EMOTION DETECTION USING FUSION OF PHYSIOLOGICAL DATA

KUSH FALDU
22BCM037

RITU KANABAR
22BCM052

GUIDE : PROF. DAIWAT VYAS

INTRODUCTION

What is Emotion and Stress?
- Emotion is a feeling like happiness, sadness, anger, or fear.
-Stress is the body's reaction to pressure or challenges, causing tension or worry

Why is Emotion and Stress Detection important?
- Real-time stress and emotion detection benefits healthcare, workplace productivity, and more.

Challenges in Traditional Methods
- Single-sensor approaches lead to inaccurate results due to noise and external factors.
-Centralized data processing in traditional ML raises privacy concerns (e.g., GDPR compliance). Federated Learning (FL) enables collaborative model training without sharing raw user data, addressing privacy issues in emotion/stress detection systems.

LITERATURE REVIEW

Multimodal Emotion Recognition (Fusion of EEG & ECG):

- Some papers explored fusion-based approaches using EEG and ECG, along with PPG, skin conductance (EDA), and facial expressions.
- Fusion models generally improved classification accuracy compared to unimodal models.
- Data preprocessing techniques such as artifact removal (ICA, bandpass filtering) and feature extraction (time & frequency domain features) were crucial for better results.
- Federated Learning (FL) frameworks (e.g., FedAvg) enhance privacy in multimodal systems by aggregating model updates locally (gradients) instead of centralizing sensitive physiological data (EEG/ECG).
- FL mitigates risks of data breaches in healthcare applications while maintaining model accuracy

Use of Budget Wearables for Emotion & Stress Detection:

- Several studies tested smartwatches and low-cost sensors for emotion recognition.
- Wearables such as Empatica E4, Shimmer sensors, and commercial smartwatches were used to track ECG, PPG, and EDA signals.
- Accuracy of wearable-based emotion detection was lower than lab-grade sensors but showed potential for real-world applications.

WHY FUSION OF PHYSIOLOGICAL SENSOR DATA?

Overcoming Sensor Limitations:
- EEG detects brain activity but is sensitive to artifacts
- EDA measures skin conductance but is affected by ambient temperature
- HRV can indicate stress but is influenced by physical activity

Fusion Enhances Accuracy:
- Combining multiple sensors improves robustness
- Machine learning models benefit from multimodal input.

FL complements fusion by enabling decentralized training across devices (e.g., wearables), preserving data privacy. For example, FL can aggregate ECG/EEG features from multiple users without exposing individual signals.

PHYSIOLOGICAL SENSOR DATA USED

EEG (Electroencephalogram)
Measures brain activity, useful for recognizing emotions.

ECG (Electrocardiogram)
Captures heart activity, linked to stress levels.

PPG (Photoplethysmography)
Measures blood volume changes, useful in stress detection.

HRV (Heart Rate Variability)
Evaluates heart rate variations, an indicator of stress.

GSR (Galvanic Skin Response)
Detects emotional arousal through skin conductance changes.

BVP (Blood Volume Pulse)
Measures cardiovascular responses to emotional stimuli.

TEMP (Temperature)
Monitors skin temperature fluctuations, correlated with stress.

SYSTEM ARCHITECTURE FLOW

Multi-Sensor Data Acquisition:
- Data collection from EEG, EDA, HRV, and Skin Temperature sensors

Preprocessing:
- Noise removal, signal normalization, feature extraction
-Optional Federated Learning Layer: Local model training on devices (e.g., wearables) with encrypted gradient updates sent to a central server for aggregation.

Feature Fusion & Classification:
- Early, Late, Hybrid fusion methods

Output:
- Stress And Emotion classification/regression results
-Privacy-preserving predictions: Global FL model adapts to individual stress/emotion patterns without storing raw data centrally

