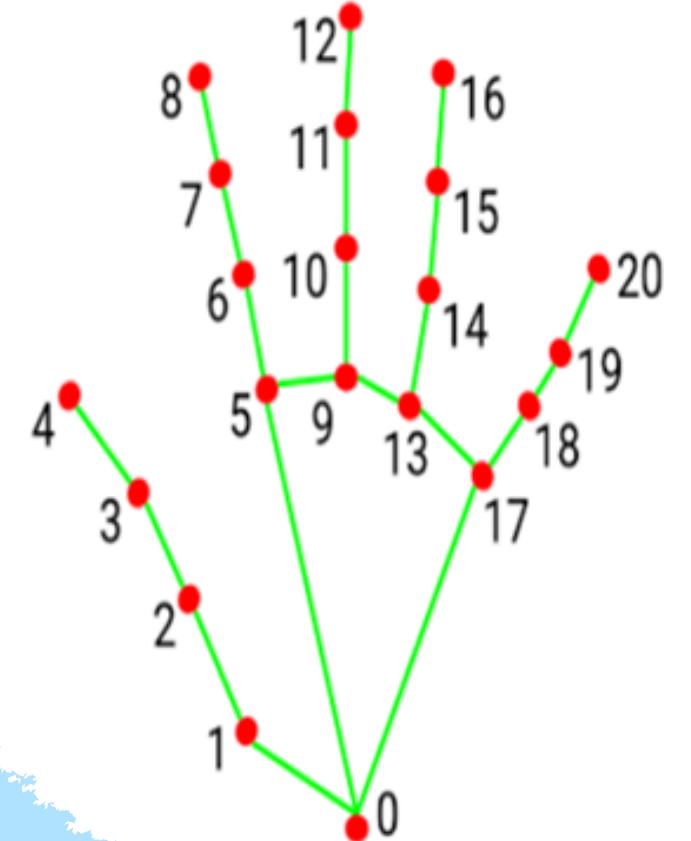
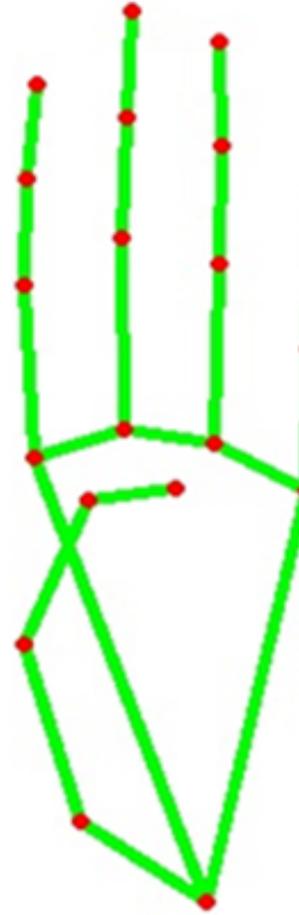
06

DATASET
&
CNN

DATASET

We use a dataset of 180 skeleton images for each letter from A to Z.

- For each letter, 144 images (80%) are used for training and 36 images (20%) are used for validation.
- Our dataset already has the MediaPipe features.
- MediaPipe works well even if the background or lighting is different.



- 0. WRIST
- 1. THUMB_CMC
- 2. THUMB_MCP
- 3. THUMB_IP
- 4. THUMB_TIP
- 5. INDEX_FINGER_MCP
- 6. INDEX_FINGER_PIP
- 7. INDEX_FINGER_DIP
- 8. INDEX_FINGER_TIP
- 9. MIDDLE_FINGER_MCP
- 10. MIDDLE_FINGER_PIP
- 11. MIDDLE_FINGER_DIP
- 12. MIDDLE_FINGER_TIP
- 13. RING_FINGER_MCP
- 14. RING_FINGER_PIP
- 15. RING_FINGER_DIP
- 16. RING_FINGER_TIP
- 17. PINKY_MCP
- 18. PINKY_PIP
- 19. PINKY_DIP
- 20. PINKY_TIP



CONVOLUTIONAL NEURAL NETWORK (CNN)

Convolutional Neural Networks (CNNs)

