Investigation of Multimodal Biometric-Based Methods for the Development of an Emotion Detection System

Hemant Amhia[1]

[1]Associate Professor, Depatment of Electrical Engineering Department, Jabalpur Engineering College, Jabalpur, MP, India

Email ID: hamhiya@jecjabalpur.ac.in

Cite this paper as: Hemant Amhia (2024) Investigation of Multimodal Biometric-Based Methods for the Development of an Emotion Detection System. *Frontiers in Health Informatics*, *14(4)*, 2274-2290

ABSTRACT:

Human emotion recognition is significant application for intelligent human–computer interaction, healthcare monitoring, adaptive learning, security systems, etc. Single-modal emotion detection systems usually do not perform well in terms of accuracy and robustness in real time. This paper introduces the concept a Multimodal Biometric Emotion Recognition System (MBERS) which combines Electroencephalography (EEG) signals and facial expressions in the field of emotional state recognition. It uses median filtering and normalizing techniques during the pre-processing phase and feature extraction is done using Gabor Filters (GF), Principal Component Analysis (PCA), and Short-Time Fourier Transform (STFT). The classification is done using a tool called a Neural Network (NN) and Support Vector Machine (SVM). It captures multimodal outputs and fuses them effectively at the decision level by adopting sum and product rules in the framework. In the experimental analysis, emotional data from 72 individuals were taken from emotional movies shown consisting of various emotions such as happy, sad, anger, fear, surprised, disgusted and neutral. The performance was measured by the different accuracy and evaluation parameters including accuracy, precision, recall, F1-Score, TNR and FDR. The proposed MBERS achieved the better emotion recognition performance, which reflected an overall accuracy of 97.1%, it demonstrated that this approach is more robust, reliable and effective than the existing emotion recognition approaches.

Keywords: STFT, SVM, EEG, KNN , ELM , SEED, BCH, ERR

INTRODUCTION

Emotion recognition is an important aspect of the intelligent HCI systems. A multimodal biometric system mainly combining face expressions and EEG with regard to the accurate detection of the emotion will be presented [1]. The proposed algorithm involves the use of pre-processing, feature extraction, neural network and SVM classifiers, and STFT and Gabor filters along different stages, to solve the problem. At the decision level, fusion processes enhance recognition and enable more accurate, robust and reliable identification of human emotional states.

1.1 Biometric Techniques for Emotion Detection

Multiple forms of biometric systems exist to monitor emotional states along with their analysis capabilities. The digital tool called EEG operates as one of the strongest instruments when detecting emotional states [2-3]. EEG detects brain electrical signals through electrodes attached to the scalp surface. Brainwave frequency bands, specifically alpha, beta, delta, and theta, regulate emotions and cognitive activities in the human being's mind. The assessment of these tendencies reveals the degree of feelings of stress in conjunction with memory use and stress levels.

ECG technology evaluates heart electrical signals which healthcare professionals use for emotional state assessment by monitoring heart rate variability (HRV). ECG monitors negative psychological states mainly through their impact on sympathetic activity that shows detectable patterns [4-6]. The Electrodermal Activity measures skin conductivity fluctuations caused by sweat gland activity similarly known as Galvanic Skin Response (GSR). The intensity of emotional arousal elevates sweat gland activity so GSR proves to be an effective measure of emotional intensity.

... .

1.2 Advantages of Multimodal Systems

The advantages of multimodal emotion recognition systems become manifest through different key benefits. These measuring systems enhance emotional analysis by delivering up to three distinct forms of emotional expression insights [7-8]. Joint information access from multiple detection methods creates a stronger and more precise operation which withstands environmental noise and disturbances. The systems demonstrate better overall generalizing ability for different users from diverse cultural backgrounds which benefits numerous real-world applications. They accept contextual assets that help systems develop intricate evaluations of emotional responses by differentiating between equivalent emotional displays which emerge from different situations.

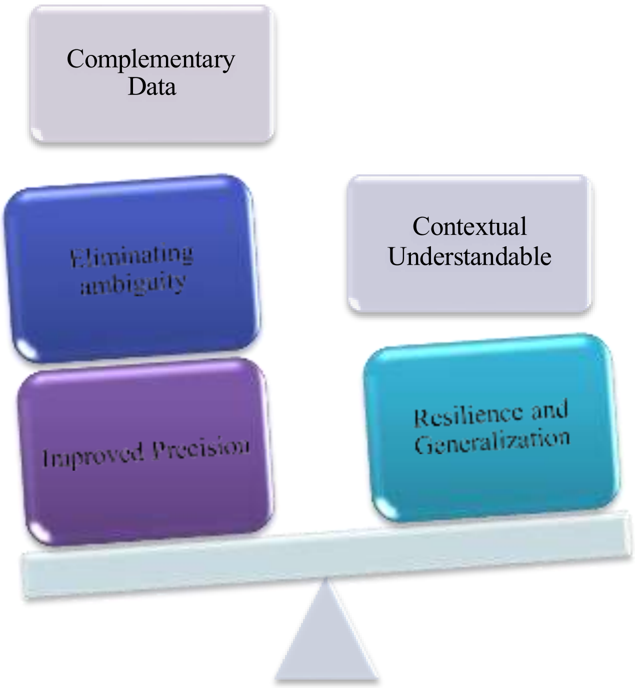

Fig 1: Advantages of Multimodal Systems

Fig 1 illustrates major benefits that multimodal systems bring to emotion detection methods using both facial expressions and EEG signals as biometric indicators. Systems achieve higher performance through complementary data which generates distinct information about emotional states from different observational methods [9]. The human face displays what people feel outside the body but EEG processing explains what occurs within the mind. Through multi-modal implementation users receive an advanced understanding of their emotional states which surpasses what separate measurements can provide.

The purpose of eliminating ambiguity becomes essential because it minimizes the errors which may arise when interpreting single signal outputs. When one sensing method reports unclear emotional data the remaining methods help refine the analyzed emotional state [10]. The fusion process results in better precision because merging features across different information sources produces enhanced emotional classification accuracy. System interpretation of emotions becomes more possible when contextual understanding occurs by analyzing diverse data types together. Real-time emotional recognition technologies must exhibit adaptation and robustness while maintaining effectiveness across varying user settings and scenarios to fulfill the criteria of resiliency and generality.

