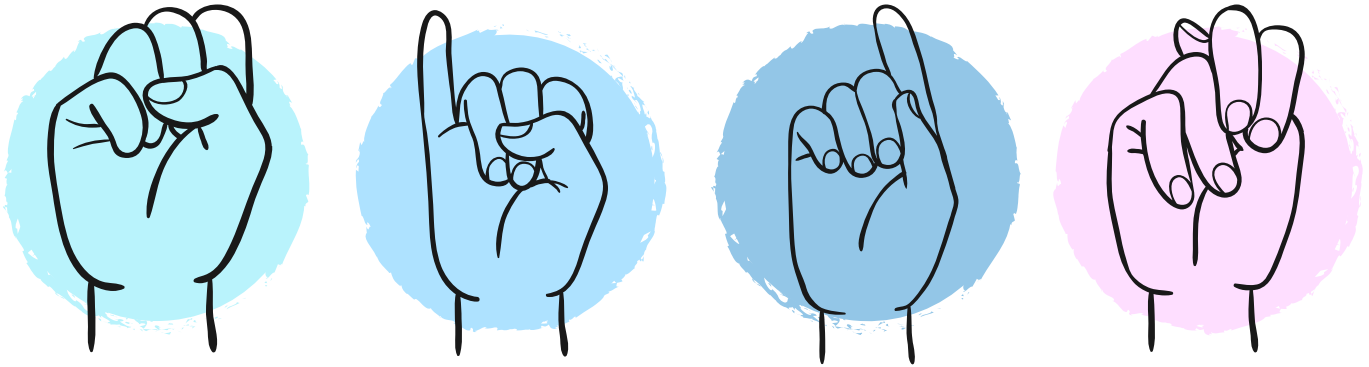
# SIGN LANGUAGE DETECTION

USING REAL TIME DATA

**GROUP – 08**



# INTRODUCTION & PROBLEM STATEMENT

**Sign Language** is a visual form of communication using hand gestures, body language, and facial expressions.
Examples include **American Sign Language (ASL), British Sign Language (BSL),** and others around the world.

- Many people cannot understand **sign language**, which creates **communication** barriers for those who rely on it.
- There is a need for a system that can detect and translate sign language **gestures** in real-time, making communication easier and more inclusive.

**Goal**

To develop a system that detects and recognizes hand gestures in **real-time** video, translating sign language into text or speech to enable better communication between signers and non-signers.
and translate sign language gestures in real-time, making communication **easier** and more inclusive.

# WORKING OF SIGN LANGUAGE DETECTION SYSTEMS-PROJECT OVERVIEW.

01

DATA COLLECTION

02

PREPROCESSING

03

FEATURE EXTRACTION

04

CLASSIFICATION

05

REAL TIME IMPLEMENTATION

06

REAL TIME RECOGNITION

# TECHNOLOGY STACK

### Machine Learning (ML)

Uses TensorFlow/Keras for gesture prediction with neural networks like LSTMs

### Computer Vision

OpenCV and MediaPipe for detecting and tracking hand landmarks in video.

### Gesture Recognition

Interprets hand shapes (static) and movements (dynamic gestures)

### Key Tools

TensorFlow, OpenCV, MediaPipe, NumPy, and Pygame.

# DATA COLLECTION & PREPROCESSING

## DATA COLLECTION

**Capture real-time video**: webcam to stream video frames in real-time.

**Extract landmarks**: Use Media Pipe to detect and extract 21 hand landmarks (x, y, z coordinates) for each frame.

**Frame segmentation**: Segment video into individual frames for easier processing.

**Environment setup**: Collect data under diverse environmental conditions, including varying lighting, backgrounds, and camera angles.