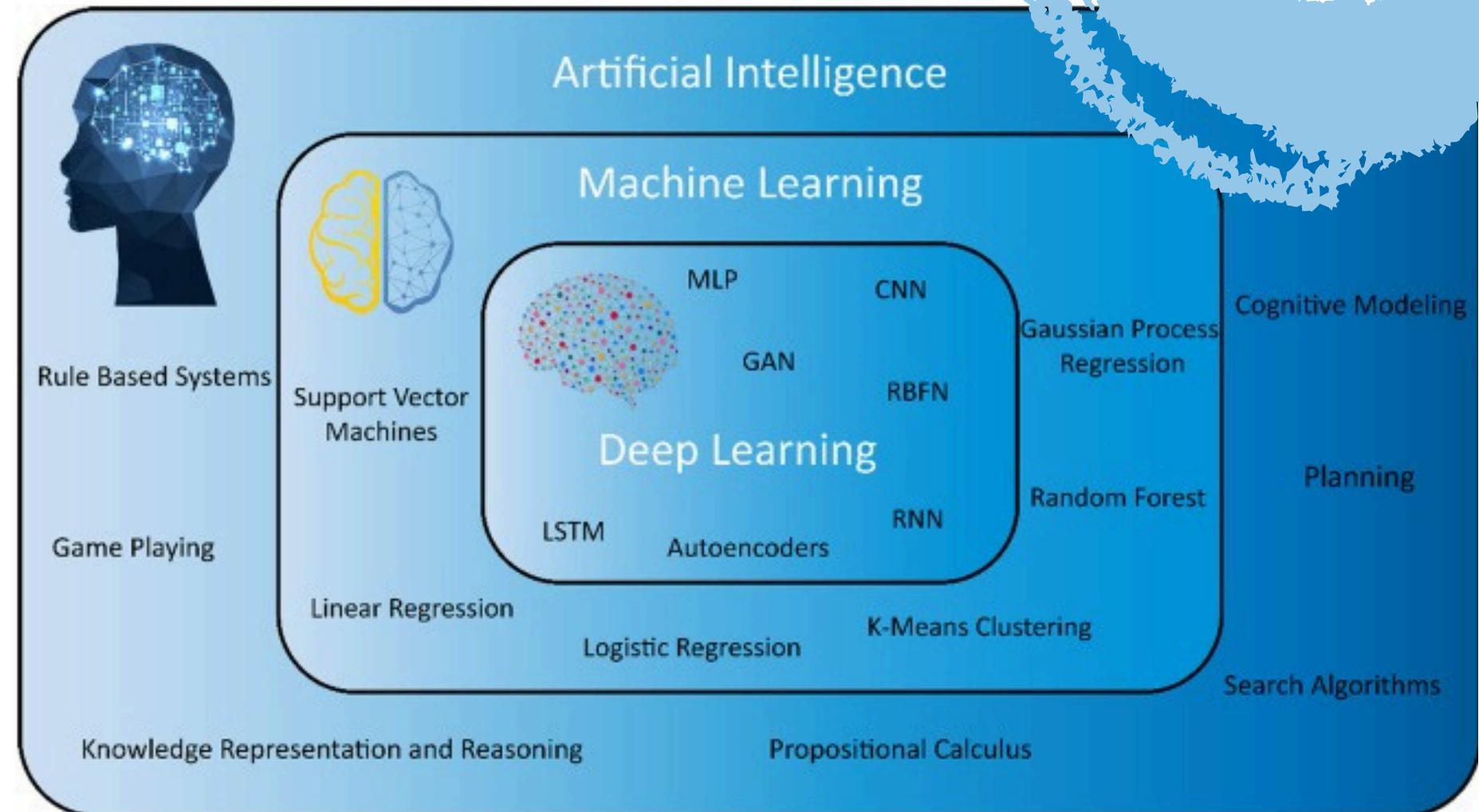
- A deep learning model specialized in image recognition.
- Detects visual patterns such as edges, textures, and shapes.

Why Use CNNs for Sign Language

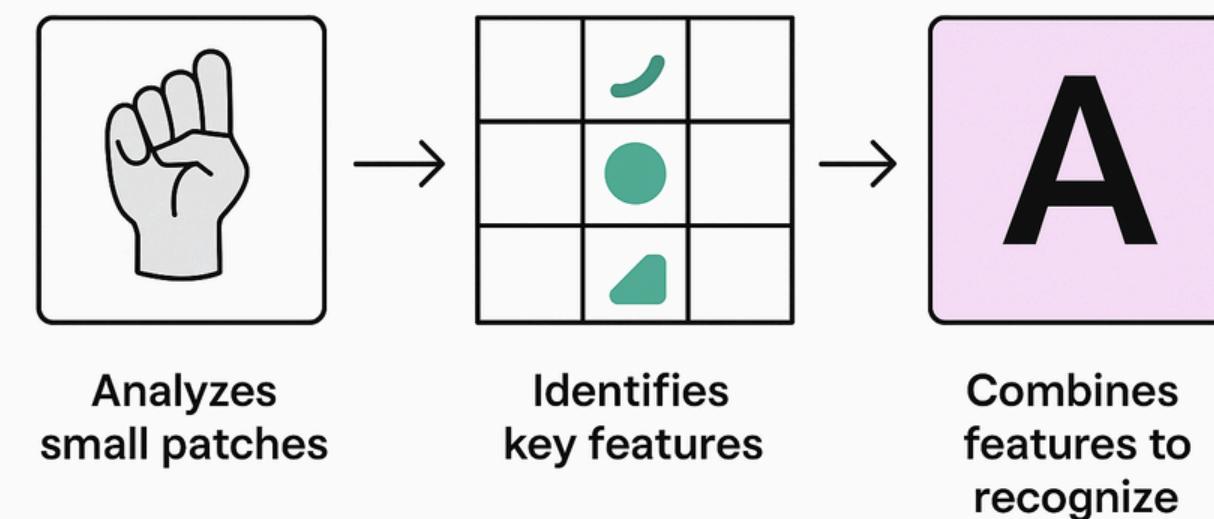
- Effective at identifying hand shapes and finger positions.
- Handles variations in gesture orientation, lighting, and scale.
- Converts visual hand gestures into recognizable text labels.

