Emotion Recognition using Images and Audio

Mr. Prathamesh Chaudhary, Mr. Shlok Das, Mr. Raghvendra Kanakdande, Mr. Anurag Kotwal

AISSMS Institute of Information Technology, Pune ,India.

prathameshchaudhary7122@gmail.com, shlokdas2002@gmail.com, anuragkotwal123@gmail.com, kanakdanderaghvendra@gmail.com

AISSMS INSTITUTE OF INFORMATION TECHNOLOGY (IOIT)
An Autonomous Institute Affiliated to Savitribai Phule Pune University
Approved by AICTE, New Delhi and Recognised by Govt. of Maharashtra
Accredited by NAAC with "A+" Grade | NBA - 5 UG Programmes

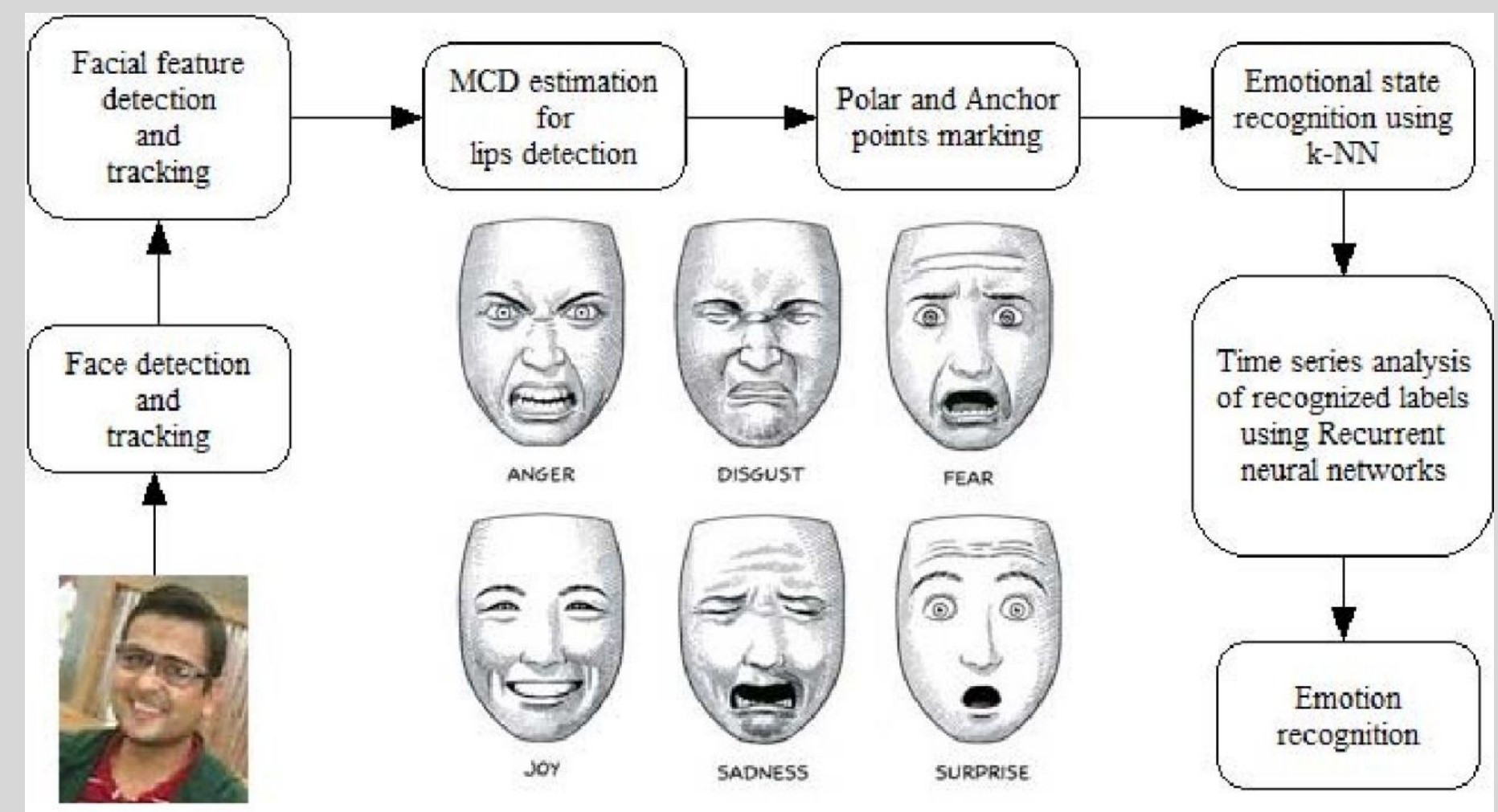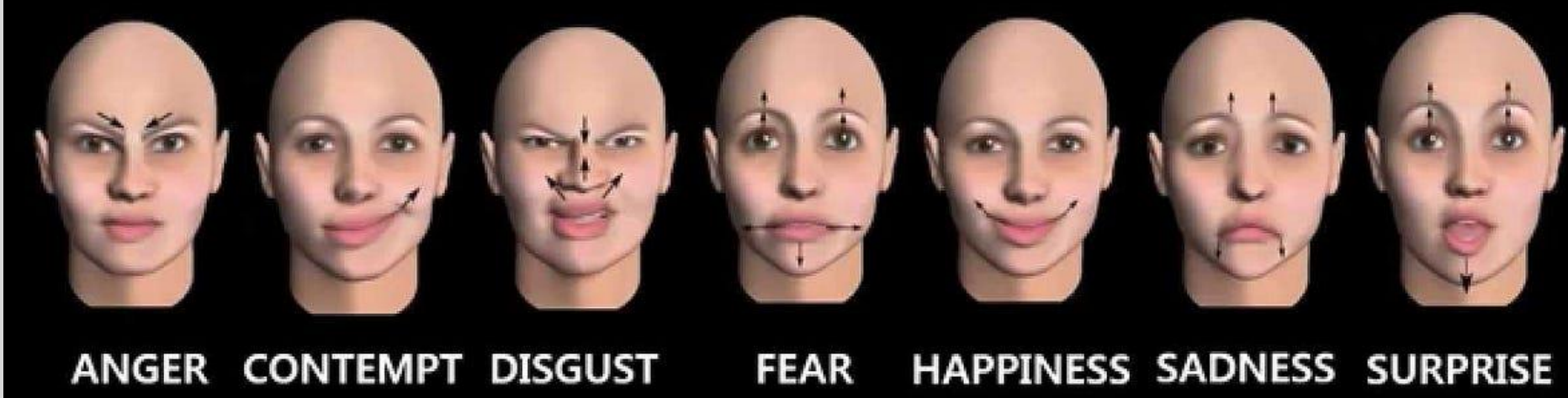Savitribai Phule Pune University

