ABSTRACT

INTRODUCTION

Depression and anxiety have become common health problems in today’s society, where the pressure of education, work, and society cannot be controlled. Mental health disorders now affect millions of people around the world. Of these, anxiety disorders are some of the most common, yet they often go undiagnosed and untreated due to factors such as stigma, ignorance, or lack of access to help. Over time, chronic stress can lead to a variety of problems, including mood disorders, productivity problems, physical problems, and even life-threatening problems like suicide. Given the high prevalence of these mental health disorders, there is an urgent need to develop new and effective strategies to diagnose and treat depression at an early stage. The Emotion Echo uses advanced emotion analysis to provide real-time feedback. Traditional methods for assessing mood rely on questionnaires or clinical assessments, which do not capture mood states as they occur or provide timely interventions. Furthermore, although many machine learning systems can handle stress, they struggle with issues such as inconsistency between classifiers and misinterpretation of human behaviour. Many also rely on automated systems, which are not always available, or only analyze data from a single source, which is real-time. Emotion Echo seeks to bridge these gaps by providing a safe, multi-sensory way to manage stress and depression. By combining various types of data—such as facial expressions, heart rate, and gestures—the platform provides an accurate representation of a user’s behaviour. Building on existing systems, Emotion Echo is committed to providing a comprehensive, accessible and user-friendly solution for mental health, empowering people to take ownership and control a good life for everyone.

RELATED WORK

### Qualitative Interview Methods

Prediction and prediction of anxiety, depression, and stress levels has been extensively explored using ML, AI, and various techniques and innovations in many fields. The first study investigates the association between depression, anxiety, and depression using the DASS 21 questionnaire, using ML algorithms such as Random Forest and Naïve Bayes, with Random t-Forest being used [1].

### Signal-based physiological methods

Heart rate variability (HRV) data collected using wearable bracelets based on Long Short-Term Memory (LSTM) networks have shown excellent accuracy, achieving an accuracy of 83% [ 2 ]. One study in young children used inertial measurement units (IMUs) in anxiety tasks to assess anxiety [ 3 ]. Another study on older adults included physiological signals (EDA, BVP) and frequency data, showing that Random Forest classifiers performed the best for accurate real-time anxiety prediction [5]. Biosignal data analysis combining EDA, temperature, and heat index showed promising results with ML algorithms such as SVM and Random Forests, reaching an accuracy of 98% [ 15 ]. The sensors used, including ECG and EEG, were used for classification with ML techniques such as SVM and Random Forests, achieving accuracies between 67% and 92% [ 14 ].

### Mobile and AI-based apps

Mobile apps such as “AnxoRelief” integrate conversational AI and standardized scales (PHQ-9, PHQ-15, GAD-7) to assess the severity of depression using ML algorithms such as SVM and an SVM model based on Random Forest 99.71% [4]. Many studies have focused on preprocessing techniques such as PCA and Synthetic Minority Oversampling to improve the performance of ML in classification [6].

### Social Media and Online Data Analysis

In vivo data have been analyzed for fatigue detection using ML models based on a hybrid physiological context, where Random Forest is combined with other algorithms for real-time detection of physiological processes [8]. Analyzing sentiment on social media posts using Bayesian ridge regression also yielded important insights, which showed a negative relationship between anxiety level and social connectedness [ 16 ].

### Human Studies

Focusing on individuals, the study in Bangladesh focused on the diagnosis of anxiety and depression through ML models, including CNNs, achieving 96.8% classification accuracy for depression [ 10 ]. For university students, Random Forests showed 78.9% accuracy in predicting anxiety [ 11 ]. For adolescents, Random Forest models showed a sensitivity of 91% when predicting anxiety and depression, accounting for the effects of sociodemographic factors similar to [our 3 study].

### Future directions and challenges

Numerous researchers have identified physiological parameters (e.g., ECG, EMG, EEG) for stress detection and developed multimodal systems for individual stress detection [ 17 ]. Finally, a systematic review of wearable technologies for mood monitoring included the combination of HRV, EEG, and EDA data together with ML algorithms to achieve 90% accuracy. However, the review revealed some challenges facing participants and highlighted the need for non-invasive, user-friendly measures for agriculture in general [ 18 ].