How CNNs Work

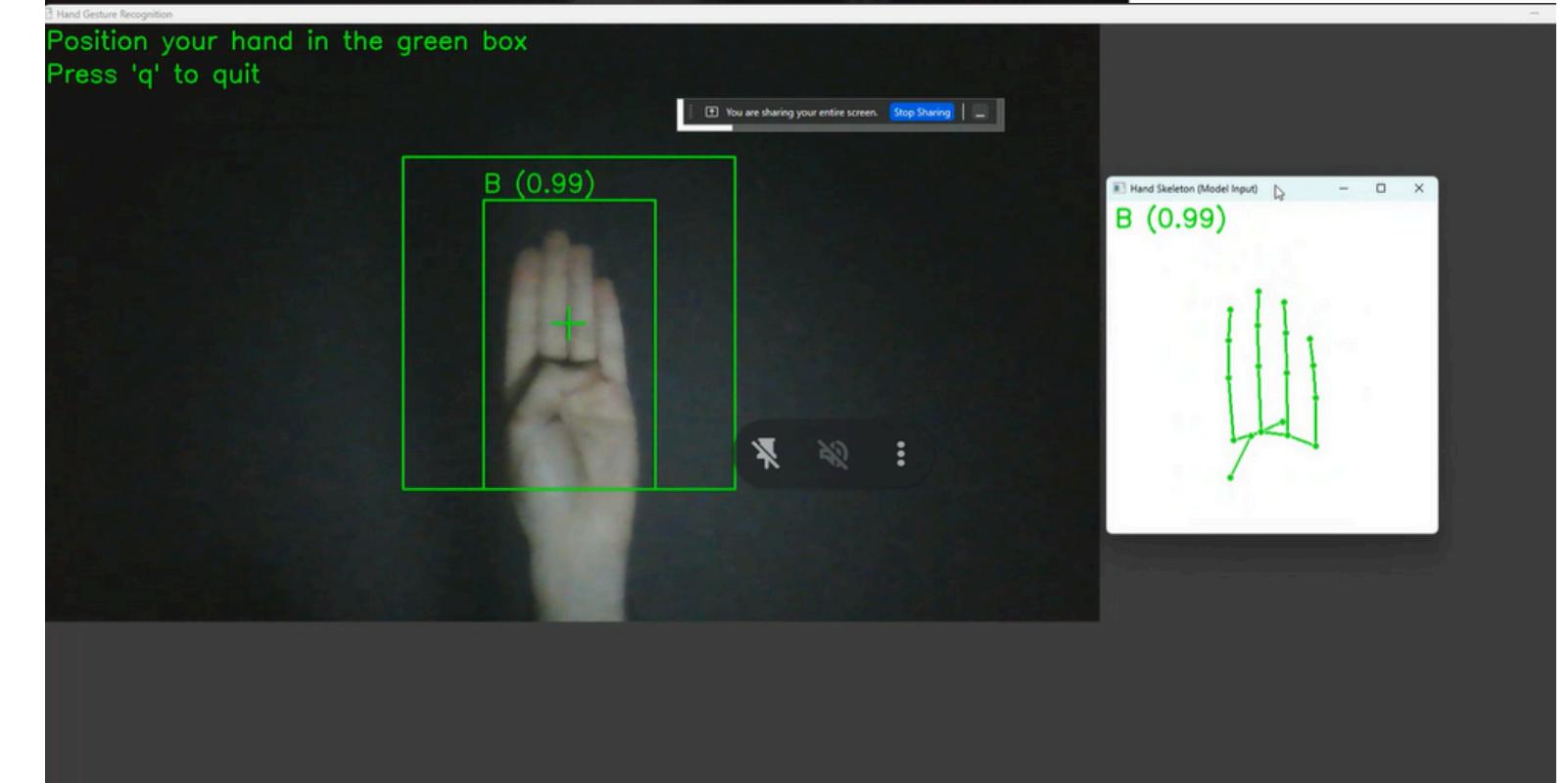
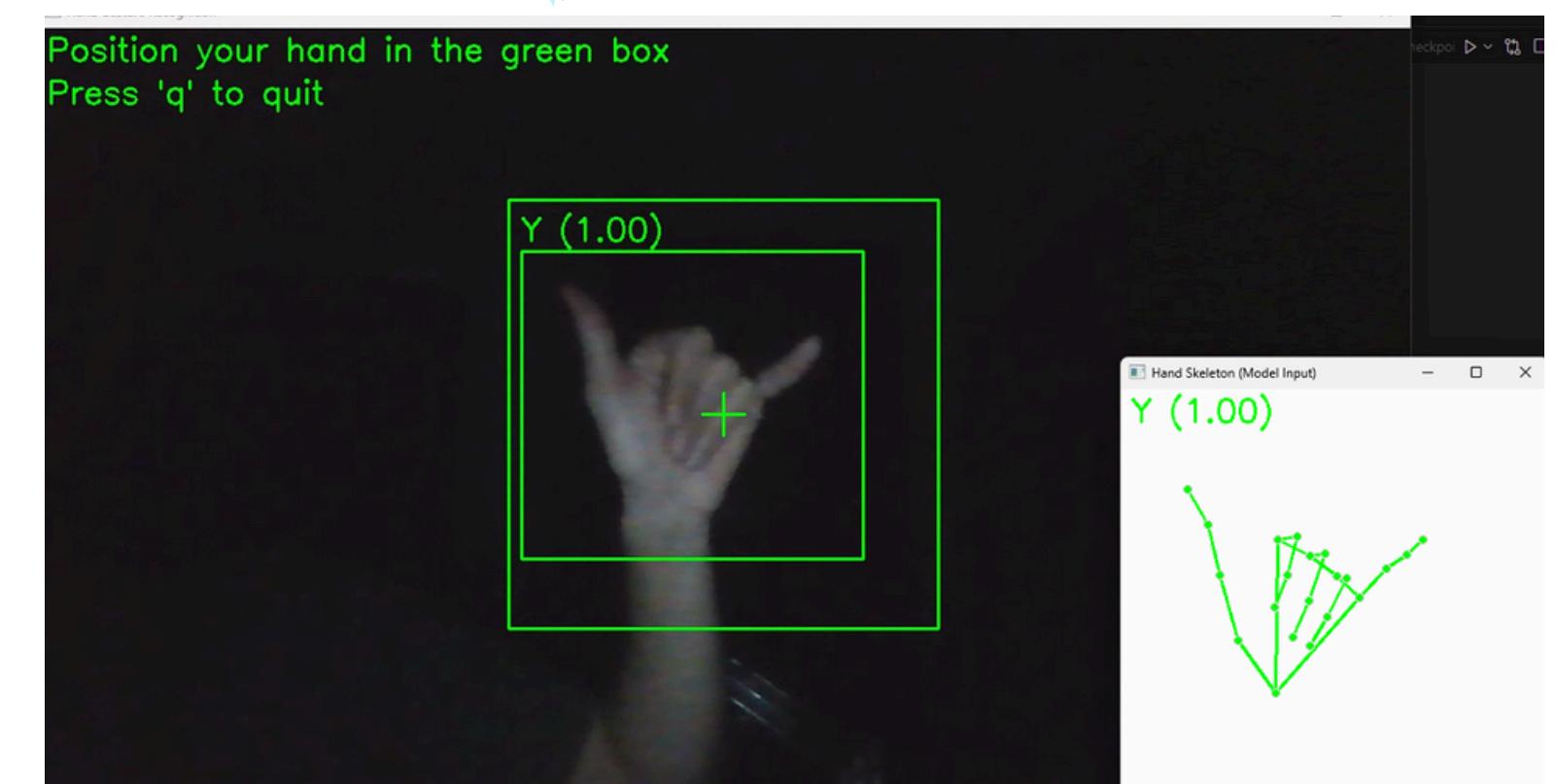