What Is Emotion Recognition?

Emotion recognition is the ability of technology to identify and understand human emotions based on facial expressions, voice tone, or other cues.

Why it is Important?

1. It enables machines to respond to human emotions, improving user experiences.
2. It can assist in detecting emotional patterns, aiding in early intervention and support for mental health issues.
3. Businesses use emotion recognition to gauge consumer reactions, optimize products, and tailor marketing strategies for better customer engagement.
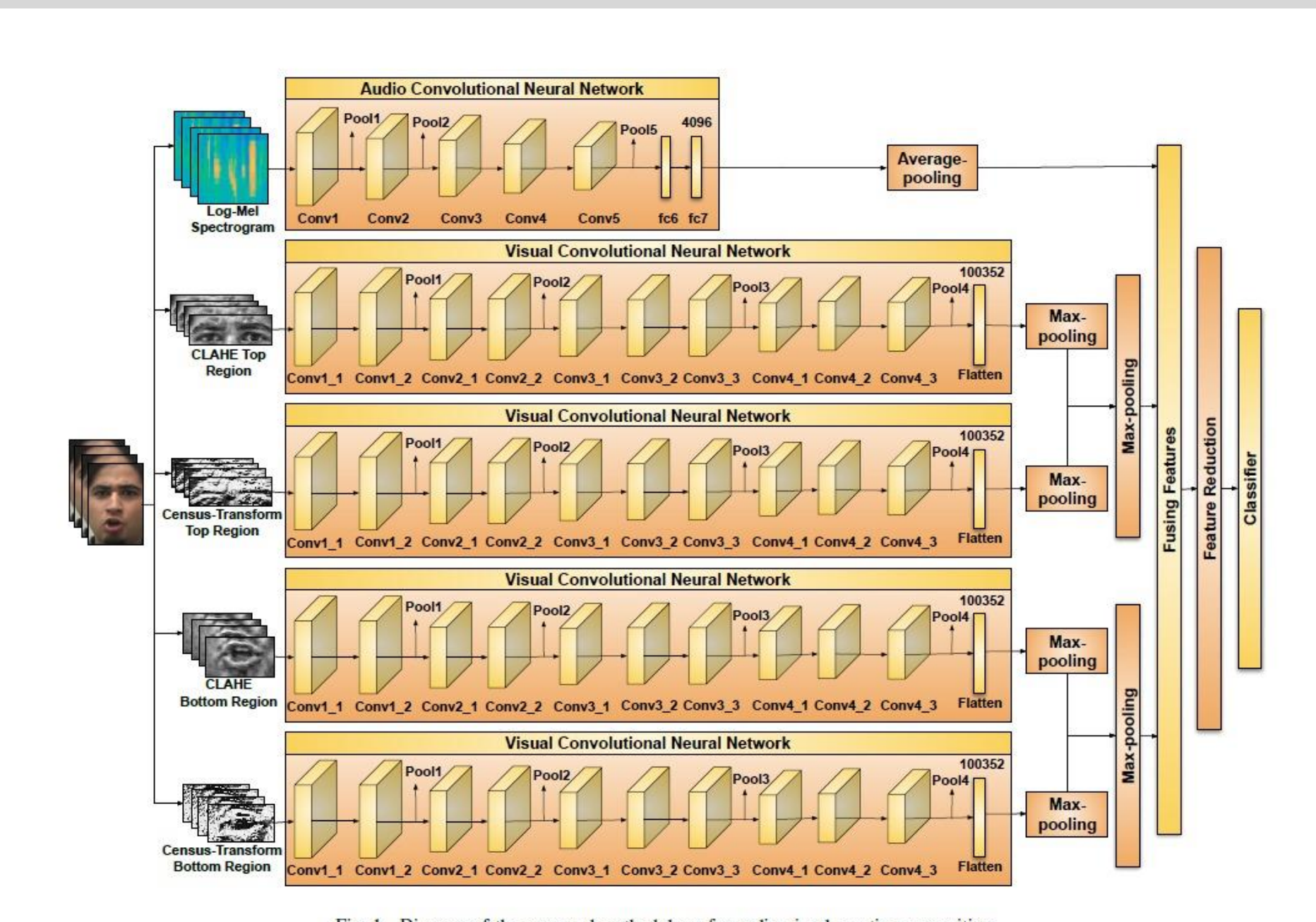
Recognition Algorithm

Fig. 1. Diagram of the proposed methodology for audio-visual emotion recognition.

www.PosterPresentations.com

How Emotion Recognition works?

Image Process

1. CNNs excel at processing image data, especially facial expressions.
2. They detect and analyze features like eyebrow position, eye openness, mouth shape, and overall facial expressions
3. CNNs focus on extracting spatial features from images or video frames, allowing them to recognize static visual cues that convey emotions.
4. Unlike RNNs, CNNs process image data in parallel, making them efficient at analyzing multiple regions of interest simultaneously, such as different facial features.

