SIGN SENSE

COMMUNICATION MADE BETTER

AGENDA

- Introduction
- Existing System
- Problem Statement
- Proposed System
- Architecture and Implementation
- Software and Hardware
- Future Scope
- Conclusion

INTRODUCTION

- Imagine a world where communication knows no limitations, where messages effortlessly traverse linguistic barriers, and gestures serve as universal connectors.

- This project seeks to craft a seamless system for translating Indian Sign Language (ISL) into speech.

EXISTING SYSTEM

[1] Real-time Conversion of Sign Language to Text and Speech

An Android application that converts real-time American Sign Language (ASL) signs into text/speech.

[2] Signtalk: Sign Language to Text and Speech Conversion

This system uses convolutional neural networks (CNN) to recognize hand gestures in American Sign Language (ASL) and convert them into text/speech.

DISADVANTAGE

- Predominant focus on American Sign Language (ASL).
- Challenges in capturing intricate structures of ISL.
- Limited adaptability to regional variations in ISL.
- Lack optimization for practical use and accessibility in the context of ISL.
- Frequent incomplete translations and misinterpretations of ISL gestures.

PROBLEM STATEMENT

- **Challenge:** Achieving real-time and accurate conversion of Indian Sign Language (ISL) to regional languages poses a substantial obstacle.

- **Impact:** The deficiency in effective translation tools adversely affects communication and constrains opportunities for individuals who are deaf and mute.

- **Limitations:** Currently employed methods, are frequently characterized by sluggishness, inaccuracy, and lack of accessibility.

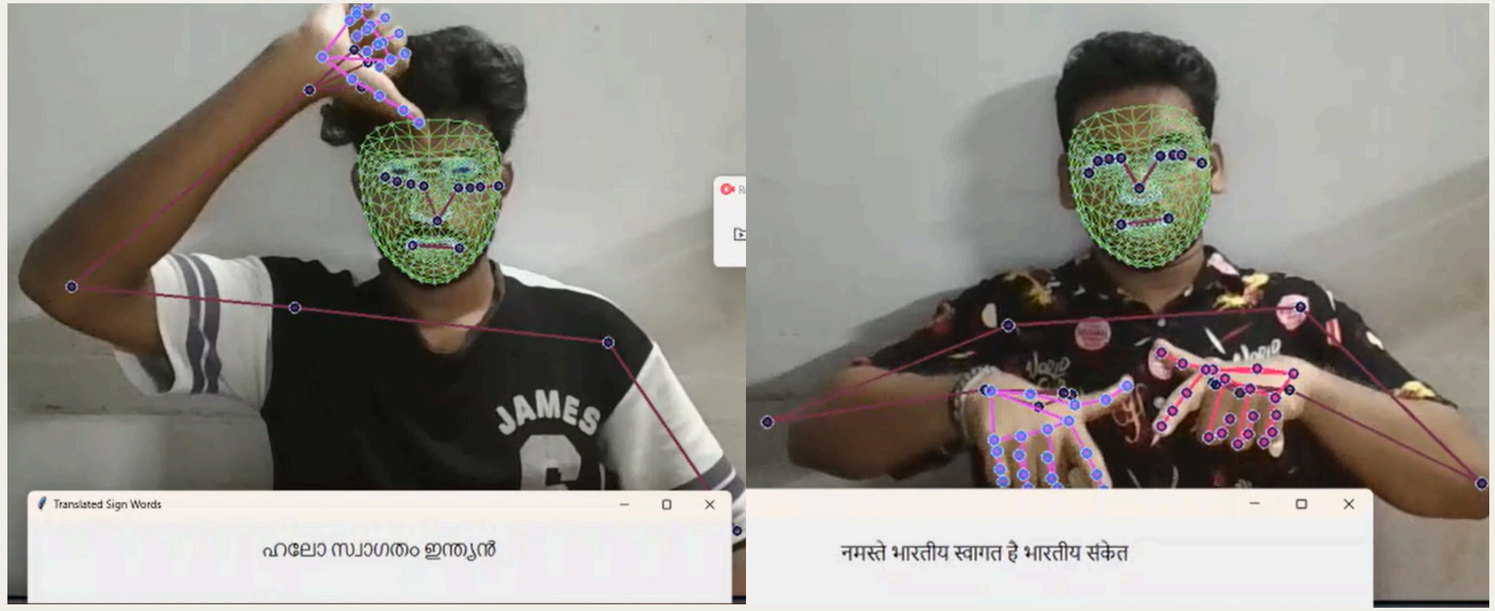
PROPOSED SYSTEM

- Addressing the limitations of the existing system, we introduce "SignSense", a system that converts Indian Sign Language (ISL) to text and vice versa across various regional languages.

- SignSense is designed to empower the speech and hearing-impaired community by facilitating seamless communication in their native languages.

