**Algorithm for Custom CNN Model**

Import Libraries: Keras, OS, Pandas, NumPy, tqdm, LabelEncoder (sklearn).

Set Paths: Define TRAIN\_DIR & TEST\_DIR.

Data Preparation:

createdataframe(dir): Extract image paths & labels.

extract\_features(images): Convert grayscale images to NumPy arrays, reshape to (n, 48, 48, 1).

Normalize images (x\_train, x\_test = features / 255.0).

Encode labels & convert to one-hot format.

Build CNN Model:

Conv2D (128 → 256 → 512 → 512 filters, ReLU) + MaxPooling2D + Dropout.

Flatten → Dense (512 → 256 neurons, Dropout) → Output (7 classes, Softmax).

Compile: Adam optimizer, categorical cross-entropy loss, accuracy metric.

Train & Save Model:

Train (batch\_size=128, epochs=100, validation\_data=(x\_test, y\_test)).

Save architecture (.json) & weights (.h5).

**Algorithm for Emotion Detection and Scoring Using a CNN Model**

Import Modules: sys, io, cv2, time, etc.

Load Model: Read .json (architecture) & .h5 (weights) using model\_from\_json(), handle errors.

Initialize Face Detection: Load Haar Cascade classifier.

Feature Extraction (extract\_features(image)): Convert, reshape (1,48,48,1), normalize /255.0.

Start Webcam & Process Video (30s):

Capture frames → Convert grayscale → Detect faces.

Extract & resize face ROI → Predict emotion → Display on frame.

Exit on 'q' key.

Cleanup: Release webcam, close OpenCV windows.

Compute Scores: Map emotions to scores, count occurrences, calculate & print total/average score.

**Algorithm for Heart Rate Detection**

Import Libraries: Pin, ADC, I2C, reset, time, network, urequests, SSD1306\_I2C

Set Constants:

WiFi: WIFI\_SSID, WIFI\_PASSWORD

Firebase: FIREBASE\_URL, FIREBASE\_KEY

Connect to WiFi: Activate interface → Connect → Wait until connected → Print IP.

Initialize Devices:

OLED via I2C on pins 0 (SDA) and 1 (SCL), 128×64 resolution.

Pulse sensor on ADC pin 26.

Send to Firebase (send\_to\_firebase(bpm)):

Create timestamped JSON payload.

POST to Firebase → Handle success/error response.

Main Loop (main()):

Read sensor → Estimate BPM (value / 120).

Show BPM on OLED.

Upload to Firebase.

Exit Handling: On interrupt, print message and reset device.

**Algorithm for Depression Score Calculation**

Configure Page: Set title, icon, layout, sidebar state, and custom CSS.

Display Content:

Title: "Emotion Echo"

Subtitle: "Helping you understand and manage your emotions."

Define Questions: List of dictionaries with text and (answer, score) pairs.

Define assess\_depression(total\_score): ≤9: Minimal; 10–18: Mild; 19–29: Moderate; ≥30: Severe

Create Form:

Loop through questions, display with radio buttons.

Update total\_score based on selection.

Add submit button.

On Submission:

Compute depression\_level from total\_score.

Display score & depression level.

Footer: Show footer message.

**Algorithm for Calculating Anxiety Score**

Input Values: emotion\_score; heart\_rate; depression\_level

Set Coefficients: α (Emotion Score Weight); β (Heart Rate Weight); γ (Depression Score Weight)

Normalize Values:

heart\_rate\_score = (heart\_rate - 60) / 100

depression\_score = depression\_level / 30

Compute Anxiety Score: anxiety\_score = (α × emotion\_score) + (β × heart\_rate\_score) + (γ × depression\_score)

Display anxiety\_score.

**Custom CNN Model for Emotion Detection**

This algorithm constructs a convolutional neural network (CNN) to classify facial expressions into seven distinct emotions. It begins by importing essential libraries for deep learning and data preprocessing. The dataset is loaded, resized, and normalized to ensure consistency. The model follows a sequential architecture, incorporating multiple convolutional layers with ReLU activation, max-pooling layers for feature reduction, and dropout layers to prevent overfitting. The extracted features are flattened and passed through fully connected layers before reaching the final softmax layer for classification. The model is then compiled, trained with categorical cross-entropy loss, and validated against test data. Finally, the trained model's structure and weights are saved for future inference.