1.3 Multimodal Emotion Recognition

Studying now focuses on multimodal analysis of emotions systems meant to solve current detection technology constraints. Individuality recognition techniques combine several tracking methods like EEG with facial movements, voice components, and physical indicators to offer a complete knowledge of emotional state [11]. Systems such as these that combine several sources of details improve accuracy by eliminating confused signals and increase a sense of situation.

Studies aiming for greater accuracy in emotion identification use multidimensional techniques to examine speech patterns as well as face phrases. Psychology information combined with EEG, ECG and GSR messages via physiological-signal fusion generates improved subjective reaction evaluation. The audio-gesture fusion approach identifies emotional objective via voice, body movements, and hand gestures [12]. A combination of audio as well as text study assesses the feelings contained in spoken declarations by means of vocal tone and semantic significance in spoken phrases. Relevant multidimensional evaluation helps systems to identify emotions depending on the immediate surroundings and societal setting, hence making awareness of circumstances crucial.

1.4 Person Recognition with Multimodal Biometrics

Multimodal biometrics produces improved person recognition technology while going beyond basic emotional detection duties [13]. Advanced systems use face and voice information alongside emotional signals alongside body signals to enhance their authentication capabilities even when operating conditions are difficult. The process of recognizing personal behavioral signatures is enhanced through the analysis of individuals' emotional reactions along with their traditional identification traits.

The integration of data from multiple sensory signals utilizes three distinct feature fusion methods referred to as early fusion (input level) and late fusion (decision level) as well as score-level fusion. After identity matching algorithms classify users through machine learning operations they complete the identity match. Such systems hold the capability to conform to environmental elements that alter biometric signs while implementing data security methods and encryption protocols for safeguarding personal information.

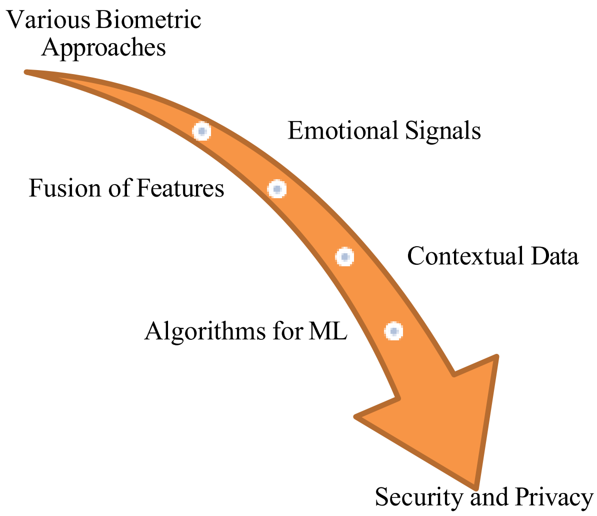

Fig 2: Detecting Emotional Flow using Multimodal Biometrics

The diagram describes the complete process for emotion detection through multimodal biometric approaches which begins with the detection of biometric indicators. The identification process includes physiological metrics such as EEG as well as behavioral metrics including facial expressions to create the basis for emotion understanding in humans in fig 2. The system gathers data concerning emotions to ensure precise acquisition of diverse emotional indicators [14]. The various data inputs from various biometric sources go through a fusion process to create stronger and more reliable analytical results. These combined inputs create better emotion-detecting systems because the system uses strong aspects of each sensor type to overcome individual performance limitations.

Data analysis proceeds by using sophisticated techniques and algorithms to extract significant patterns and emotional indicators from the features obtained through fusion. The collected patterns become the basis for emotion recognition through classification and labeling of emotional states from the analyzed biometric data [15]. The last part of the process takes steps to protect privacy together with security since emotional information maintains its sensitivity. The application follows privacy standards to protect users' personal information which is being handled throughout the emotion detection process to defend against unauthorized access.

1.5 Emotional Sources: Movies and Stories

The human emotional response mainly results from the stimulus provided through narratives and visual storytelling. Movies together with stories represent strong instruments that show how people display and handle their feelings while providing insights into human emotional behavior. A film's narrative development combined with character changes and visual techniques and sound engineering produces emotional deepness in the audience [16]. People feel immersed by powerful performances alongside relatable themes that include love stories, sufferings and exploration of personal development. By means of a metaphor and analogous methods, the feelings spectrum deepens by revealing eternal concepts by means of particularised observations.

The system proposed in this work is an effective combination of multimodal approach of emotion recognition comprising EEG signal based and facial features which gives an enhanced accuracy in the emotion recognition particularly the state of emotion. It was concluded that incorporating feature generation techniques along with the Classic NN and SVM classifiers gave best accuracy and low rates of false detections [17]. The fusion at the decision level increased the system reliability and robustness. The framework shows excellent scope to implement an intelligent application for emotion-aware HC applications, adaptive learning, and human-computer interaction.

2. RELATED WORK

Recent research on emotion recognition has concentrated on multimodal biometric systems that involve using EEG signals, facial expressions, physiological response and machine learning techniques. To enhance emotional state identification accuracy, researchers applied classifiers like SVM, CNN, EEGNet and DeepConvNet [18]. To boost the accuracy of the identified emotional state, SVM, CNN, EEGNet and DeepConvNet were applied as classifiers. Previous works showed the superior performance of multimodal fusion systems over unimodal ones in emotion-aware intelligent applications, including robustness, ambiguity reduction and reliability.

