



# EmoDetect

*Emotion Recognition Using Facial and Speech Analysis*

Artificial Intelligence (CSE3705)

# Group-14

Serial Number	Name	Enrollment Number
1	Ananya Aggarwal	220378
2	Aastha Singh	220387
3	Siddhika Sinha	220388
4	Peehu	220623



# Problem Statement



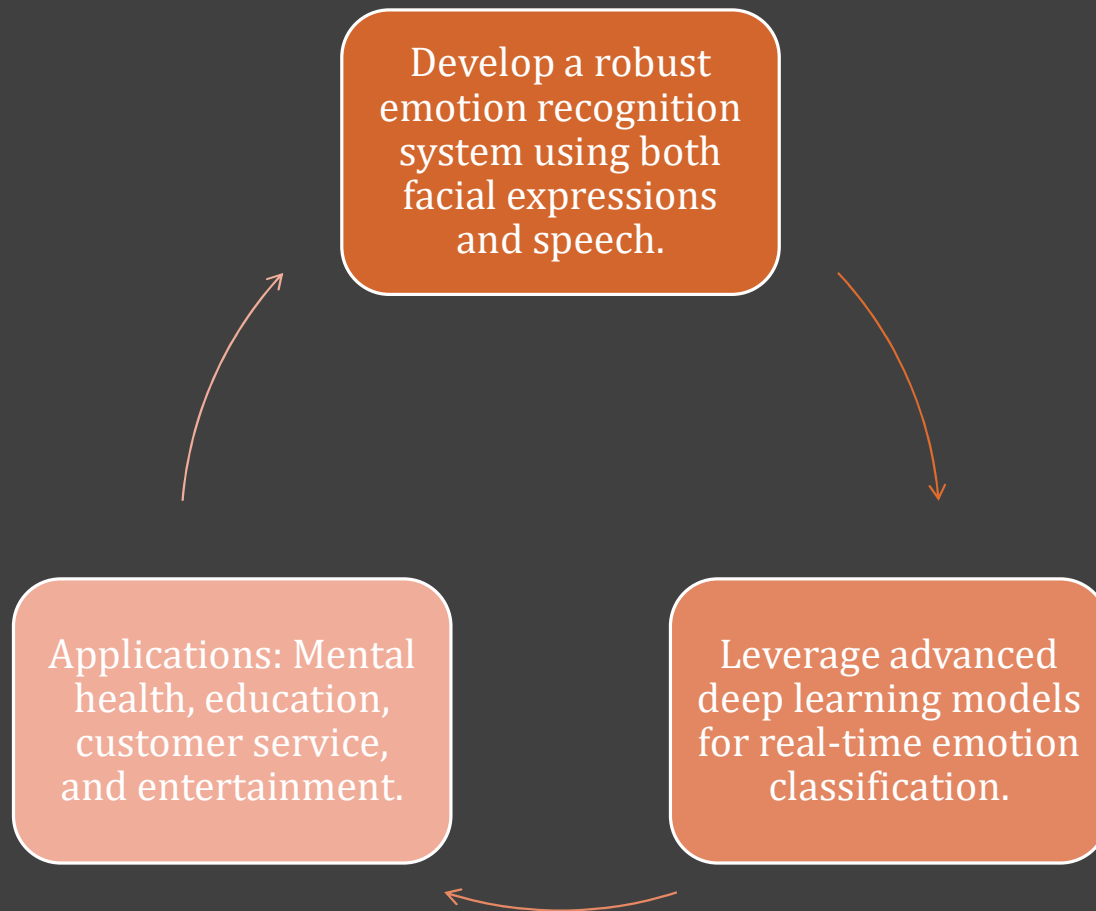
Understanding emotions is crucial for improving human-computer interaction.



Unimodal systems struggle in real-world scenarios due to noise or occlusions.



Need for a multimodal system combining speech and facial cues to increase accuracy and robustness.