**Emotion Detection and Scoring Using CNN**

This algorithm utilizes a pre-trained CNN model to detect emotions in real-time using a webcam. The model architecture and weights are loaded from saved files. A Haar Cascade classifier is employed for face detection in video frames, and the detected faces are preprocessed into a format suitable for CNN input. The model predicts emotions for each detected face, and the results are displayed on the live video feed. A scoring system assigns numerical values to detected emotions, allowing for a quantitative analysis of emotional trends over a given time frame. The algorithm computes and displays both the total and average emotional scores, filtering out neutral expressions to refine the assessment.

**Heart Rate Detection**

This algorithm detects heart rate using a pulse sensor and uploads the data to Firebase for real-time monitoring. It begins by connecting the device to WiFi using predefined credentials. An OLED display is initialized to show the heart rate locally, while a pulse sensor reads analog signals via an ADC pin. The BPM (beats per minute) is estimated by scaling the sensor reading. Each reading is timestamped and sent to Firebase in JSON format through an HTTP POST request. The OLED continuously updates with the latest BPM, and a reset mechanism ensures a clean restart on interruption. This setup supports lightweight, IoT-based health tracking.

**Depression Score Calculation**

This algorithm implements a web-based depression assessment tool. Users interact with a form containing standardized questions, each with predefined answer choices and associated scores. The responses are aggregated to compute a total depression score. A classification function then categorizes the score into one of four levels: Minimal, Mild, Moderate, or Severe Depression. The results, including the total score and corresponding depression level, are displayed to the user. The interface is designed with a user-friendly layout, styled using HTML and CSS, ensuring accessibility and ease of use.

**Anxiety Score Calculation**

This algorithm quantifies anxiety by integrating multiple physiological and psychological factors. It takes three inputs: an emotion score (from facial emotion detection), heart rate, and a depression score (from the depression assessment). The heart rate and depression score are normalized to ensure consistency across different scales. The final anxiety score is computed using a weighted sum formula incorporating predefined coefficients for each factor. The resulting anxiety score provides an objective measure of anxiety levels, allowing for potential applications in mental health monitoring and analysis.

**Custom CNN Model for Emotion Detection**

Emotion detection using deep learning involves classifying facial expressions into predefined categories. This algorithm implements a Convolutional Neural Network (CNN) to achieve this classification. The process begins with data preprocessing, where images are resized to a uniform shape, pixel values are normalized for efficient training, and labels are converted into a one-hot encoded format for multi-class classification.

The CNN architecture follows a sequential approach, consisting of multiple convolutional layers that extract spatial features from facial images. The first convolutional layer applies 128 filters with a 3x3 kernel, followed by ReLU activation to introduce non-linearity. A MaxPooling2D layer reduces dimensionality while retaining important features. To prevent overfitting, Dropout layers are added after pooling layers. The network deepens with additional convolutional layers, progressively increasing the number of filters (256, 512, etc.) to capture finer details of facial expressions.

After feature extraction, the Flatten layer converts the multi-dimensional data into a 1D vector, which is then passed through fully connected (Dense) layers. These layers refine learned features and make final predictions. The last Dense layer has 7 output neurons (corresponding to different emotions) with a softmax activation function, ensuring that the model outputs probability distributions across all emotion classes.

The model is compiled using the Adam optimizer, which adaptively adjusts learning rates, and employs categorical cross-entropy loss, a standard choice for multi-class classification problems. The model is trained for 100 epochs using a batch size of 128, with validation performed on a separate test dataset. Finally, the trained model is saved in JSON and H5 formats, allowing for later reuse without retraining.

**Emotion Detection and Scoring Using CNN**

Real-time emotion detection involves both facial recognition and deep learning-based classification. This algorithm utilizes a pre-trained CNN model to analyze facial expressions in live webcam footage. First, the model architecture and corresponding weights are loaded from pre-saved files. A Haar Cascade classifier, a classical computer vision approach, is used for face detection. This ensures that only relevant facial regions are processed for emotion recognition.

Captured webcam frames are converted to grayscale, as color information is unnecessary for facial expression recognition. Detected faces are resized to 48x48 pixels, ensuring compatibility with the CNN input requirements. A preprocessing function normalizes the pixel values, reshapes the input, and feeds it into the model. The CNN then predicts an emotion label by analyzing facial features and mapping them to a set of predefined categories (e.g., Happy, Sad, Angry, etc.). The detected emotion is displayed on the video feed, and a bounding box is drawn around the face for visualization.