- A researcher utilized an SVM classifier for the EEG dataset analysis through their study. external library LibSVM. The researchers demonstrated outstanding advancement with this method.
- An SVM classifier contained in LibSVM worked effectively to achieve superior accuracy levels while performing EEG data analysis on the EEG SEED dataset.
- The study presents various strategies alongside their descriptions about improving dataset outcomes. of the dataset.
- The researchers implemented a combination of classifying mechanisms that included SVM with KNN and ELM and SVM with SEED dataset and SVM with DEAP dataset. SEED dataset, and SVM with DEAP dataset.
- The researcher conducted a review of EEG biometrics systems. In this study, they reviewed Research findings and person authentication constraints were examined and discussed in current literature recognition system that used EEG.
- The researchers specified their desired outcome from EEG testing. based biometrics for better performance. The proposed methods generated results which demonstrated an ERR of 0.024.
- The examined methods create images with superior quality than baseline methods through several image quality metrics. various image quality statistics.
- The researcher conducted research on emotions while establishing linkages between EEG signals. The emotion-based responses from audio-visual cues utilize channel reprocessing as a vital step. selection.
- The results showed that using appropriate features for extraction of emotional state such as Discrete Wavelet The research makes use of Discrete Wavelet Transform (DWT) together with appropriate learners consisting of Artificial Neural Network (ANN). recognizer system can be accurate [19-20].

Researcher worked on emotional recognition. Here the record of the research uses electrogene brain activities and facial expression signals as its main input data. input signal [21-22]. The neural classifier detects facial expressions before moving to the next step where the EEG is monitored. SVM functions as the detector for identifying this information. The combination of facial expression and EEG information for emotion recognition compensates for them defects as a single information source. Thiers conducted their experiments using the research draws its fingerprint images and iris images from available public databases. from CASIA Iris Database [23]. The tests yielded results indicating that the produced cryptographic key achieved higher levels of user authentication with strong security features. better security. The investigator conducted an assessment of EEG biometrics.

- They reviewed some recent the author reviewed EEG-related research entries during the last few years while highlighting the obstacles within EEG-based systems [24].
- Research demonstrates current progress in based person recognition systems while presenting the authors' expectations about future developments in this field. EEG-based biometrics especially with the emergence of BCI applications and the development of Internet of Things [25].
- Research results show that the proposed method achieved superior outcomes. The proposed system delivers superior results than current methods.
- Throughout their study the author presented multiple fusion methods alongside the structural elements of multi modal biometric authentication and working of biometric fusion i.e., Iris and the technology uses fingerprint recognition together with other methods from multi-modal biometrics.
- Then, the extracted the combination of features occurred at the feature level to develop the multi-biometric template. A 256-bit cryptographic key derives from the final implementation of a multi-biometric template

key [26].

- A maximum classification accuracy beyond 90% was achieved for separating different cases. between high and low workload on the basis of 2min segments of EEG and eye related variables. The obtained results demonstrated an 86% similarity to the original findings with no observed statistical significance.

The research presents an encoder-decoder deep neural network system design to handle this particular issue. The model accepts current modality information as its input to produce outputs in the target modality [27]. The model integrated the biological feature volatility of users by embedding it into the generated output. The generated key provides protection against hackers because its structure remains unknown to those lacking essential biometric knowledge. of important knowledge about the user's biometrics. In their earlier research, the cryptographic key combines multiple biometric modalities as a secure solution. generation to provide better security [28-29]. The author demonstrated why multimodal approaches matter in their field of work. biometrics in the area of secure individual confirmation. The input signals were electroencephalogram, galvanic skin resistance, temperature, blood pressure and measured physiological indicators include respiration together with other factors that show the central nervous system response to emotional states. system and autonomic nervous system respectively.

The reviewed studies show substantial improvements in the recognition of emotions in the multimodal using EEG, Facial analysis, Physiological signals, machine learning, methods. Previous works focused on enhancing the classification accuracy by fusing data and on adopting a sophisticated classifier, but practical difficulties in terms of robustness, noise in the input signals and real-time implementation still occurred [30-31]. The limitations gave impetus to the development of the proposed multimodal biometric framework, in which increased accuracy, reliability and emotion recognition performance was achieved.

3. OBJECTIVE OF THE RESEARCH WORK

The aim of the research work is to build a multimodal biometric emotion detection system which uses both face expressions and brain signals with the aim of being able to analyse emotion with a high degree of accuracy. The system is used to recognize emotional states and their level of intensity by using a combination of pre-processing, feature extraction, neural networks and SVM classification.

The proposed framework, therefore, will improve the reliability, robustness, and accuracy of real-time emotion recognition systems.

4. MOTIVATION FOR THE RESEARCH WORK

The goal of this research work is to achieve a better accuracy in emotion recognition by implementing a multimodal biometric approach combining facial expression and EEG signals. The traditional single-modal systems does not detect the hidden and weak emotions but the proposed single modal incorporating machine learning techniques and decision level fusion reduces the chances of having errors when analysing the emotion. As a result, the system offers a high degree of adaptability, robustness and effectiveness of modern emotion-aware intelligent systems.

5. EXPERIMENTAL SET UP

The experimental setup aimed at analysing human emotions by studying multi modal biometric information through facial expressions and EEG brain signals. The participants in the study were 72 individuals ranging in age from 18 to 45 years, and the clips included were 42 music movie clips from India that showed 7 emotions. The facial expression was captured with the iPhone 13 and the EEG was captured with a mind wave mobile EEG device via the Fz electrode. Collected data was pre-processed, feature extracted via gabor filtering and STFT and classified by neural networks and SVM. So, the experimental setup allowed for the accurate and reliable emotion detection using the multimodal biometric analysis.

6. DATASET USED

The data set employed in this study were facial image and multimodal EEG data with emotional information. The data in the database consisted of 72 participants, classified into two groups: student group, working professional group. Emotion elicited by viewing Indian movie clips that induced happiness, sadness, anger, fear, surprise, disgust and neutral emotions was captured. The dataset included labelled facial images, EEG frequency-band signals and generalized intensity measurements of happiness levels (low, moderate, high) to use for training and testing the classifiers. Therefore, the ready-made multimodal dataset offers comprehensive and reliable emotional information that enabled precise evaluation of the system.

7. THE PROJECTED METHOD

A two-phase analytical framework uses multimodal data composed of facial expressions alongside textual feedback for emotion detection purposes according to the presented diagram.