*Our Aim*



# What Has Been Done?


## **Data Acquisition:**

- TESS Dataset for speech-based emotion recognition.
- FER-2013 Dataset for facial emotion recognition.

## **Feature Extraction:**

- Speech: MFCCs, pitch variations.
- Facial: Convolutional Neural Networks (CNNs).

## **Model Development:**

- LSTM for speech data processing.
  - CNN for facial data processing.
- 

---

# Datasets for Analysis

---

## **TESS Dataset (Speech):**

- 2 female actors recording 7 emotions: happiness, sadness, anger, fear, disgust, surprise, neutral.
- High-quality WAV files, 16 kHz sampling rate.

## **FER-2013 Dataset (Facial):**

- 48x48 grayscale images of facial expressions.
- Covers 7 emotion classes with 35,887 labeled images.



# Features for Analysis

## **Speech Features:**

- MFCCs
- Pitch variations
- Spectral contrast

## **Facial Features:**

- Hierarchical features learned using CNNs.



# Technology Stack

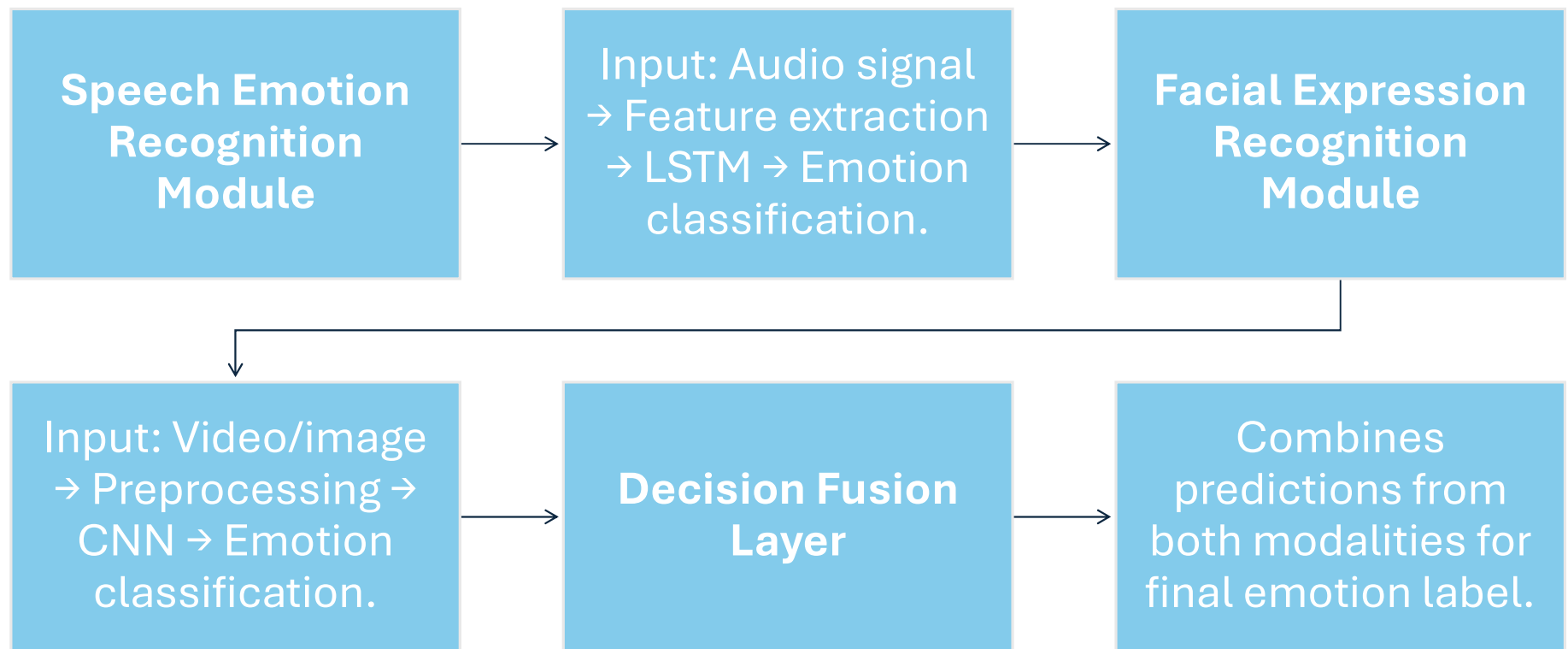
---

Programming Language	Python
Libraries	TensorFlow, Keras, Librosa, OpenCV, Seaborn, Matplotlib
Models Used	LSTM for speech processing. CNN for facial expression recognition.

---



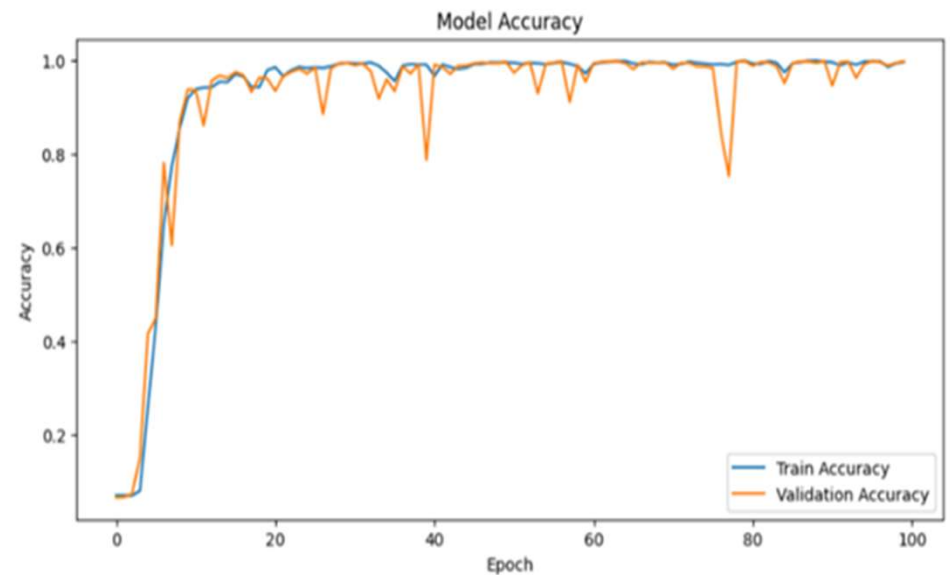
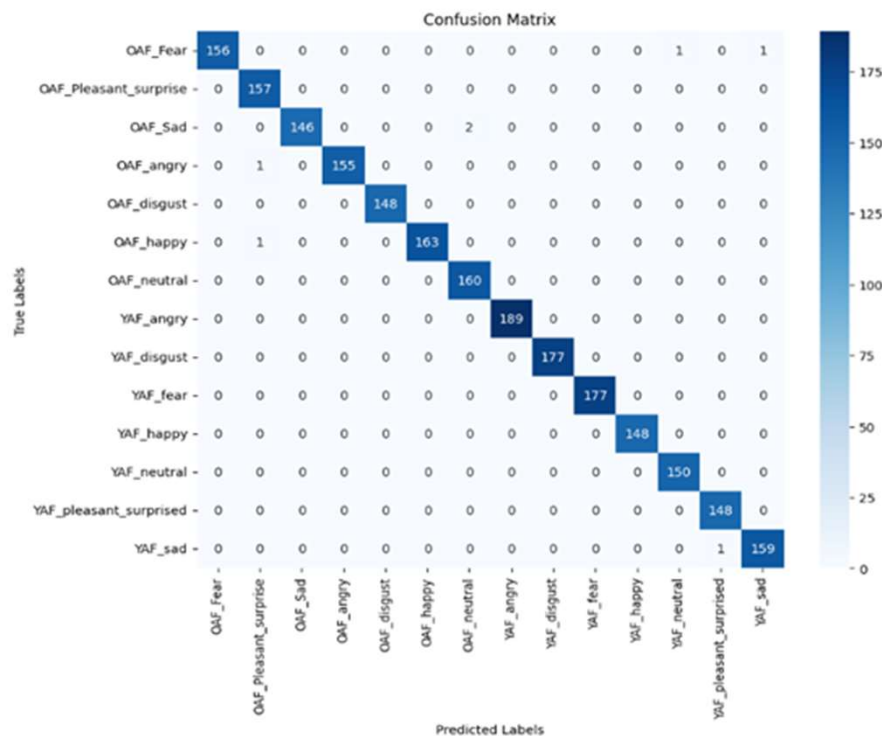
# System Architecture