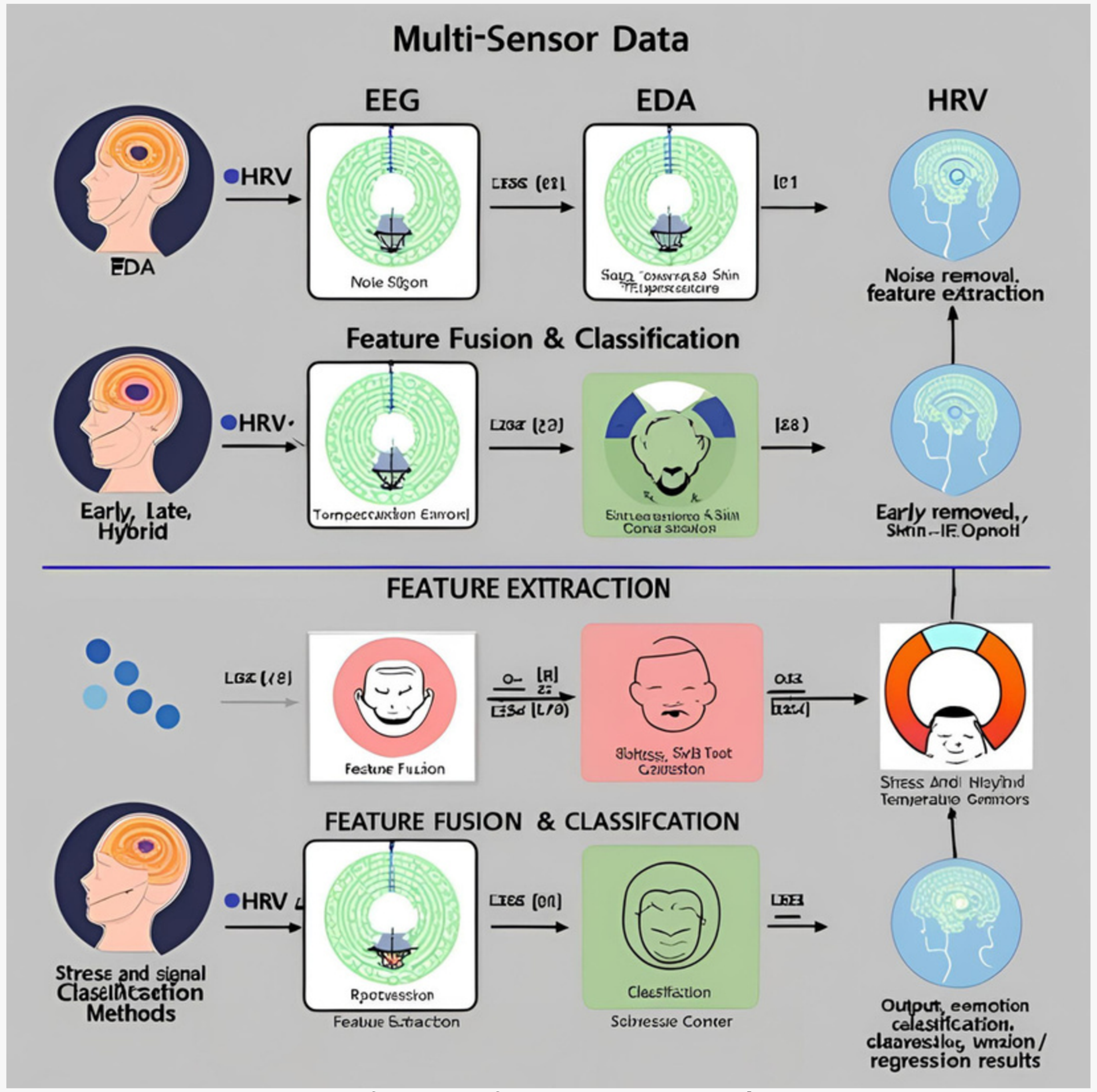


Fig-1 Ai Generated

DREAMER DATASET

Key Features

- Physiological Signals: EEG (128 hz), ECG (256 hz)
- Participants: 23 individuals (13 males, 10 females)
- Data Collection Method: 18 video clips were used as stimuli to elicit emotional responses.
- Valence, Arousal & Dominance (VAD) rating (0-5)

Why We Chose DREAMER Dataset?

- Provides both EEG and ECG, making it ideal for fusion-based emotion detection.
- The dataset is widely used in research, allowing comparison with previous studies.
- Covers a structured emotion representation (Valence-Arousal-Dominance model) for precise classification.