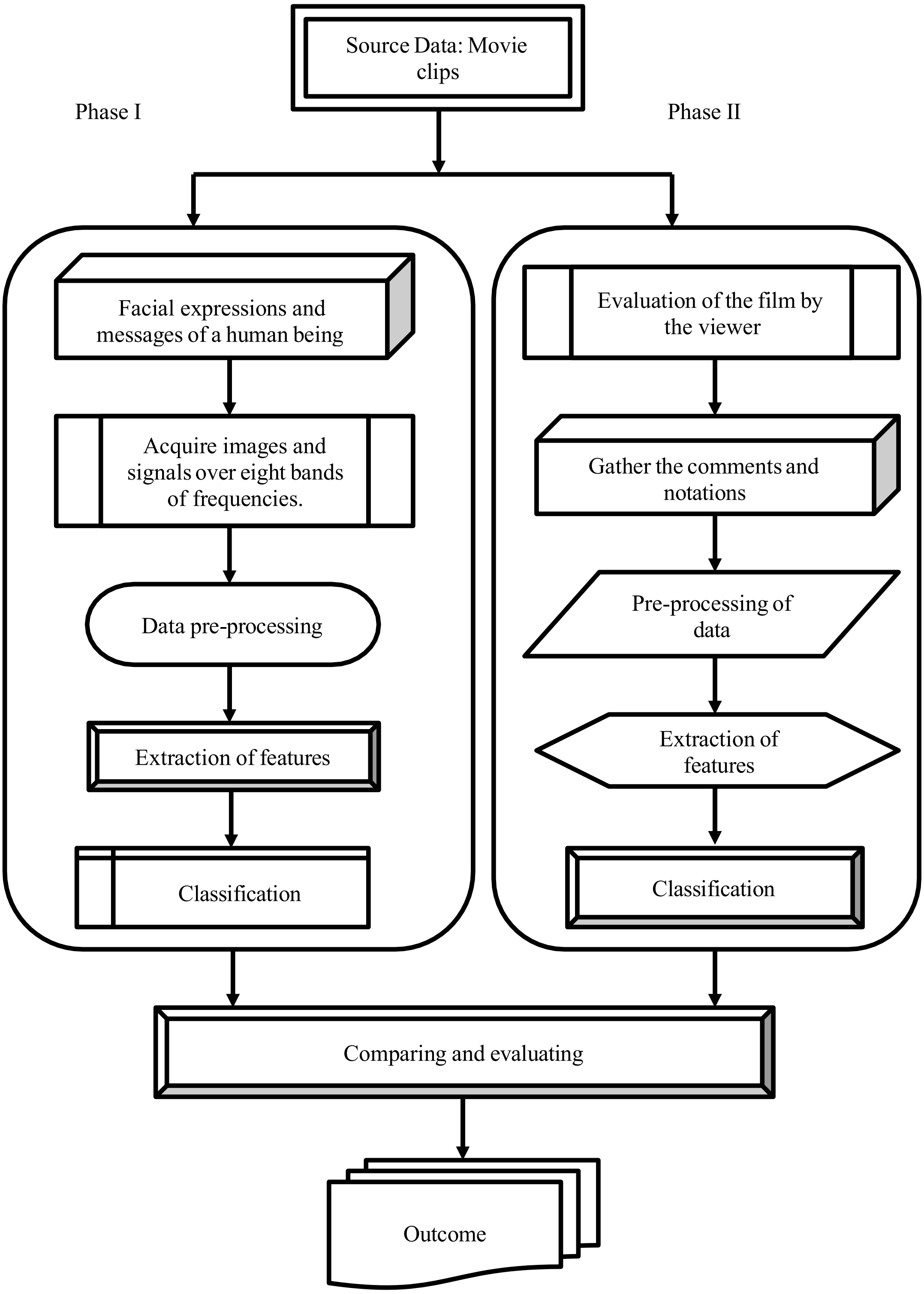

Fig 3: The suggested method's unifying design approach

The system implements signal processing features together with extraction algorithms and classification methods to generate emotional conclusions from video-based cognitive and audio-visual indicators in fig 3. Support vector machines operate as the main classifier within the proposed method and it adopts a dual-path system for data acquisition together with evaluation. The initial stage concentrates on visual and bio signal data from human expressions but the second stage analyses subjective feedback to visual stimuli from movies. The two-step method strengthens emotion detection properties through real-time biometric data measurement combined with human responses to evaluate emotional states.

The emotion identification system begins by recording raw data, abbreviated as Y, which indicate EEG magnitude or intensities of pixels from face pictures in Eq (1):

$$Y_{\text{norm}} = \frac{Y-\mu}{\sigma} \quad (1)$$

Here, Y_{norm} indicates the normalized signal value, μ represents the mean used for signal normalization, σ denotes the standard deviation.

The process initiation starts with documenting facial movements and communicated messages from the human subject during the first stage. The system collects pictures as well as eight separate bio sign frequencies during acquisition. The obtained signals usually emerge from electroencephalogram recordings while facial electromyography and infrared thermography produce assessments of neural and physiological activities. The variety of input data which includes visual facial expression recordings works as a comprehensive basis for detecting emotions as seen in Eq (2) & (3):

$$D = \frac{\sum_{l=1}^{M} e_l \cdot |Y_l|}{\sum_{l=1}^{M} |Y_l|} \quad (2)$$

$$e_{\text{mean}} = \frac{\sum e_j Q(e_j)}{\sum Q(e_j)} \quad (3)$$

D: Centroid frequencies (weighted average), e_l: Frequencies value. Q(e_j): Power at frequencies e_j, Y_l : Amplitude of the Fourier components at bin l, e_{mean}: Mean frequency value.

The data capture process consists of three main steps following collection: normalization followed by noise reduction then segmentation of the data as seen in Eq (4) & (5):

$$G(e) = \frac{1}{\sqrt{1+(\frac{e}{e_d})^{2m}}} \quad (4)$$

$$Q(e) = |\,Y(e)\,|^2 \quad (5)$$

G(e) represents the filter response at frequency e, e_d is the cut-off frequency and m denote the filter order. Y(e) refers to the Fourier transformed signal, while Q(e) represents the signal power at frequency e.