Audio Process

5. RNNs process audio data, which is often represented as spectrograms or Mel-frequency cepstral coefficients (MFCCs).
6. They analyze the sequence of audio frames to identify changes in tone, pace.
7. RNNs process audio data in a sequential manner, making them suitable for analyzing audio signals
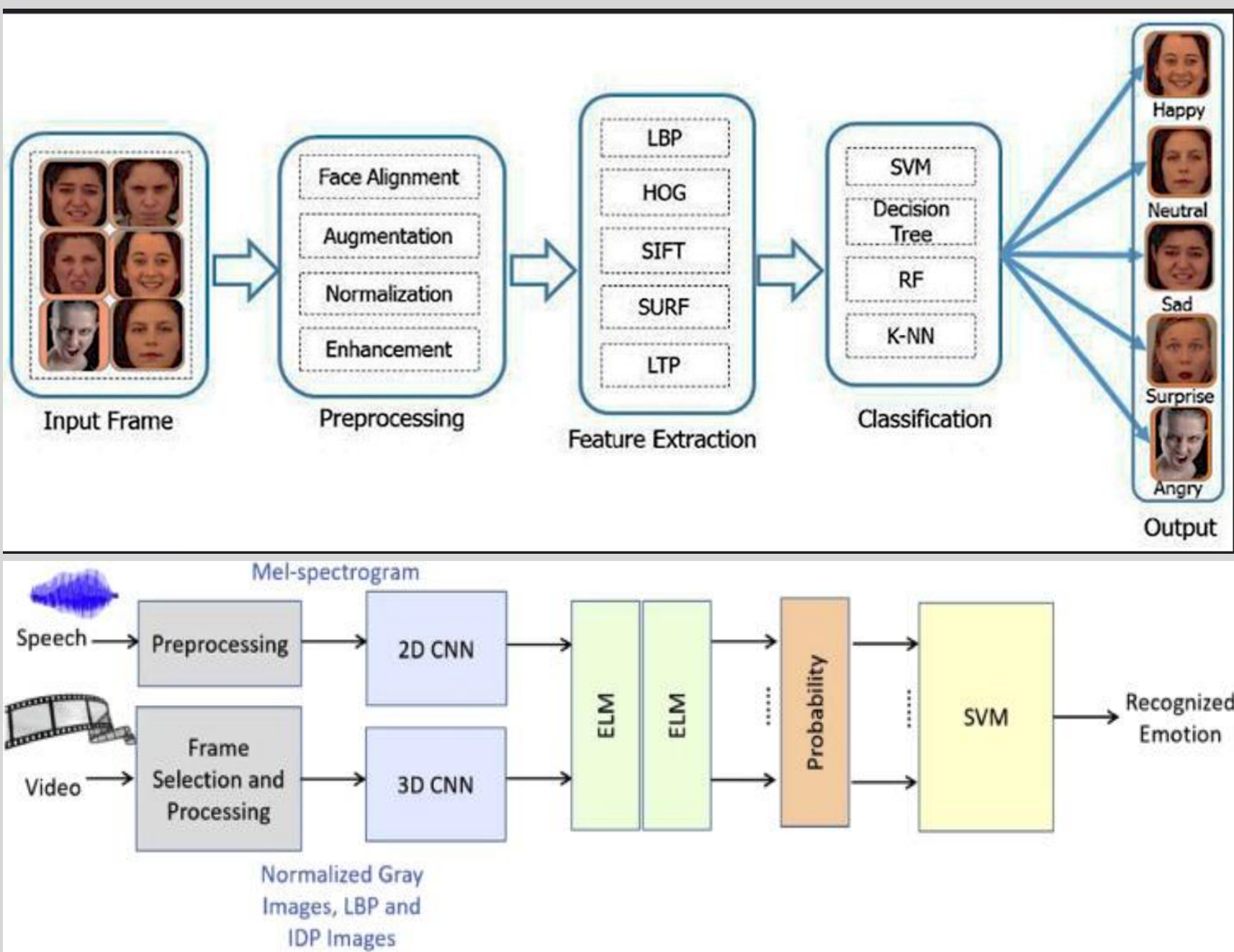