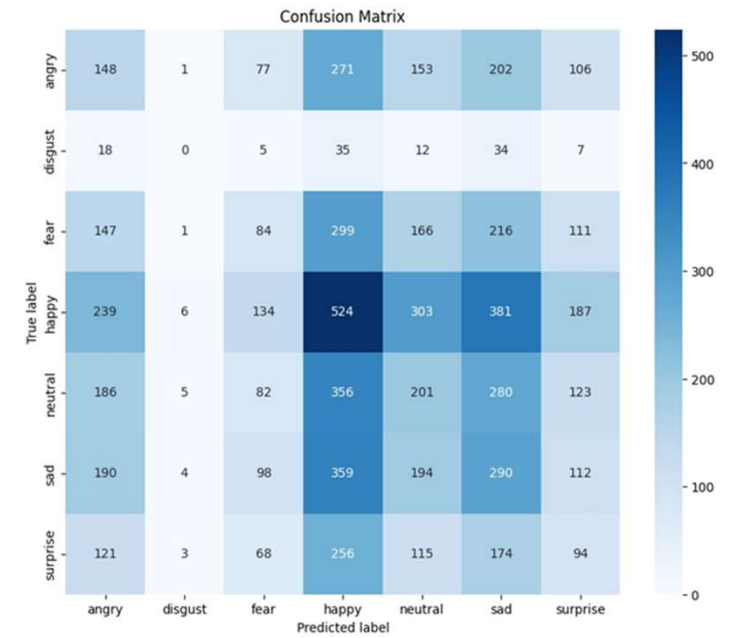
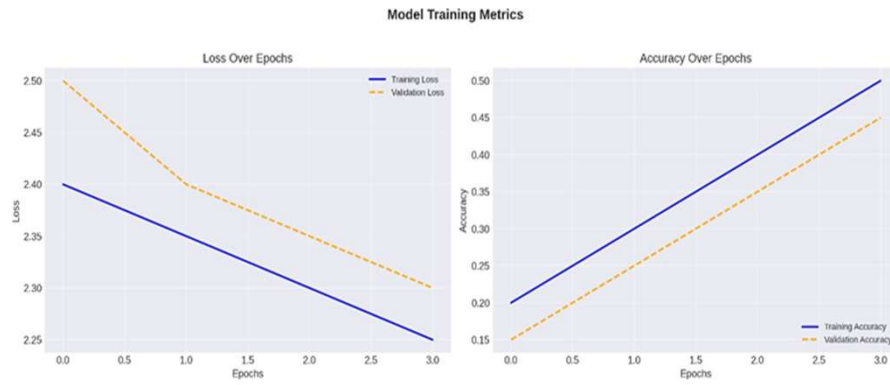