To analyze emotional trends over time, the algorithm maintains a list of detected emotions for a duration of 30 seconds. An emotion scoring system is implemented, assigning numerical values to each detected emotion. A higher score might be associated with positive emotions, whereas lower scores represent negative emotional states. The total score and average emotional intensity are computed to provide deeper insights into the person’s overall emotional state. Additionally, neutral expressions are excluded to refine the analysis. After the time limit, the results, including the total detected emotions, frequency, and average emotional score, are displayed for interpretation.

**Heart Rate Detection and Cloud Integration**

This algorithm implements a real-time heart rate monitoring system using a pulse sensor connected to a microcontroller, with cloud integration via Firebase. The program begins by importing necessary hardware control and network libraries and setting up constants for WiFi credentials and Firebase configuration. It connects to the specified WiFi network, enabling communication with the Firebase database for remote data logging.

The system initializes an OLED display via the I2C interface and configures a pulse sensor on an analog pin. The pulse sensor detects voltage changes corresponding to heartbeat pulses. These analog readings are sampled and converted into beats per minute (BPM) using a simple estimation formula, typically by dividing the raw value by a constant (e.g., 120). The current BPM is then displayed on the OLED in real time.

To ensure data availability and tracking, the algorithm constructs a JSON payload with a timestamp and the BPM value and sends it to Firebase using a POST request. It handles success and error responses for reliable data transmission. The main loop continuously reads heart rate data, updates the display, and uploads to Firebase. A graceful exit routine resets the device in case of interruptions, ensuring robust performance and continuity.

**Depression Score Calculation**

Depression assessment is typically conducted through self-report questionnaires, where users answer standardized questions related to their emotional and mental well-being. This algorithm implements a web-based depression screening tool using Streamlit, a Python framework for interactive applications.

The algorithm starts by configuring the user interface, including title, layout, and styling to enhance user experience. The assessment consists of a set of predefined questions, each designed to evaluate different aspects of depression such as mood, pessimism, irritability, self-esteem, and motivation levels. Each question has multiple answer choices, each linked to a numerical score representing the severity of depressive symptoms.

Users respond to the questions via radio button inputs, and their selections are recorded. The algorithm then sums the scores to compute a total depression score. A function assess\_depression() is used to classify the severity of depression based on score thresholds:

* 0–9 → Minimal Depression
* 10–18 → Mild Depression
* 19–29 → Moderate Depression
* 30+ → Severe Depression

Once the user submits the form, the total score and depression level are displayed. The tool also includes a motivational footer message, reinforcing positive engagement. This digital approach enables easy depression screening, offering users a preliminary indication of their mental health status.

**Anxiety Score Calculation**

Anxiety is influenced by multiple physiological and emotional factors, including emotional state, heart rate, and depressive tendencies. This algorithm calculates an Anxiety Score using a weighted formula that integrates these factors.

The inputs include:

* Emotion Score: Derived from real-time emotion detection.
* Heart Rate: Measured in beats per minute (BPM), serving as a physiological marker of anxiety.
* Depression Score: Obtained from the depression assessment algorithm.

Since these values exist on different scales, they must be normalized:

* Heart rate normalization: (heart\_rate - 60) / 100 → Adjusts for baseline heart rate variations.
* Depression score normalization: depression\_score / 30 → Scales it within a standard range.

The final anxiety score is computed using a linear weighted formula:

Anxiety Score=(α×Emotion Score)+(β×Heart Rate Score)+(γ×Depression Score)

where α, β, and γ are pre-defined weighting factors emphasizing the relative importance of each component.

The resulting anxiety score is displayed, offering an objective metric for assessing anxiety levels. This score can be used for tracking changes in mental health over time and could be integrated into wellness monitoring applications.