Convolutional Neural Network
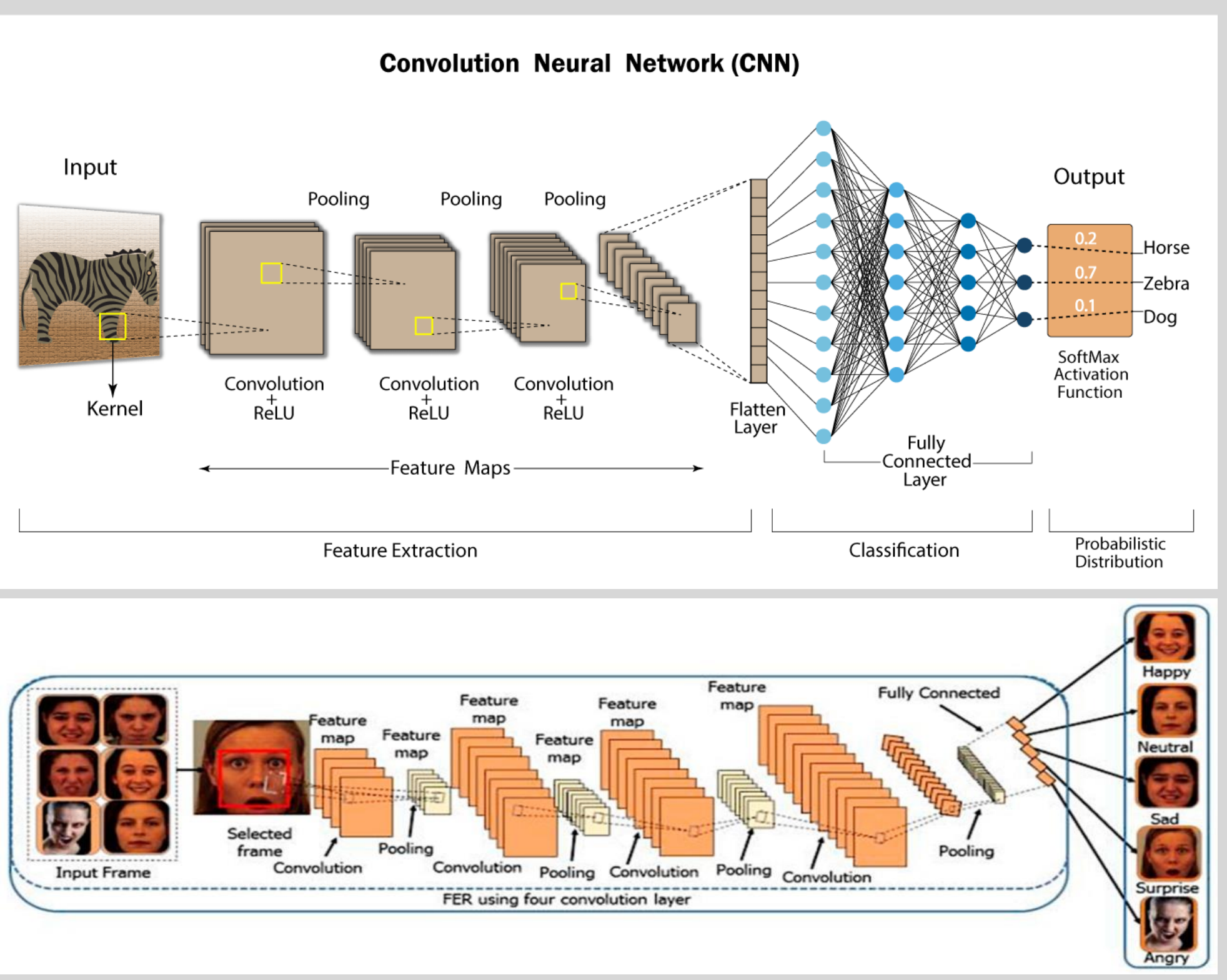
Recurrent Neural Network
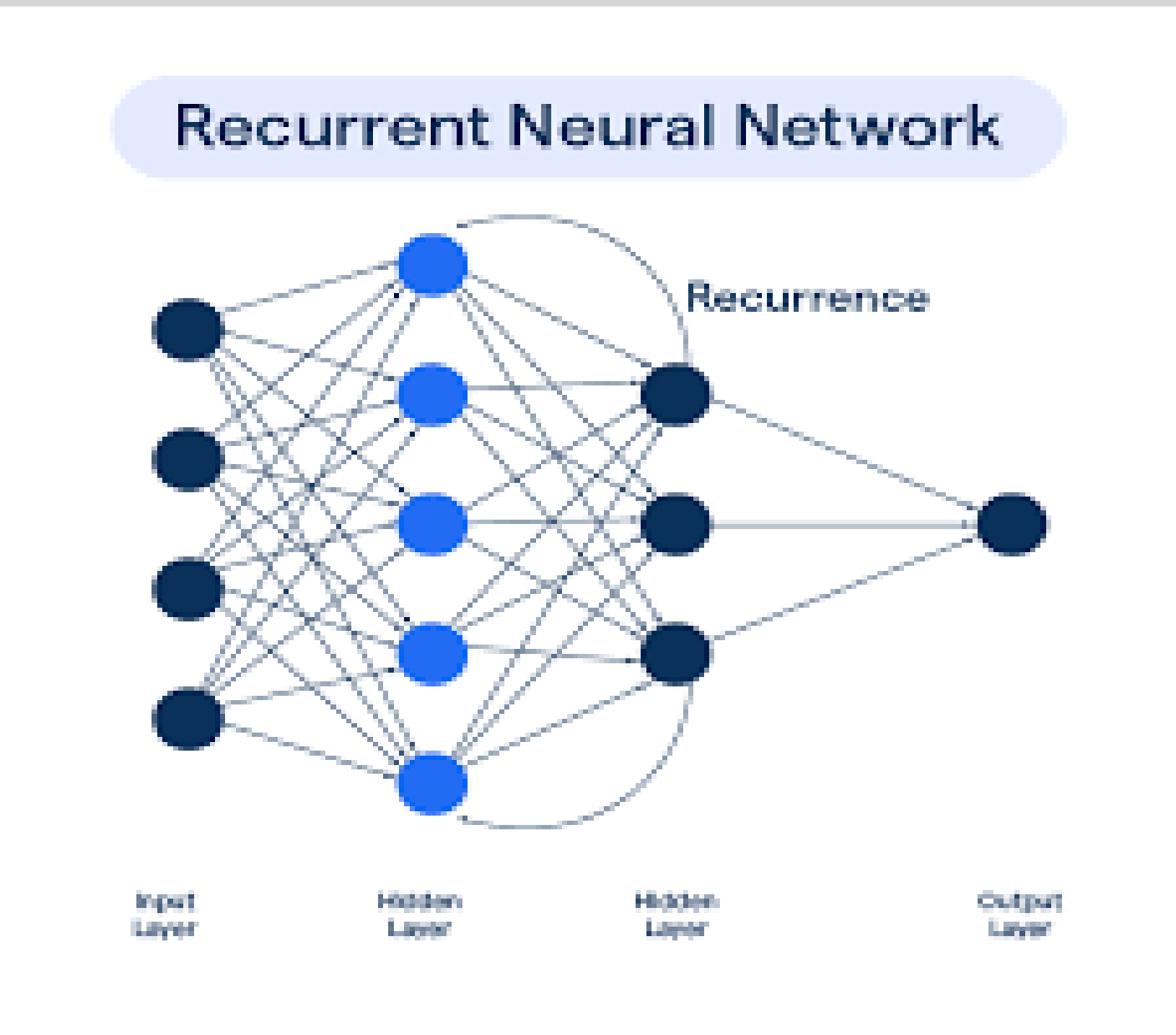
Emotion Recognition & Its features


