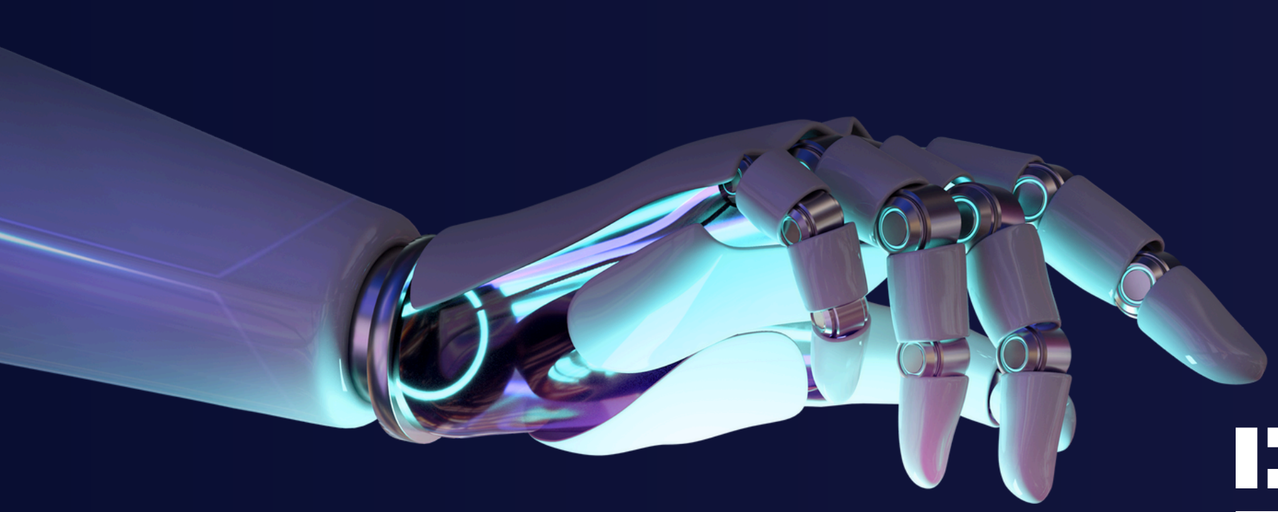


SIGN LANGUAGE

Real-Time Sign Language Recognition in Embedded Systems

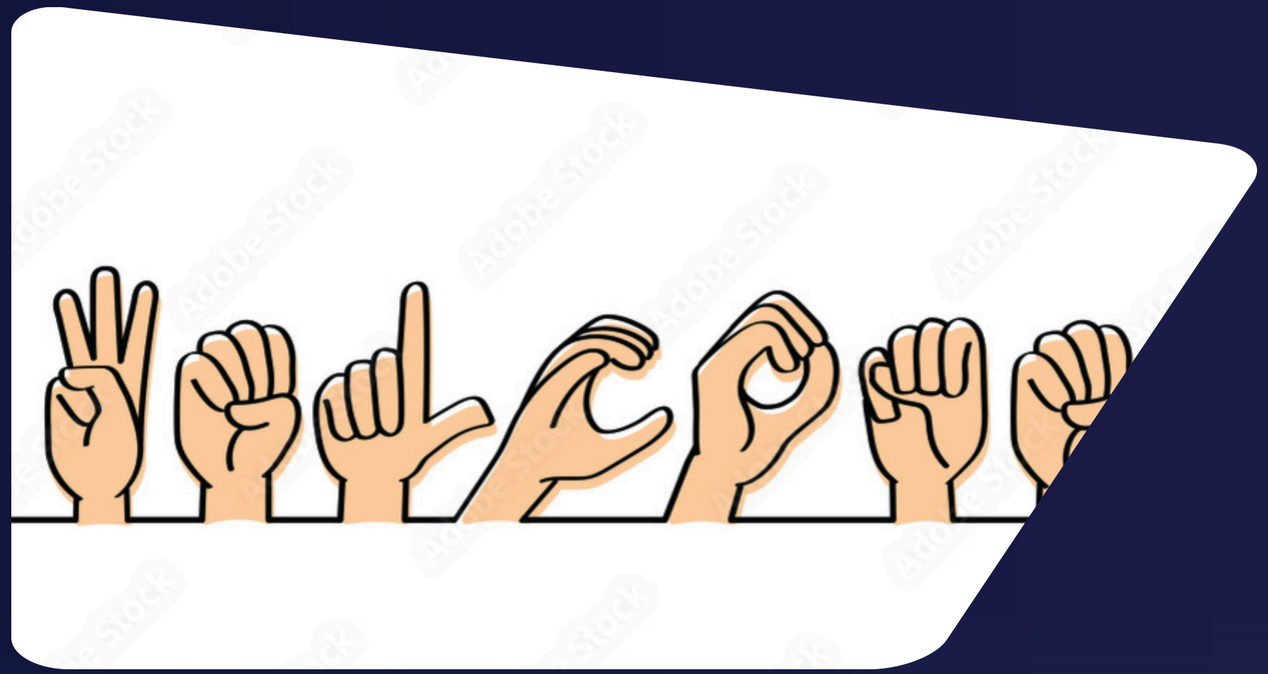
Alireza Mahdizadeh

Course: Embedded systems

INTRODUCTION

WHY _____?

HOW ?



APPROACHES IN SIGN LANGUAGE TRANSLATION

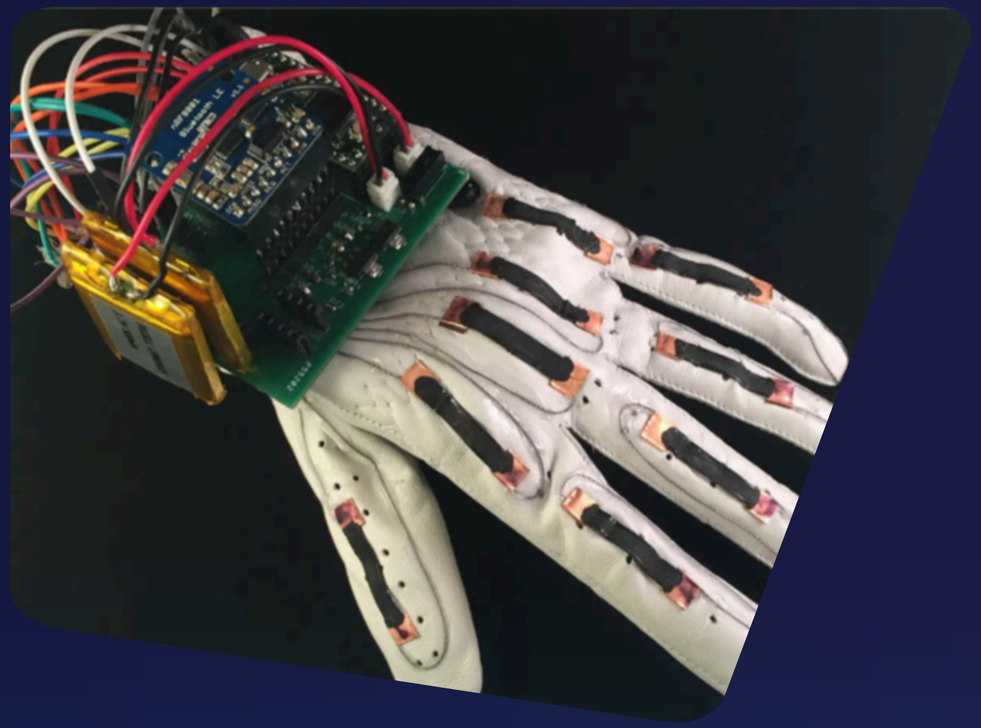
- Speech-to-Text Integrated Approaches
- Sensor-Based Approaches
- Vision-Based Approaches
- Hybrid Approaches

SPEECH-TO-TEXT INTEGRATED APPROACHES

Uses speech recognition to capture spoken language, translating it into sign language or text for interaction with deaf and hard-of-hearing individuals.