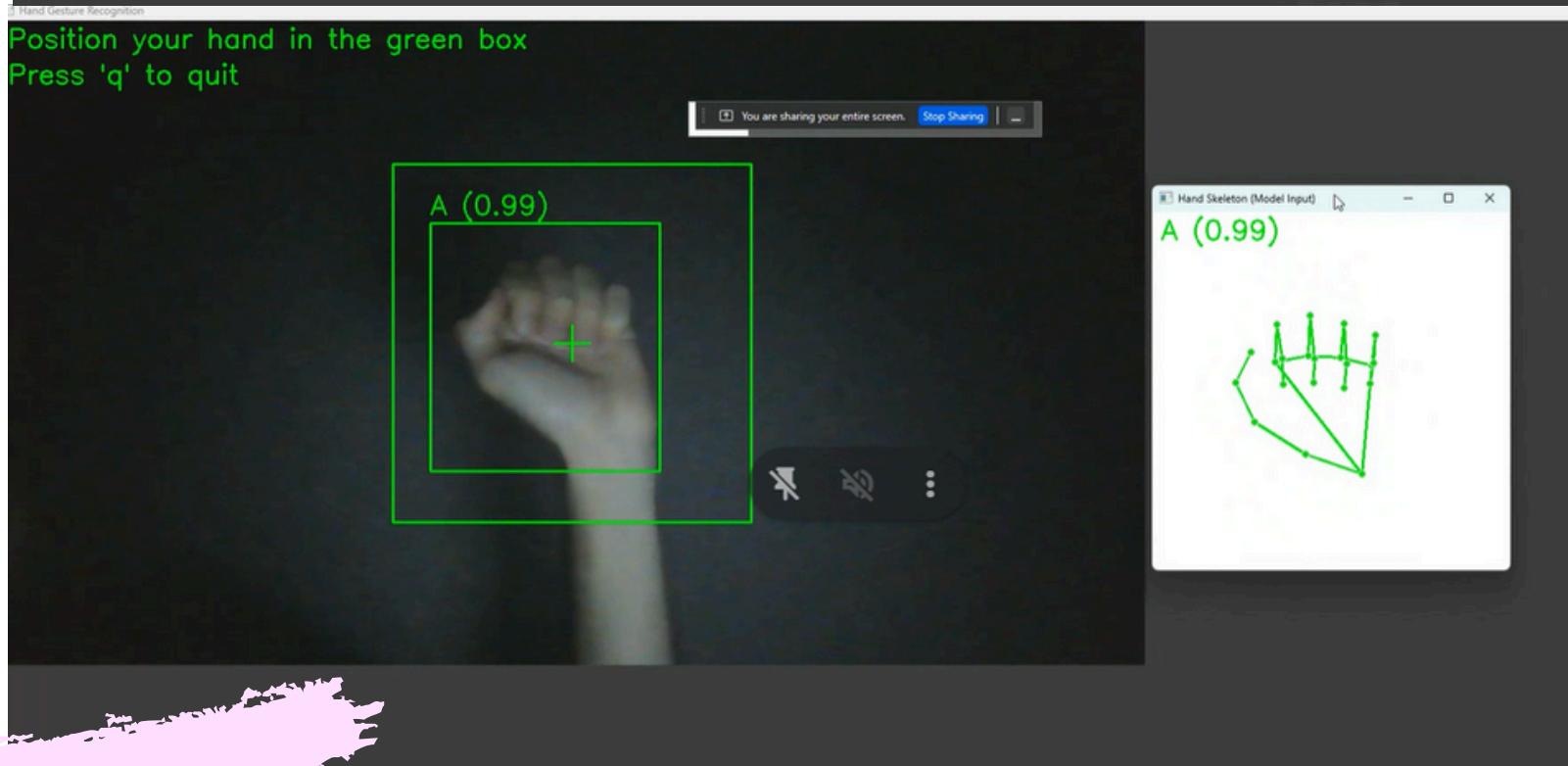
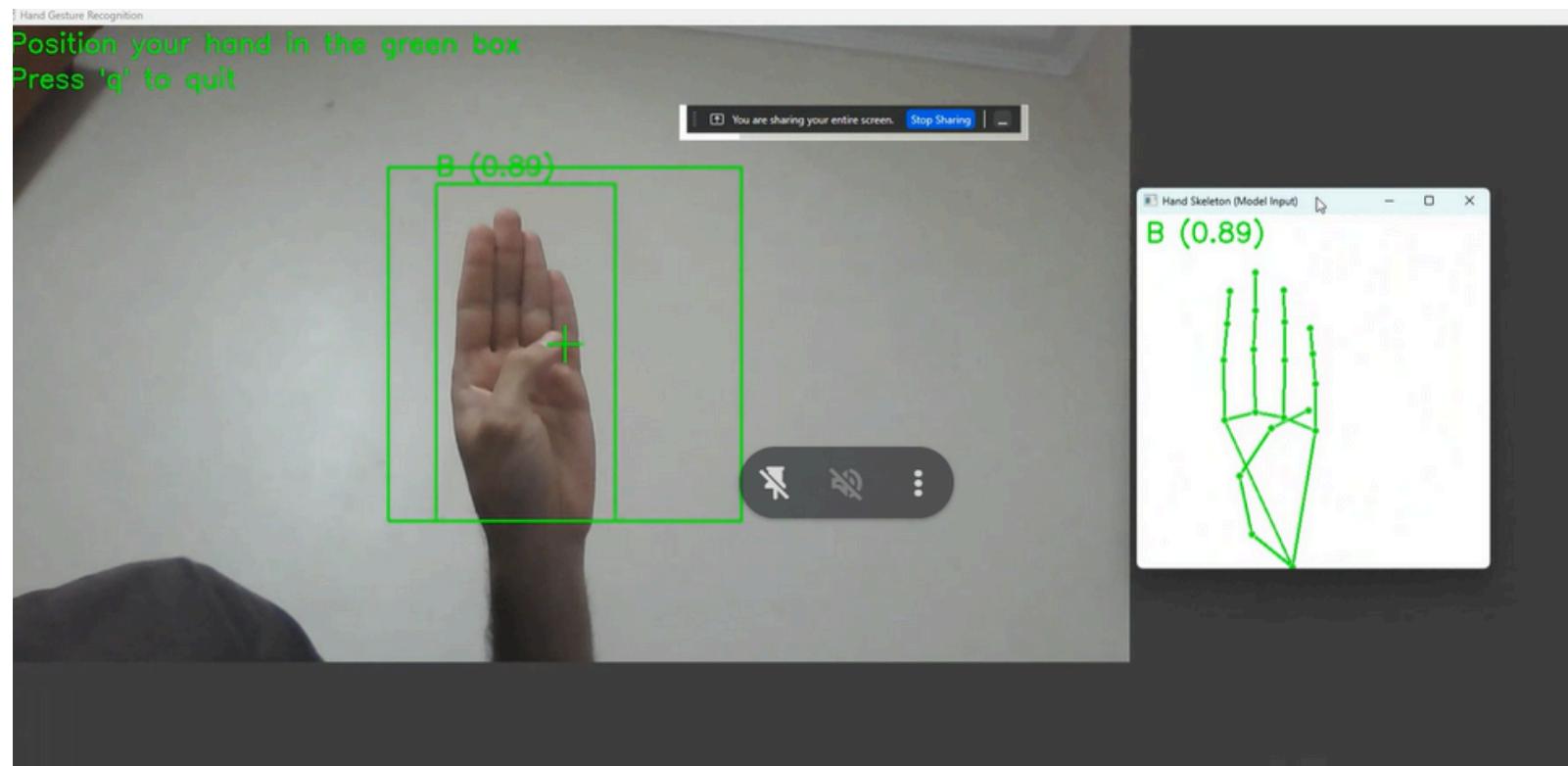
- Start by scanning small regions of the image.
- Extract important features at each layer.
- Combine features to form a high-level understanding of the gesture.