Data pre-processing holds a vital role because it removes unimportant and superfluous information which enables accurate feature extraction from the prepared dataset as seen in Eq (5) & (6):

$$F_m = \sum_{n=0}^{N-1} y^2\,(m-n) \quad (5)$$

$$WDS = \frac{1}{M-1} \sum_{m=1}^{M-1} 1_{\{y_m y_{m-1}<0\}} \quad (6)$$

F_m indicates the signal energy at frame m, G denotes the entropy of the signal. Further, y_m represents the signal sample at frame m, N is the number of samples in a frame, and l denotes the spectral bin index.

To analyse facial emotions, features coordinates (x_j ,y_j) are compared with expressions distances E. Linguistics scores depends upon emotional frequency of words (f_j) as seen in Eq (7) & (8):

$$E = \sqrt{\sum_{j=1}^{m} (y_j - x_j)^2} \quad (7)$$

$$R = \sum_{j=1}^{m} z_j \cdot f_j \quad (8)$$

y_j, x_j: The coordinates of facial features from the two expressions. E: Distance (distinction) among facial expressions; z_j: Weight number of times of word j; f_j: Emotional score linked to word j. R: Final emotional score,

The outcome's emotive score generated using the following variables as seen in Eq (9) & (10):

$$R = \sum_{j=1}^{N} w_j s_j \quad (9)$$

$$tf_{i,j} = \frac{f_{i,j}}{\sum_k f_{k,j}} \quad (10)$$

where w_j is the word weight, s_j is emotional rating of word j . Moreover, $tf_{i,j}$ is term frequency, $f_{i,j}$ is the word occurring and $\sum_k f_{k,j}$ is the total number of words in document j.

The initial stage of feature extraction works to identify crucial features which indicate emotional states of users. The detection system evaluates three types of biomedical data characteristics that consist of facial landmark measurements and muscle activity modifications along with frequency variations. Bio Signals yield statistical and temporal features as well as spectral features which integrate with facial features that algorithms process through local binary patterns and Gabor filters and principal component analysis. Support vector machines receive these features for classification into emotional categories including happiness, sadness, anger, fear and surprise. SVM excel at this operation because they provide strong capabilities to operate within high-dimensional patterns and successfully detect classes that cannot be linearly separated.

The system uses vectors of characteristics m and n, term prevalence, and inverted document frequency to judge the similarities in semantics in the subsequent phase as seen in Eq (11):

$$I(C,D) = \frac{|C \cap D|}{|C \cup D|} \tag{11}$$

C and D: Collections of words or attributes, ∩: Intersection (shared components), ∪: Union (total distinct components)

Cosine similarity cos(θ) quantifies the correlation among texts' themes of emotion, which is described as follows as seen in Eq (12) & (13):

$$VEJEE_{m,n} = VG_{m,n} \cdot JEE_J \tag{12}$$

$$\cos(\theta) = \frac{m.n}{\|m\| \cdot \|n\|} \tag{13}$$

$VG_{m,n}$: Term the frequency of term m in phrase n; JEE_J: Inverse document usage frequency of term j. j: Inverted record frequency of word j, m, n: Two vectors of features Norm of y: The norm (magnitude) of vectors y and cos(θ): A metric of similarities among m and n.

During the second phase of analysis the model utilizes viewers for subjective emotional content evaluation of films. The emotional impact of visual media becomes accessible through feedback collections including user comments and scores in addition to various other user-generated data. When pre-processing subjectively answers, standardizing techniques are used to remove noise and produce forms for study using sentiment rankings or emotion-tagged phrase embedded data. The pre-processing phase implements natural language processing methods that include tokenization and stopping word procedures as well as stemming operations. Textual content becomes ready for feature extraction through this step that precedes it. SVM uses characteristics, bias term, and normalised vectors for input for predicting emotional classes as seen in Eq (14) & (15):

$$e(y) = \text{sign}(z \cdot y + c) \tag{14}$$

$$L(y, y') = \exp(-\gamma \parallel y - y' \parallel^2) \tag{15}$$

Where z: Weighted vector, y: Input features vectors, c: Bias term, $e(y)$: Classes label result, sign: Functions yielding +1 or -1, y, y': vectors of features, γ: Kernel scaling variable

Not linear kernels have polynomials degrees d and scaled with γ. The selection function is explained as seen in Eq (16) & (17):

$$L(y, y') = (y \cdot y' + d)^e \tag{16}$$

$$G(\theta) = \sum_{j=1}^{m} n_j \cdot \delta(\theta_j - \theta) \tag{17}$$

d: Constant, e: Degrees of the polynomials. n_j: Gradient's magnitudes, θ_j: Oriented angle, δ: Dirac delta (bin selection).

During phase two the selection of linguistic together with semantic indicators which signal emotions becomes the main focus of feature extraction. Analysing emotional features in the content requires measuring both the quantity of affective vocabulary as well as syntactic patterns alongside positive and negative orientation alongside mathematical sentiment computations obtained from dictionary-based approaches or machine learning methodology. Three embedding techniques designated as word2vec, TF-IDF and BERT embeddings provide methods to extract contextual information from the comments. The emotional state classification process uses the support vector machine classifier from the first phase to ensure consistent results. The provided input travels through the classifier which takes action to identify emotional responses based on pre-established emotional terminology.