## PREPROCESSING

**Normalization:** Landmark coordinates are likely normalized to a consistent range.

**Relative Coordinates**: Coordinates may be adjusted relative to a key point to account hand position.

**Sequence Creation :** Frames may be grouped into temporal sequences to prepare data for models like LSTMs.

**Noise Filtering**: Landmark Data may be smoothed using filters to reduce jitter in detection

# DATASET PREPARATION

| | | |
|---|---|---|
|  HELLO |  I LOVE YOU |  THANK YOU |
|  LIVE LONG AND PEACE |  WHAT ARE YOU DOING |  NO |
|  ALL DONE |  YES |  THAT'S IT |

Create sequences of 14 frames for temporal analysis.
**Label sequences** with gesture names (e.g., "Yes," "No").
Split the labeled data into training, validation, and test sets.

# MODEL SELECTION – LSTM

**Best for Sequential Data**: LSTM processes **temporal patterns**, making it ideal for gestures that evolve over time.

**Input**: Preprocessed hand landmarks from Media Pipe as sequences of 10 frames.

**Output**: Predicts gesture classes like “Go Coogs" or “Thank you."

Media pipe extracts **21 Hand landmarks** per frame providing the co-ordinates for each joint. This turns the video frames into numerical data, ideal for LSTM.

LSTM Model Processes input sequences of shape**(10,126).**

10 Frames

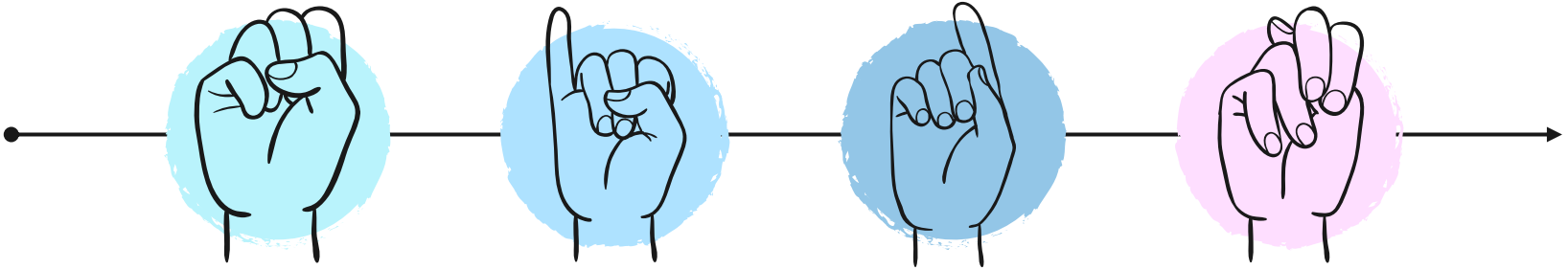
126 Features(21 Landmarks x 2 Hands x 3 Coordinates