DREAMER DATASET - APPROACH

Bandpass Filtering

- Making Frequencies between lowcut and highcut
- EEG Filtering : 0.5 Hz to 45 Hz.
- ECG Filtering : 0.05 Hz to 100 Hz.
- Highcut frequency should be lower than nyquist frequency (0.5 * freq)

Feature Extraction

- Extract Time and Freq Domain features from filtered data
- Statistical Features : mean, standard deviation, min, max, and power for each channel (14 for EEG, 2 for ECG)
- Frequency-Domain Features (FFT Power Spectral Analysis) : delta: 0.5–4 Hz, theta: 4–8 Hz, alpha: 8–12 Hz, beta: 12–30 Hz, gamma: 30–45 Hz

DREAMER DATASET - APPROACH

Data Normalization and Spliting

- Normalize data
- 80% - 20% Train Test Split

Regression Model

- Random Forest regressor
- Gradient Boosting
- MLP Regressor

  Worst than Baseline model (that predicts mean of target variables)
- SHAP (SHapley Additive exPlanations) to get idea of best features

Classifcation Model

- Random Forest Classifier
- SVM
- XGBoost

DREAMER DATASET - RESULTS

~	Random Forest	SVM	XGBoost
Valence	67.47 %	73.49 %	78.31 %
Arousal	72.29 %	74.7 %	78.31 %
Dominance	78.31 %	83.13 %	84.34 %

DREAMER DATASET - EMOTION MAPPING

Valence	Arousal	Dominance	Emotion(s)	Description
Low	Low	Low	Sadness, Depression, Melancholy	Negative mood with low energy and feeling powerless or submissive. Example: Feeling defeated or hopeless.
Low	Low	High	Boredom, Apathy	Negative but calm state with a sense of control or indifference. Example: Feeling uninspired but not helpless.
Low	High	Low	Fear, Anxiety, Distress	Intense negative emotion with lack of control. Example: Panic or nervousness in a threatening situation.
Low	High	High	Anger, Frustration	Intense negative emotion with a sense of control or defiance. Example: Feeling outraged or irritated with agency to act.
High	Low	Low	Contentment, Serenity	Positive, calm state but feeling passive or submissive. Example: Quiet satisfaction or peacefulness without strong agency.
High	Low	High	Happiness, Calm Confidence	Positive, relaxed state with a sense of control. Example: Feeling joyful and self-assured in a stable environment.
High	High	Low	Excitement, Elation	Intense positive emotion but feeling somewhat out of control. Example: Thrill or euphoria, like during a surprising event.
High	High	High	Joy, Pride, Triumph	Intense positive emotion with strong control. Example: Feeling ecstatic and empowered, like achieving a major goal.

PHYMER DATASET

Key Features

- Physiological Signals: EEG, EDA, BVP, Peripheral Skin Temperature(TEMP)
- Emotions Considered: Valence (positive/negative), Arousal (high/low), Categorical labels
- Personality Assessment: Big Five personality dimensions(Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism)
- Data Collection Method: Participants watched emotionally stimulating videos, and their physiological responses were recorded.

Why We Chose Phymer Dataset?

- Provides a diverse set of physiological signals, allowing multimodal fusion.
- Includes BVP, EDA and personality traits, which are useful for stress detection.
- Offers labeled emotion categories, enabling supervised learning.

PHYMER DATASET - APPROCH

Feature Extraction

- Already Processed
- 6 (mean, standard deviation, min, max, range, median) per signal
- Features: EDA, BVP, Temp, HR
- 6*4 = 24 features
- ('eda_mean', 'eda_std', 'eda_min', 'eda_max', 'eda_range', 'eda_median', 'bvp_mean', 'bvp_std', 'bvp_min', 'bvp_max', 'bvp_range', 'bvp_median', 'temp_mean', 'temp_std', 'temp_min', 'temp_max', 'temp_range', 'temp_median', 'hr_mean', 'hr_std', 'hr_min', 'hr_max', 'hr_range', 'hr_median')

Classification Model

- Random Forest
- SVM
- XGBoost

PHYMER DATASET - RESULTS

Random Forest	64.84%
SVM	75.23%
XGBoost	71.23%

FER-2013 DATASET

Key Features

- Data Type: Facial expression images (grayscale, 48×48 pixels).
- Emotions: 7 classes (Angry, Disgust, Fear, Happy, Sad, Surprise, Neutral).
- Size: 35,887 images (train: 28,709, test: 3,589, validation: 3,589).
- Collection Method: Crowdsourced via the Google search engine.

Why We Chose FER2013?

- Provides standardized facial expression data for emotion detection, complementing physiological signals (EEG/ECG).
- Enables multimodal fusion (e.g., combining facial features with EDA/HRV for robust classification).
- Widely used in research, allowing direct comparison with existing models.