CNN

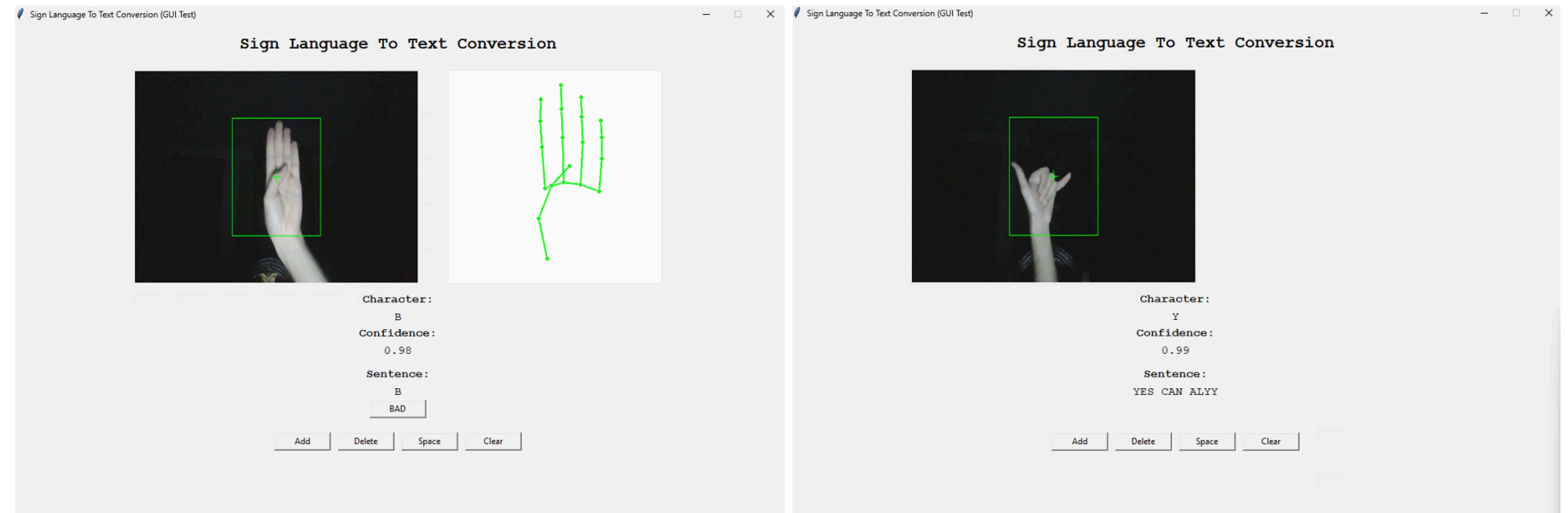


TESTING



These gestures were captured under different backgrounds and lighting conditions, showcasing the system's robustness and accuracy in diverse environments.