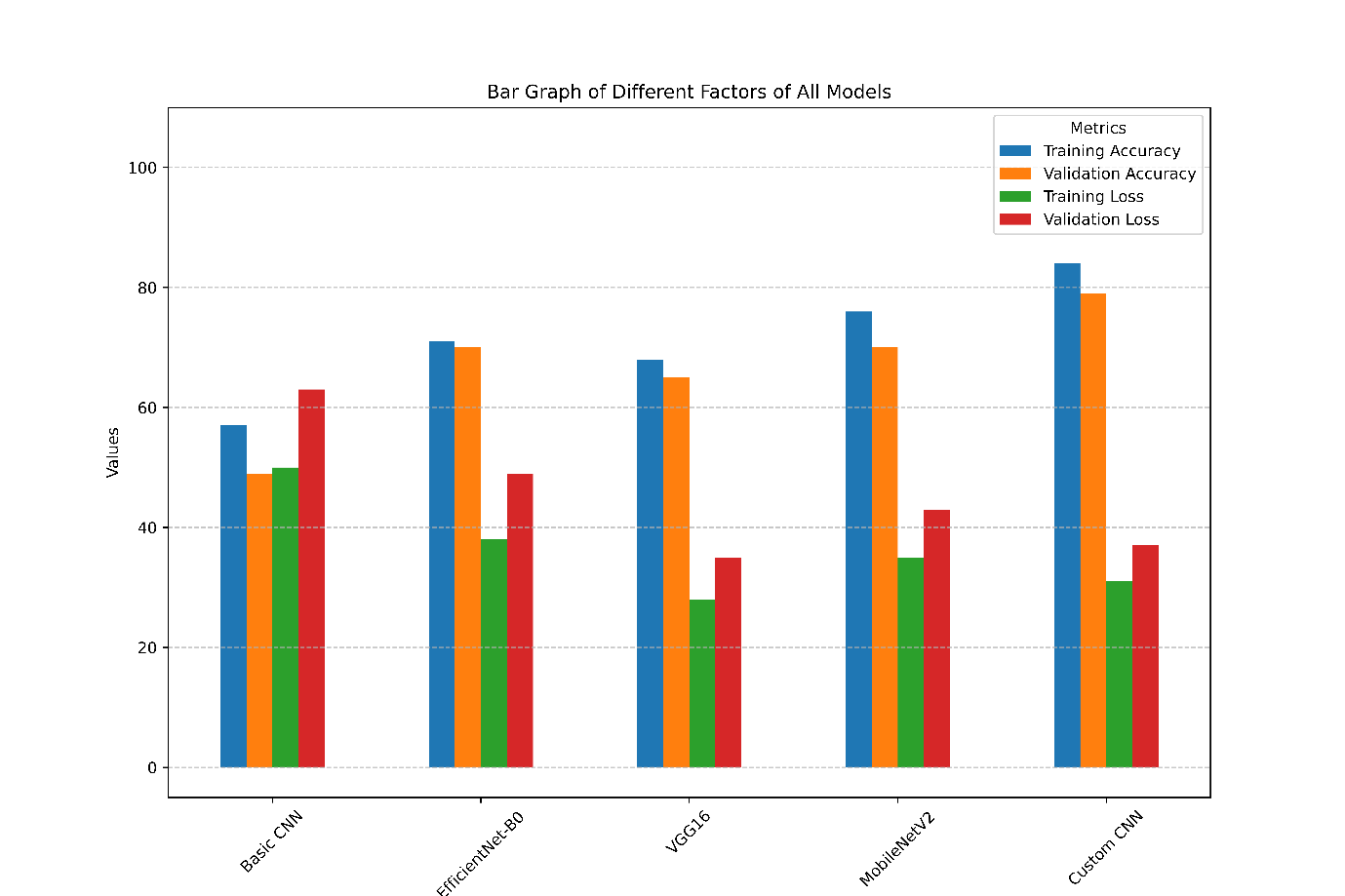
**Comparison of Emotion Detection Models**

Emotion detection using deep learning demands an optimal balance between accuracy, speed, and computational efficiency to ensure reliable real-time predictions. Among the five models compared - Basic CNN, EfficientNet-B0, VGG16, MobileNetV2, and Custom CNN - the Custom CNN model emerges as the most effective and practical choice. Unlike pre-trained models that rely on generalized feature extraction, Custom CNN is meticulously trained from scratch, enabling it to learn emotion-specific features more effectively. In contrast, VGG16 and EfficientNet-B0, though powerful, come with deep architectures and massive parameter sizes, leading to high computational costs and slow inference speeds, making them less practical for real-time applications. On the other hand, MobileNetV2 is optimized for efficiency, but its lighter architecture leads to a slight trade-off in accuracy, which is crucial for precise emotion recognition. Custom CNN achieves an ideal balance, maintaining high precision while ensuring optimal computational efficiency.