FER-2013 DATASET

Class Distribution

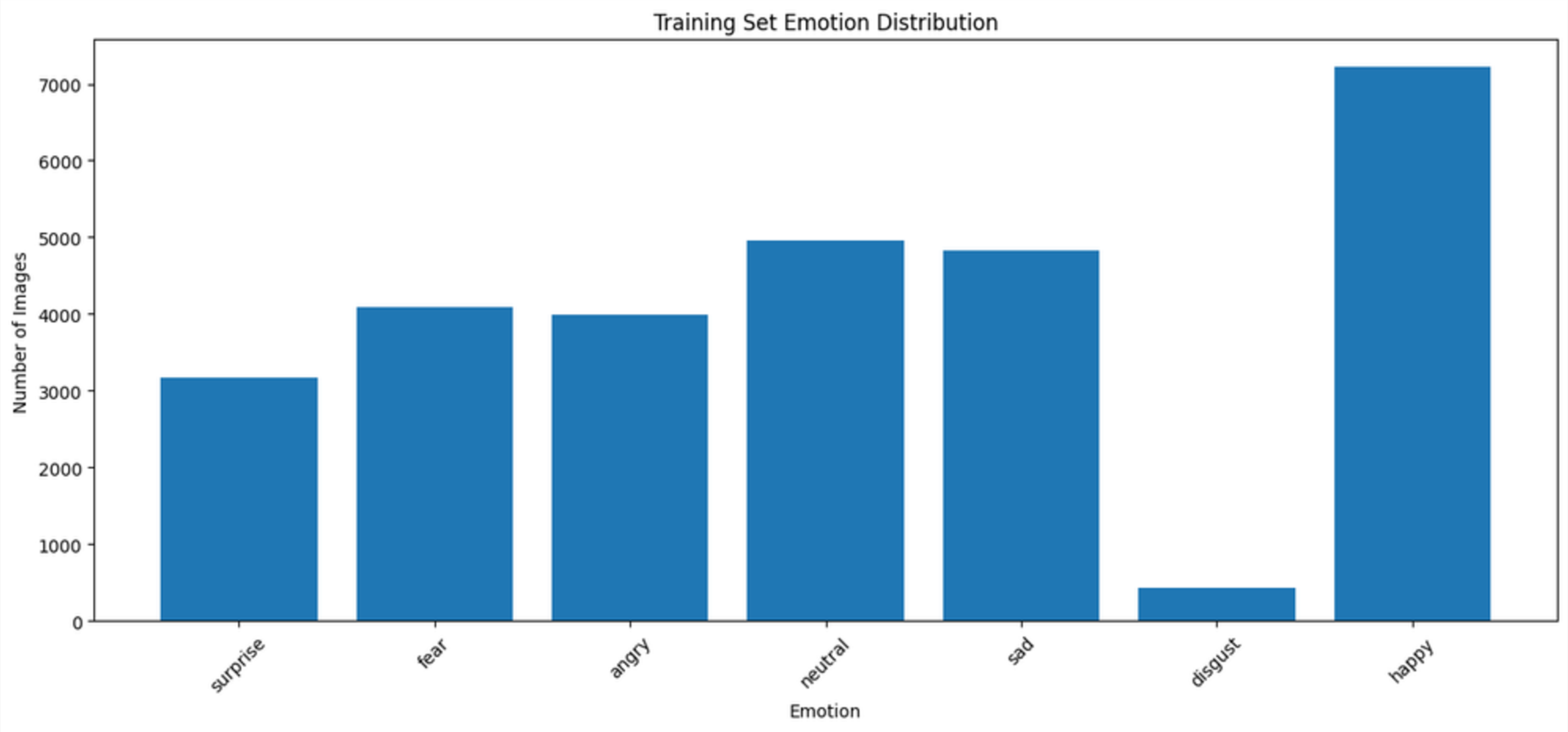


Fig-3