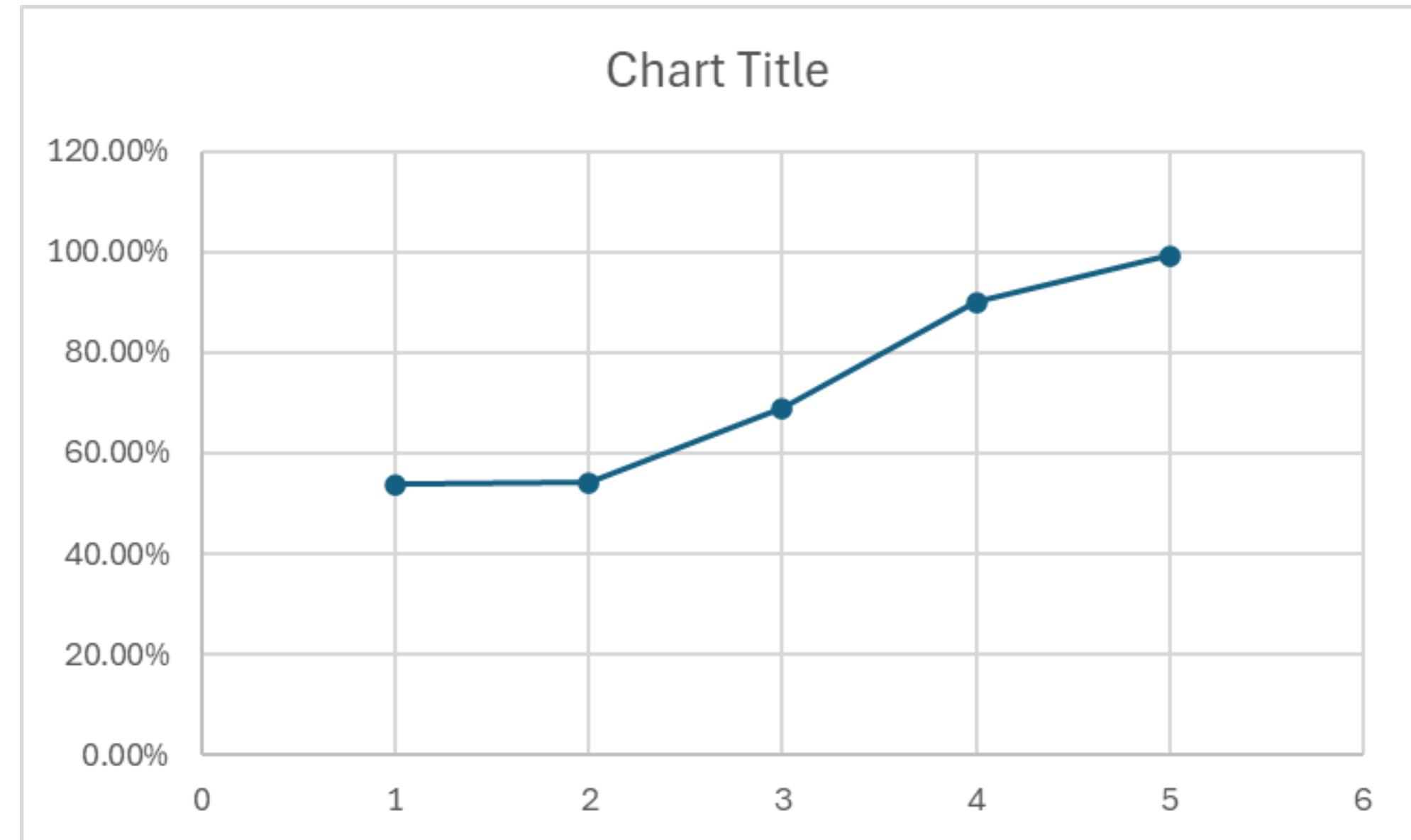
GUI



The GUI offers a simple, real-time interface that instantly shows users their detected hand gesture and its recognized letter.

Accuracy

99.23%



(X-AXIS) SHOWS HOW MANY TIMES WE TRAINED THE MODEL.
THE SIDE LINE (Y-AXIS) SHOWS HOW ACCURATE THE MODEL IS IN PERCENT.

07

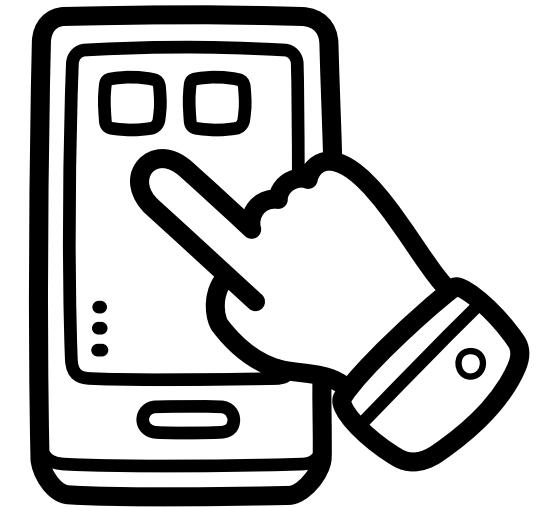
FUTURE WORK

FUTURE WORK



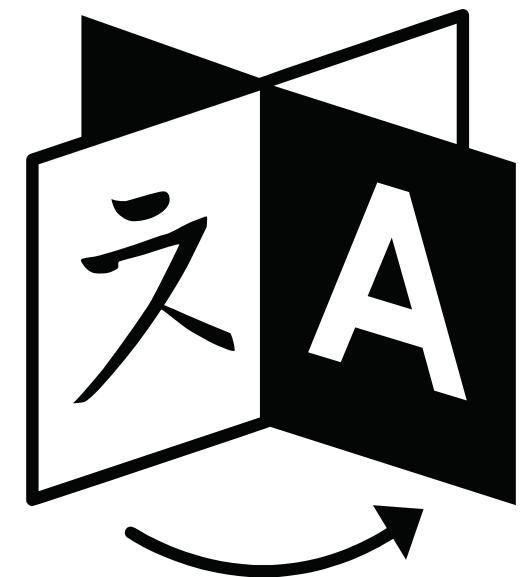
VOICE OUTPUT

The app will speak the translated text out loud.