FEATURES

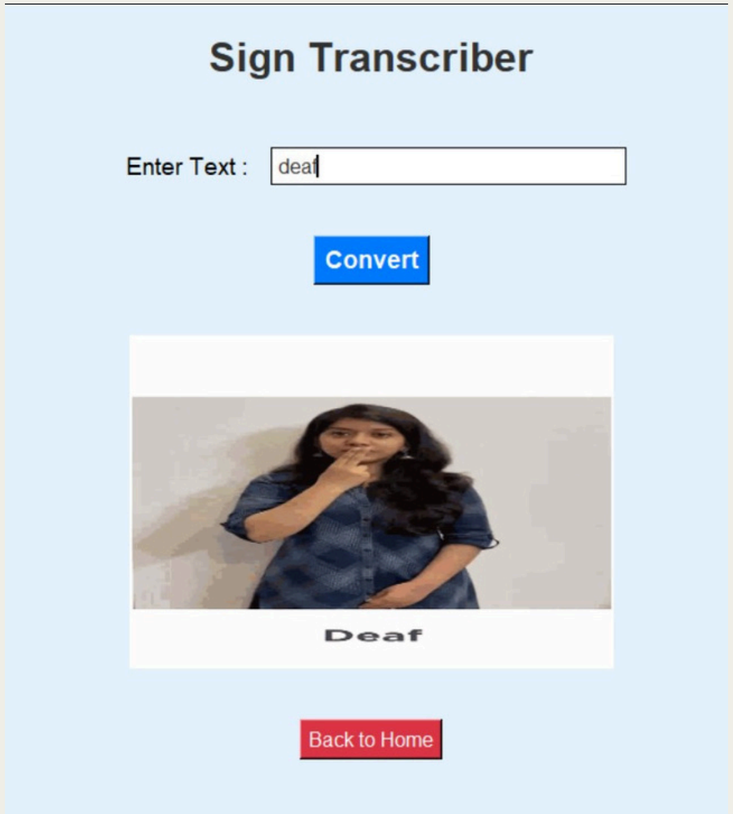
Sign to Text Conversion:



Users can communicate using sign language, and the system translates these gestures into written text.

FEATURES

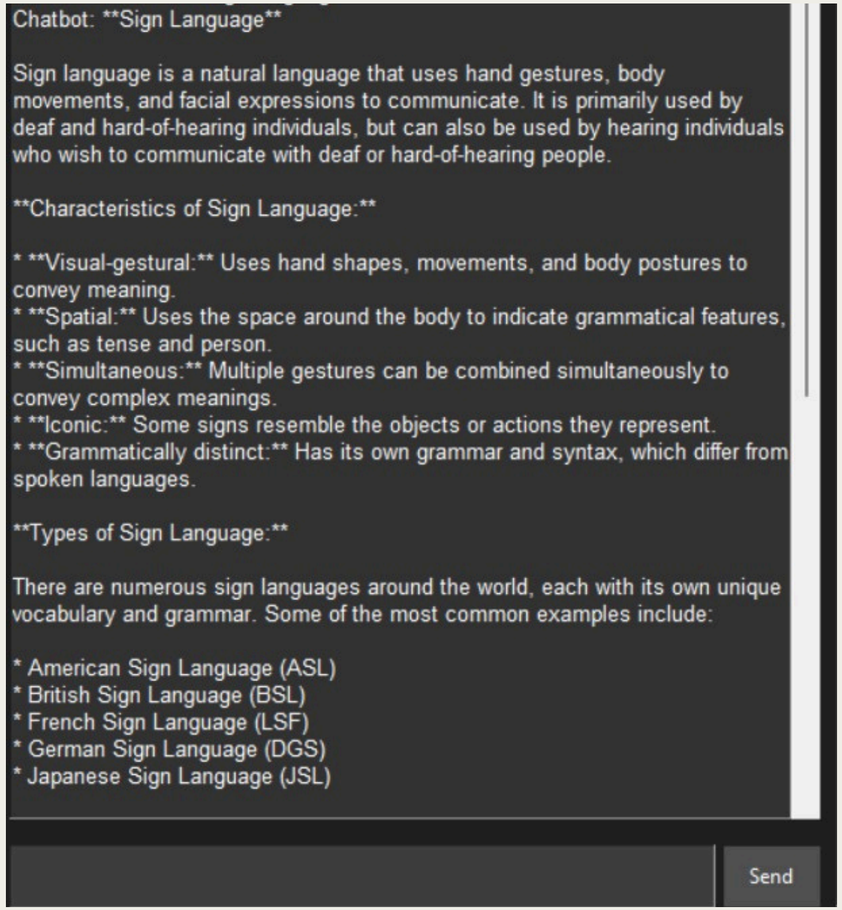
Text to Sign Conversion:



This feature allows users to input text, and the system generates corresponding sign language gestures in real-time.

FEATURES

AI Chat Bot:



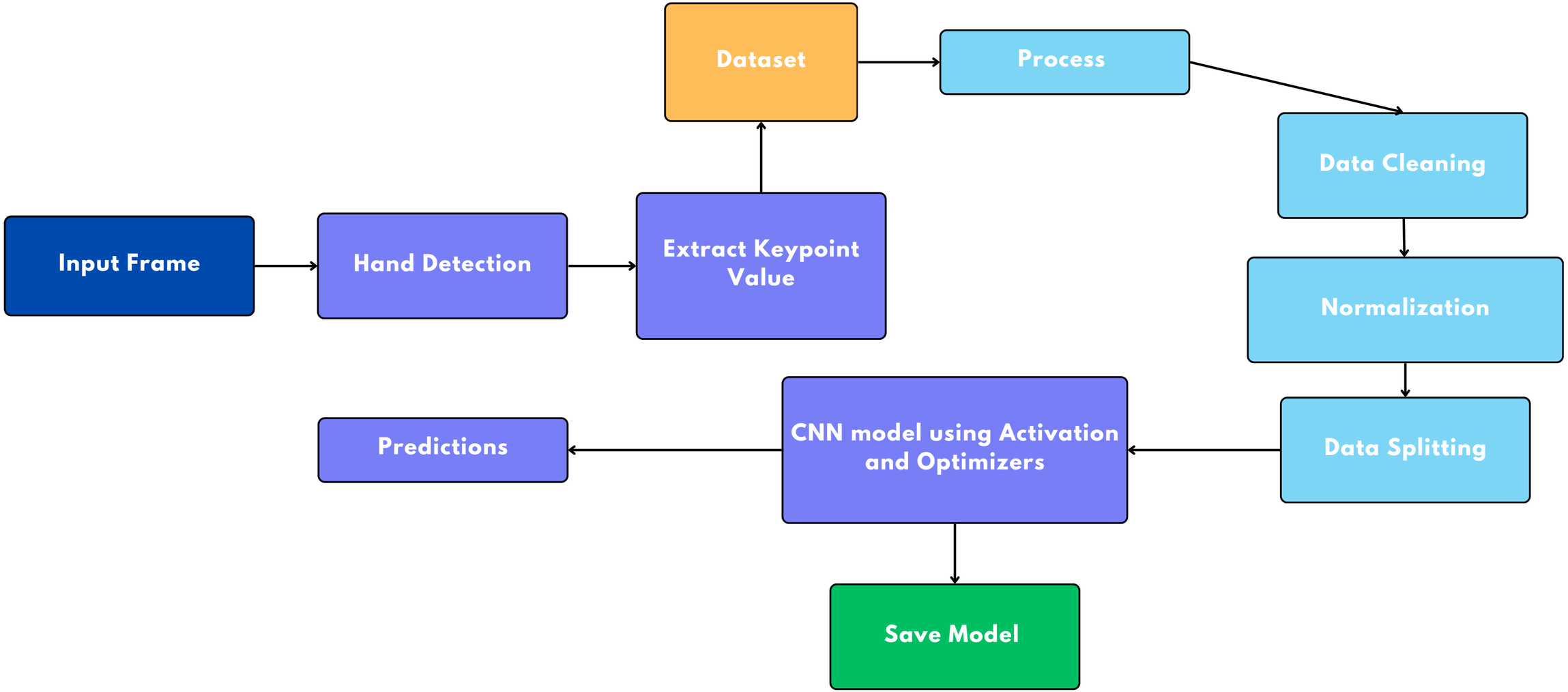
Integration of a chat bot feature facilitates seamless communication between users and the system

WHAT MAKE'S OUR PROJECT DIFFERENT ?

1. The one and only tech that transcribes Indian Sign Language.
2. Supports regional languages.
3. Custom made sign language data set.