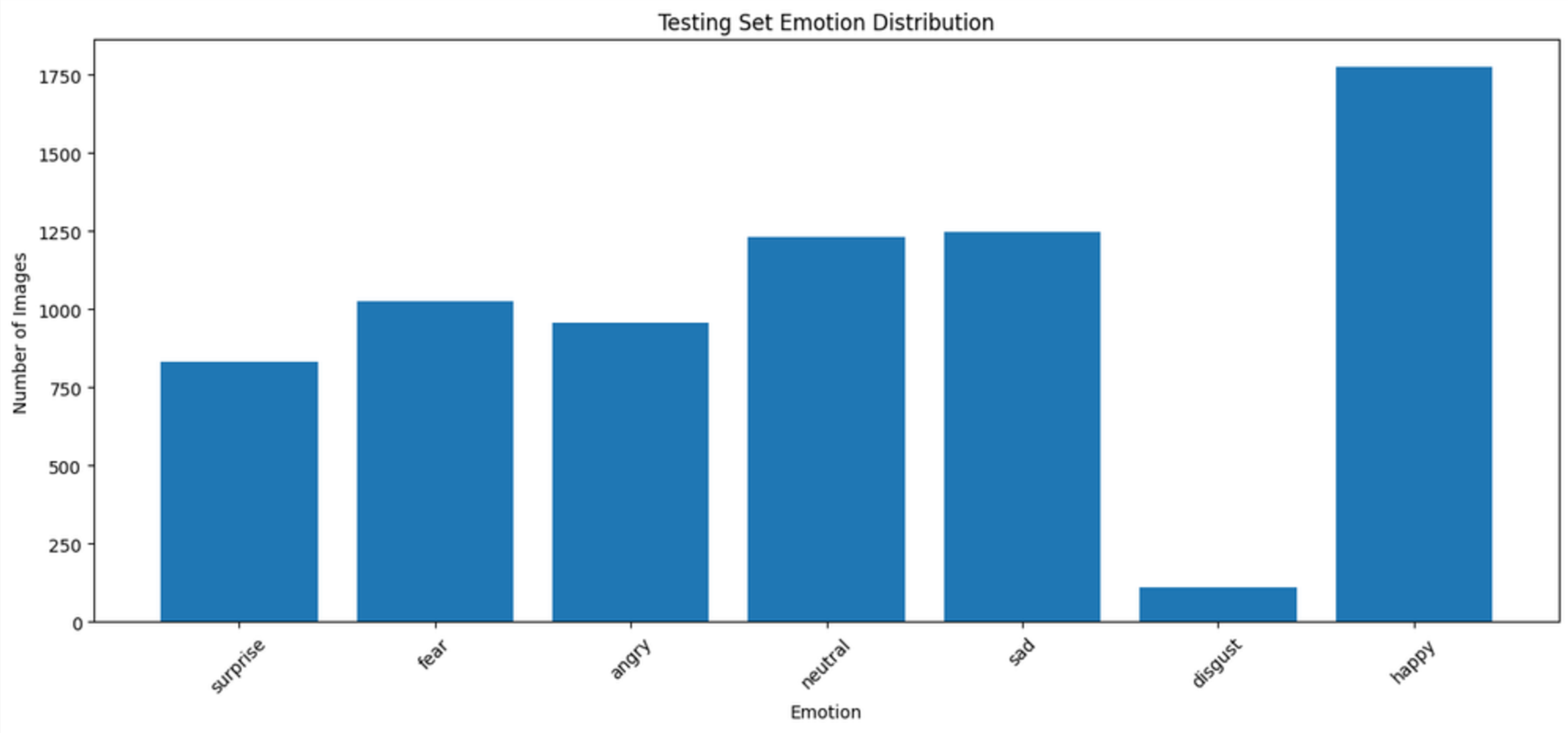


Fig-4

FER-2013 DATASET - MODEL

CNN Model Architecture