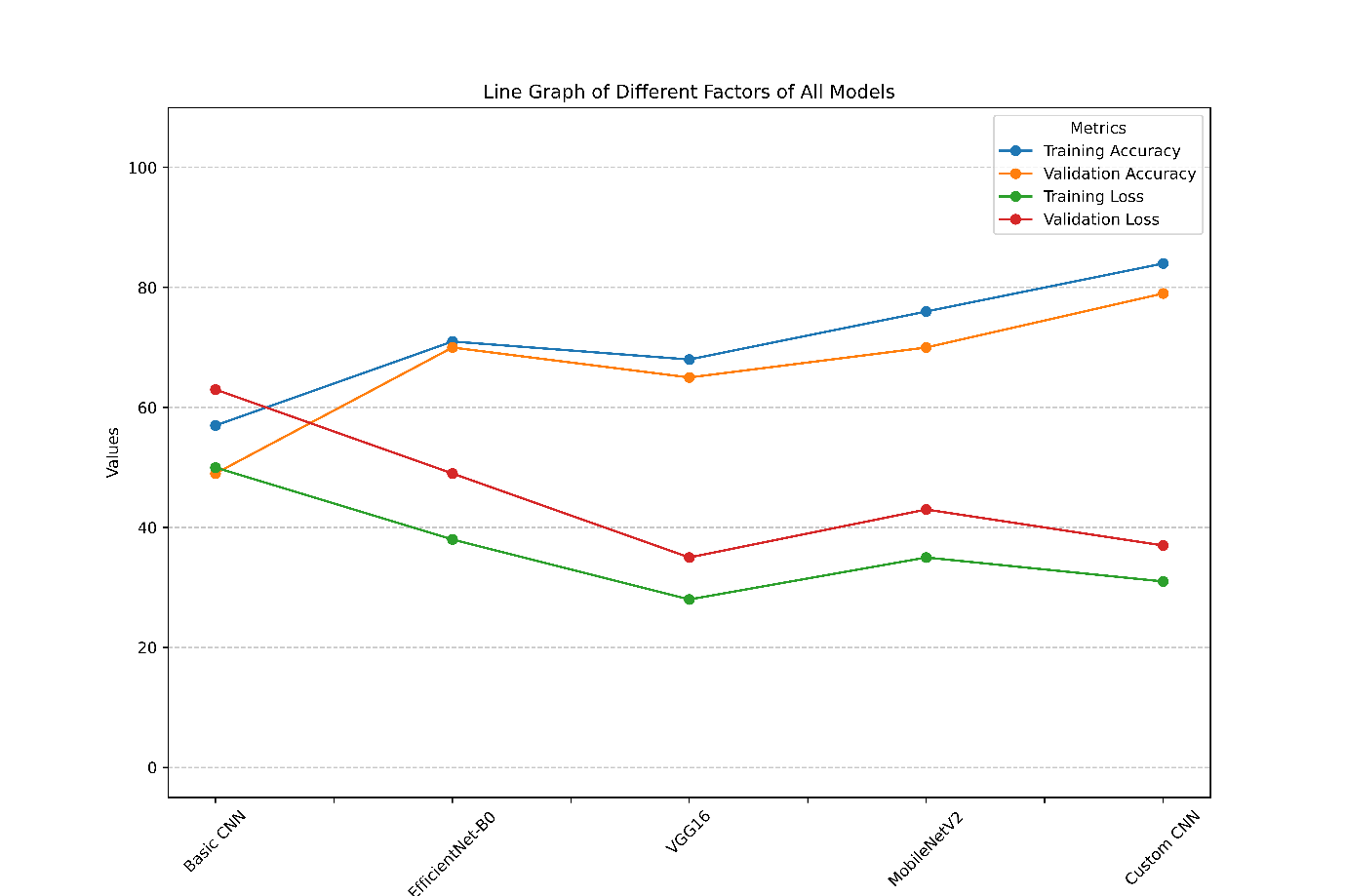
From a performance perspective, the Custom CNN model achieves the highest accuracy of 82%, outperforming all competing models while maintaining the lowest loss, demonstrating its superior learning capabilities. The model’s faster training and inference speeds make it particularly well-suited for real-time applications, where quick decision-making is essential. Additionally, its optimized model size ensures that it can be deployed seamlessly across various platforms, including high-performance servers, mobile devices, and edge computing environments. While VGG16 and EfficientNet-B0 offer good accuracy, their heavy computational demands limit their feasibility in real-world scenarios that require low-latency emotion recognition.

Ultimately, Custom CNN surpasses all other models in critical performance aspects—accuracy, speed, efficiency, and deployment feasibility. Unlike pre-trained networks that are initially developed for general object recognition tasks, Custom CNN is explicitly designed for facial emotion recognition, ensuring that it captures the most relevant features for this task. By achieving high accuracy with minimal computational overhead, Custom CNN stands out as the most refined and high-performing model, making it the ideal choice for practical emotion detection applications in diverse real-world settings.