SENSOR-BASED APPROACHES

These methods rely on wearable devices that capture movement, orientation, and other physical properties of hand gestures.



VISION-BASED APPROACHES

Vision-based methods use computer vision techniques to capture and interpret sign language gestures without any physical contact with the user.

HYBRID APPROACHES

Combines vision-based and sensor-based data to increase accuracy and robustness of sign language recognition.

TYPES OF VISION-BASED ALGORITHMS

Feature-Based Algorithms

Template Matching

3D Vision and Depth-Based Algorithms

Pose Estimation Algorithms

Optical Flow-Based Algorithms

Deep Learning Algorithms

FEATURE-BASED ALGORITHMS

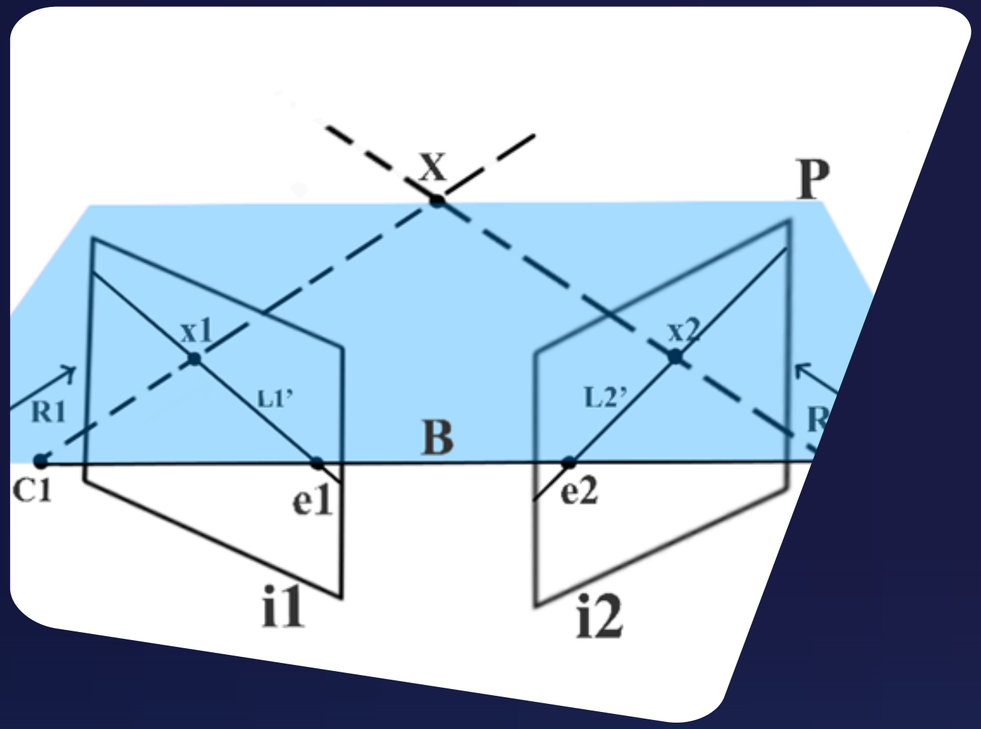
Use predefined features (e.g., hand shape, edges) to recognize gestures.

TEMPLATE MATCHING

Compares input images to a set of stored templates and finds the closest match.

