UNIVERSITY OF BAHRAIN
COLLEGE OF INFORMATION TECHNOLOGY
DEPARTMENT OF COMPUTER ENGINEERING

SIGN L ANGUAGE GESTURE RECOGNITION USING COMPUTER VISION IN REAL TIME

Zainab Redha Ebrahim Matrook 20195982
Puja Karmakar 20194981
Afifa Mohamed Ismaeel Altorabi 20183087

Supervised By: Dr. Jalal Mohamed Mufleh Khlaifat

ABSTRACT

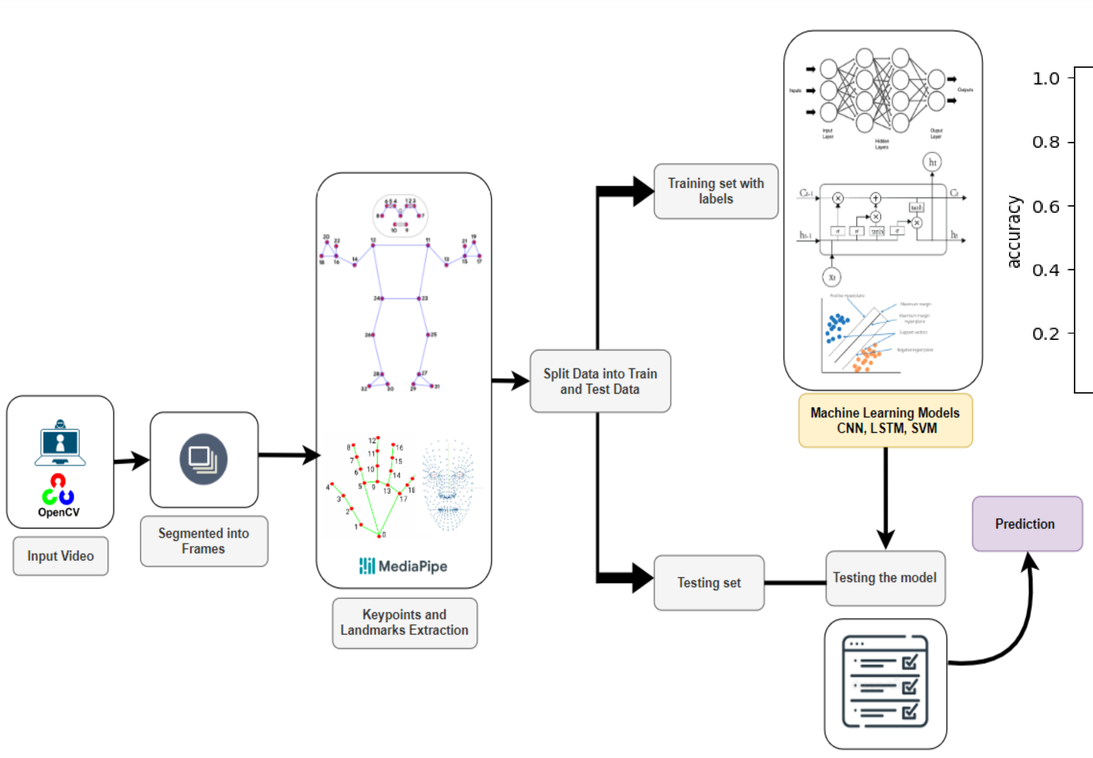
Sign Language Gesture Recognition (SLGR), with its ability to decipher hand and body movements into meaningful signs, serves as a bridge of communication between the deaf and the hearing. Our project proposes Machine Learning (ML) models for Sign Language Gesture Recognition (SLGR) using computer vision in real time. The model employs a deep learning-based approach, specifically a CNN-LSTM-SVM model, to recognize and classify various sign language gestures in real time. The proposed model achieved an accuracy of 95% for 45 Word-Level dataset of American Sign Language (ASL) gestures and can be used as an educational tool to promote social inclusion and diversity. We used various sign language gestures in real time. The ML models were using OpenCV and Media Pipe Feature Engineering. It can be used as an educational tool for individuals who wish to learn sign language, thereby increasing social inclusion and promoting diversity.