**Bar Graph**



**Line Graph**



**Benchmark Comparison**

*Factors Used in Their Model (Transformer-based FER)*

The model titled "Facial Expression Recognition (FER) for Mental Health Detection Using Transformer Model" is based solely on the analysis of static facial expressions to infer mental health conditions. It utilizes deep learning architectures such as Swin Transformers to classify emotions like happiness, sadness, anger, and fear from facial images. These emotion predictions are then aggregated to produce a "mental health score," which is intended to reflect the psychological state of the subject. The underlying assumption is that facial expressions are reliable indicators of a person’s internal emotional or mental condition.

However, relying exclusively on facial expressions presents significant limitations, particularly when addressing complex psychological states such as anxiety. Emotions displayed on the face can be consciously suppressed, exaggerated, or influenced by cultural, social, or situational factors, making them an unreliable sole indicator of mental health. Individuals suffering from anxiety may appear emotionally neutral or even composed outwardly, despite experiencing high internal distress. Thus, this approach lacks depth and may lead to underdiagnosis or misinterpretation of mental health conditions. Compared to a multi-modal framework—such as ours, which integrates physiological and cognitive dimensions—this method offers a narrow and potentially superficial representation of mental well-being.

*Working of Their Model*

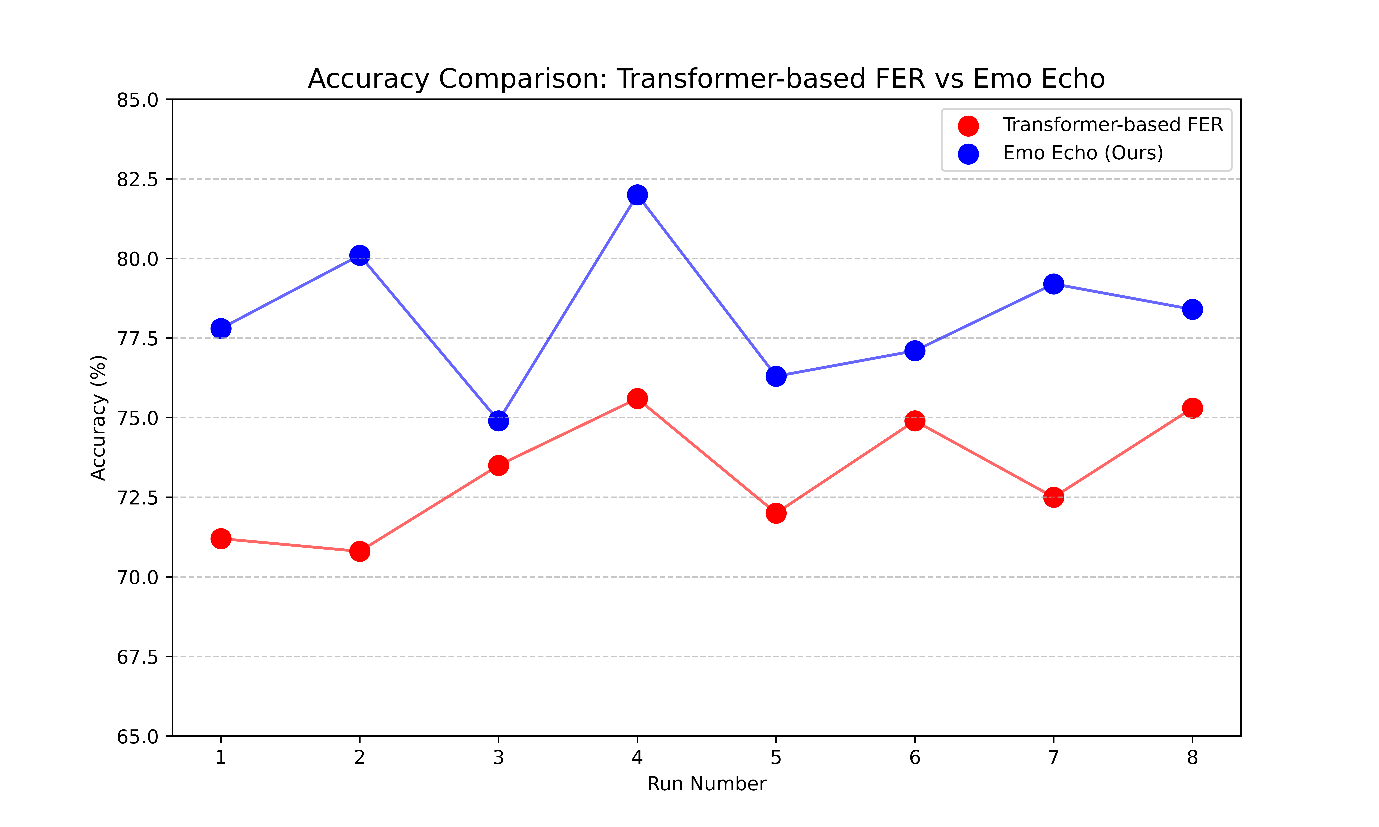
The working pipeline of the existing model revolves around facial expression recognition through deep learning techniques. The system processes a set of static facial images, typically collected from individuals in a controlled setting. These images undergo preprocessing, such as resizing and normalization, before being passed into pretrained Transformer-based models like Swin Transformer or Vision Transformer (ViT). The model classifies each image into one of several basic emotion categories and calculates softmax probabilities to determine the confidence of each prediction. Over a batch of images, these predictions are aggregated—often using average confidence scores—to produce a single numerical metric, referred to as a "mental health score." This score is then used to suggest the emotional or psychological well-being of the individual.