PROBLEMS

* Many studies are based on small or geographically limited datasets, which limit the generalizability of samples to the general population.
* Datasets of barriers often suffer from class imbalance, resulting in biased model predictions in favour of the dominant class.
* Although some research focuses on real-time problem detection, many ML models still suffer from latency and computational difficulties.
* Using physical and social media devices to detect problems raises privacy concerns that require strong security and ethical considerations.
* Several studies based on wearable devices have shown that users are indifferent to uncomfortable devices, lack of battery life, and the need for constant monitoring.
* Current diagnostic models for schizophrenia exclude physical deficiencies, making it difficult for individual differences in physical and cognitive deficits.
* Many studies rely on single sources of data (e.g., HRV or EEG alone), but combinations of multiple physiological markers can improve accuracy.
* Most studies document stress at a single point in time rather than examining change over time, which is important for effective mental health treatment.

ADDRESSING PROBLEMS

Emotion Echo addresses these challenges by combining vast amounts of data and continuously improving the model. To tackle the class imbalance, it uses data augmentation and weight reduction functions to make predictions. Real-time anxiety detection is based on a lightweight model that adapts to the real-time face and chest detection process. Privacy concerns are mitigated by encryption, anonymity, and user-controlled data sharing. To improve user compliance, the system allows for manual heart rate input, reducing reliance on wearable devices. Emotion Echo also ensures personalization by learning from user interactions and adjusting its recommendations over time. Unlike single-source studies, it integrates multiple physiological inputs—facial expressions, heart rate, and user feedback—to improve accuracy. Furthermore, a longitudinal tracking system allows users to track their anxiety trends over time, ensuring ongoing assessment and effective intervention strategies.

PROPOSED SYSTEM

The Emotional Echo system is an application designed to measure and manage emotional mood by combining deep learning, computer science, and psychological assessment techniques. The proposed system comprises four main components: Facial Emotion Detection, Emotion Scoring, Depression Assessment, and Anxiety Calculation. Using machine learning models and structured scoring techniques, the system can identify a person's emotional state in real time.

### 1. Facial Emotion Detection Using a Custom CNN Model

The first step in the system involves recognizing human emotions through facial expressions. A Custom Convolutional Neural Network (CNN) is used to classify facial expressions into seven emotion categories: Happy, Sad, Angry, Fearful, Surprised, Disgusted, and Neutral.

* The facial expression images are grayscaled, normalized, and rescaled to 48×48×1 size for optimal accuracy during training and estimation.
* The CNN model with four convolutional layers (128, 256, 512, 512), max-pooling and dropout layers extracts features and divides emotions into seven categories with the help of softmax output layer.
* The trained model is integrated with a webcam-based system, where the Haar Cascade Classifier detects faces, preprocesses them, and predicts emotions in real time.

**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).

### 2. Emotion Detection and Scoring Using CNN

This module detects and quantifies emotions from live webcam feeds using a built-in CNN model.

* Each detected emotion is scored based on its intensity and psychological impact.
* The system records and aggregates emotion scores every 30 seconds to track mood trends and trends.
* Emotional scoring can help identify patterns of emotional distress and help describe levels of dissatisfaction in the system.

**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.

### 3. Depression Score Calculation

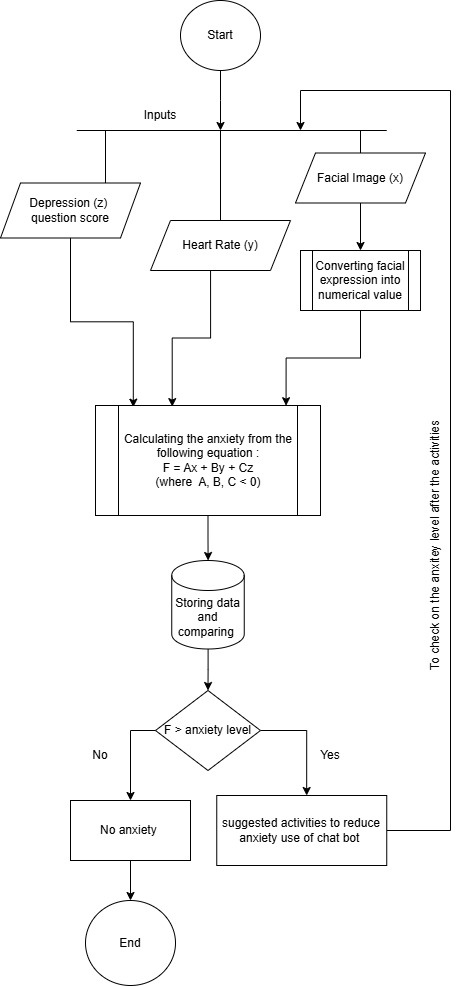
The system includes a self-assessment module to evaluate depression levels using a structured questionnaire.