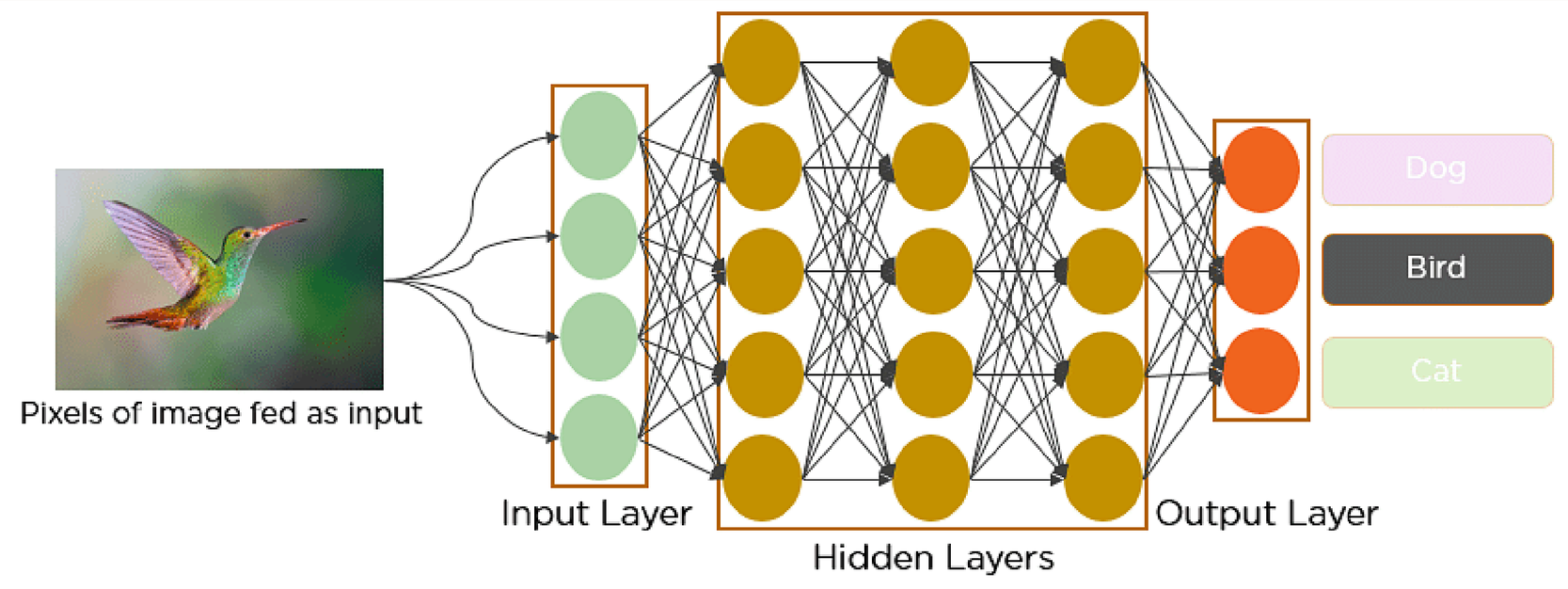
SIGN SENSE

ARCHITECTURE



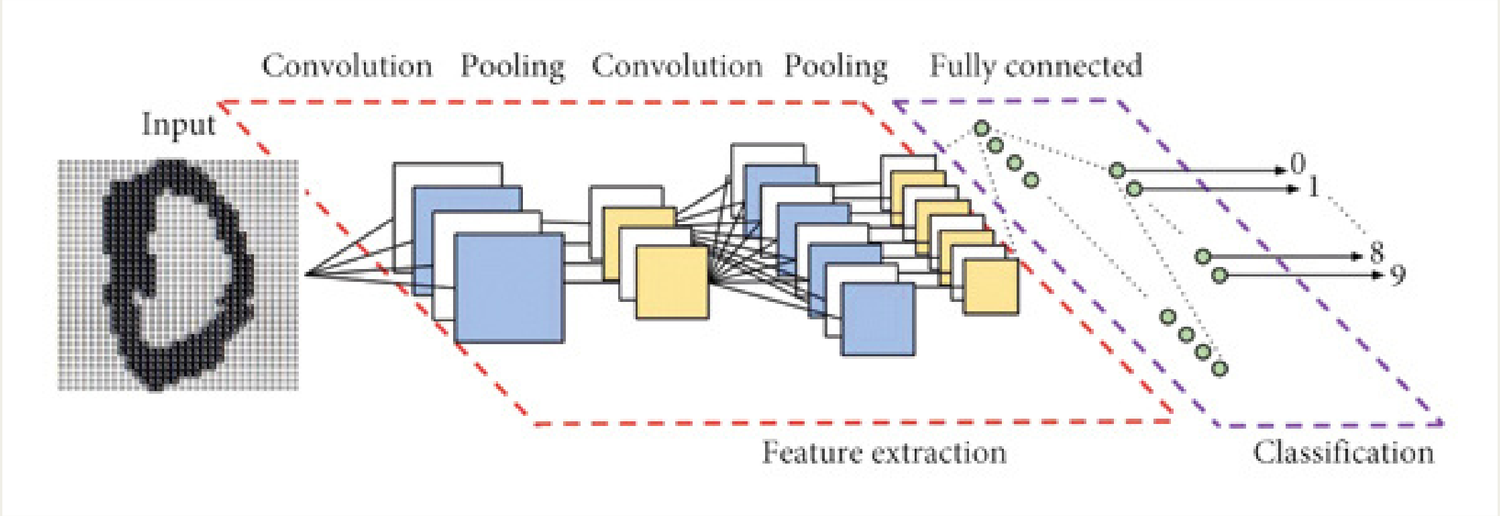
ALGORITHM

1. Convolutional Neural Network(CNN)



An artificial neural network used mainly for image recognition and processing task. It's main functions are feature extraction and classification. CNN is used in our project for gesture classification by spotting similar features in the dataset.

ALGORITHM

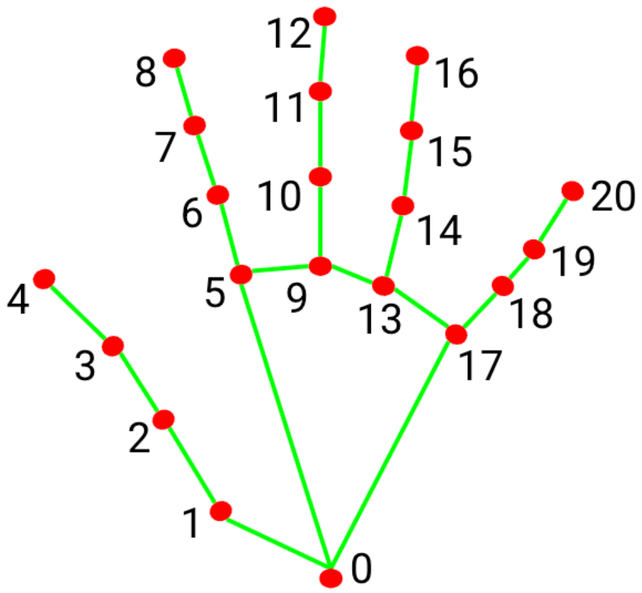


- Utilizes a multi-layer neural network to classify sign language gestures.
- Input Handling: Accepts a feature vector derived from landmark coordinates.
- Network Layers:
  - Input Layer: Receives the initial data.
  - Hidden Layers: Multiple layers where data processing occurs.
  - Output Layer: Produces classification results within predefined classes.

- Utilizes a multi-layer neural network to classify sign language gestures.
- Input Handling: Accepts a feature vector derived from landmark coordinates.
- Network Layers:
    - Input Layer: Receives the initial data.
    - Hidden Layers: Multiple layers where data processing occurs.
    - Output Layer: Produces classification results within predefined classes.