**By directly using landmarks instead of raw frames, the system skips the computational overhead of CNNs.**

**LSTM’s ability to capture long-term dependencies ensures accurate recognition of complex gestures.**

# WORKFLOW – REAL TIME IMPLEMENTATION

Webcam captures video using OpenCV

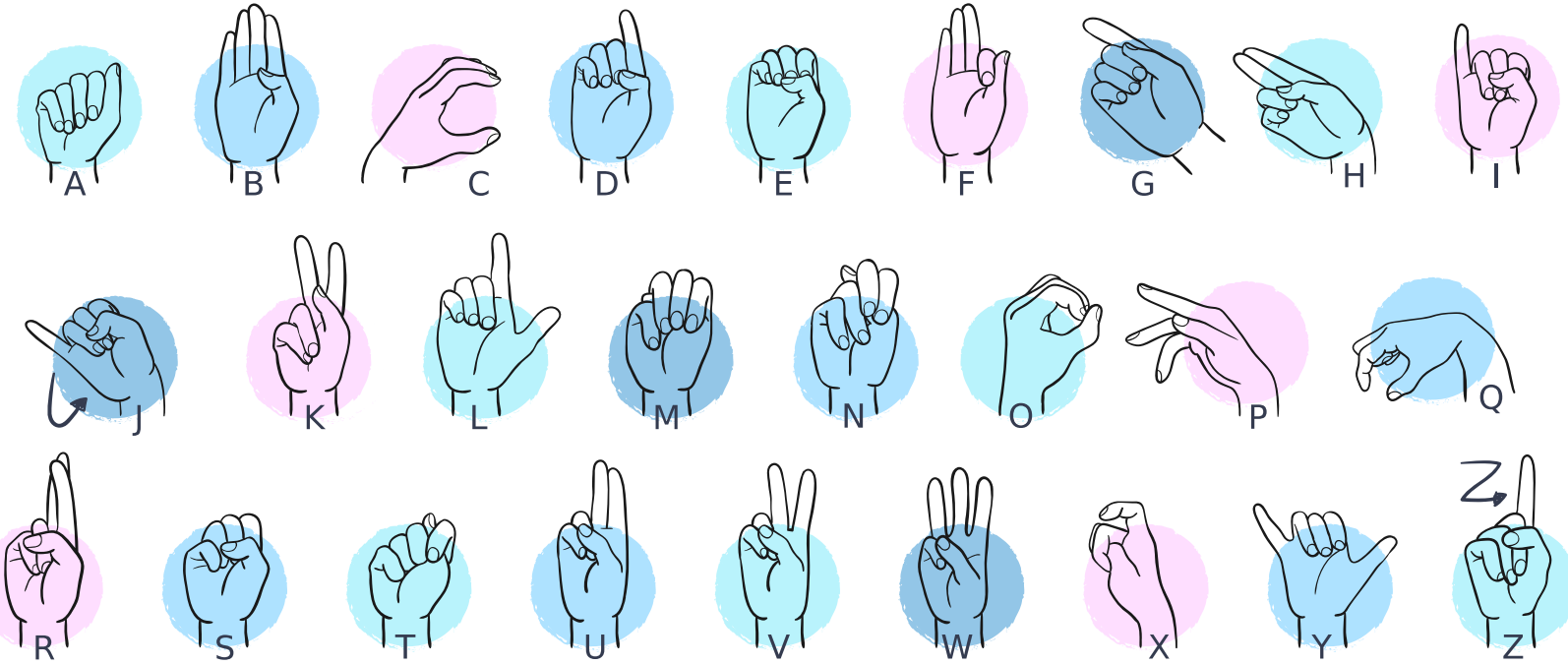


Media pipe detects hand landmarks in real-time

Sequence of landmarks is passed to the LSTM Model

The Model predicts gestures

# SIGN LANGUAGE DETECTION GAME USING LSTM AND CNN



# MODEL SELECTION FOR SIGN DETECTION GAME – LSTM AND CNN

**ROLE OF CNN IN THE GAME**

Identifies Key Features like landmarks from raw video frames.

Preprocesses input for temporal analysis by LSTM.

**ROLE OF LSTM IN THE GAME**

Understands gesture patterns over time(Eg. Forming the Number 3)

Predicts the correct gesture based on the sequence of spatial features extracted by CNN.