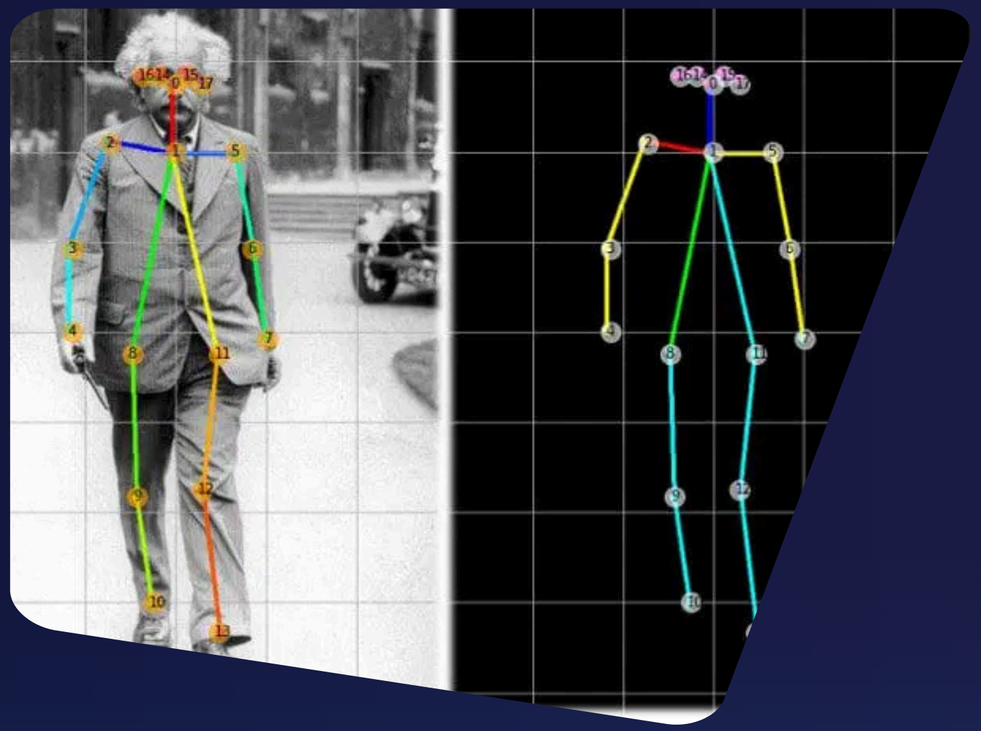
3D VISION AND DEPTH-BASED ALGORITHMS

Use depth information in addition to RGB data for more accurate gesture recognition.



POSE ESTIMATION ALGORITHMS

Detect key points (e.g., joints, hand landmarks) and infer pose from these points.