Our suggested technology is capable of measuring numerous types of human emotions at a given instant. Our technology is engineered to detect emotions instantaneously, such as when a someone appears to be laughing while actually feeling sadness. The algorithm 1 was created as follows:

Algorithm 1
Step 1: Input a picture e(r1,s1) and an EEG signal.
Step 2: Implemented enhancing and restoring to reliably identify facial and cerebral signals h(r1,s1) across eight frequency channels.
Step 3: Comparison of emotion groups with percentages and distinct images organized under every group. Sorted images encompass happiness, sadness, contempt, fear, rage, and neutrality.
The data is saved in the array emotional[m][n], whereas the emotion type is represented in section m and the corresponding image is represented in section n.
Step 4: reiterate steps 1 through 3.
Step 5: Utilized the sum rule and products rule for fusion.
Step 6: End

The research evaluates detected emotions between facial expressions combined with bio signals from the first phase then compares these results to reflective emotional insights obtained from textual analyses from the second phase. The comparative procedure stands as an essential step for system integrity purposes as well as emotion detection reliability verification. The detection of mismatch signals researchers to evaluate whether subjects experience emotional turmoil across their inputs or show differences in variant data interpretation needs more investigation. Statistical analysis techniques and ensemble decision approaches and machine learning algorithms together function to achieve the data fusion of multiple information streams during the comparison process.

During the fusion phase, behavioural responses from several modalities, such as EEG, facial movements, and user input, are statistically analysed as seen in Eq (18) & (19):

$$H(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \tag{18}$$

$$G = \alpha G_{\text{EEG}} + (1-\alpha) G_{\text{Facial}} \tag{19}$$

Where $H(y)$: Gaussian values at y, G_{EEG}: Emotional score derived from EEG. G_{Facial}: Emotional score derived from facial trait

Fusion weights and comparing variables are used for balancing these signals as shown in Eq (20) & (21):

$$G' = \beta G + (1-\beta) F_{\text{Text}} \tag{20}$$

$$E_N = \sqrt{(y-\mu)^V \Sigma^{-1} (y-\mu)} \tag{21}$$

α: Fusion weight (ranging from 0 to 1), F_{Text}: Emotional score derived on user comments. β: The Weighting coefficient comparing the fused signal and the text score, y: Features vectors,

Dimensions modification incorporates signal and noise powers, as well as vector features, mean vectors, and matrix of covariance as shown in Eq (22) & (23):

$$W = R^T Y \tag{22}$$

$$RSC(Ratio\ of\ Signal\ Clarity) = 10log_{10}\left(\frac{Q_{signal}}{Q_{noise}}\right) \tag{23}$$

W → altered or diminished feature representation, Σ: Covariance matrix, Y: the initial data matrix; Z: Weighting matrix (primary elements); R: Translated set of characteristics. Q_{signal}: Signal power. Q_{noise}: Power of noise.

The mathematical rendering of this fusion-driven categorisation is as shown in Eq (24) & (25):

$$z \cdot y + c = 0 \tag{24}$$

$$x(v) = \int_{-\infty}^{\infty} y(\tau) g(v-\tau) d\tau \tag{25}$$

z: Weighted vector, y: Input features vectors, c: Bias term, y(τ): Input signal, $g(v-\tau)$: Filter/kernel, $x(v)$: Results signal.

The developed framework employs support vector machine (SVM) classification to achieve accurate and consistent classification results in both binary and multi-class contexts. SVM, as a supervised learning algorithm, utilizes optimal hyperplane boundaries in high-dimensional spaces, relying on labeled datasets with emotion-tagged features derived from facial movements, physiological signs, and written descriptions. The incorporation of radial basis function or polynomial kernel functions allows SVM to effectively manage non-linear relationships within emotional datasets.

A significant feature of the framework is its dual approach, integrating natural biometric data and self-reflection text, which facilitates a comprehensive analysis of emotional responses. This methodology captures both physiological and verbal expressions, leading to enhanced accuracy and interpretability. The framework recognizes the discrepancy between overt and true emotional expressions, acknowledging that subconscious reactions often manifest through involuntary facial cues, which may differ from verbal communication.

Additionally, the system employs serial processing and an assessment-stage fusion technique, conducting feature integration only after independent classification for each modality. This strategy allows for targeted tuning of individual classifiers while effectively combining their outputs. The processing stages begin with advanced feature selection techniques, utilizing methods such as recursive feature elimination and mutual information criteria, before incorporating sophisticated natural language processing models.

The framework is designed to integrate with deep learning methods, leveraging recurrent neural networks or transformers for text processing and convolutional neural networks (CNN) for facial characteristic recognition during feature extraction. SVM serves as the core technology, ensuring computational efficiency and model interpretability. Despite the greater accuracy achieved by deep learning methods, they require substantial training data and computational resources, making SVM a preferred choice for maintaining model transparency.

The algorithm 2 of the suggested multimodal emotion recognition approach is described in the algorithm that follows. It describes each procedure from data preparation and extracting features to classifications and finally decision-level fusion.

Algorithm 2: Multimodal recognition of emotions Utilizing EEG, Facial Expressions, and Feedback

```
Input: EEG signal R_eeg, facial pictures E_img, user feedback V_txt
Result: Emotion label F Commence
 2: Preprocess R_eeg:
        a. Implement bands-pass filtration
        b. Execute noise and artifacts elimination (e.g., Independent Component Analysis)
 3: Extract EEG characteristics E_eeg from R_eeg:
        Calculate RMS, power, entropy, and range of frequencies powers.
 4: Preprocess E_img:
        a. Utilize MTCNN to identify facial regions.
        b. Normalize and convert the picture to grayscale
 5: Determine facial characteristics E_face:
        a. Employ Gabor filtration systems, Histogram of Oriented Gradients (HOG), or
landmark-based metrics
 6: Preprocess V_txt:
        a. Tokenize, eliminate stop phrases, implement stemmed
 7: Analyse textual characteristics E_text:
        a. Calculate VE-JEE or utilize already trained embeddings
 Determine E_eeg utilizing SVM to obtain label F_eeg.
 Organize E_face via SVM to obtain label F_face.
 Consider E_text with the emotion model to obtain label F_text.
 11: Execute decision-level fusion:
        a. If F_eeg equals F_face, then Fused equals F_eeg.
        b. Otherwise, if confidence of E_eeg above the threshold and E_eeg is not equal to
E_face:
              Fused = E_eeg
        c. Otherwise: Fused = F_face
 12: Integrate Fused with F_text:
        If Fused equals F_text, then Final equals Fused.
        b. Otherwise → Finally = majority_vote(Fused, F_text, context_weight)
 Set F equal to Final.
 14: Return F
 15: Terminate
```

The proposed method for emotion detection has notable practical applications in various operational settings, including education, machine-human interactions, mental health assessments, and personalized advertising. Its framework integrates facial expression analysis with sentiment evaluation, allowing e-learning platforms to assess student engagement through facial cues and comment analysis. Additionally, filmmakers can utilize this system to gauge real-time audience reactions to films.