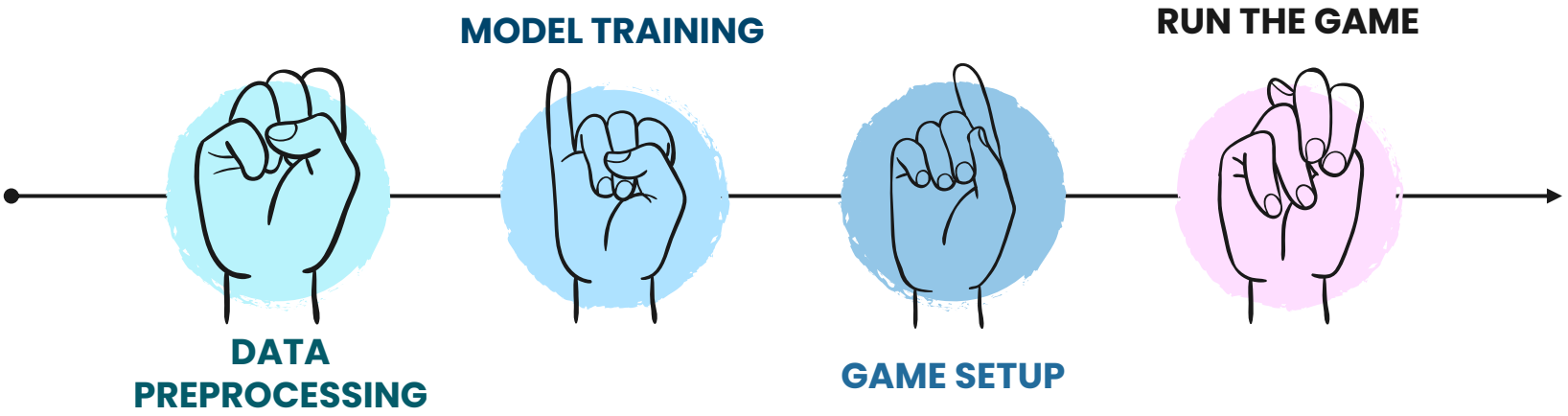
**WHY BOTH CNN AND LSTM?**

Gestures involve both spatial(hand shape) and temporal(movement over time) data.

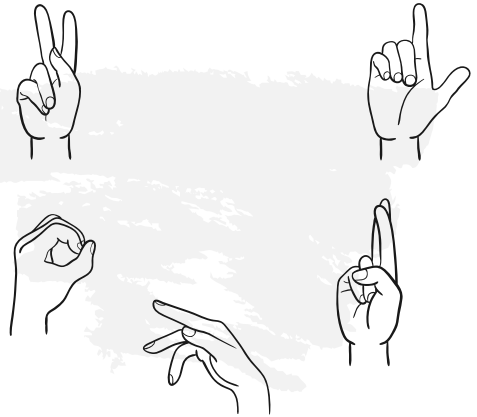
CNN Handles spatial analysis and LSTM Processes sequential patterns ,ensuring accuracy

CNN + LSTM provides the perfect balance for real-time gesture detection, enabling an interactive and rewarding gameplay experience.

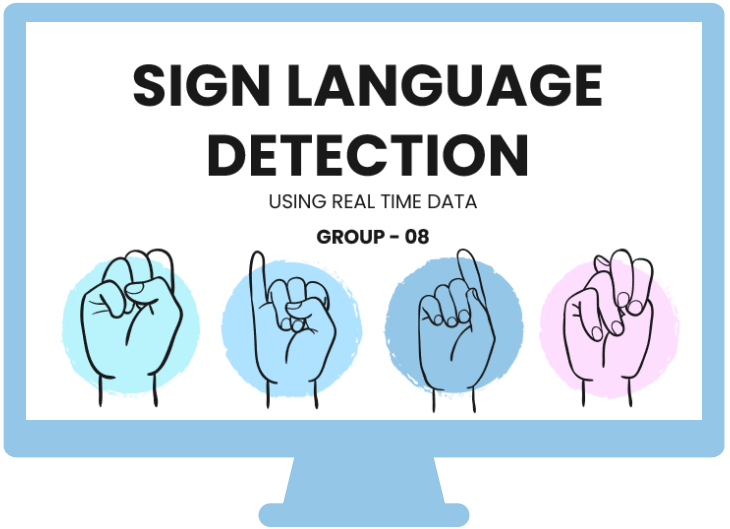
# WORKFLOW – SIGN DETECTION GAME



# WHAT'S NEW ?



Our project stands out for combining **real-time gesture** recognition with an **interactive game**, making it both **engaging and educational.**




# DEMO