ALGORITHM

2. MediaPipe Hand landmarks detection

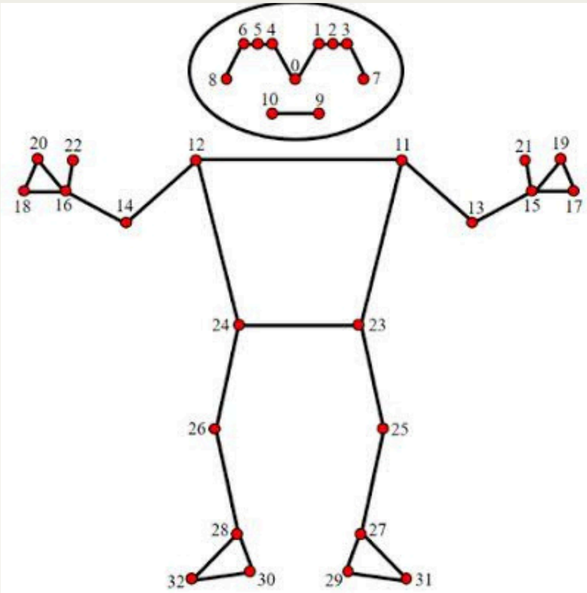


0. WRIST
1. THUMB_CMC
2. THUMB_MCP
3. THUMB_IP
4. THUMB_TIP
5. INDEX_FINGER_MCP
6. INDEX_FINGER_PIP
7. INDEX_FINGER_DIP
8. INDEX_FINGER_TIP
9. MIDDLE_FINGER_MCP
10. MIDDLE_FINGER_PIP
11. MIDDLE_FINGER_DIP
12. MIDDLE_FINGER_TIP
13. RING_FINGER_MCP
14. RING_FINGER_PIP
15. RING_FINGER_DIP
16. RING_FINGER_TIP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

By pinpointing the location of 21 key points on the hand , we can understand the hand's posture and movement in detail. This helps in recognizing sign language easily, and even finger tracking for gaming applications.

ALGORITHM

BodyPose Landmarker

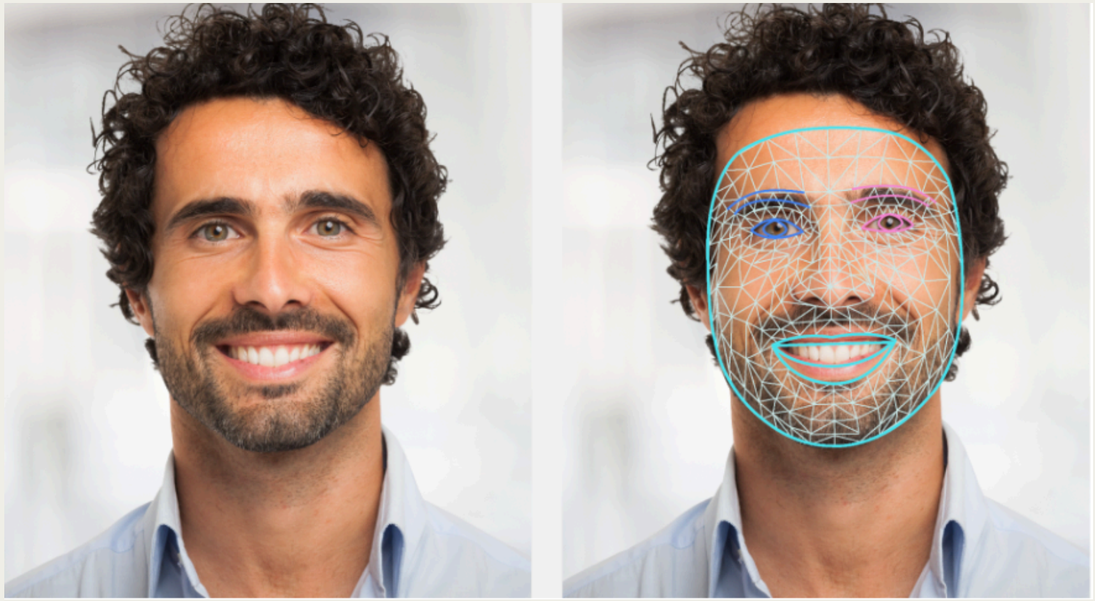


0. Nose
1. Left_eye_inner
2. Left_eye
3. Left_eye_outer
4. Right_eye_inner
5. Right_eye
6. Right_eye_outer
7. Left_ear
8. Right_ear
9. Mouth_left
10. Mouth_right
11. Left_shoulder
12. Right_shoulder
13. Left_elbow
14. Right_elbow
15. Left_wrist
16. Right_wrist
17. Left_pinky
18. Right_pinky
19. Left_index
20. Right_index
21. Left_thumb
22. Right_thumb
23. Left_hip
24. Right_hip
25. Left_knee
26. Right_knee
27. Left_ankle
28. Right_ankle
29. Left_heel
30. Right_heel
31. Left_foot_index
32. Right_foot_index

The MediaPipe Pose Landmarker task lets you detect landmarks of human bodies in an image or video. You can use this task to identify key body locations, analyze posture, and categorize movements.