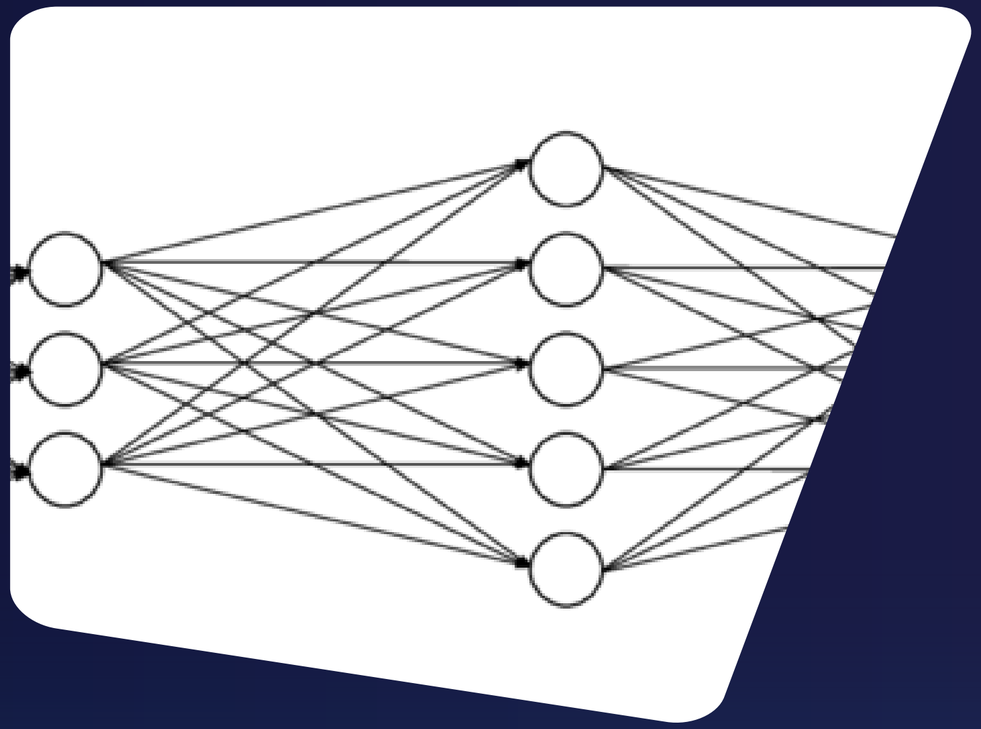


OPTICAL FLOW-BASED ALGORITHMS

Track movement by analyzing the flow of pixels between frames.

DEEP LEARNING ALGORITHMS

Learn complex features from large datasets, automatically recognizing patterns in images.



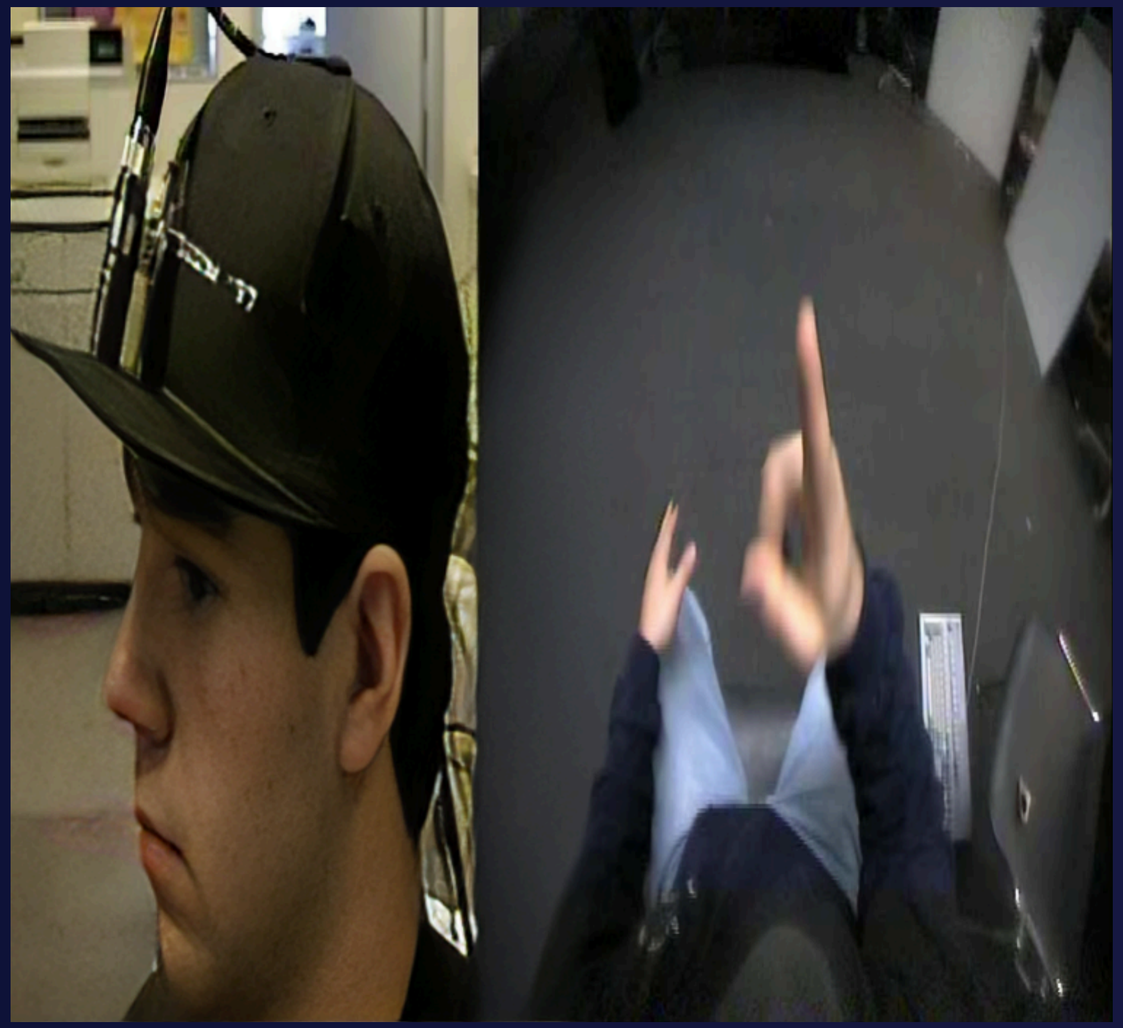
- Examples:
- Convolutional Neural Networks (CNNs): Capture spatial hierarchies in images and are widely used for image classification tasks.
- Recurrent Neural Networks (RNNs) and LSTMs: Used in combination with CNNs to capture temporal patterns in continuous gestures.