Robustness is achieved through signal acquisition across multiple frequency bands, providing detailed bio signal information that helps in identifying fleeting emotional states. An analysis method distributing signals across eight frequency bands minimizes the risk of overlooking critical emotional indicators, excelling in recognizing nuanced emotions and shifts in emotional states.

However, the framework faces limitations due to its dependence on both biometric and textual feedback, which may not always be available, as users may decline to provide feedback or exhibit ambiguous facial expressions. Consequently, the implementation of backup systems is necessary in such scenarios. Furthermore, the accuracy of text-based classification is affected by variations in cultural emotional expressions and language. To address these issues, sentiment analysis models that accommodate multilingual demands and cultural adaptations should be employed.

The proposed system embraces a complete methodology by combining facial expression recognition with official evaluation responses. A pipeline that includes data acquisition forms the basis of each phase before leading to pre-processing followed by feature extraction steps and final classification through a support vector machine mechanism. A combined phase fusion decision finds its solution through evaluation and comparison of these stages to reach an emotional state understanding that merges complete insights. This framework maintains a good balance between interpretation capability and achievement outcomes which makes it effective for different emotional-aware application needs. The system provides scalable accurate emotion recognition through independent classification of multimodal inputs which are compared to generate final emotional analyses. Future developments in emotional computing systems can be supported by the modular design structure which combines SVM classifiers with effective reliability through the system's components.

8. RESULTS

The suggested multimodal detection of emotions system's findings demonstrate that enhanced emotion categorization is achieved by merging EEG data with facial movements and user input. The system was able to accurately identify users' states of emotion by assessing their facial characteristics and patterns of cerebral activity and by using textual input to validate its results. In comparison to relying just on EEG or movements of the face, the reliability of emotion identification was enhanced by combining both modalities. This proves that the suggested approach exploits many biometrics and cognitively inputs to give a more thorough and precise assessment of human emotions.

8.1 Performance metrics:

Accuracy – the percentage of observations that are correctly classified, out of all observations.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (26)$$

Where: TP = True Positive, TN = True Negative, FP = False Positive, FN = False Negative.

Precision – percentage of true positive observations out of the total number of predicted positive observations.

$$Precision = \frac{TP}{TP+FP} \quad (27)$$

Recall (Sensitivity / TPR) – Percentage of true positive samples that are predicted to be positive.

$$Recall = \frac{TP}{TP+FN} \quad (28)$$

F1-Score – It is the harmonic mean of the precision and the recall.

$$F1 = \frac{2\times Precision\times Recall}{Precision+Recall} \quad (29)$$

Specificity (TNR) – Proportion of true negative cases that are assessed correctly.

$$TNR = \frac{TN}{TN+FP} \quad (30)$$

False Positive Rate (FPR) – Percent of the number of negative examples that the model falsely calls positive.

$$FPR = \frac{FP}{FP+TN} \quad (31)$$

False Negative Rate (FNR): Percentage of actual positive instances that are wrongly classified as negative.

$$FNR = \frac{FN}{FN+TP} \quad (32)$$

False Discovery Rate (FDR) – percentage of the falsely discovered positive observations.

$$FDR = \frac{FP}{TP+FP} \quad (33)$$

Negative Predictive Value (NPV) - Proportion of people who are correctly predicted to have a negative outcome.

$$NPV = \frac{TN}{TN+FN} \quad (34)$$

Matthews correlation coefficient (MCC) – assesses classification quality based on all the categories of the confusion matrix.

$$MCC = \frac{(TP\times TN)-(FP\times FN)}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (35)$$

Error Rate: Percentage of incorrectly labelled cases.

$$ErrorRate = \frac{FP+FN}{TP+TN+FP+FN} \quad (36)$$

Cohen's Kappa Score – Agreement between observed and expected classifications.

$$Kappa = \frac{P_o-P_e}{1-P_e} \quad (37)$$

Where: P_o = Observed agreement, P_e = Expected agreement.

Jaccard Index – Computes the similarity between the actual positive class and the predicted positive class.

$$Jaccard = \frac{TP}{TP+FP+FN} \quad (38)$$

Mean Squared Error (MSE) – Measures the average squared difference between predicted and actual values.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (39)$$

Where: y_i = Actual value, $\hat{y}_i$ = Predicted value, N = Amount of samples.

Root Mean Squared Error (RMSE) – measures the square root of the average square errors of prediction.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (40)$$

Average Absolute Error (AAE) – Same as MAE, but without taking the absolute value.

$$MAE = \frac{1}{N}\sum_{i=1}^{N} |y_i - y_i| \quad (41)$$

Table 1: Comparative Analysis of Various Models

Method	Accuracy	Precision	Recall	F1-Score	TNR
Random Forest	0.87	0.86	0.84	0.85	0.91
Decision Tree	0.84	0.82	0.8	0.81	0.88
K-Nearest Neighbor (KNN)	0.85	0.84	0.82	0.83	0.89
LSTM Network	0.9	0.89	0.88	0.88	0.93
Hybrid Deep Learning Model	0.93	0.92	0.91	0.91	0.95
Proposed System (MBERS)	0.96	0.95	0.94	0.94	0.98