ALGORITHM

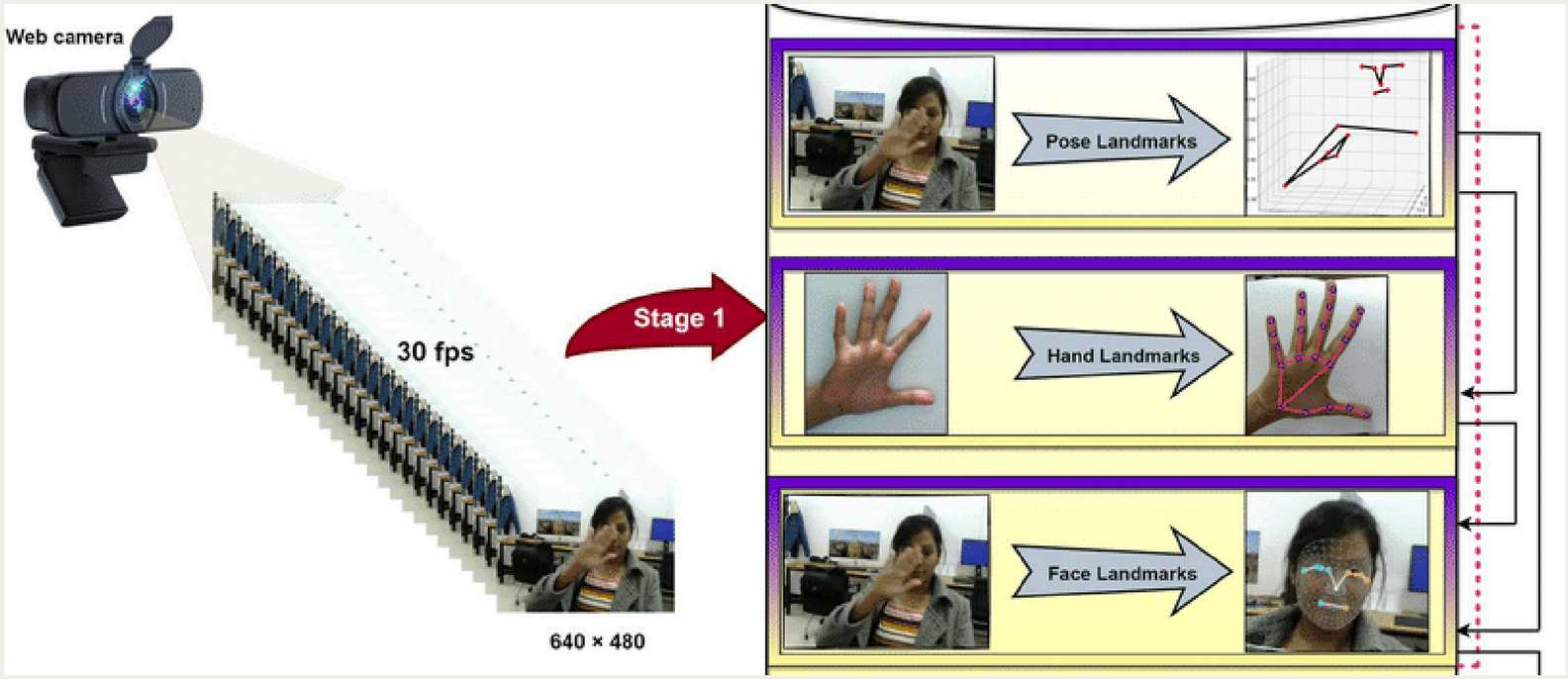
Face Mesh



The MediaPipe Face Landmarker task lets you detect face landmarks and facial expressions in images and videos. Here it is used to detect our head in our body.

IMPLEMENTATION

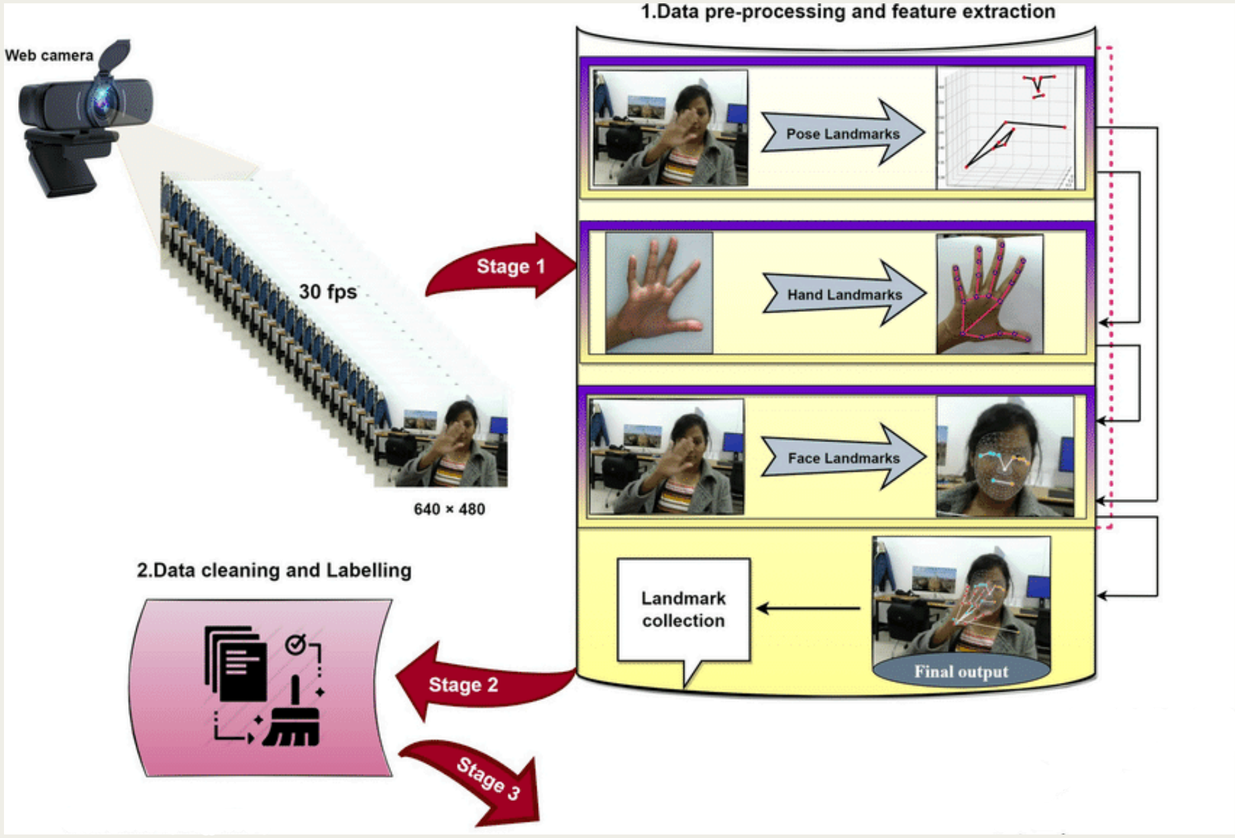
1. Data pre-processing and feature extraction