I. **Multisensory Data:** Utilizes audio, visual, and other cues for emotion analysis.
II. **Real-World Applications:** Useful in fields like healthcare, human-computer interaction, and marketing.
III. **Adaptive Systems:** Enables technology to respond appropriately to human emotions.
IV. **Ethical Considerations:** Raises concerns about privacy and consent in emotion data collection.
V. **Ongoing Research:** A dynamic field evolving with advancements in AI and machine learning.

Possible Uses

I. Mental Health Monitoring.
II. Virtual Assistants: Enhancing human-computer interactions.
III. Creating immersive gaming experiences and emotionally responsive content.
IV. Analyzing consumer reactions to products and advertising for better marketing strategies.

Face Detection — Facial Expression Detection — Classification to Emotional State "Joy"

Performance Analysis
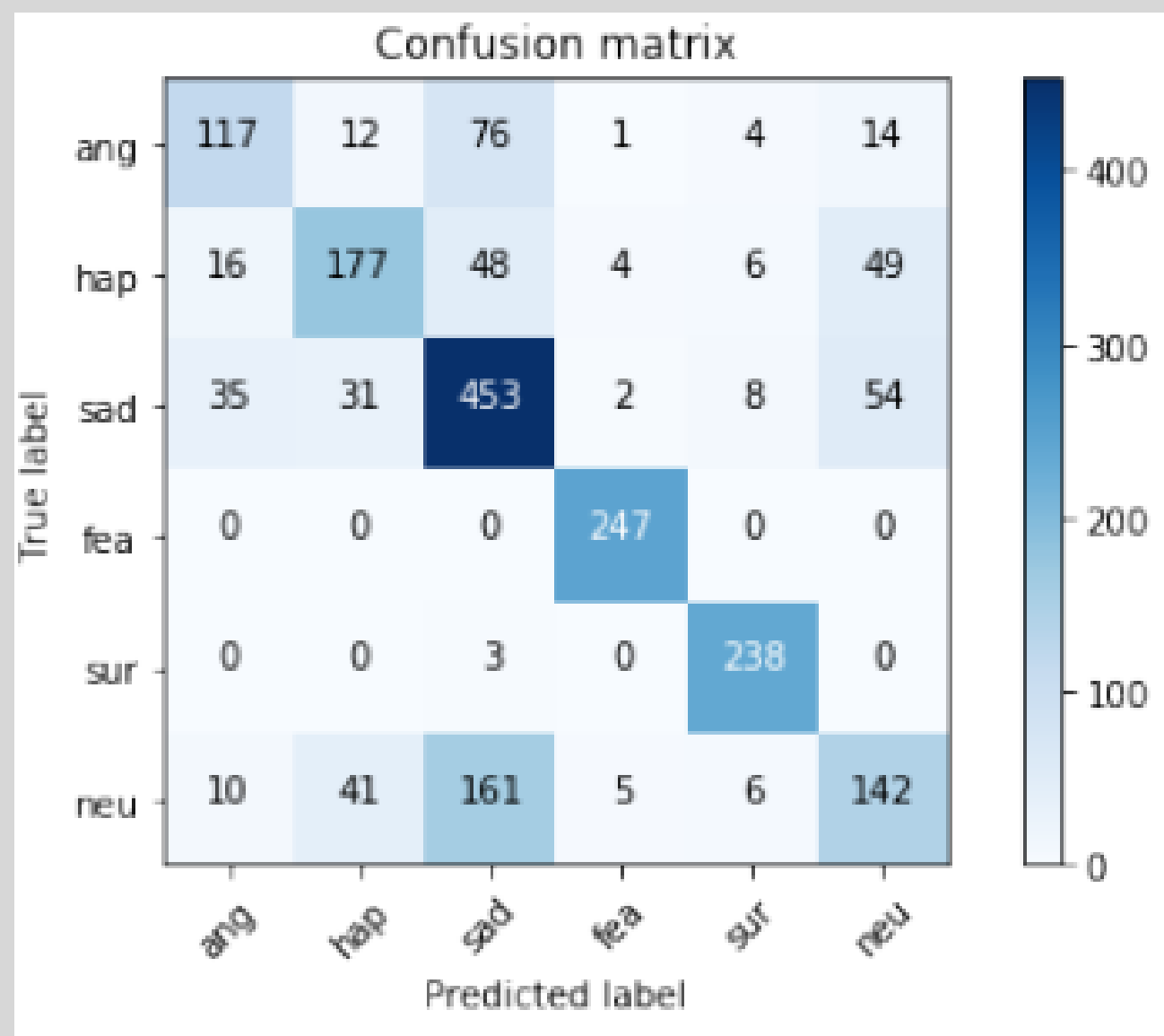
I. False Positive / False Negative.
II. Precision and Recall.

Conclusion

Our ongoing research aims to develop an effective multimodal emotion recognition system using audio and images data. We anticipate that our work will enhance emotion recognition accuracy and have significant applications in various domains.

References

1. [1] Arimura, K., Hagita, N., "Feature space design for image recognition withimage screening," in Pattern recognition, vol. 2, 1996, pp. 261–265.
2. [2] Lijiang Chen a, Xia Maoa, Yuli Xue, Lee Lung Cheng, "Speech emotionrecognition: Features and classification models", in Digital SignalProcessing, vol.22, 2012, pp.1154-1160.
3. [3] Dai Fang, He Haimei, Han Wei, "Integrating multi-feature of image basedon correspondence analysis".