# *RESULTS*

# Speech Analysis Model Results





# Facial Analysis Model Results





*DEMO*

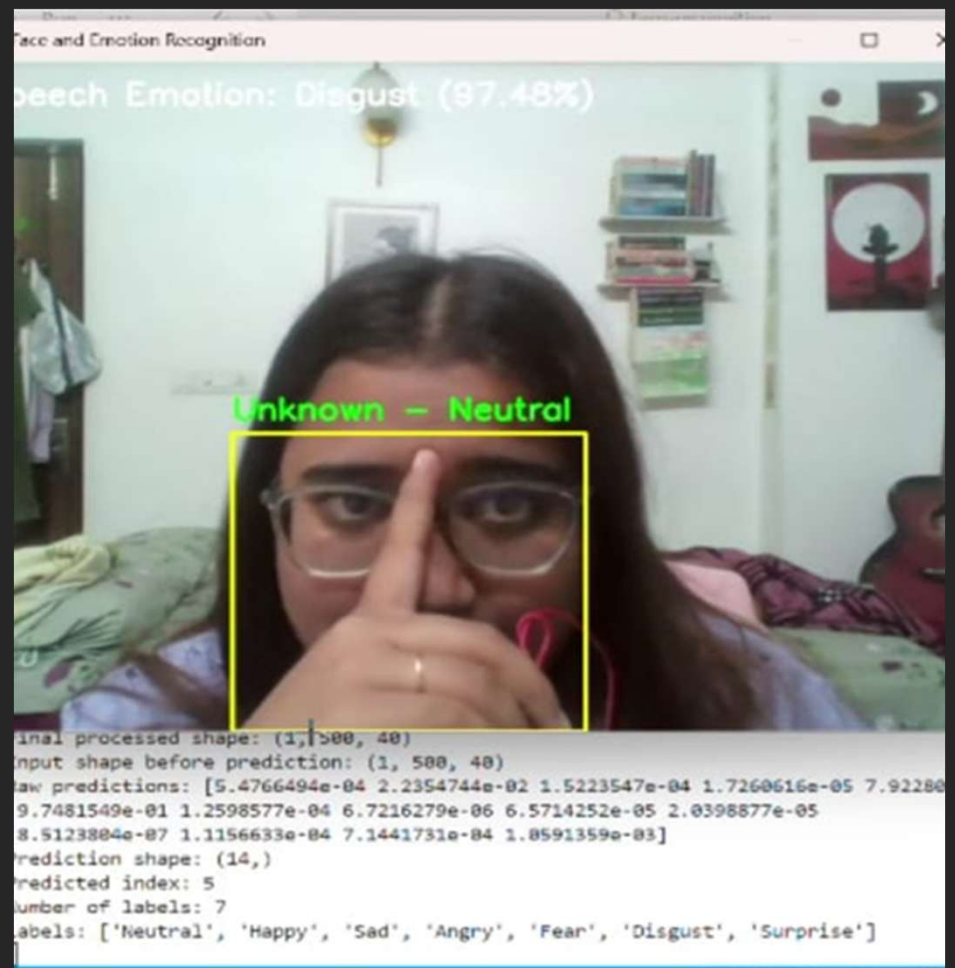
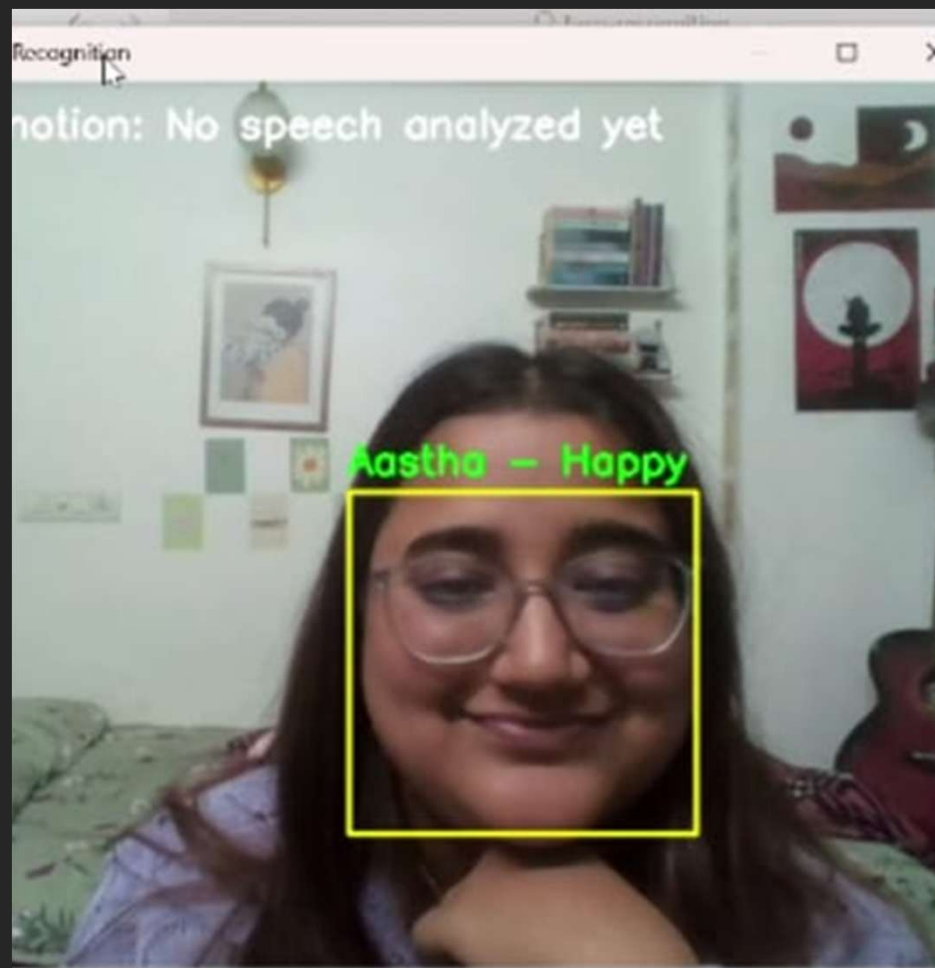


```
Go Run ... ← → Face recognition
face_recognition.py import tensorflow as tf.py 1 face_emotion_recognition[1].py
FACE SPEECH EMO.py > main
292 def main():
382
383     except Exception as e:
384         print(f"Error in main loop: {str(e)}")
385
386     finally:
387         running = False
388         cap.release()
389         cv2.destroyAllWindows()
390         # Wait for threads to finish
391         keyboard_thread.join(timeout=1)
392         audio_thread.join(timeout=1)
393         print("\nProgram terminated.")
394
395
396 if __name__ == "__main__":
397     main()

PROBLEMS 30 DEBUG CONSOLE TERMINAL OUTPUT PORTS JUPYTER
Final processed shape: (1, 500, 40)
Input shape before prediction: (1, 500, 40)
Raw predictions: [6.1641382e-03 5.5059133e-04 4.0822248e-03 4.3322626e-04 2.2493205e-03
1.5153948e-03 1.4711329e-04 6.6922158e-03 2.9967401e-02 2.9664708e-02
1.8276734e-02 2.8386363e-01 3.4857117e-02 5.8153623e-01]
Prediction shape: (14,)
Predicted index: 13
Number of labels: 7
Labels: ['Neutral', 'Happy', 'Sad', 'Angry', 'Fear', 'Disgust', 'Surprise']

Program terminated.
PS C:\Users\seeth\OneDrive\Desktop\Face-recognition>
```

Ln 393, Col 30 Spaces: 4 UTF-8



# Applications



**MENTAL  
HEALTH**



**EDUCATION**



**CUSTOMER  
SERVICE**



**GAMING**





Multilingual adaptability.



Integration with IoT devices  
for real-time applications.



Advanced fusion techniques  
like transformers.

***Future Scope***



## *Conclusion*

The developed system is a robust multimodal emotion recognition model that effectively combines speech and facial data to analyze and classify emotions. By leveraging advanced deep learning techniques, the system achieves enhanced accuracy and demonstrates real-time processing capabilities. This integration of modalities ensures reliable performance across diverse scenarios, making it a significant advancement over traditional unimodal approaches.

---

# Thank You

---