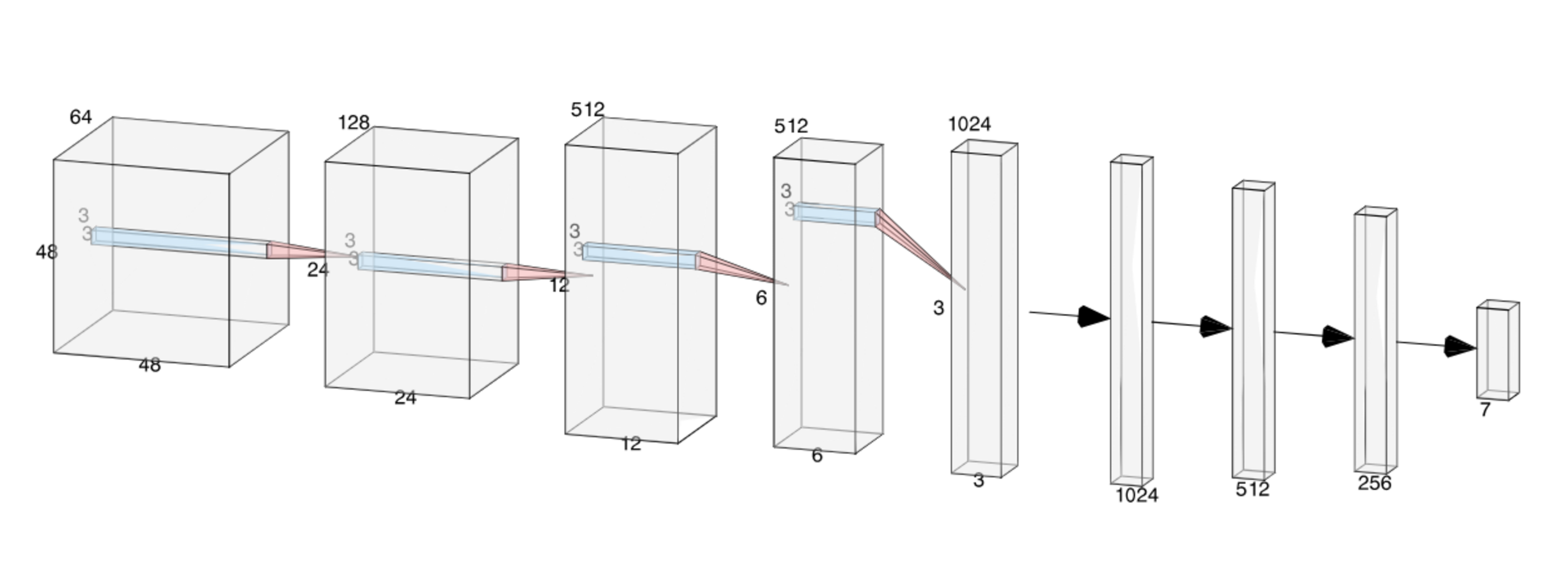
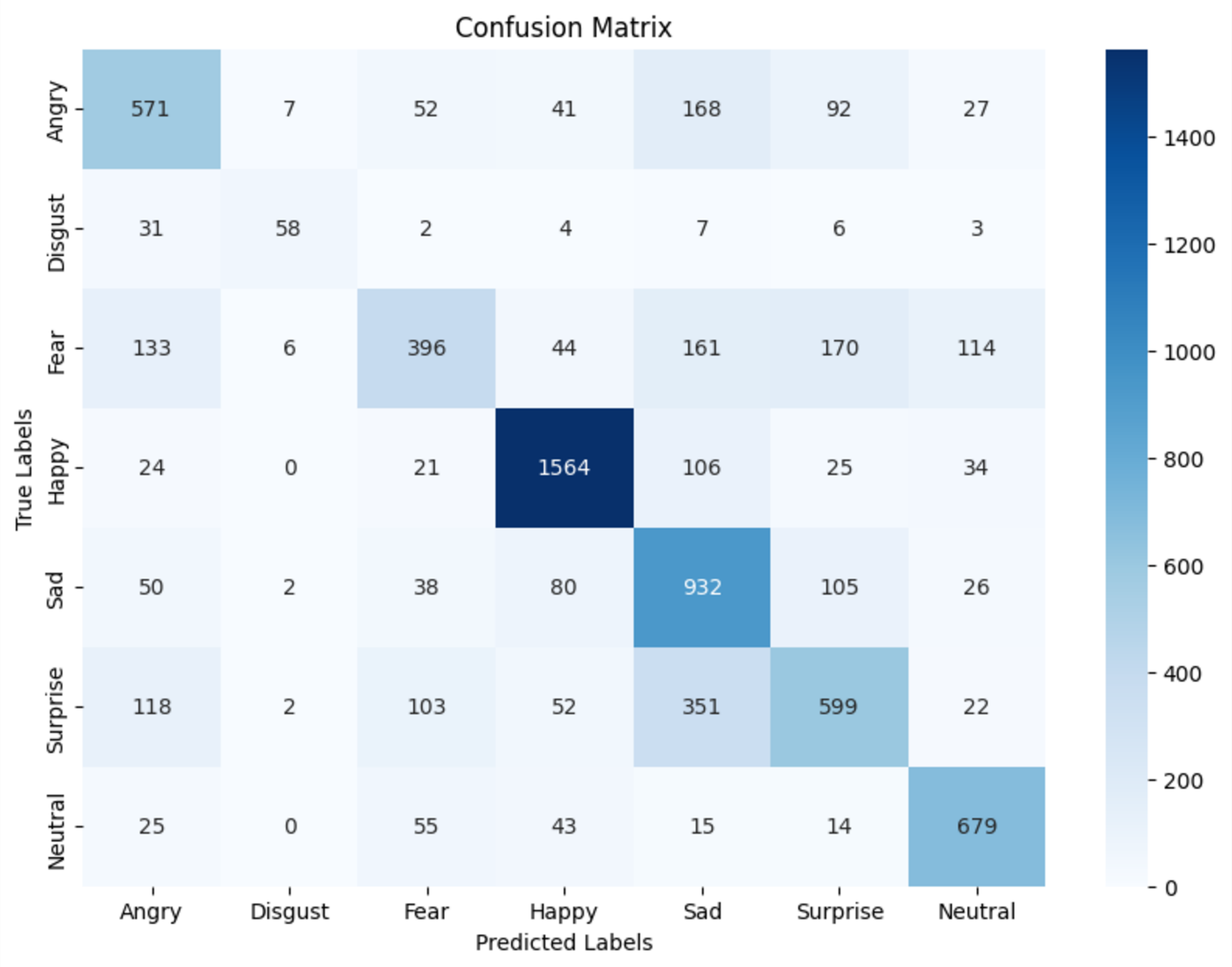