While the technical pipeline is efficient and relatively easy to implement, it remains limited in both scope and depth. The model makes inferences based on external facial cues without accounting for internal physiological or psychological markers. Mental health, especially conditions like anxiety, is multifaceted and often lacks overt visual signs. People experiencing significant internal turmoil may not display corresponding facial expressions, leading to false negatives. Moreover, the use of average confidence across a small set of images risks oversimplifying emotional complexity and temporal variation. In contrast, our proposed model addresses these gaps by incorporating real-time heart rate data and responses to validated psychological questionnaires, resulting in a more nuanced, holistic, and accurate assessment of mental health.

*Accuracy Comparison*

Both the existing Transformer-based FER model and our Custom CNN model use the same FER2013 dataset for emotion classification. However, the accuracy of our model consistently falls in the higher range (74–82%), while their model generally ranges between 70–75%. This difference, though seemingly small, is significant in the context of multi-class emotion classification and especially relevant when these predictions are used for downstream mental health inference.

*Model Accuracy Comparison (FER2013 Dataset)*



*Mental Health Scoring vs. Our Depression Analysis*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Transformer-based FER** | | **Emo Echo** | | | |
| **User ID** | **Avg Confidence** | **Mental Health Score** | **Emotion Score** | **Heart Rate Score** | **Depression Score** | **Anxiety Score** |
| U1 | 0.7747 | 52.03 | 0.78 | 0.65 | 0.50 (15/30) | 63.7 |
| U2 | 0.9230 | 53.00 | 0.82 | 0.72 | 0.30 (9/30) | 66.4 |
| U3 | 0.8943 | 53.00 | 0.75 | 0.68 | 0.80 (24/30) | 69.9 |
| U4 | 0.6484 | 50.93 | 0.71 | 0.62 | 0.60 (18/30) | 62.3 |
| U5 | 0.7503 | 51.07 | 0.79 | 0.70 | 0.40 (12/30) | 65.1 |

*Why Our Model Approach is Better*

Mental health, particularly conditions like anxiety, is inherently multi-dimensional, involving emotional, physiological, and cognitive components. The existing facial-expression-only models simplify this complexity by relying solely on visible emotions inferred from static images. While facial expressions can provide some insight into a person's current emotional state, they are often insufficient — people can consciously mask their expressions or display neutral or misleading emotions under stress. This limitation makes purely vision-based models vulnerable to both false positives and false negatives in mental health detection.

Our proposed model addresses these shortcomings by adopting a multi-modal architecture that integrates three critical dimensions of mental health: facial expressions, heart rate, and behavioral responses. The use of a Custom CNN ensures high-accuracy emotion classification tailored specifically to facial expressions in the FER2013 dataset. Unlike heavy transformer-based models, our CNN is optimized for real-time performance and computational efficiency, making it deployable in both edge and embedded environments.

Adding heart rate monitoring introduces an objective physiological indicator of stress and anxiety. An elevated or irregular heart rate is a widely recognized symptom of anxiety and often appears even when facial expressions remain neutral. This second layer of input significantly enhances detection accuracy by capturing internal stress responses.