MOBILE APPLICATION

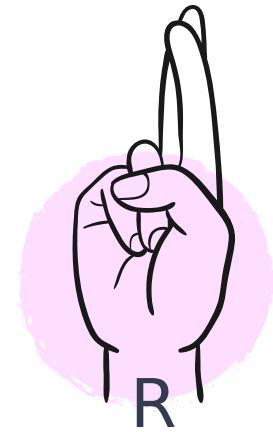
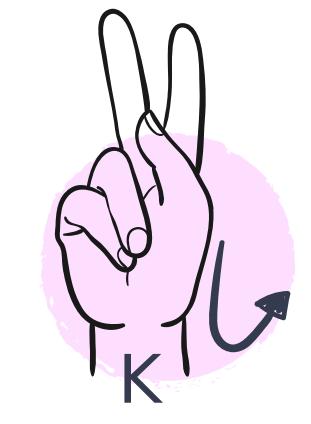
Easily use the app on your phone to translate sign language in real-time.

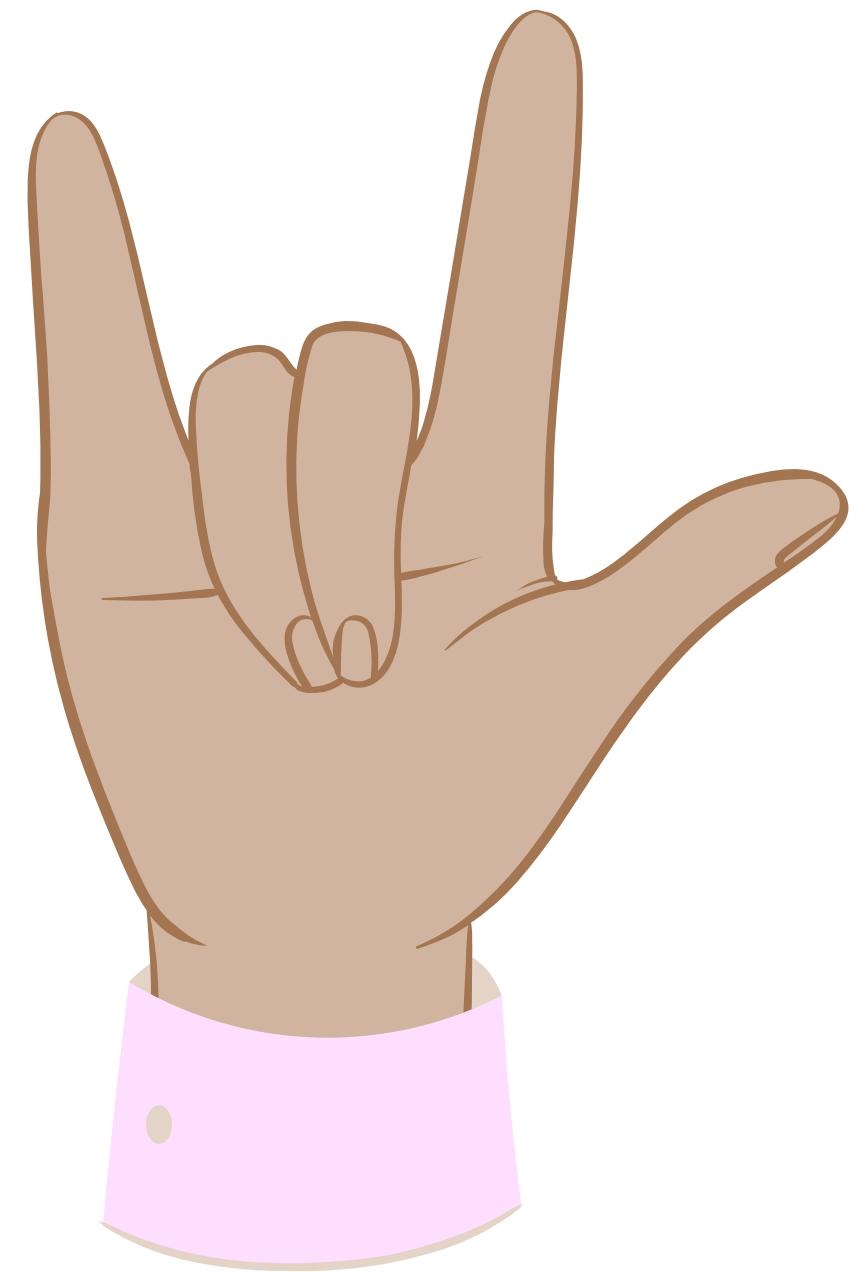


LANGUAGE OPTIONS

Translate signs into other languages like Spanish, French, and more.

THE ALPHABET





THANK YOU!