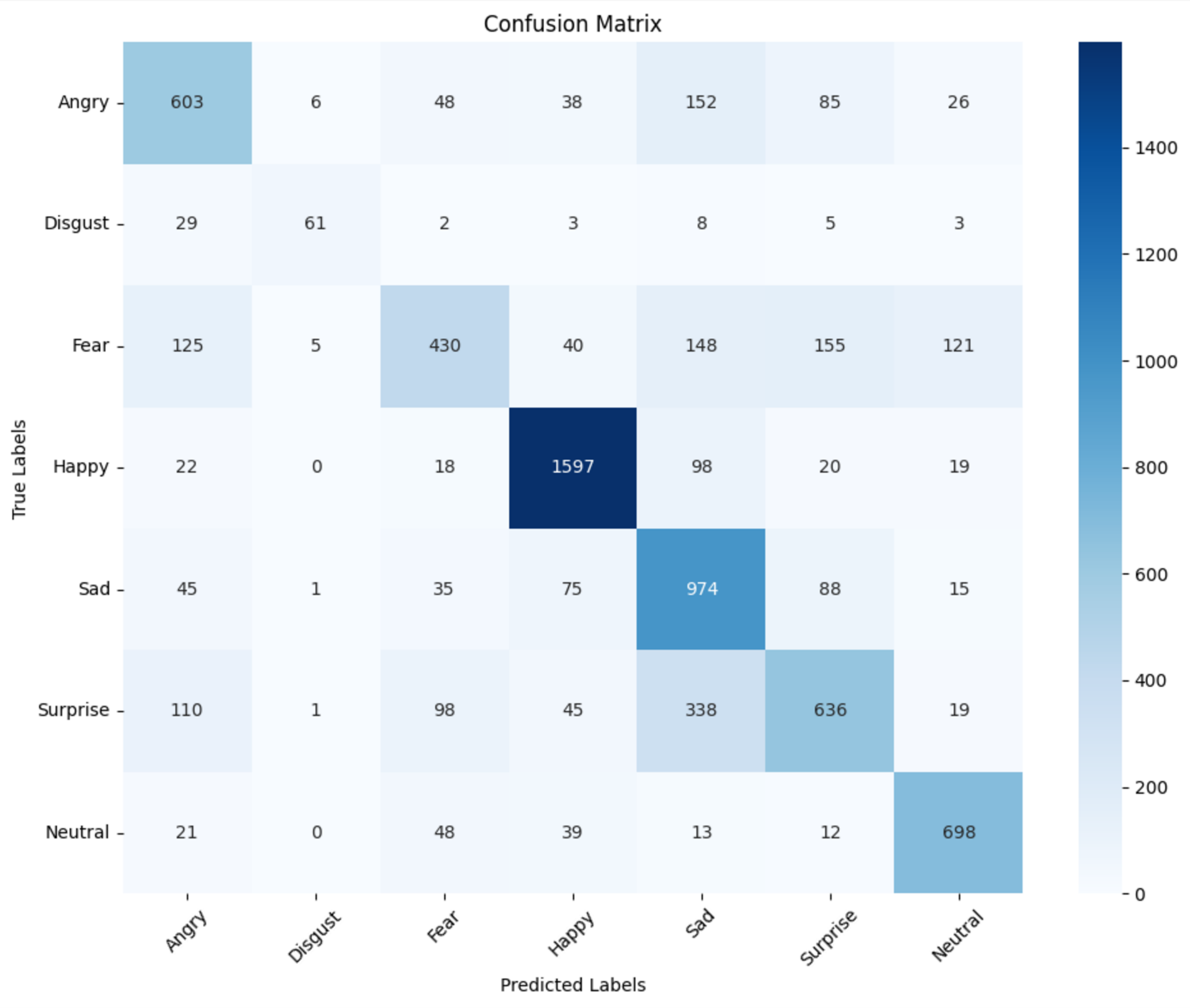
Fig-5 CNN Model Used for Classification (Source: alexlenail.me)