- Capture video from webcam at 30 fps with resolution 640*480.
- Extract landmarks from each frame.

IMPLEMENTATION

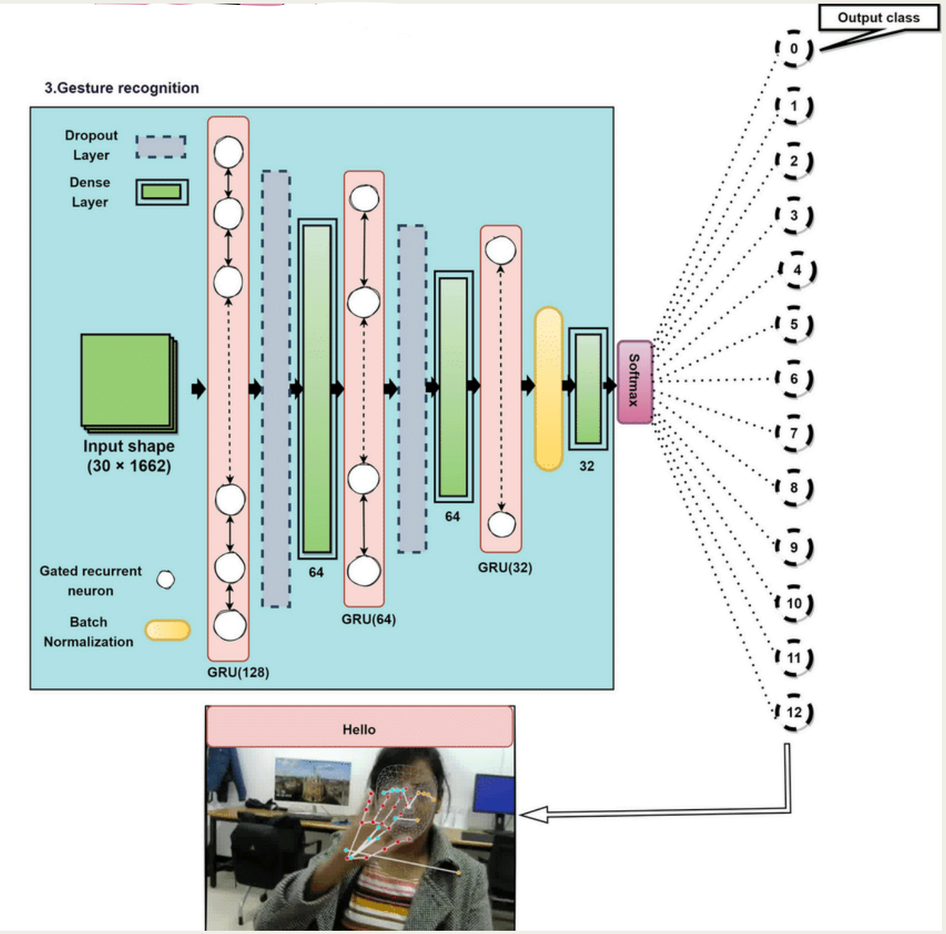
2. Data cleaning and Labelling



- Collect and label landmark data.
- Clean the data to remove noise or errors.