ARTICLE APPROACH

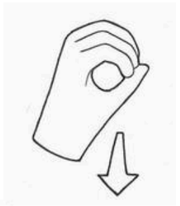
VISION-BASED SIGN LANGUAGE TRANSLATION DEVICE

mobile system developed to translate Indian Sign Language into spoken English using real-time vision processing. Designed as an application on a mobile device with LabVIEW software, the system captures sign gestures through the phone's camera, processes the images, and converts recognized gestures into audio output. By recognizing one-handed alphabet signs and numbers, the device aims to bridge communication gaps for the Deaf community without requiring a human interpreter.

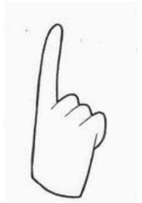
FIGURE OF ENVIRONMENT SETUP



THINGS THAT THE EMBEDDED SYSTEM DEVICE GIVES



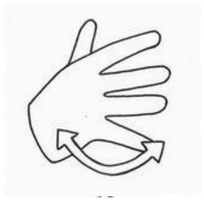
0



1

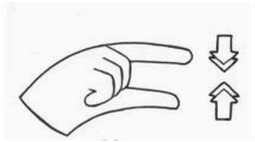
5

6



10

11



12

13

2

3

4

7

8

9

METHODOLOGY

a real-time vision-based system for recognizing finger spelling continuous Sign Language (ASL) using a single camera to track the user's unadorned hands. This system is broken down into three main parts starting with the **image acquisition** followed by **image processing** to extract features for recognition and last comes the recognition stage where signs are identified and audio output is given.

PROBLEMS

THERE IS NO GOOD WITHOUT BAD!!!

01

Limitation:
The number of characters it recognizes is limited to "letters" and "numbers".

02

Accuracy:
Due to the selected model and the use of a mobile phone for processing, it is less accurate (needs to create a minimum for the accuracy of the work)

03

Limited Dataset:
The system is trained on a small dataset, which restricts its accuracy and effectiveness, especially when recognizing signs from different users or slightly varied gestures.

IMPLEMENTATION

OBJECTIVE

Implementing real-time sign language recognition using vision-based methods.

IMPLEMENTATION

APPROACH

- Leveraged MediaPipe Holistic model to capture body, hand, and facial landmarks.
- Used OpenCV for real-time video processing and display.
- Collected data by capturing key points from specific gestures (“Salam,” “mamnoon,” and “khoda hafez”).

IMPLEMENTATION

PROCESS

- Landmark Detection: Facial, hand, and body landmarks are detected using MediaPipe.
- Keypoint Extraction: Extracted key points for each landmark to represent gestures numerically.
- Data Storage: Saved key points in structured folders, organized by gestures and sequences.

TIT FOR TAT