FER-2013 DATASET - RESULTS



Testing Accuracy : 67%

FER-2013 DATASET - RESULTS AFTER FEDAVG



Testing Accuracy : 71%

CREMA-D DATASET

Key Features

- Data Type: Audio-visual (videos + speech recordings).
- Emotions: 6 classes (Happy, Sad, Anger, Fear, Disgust, Neutral).
- Size: 7,442 clips from 91 actors (diverse age/gender/ethnicity).
- Collection Method: Actors performed scripted emotional scenarios.

Why We Chose CREMA-D?

- Captures vocal stress cues (e.g., pitch, intensity) linked to arousal levels.
- Supports temporal fusion with physiological signals (e.g., synchronizing BVP with speech patterns).
- Includes demographic diversity, reducing bias in emotion/stress models.

CREMA-D DATASET

Feature Extraction

- Mel-Frequency Cepstral Coefficients (MFCCs) from audio files
- MFCCs are based on the Mel scale
- MFCCs reduce complex audio signals into a small set of coefficients (e.g., 13 in this case)
- In speech emotion recognition, MFCCs capture variations in vocal timbre, pitch, and energy that correlate with emotional states (e.g., angry speech may have sharper spectral peaks)

CREMA-D DATASET - MFCC

How MFCCs are Computatied?

- **Pre-emphasis**: Boost high-frequency components of the audio to balance the spectrum.
- **Framing**: Split the audio into short overlapping frames (e.g., 20–40 ms) to analyze short-term changes.
- **Windowing**: Apply a window (e.g., Hamming) to each frame to reduce edge effects.
- **Fourier Transform**: Compute the Fast Fourier Transform (FFT) to get the power spectrum of each frame.
- **Mel Filterbank**: Apply a set of triangular filters spaced according to the Mel scale to emphasize perceptually relevant frequencies.
- **Log Energy**: Take the logarithm of the filterbank energies to mimic human loudness perception.
- **Discrete Cosine Transform (DCT)**: Apply DCT to the log energies to obtain cepstral coefficients, which decorrelate the features and focus on the most significant components.
- **Select Coefficients**: Keep the first few coefficients (e.g., 13) as the MFCCs, discarding higher ones that capture less relevant details.

CREMA-D DATASET - RESULTS

LSTM MODEL

- 2 LSTM, 2 Dense

Results

- Test Accuracy: 63.43%

Results after FedAvg

- Test Accuracy: 67.25%

APPLICATIONS

Mental Health Monitoring – Early stress detection

Human-Computer Interaction – Adaptive interfaces

Workplace Stress Management – Employee well-being

Gaming & VR – Real-time emotion-adaptive gaming

FUTURE SCOPE

- Expanding Multimodal Stress Detection
- Improving Signal Processing and Feature Extraction
- Deep Learning and Transformer-based Models
- Real Time Implementation using wearables
- Dataset Expansion and Generalization
- Application in Mental health Monitoring
- Scalable, privacy-aware stress detection by training models across distributed devices (e.g., Empatica E4) with FL frameworks.

CONCLUSION

- A detailed review of physiological signal properties, collection mechanisms, and preprocessing techniques.
- An analysis of 16 research papers summarizing existing methodologies.
- A feasibility study on budget wearables for real-time physiological tracking.
- Implementation of two multimodal datasets (DREAMER and Phymer) for emotion detection.
- Development of a machine learning pipeline for feature extraction and classification.

THANK YOU