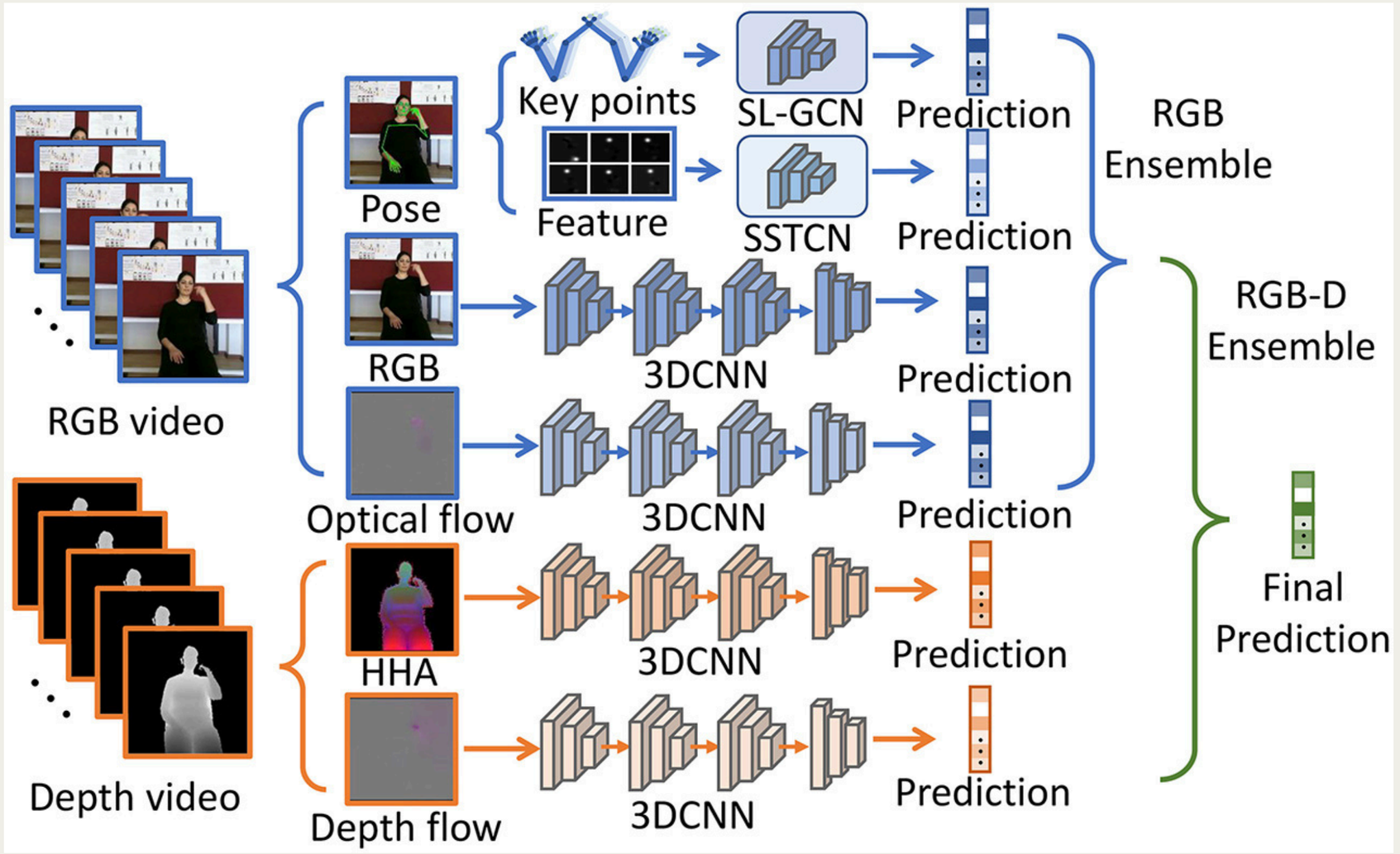
IMPLEMENTATION

3. Gesture Recognition



- Define gestures to be recognized.
- Create a ML model using CNN.
- CNN with ReLU layers.
- Train the model on the labelled data.
- Evaluate model performance on unseen data.

BEHIND THE SCENE



BEHIND THE SCENE

Activation Layers (ReLU): An activation function like ReLU (Rectified Linear Unit) is applied to introduce non-linearity, allowing the model to learn more complex patterns.

Pooling Layers: These layers reduce the spatial size (width and height, not depth) of the input volume for the next convolutional layer.

Fully Connected Layers: Towards the end, CNNs have one or more fully connected layers. These layers take the high-level filtered images after several convolutional and pooling layers and translate them into final predictions.

SOFTWARE AND HARDWARE

Tkinter

OpenCV

MediaPipe

TensorFlow

Processor Type: Multicore CPU (Intel i5/i7 or AMD Ryzen 5/7).

RAM Capacity: Minimum of 8GB.

Camera Specifications Resolution: Minimum 720p (1280x720).

Graphics Processing Unit (GPU) Model: NVIDIA GPU with CUDA cores.

FUTURE SCOPE

- Implementing the system in classrooms.
- Inclusion of more local languages.
- Translation to other sign languages (ASL, FSL etc.)
- Mobile Application.
- Integration of the system in Government offices.
- plugin for video conferences.

CONCLUSION

- In conclusion, Sign Sense is a groundbreaking tool that enhances communication for the speech and hearing-impaired by providing real-time ISL to text translation and vice versa. Its AI-powered chatbot assistance feature support in doubt clearance, promoting confidence and inclusivity.

- Sign Sense is more than a technological innovation; it is a step towards a more accessible and connected world.

REFERENCES

[1] Kohsheen Tiku, Jayshree Maloo, Aishwarya Ramesh, Indra R, “Real-time Conversion of Sign Language to Text and Speech,” 2020 Second International Conferenceon Inventive Research in Computing Applications, Coimbatore, India, 2020, pp. 346-351.

[2] C. Uma Bharti, G. Ragavi, K. Karthika \"Signtalk: Sign Language to Text and Speech Conversion,\" 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), Coimbatore, India, 2021, pp. 1-4, doi: 10.1109/ICAECA52838.2021.9675751.

Thank you!

HAVE A GREAT DAY AHEAD...

SIGN SENSE