* Users answer a standardized questionnaire, with each response assigned a predefined score.
* The total score classifies depression into Minimal (0–9), Mild (10–18), Moderate (19–29), or Severe (30+) categories.
* A web-based interface records responses instantly, providing immediate feedback.
* The depression score contributes to anxiety level calculations in the system.

**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.

### 4. Anxiety Score Calculation

The final component of the system is an anxiety assessment, which integrates three major factors: emotional, physiological, and psychological to provide a holistic evaluation of an individual’s mental state.

* The anxiety score is an emotional score, heart rate, and the depression score is a non-calculator, and the heart rate is automatically calculated from the input data.
* The generalization is confirmed by a single-weighted formula containing the proportional contribution of each parameter to the final anxiety level.
* For high anxiety disorders, relaxation tips such as meditation, music therapy, and music therapy are recommended, and continued high-level professional help is recommended.

**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.

FLOW CHART

STUDY OF VARIOUS EMOTION DETECTION MODELS

Emotion detection models are evaluated based on accuracy, computational efficiency, and real-time feasibility. Several deep learning architectures, including Basic CNN, VGG16, EfficientNet-B0, MobileNetV2, and Custom CNN, have been tested to determine the most suitable model for our system.

### Basic CNN

Initially, a basic convolutional neural network (CNN) was considered for emotion detection. While it offers rapid training and inference, its low accuracy (~65%) and inability to effectively capture intricate facial features render it unsuitable for practical applications.

### VGG16

VGG16, a pre-trained deep learning model with a well-defined architecture, demonstrates good accuracy (~79%). However, its substantial model size and high computational demands make it unfeasible for real-time emotion detection. The complexity of deeper architectures slows down inference, which is a significant disadvantage for scenarios that require immediate responses.

### EfficientNet-B0

EfficientNet-B0 outperforms VGG16 while maintaining comparable accuracy (~80%). It employs an efficient scaling method for enhanced feature extraction with a reduced number of parameters. Nevertheless, despite the considerable performance gains, it continues to have high computational requirements, which hinders its real-time applicability, particularly on edge devices.

### MobileNetV2

MobileNetV2 is engineered for lightweight solutions and rapid inference speeds, making it highly effective for mobile devices and embedded systems. However, its accuracy (~76%) falls short compared to deeper models, which diminishes its reliability for emotion detection, where accurate classification is crucial.

|  |  |  |
| --- | --- | --- |
| **Model** | **Accuracy & Performance** | **Computational Efficiency & Suitability** |
| **Basic CNN** | Low accuracy (~65%), lacks deep feature extraction. | Fast training & inference but poor generalization. |
| **VGG16** | High accuracy (~79%) but requires significant resources. | Large model size, slow inference, unsuitable for real-time use. |
| **EfficientNet-B0** | Similar accuracy to VGG16 (~80%), optimized feature extraction. | Lower computational cost than VGG16 but still resource-intensive. |
| **MobileNetV2** | Moderate accuracy (~76%), optimized for efficiency. | Fast inference, lightweight, but lower precision in emotion detection. |
| **Custom CNN** | Highest accuracy (82%), trained specifically for emotions. | Balanced speed, efficiency, and deployment feasibility. |

### Why Our Custom CNN Model is Better?

The Custom CNN model is specifically tailored for facial emotion recognition, ensuring an optimal mix of accuracy, speed, and efficiency. Unlike previously established models like VGG16 and EfficientNet-B0, which were primarily developed for general image classification, our model is trained from the ground up on emotion datasets, enabling it to effectively capture emotion-specific features.

* It achieves the highest accuracy (82%), surpassing all other mentioned models.
* It maintains a quick inference speed, making it ideal for real-time emotion detection.
* With a moderate computational cost, it is compatible with a wide range of devices, including mobile and edge computing environments.
* The optimized architecture eliminates unnecessary complexity while preserving strong feature extraction capabilities.