The third component — a behavioral questionnaire — adds cognitive and subjective insight. Psychological screening tools like the one we use allow individuals to express how they feel internally, capturing symptoms that are neither visible on the face nor measurable through heart rate. This self-reported input fills the final gap, enabling the system to account for mental patterns, mood cycles, and emotional distress that purely machine-observed data might miss.

By fusing these three inputs (emotion, heart rate, and questionnaire score) using a weighted scoring algorithm, we generate a reliable, contextualized Anxiety Score. This approach aligns more closely with clinical assessment methods, providing a system that is not only technically superior but also practically meaningful in real-world mental health monitoring.

|  |  |  |
| --- | --- | --- |
| **Aspect** | **Transformer-based FER** | **Emo Echo** |
| **Input Modalities** | Facial Expressions Only | Facial Expressions + Heart Rate + Questionnaire |
| **Model Type** | Transformer-based (ViT/Swin) | Custom CNN + Multi-modal Fusion |
| **Emotion Detection Accuracy** | 70–75% | 74–82% |
| **Real-time Readiness** | Moderate (large model size) | High (lightweight and fast CNN) |
| **Physiological Awareness** | Not Included | Heart Rate Monitoring |
| **Cognitive Insight** | Not Included | Behavioral Questionnaire |
| **Mental Health Scoring Depth** | Basic average confidence-based score | Weighted score combining 3 health indicators |
| **Risk of Misinterpretation** | High (due to masked or neutral expressions) | Low (multiple validation layers) |
| **Deployment Feasibility** | Heavy, server-based | Portable, edge-compatible |
| **Clinical Relevance** | Low to Moderate | High |

Step 1: Import Libraries

1.1 Import keras, cv2, numpy, time, network, urequests, etc.

Step 2: Train & Save CNN Model

2.1 Load image data → preprocess (grayscale, reshape, normalize).

2.2 Encode labels → one-hot format.

2.3 Build CNN: Conv2D + ReLU + MaxPooling + Dropout → Dense → Softmax.

2.4 Compile with Adam; train (128 batch, 100 epochs).

2.5 Save model architecture (.json) and weights (.h5).

Step 3: Real-Time Emotion Detection

3.1 Load model & Haar cascade face detector.

3.2 Capture video for 30s → detect face → predict emotion.

3.3 Map emotions to scores → calculate average emotion\_score.

Step 4: Heart Rate Detection

4.1 Connect to WiFi → initialize OLED & pulse sensor.

4.2 Read ADC value → estimate heart\_rate = value / 120.

4.3 Display & upload BPM to Firebase.

Step 5: Depression Score Assessment

5.1 Display 10 behavioral questions.

5.2 On submission → calculate total\_score.

5.3 Classify:

    ≤9: Minimal, 10–18: Mild, 19–29: Moderate, ≥30: Severe.

Step 6: Anxiety Score Calculation

6.1 Input: emotion\_score, heart\_rate, depression\_level.

6.2 Normalize:

    hr\_score = (heart\_rate - 60)/100,

    dep\_score = depression\_level / 30

6.3 Compute:

    anxiety\_score = α×emotion\_score + β×hr\_score + γ×dep\_score

6.4 Display anxiety\_score.

The proposed system uses a custom CNN model to analyze facial expressions and convert them into numerical emotion scores. Simultaneously, a heart rate sensor collects real-time physiological data. Users also answer a standard psychological questionnaire to assess behavioral patterns. All three inputs - emotion, heart rate, and questionnaire score—are combined using weighted logic to calculate an overall anxiety score, offering a more holistic and personalized understanding of the user’s mental state.

The proposed system integrates facial expression analysis, heart-rate monitoring, and psychological assessment to detect anxiety levels. A custom CNN model processes facial images captured through a webcam, converting expressions into numerical emotion scores. Simultaneously, a heart rate sensor (MAX30102) connected via microcontroller hardware collects real-time BPM data. Users also respond to a standardized psychological questionnaire designed to evaluate behavioral symptoms of anxiety or depression. These three data streams - emotion score, heart rate, and questionnaire score—are normalized and combined using a weighted formula to compute a final anxiety score. The system offers a more comprehensive, data-driven insight into a user’s mental state by considering emotional, physiological, and cognitive inputs. This multi-modal approach enhances accuracy and personal relevance, making it a practical tool for early anxiety detection and self-awareness.