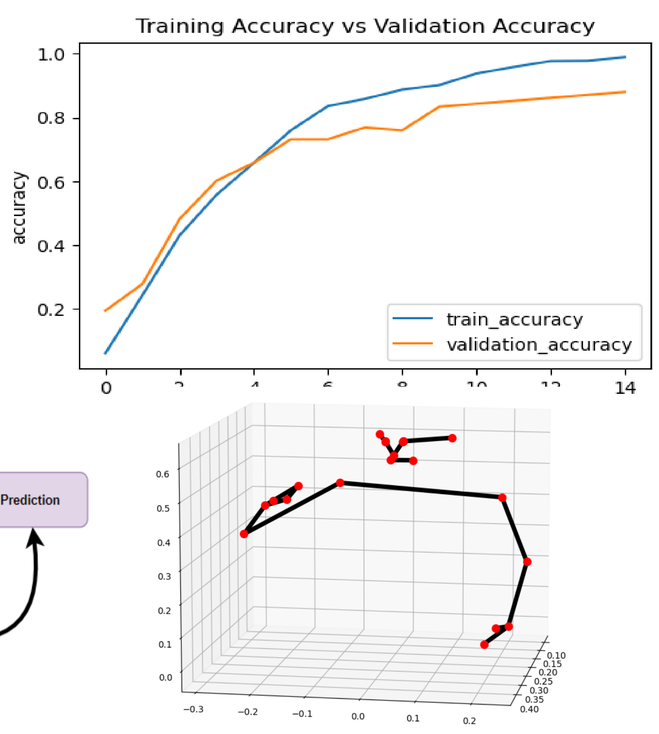
OBJECTIVES

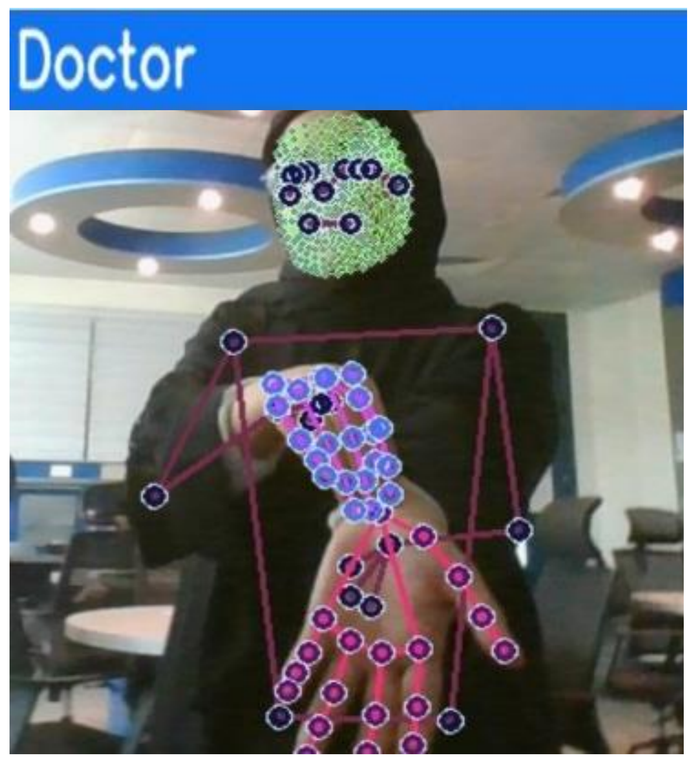
The primary goal of this project is to develop real-time 45 word-level Sign Language Action Gesture Detection using Computer Vision techniques, MediaPipe and OpenCV.

- Investigate and develop Machine Learning (ML) models **SVM-LSTM-CNN** and evaluate feature extraction.
- Enhance precision and dependability.
- Optimize the model performance with respect to speed and computational efficiency for real-time operation.
- Train models using TensorFlow with video streams on various numbers of sequences, frames, and words .
- Determine model's best accuracy after rigorous testing and evaluation.
- Hardware implementation and Deployment on Raspberry Pi3.
- Critical appraisal of each model's strengths and limitations.

METHODS/DIAGRAMS/FIGURES







RESULTS

Final ML Model Accuracy

Dataset	Model	Accuracy	Time
45 word-level SL Gestures 30 frames Test size = 0.2	SVM	Train: 87.59% Test: 74.44%	Train: 69.30 seconds Test: 16.73 seconds
	CNN (3-Layer)	Train: 98.15% Test: 80.37%	Train: 584.44 seconds Test: 2.71 seconds
	LSTM	Train: 80.56% Test: 65.93%	Train: 82.29 seconds Test: 0.61 seconds

CONCLUSION & FUTURE WORK

The **CNN model had the highest accuracy of over 95% on the ASL dataset**, but the LSTM model performed better in terms of parameter handling. However, accuracy alone may not reflect true performance in real-world settings. Future work includes optimizing the model for real-time operation on the Raspberry Pi and exploring a hybrid model for better operability and accuracy.