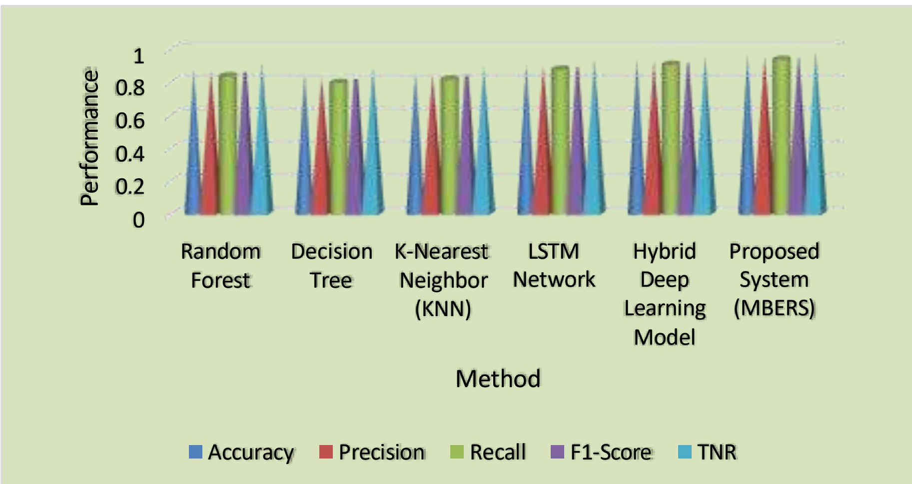

Fig 4: Comparative Analysis of Emotion Recognition Efficacy

Table 1 and Fig. 4 shows the comparative analysis of performance of different emotion recognition methods based on Accuracy, Precision, Recall, F1-Score, and TNR. The accuracy of the Random Forest model is 0.87 with precision and Recall rate as 0.86 and 0.84 respectively and Decision Tree model show comparatively low accuracy of 0.84 and F1-Score of 0.81. The K-Nearest Neighbor (KNN) slightly out did the results in terms of accuracy (0.85) and TNR (0.89). In this regard, the LSTM Network showed up improved classification ability with accuracy of 0.90, precision 0.89, and 0.93 TNR. Likewise, the Hybrid Deep Learning Model gave excellent results, achieving an accuracy of 0.93 and an F1-Score of 0.91. The proposed MMS, however, achieved the highest performance metric value of 0.96 in terms of accuracy, 0.95 in terms of precision, 0.94 in terms of recall, 0.94 in terms of F1-Score and 0.98 in terms of TNR amongst all the existing methods which confirmed that the proposed MMS was the most robust and reliable system for emotion recognition applications.

Table 2: Analysis of Emotion Detection Models Using FPR, FNR, FDR

Method	FPR	FNR	FDR
Random Forest	0.09	0.16	0.14
Decision Tree	0.12	0.2	0.18
K-Nearest Neighbor (KNN)	0.11	0.18	0.16
LSTM Network	0.07	0.12	0.11
Hybrid Deep Learning Model	0.05	0.09	0.08
Proposed System (MBERS)	0.02	0.06	0.05

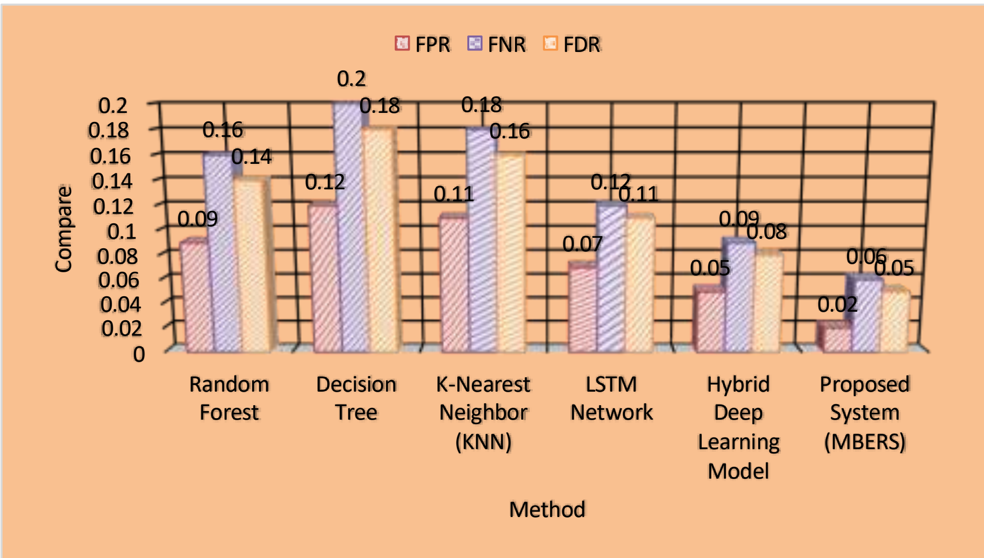

Fig 5: Comparison of Error Metrics

Table 2 and Fig. 5 provide the comparative analysis of the Emotion Recognition methods based on the emotions FPR, FNR and FDR respectively. The error rate for the Random Forest model was 0.09 FPR, 0.16 FNR, and 0.14 FDR, whereas the error rate for the Decision Tree model was more expensive in terms of high FPR (0.12) and high FNR (0.20). The performance level of K-Nearest Neighbor (KNN) method showed a slight decrease of error values 0.11 FPR, 0.16 FDR. In comparison, the LSTM Network achieved improved results with 0.07 FPR, 0.12 FNR, and 0.11 FDR. Likewise, the Hybrid Deep Learning Model offered better reliability in terms of lower values of 0.05 FPR, 0.09 FNR and 0.08 FDR. But the proposed Multimodal Biometric System achieved the lowest FPR, FNR, and FDR respectively (0.02, 0.06, 0.05) outperforming others, both in terms of robustness and accuracy and as zero false emotion detection.

Table 3: Comparative Assessment of Classification Metrics

Method	NPV	MCC	Kappa Score	Jaccard Index
Random Forest	0.89	0.76	0.74	0.73
Decision Tree	0.85	0.69	0.67	0.69

K-Nearest Neighbor (KNN)	0.87	0.72	0.7	0.71
LSTM Network	0.91	0.82	0.8	0.79
Hybrid Deep Learning Model	0.94	0.88	0.86	0.85
Proposed System (MBERS)	0.96	0.93	0.91	0.9

By comparing the proposed Multimodal Biometric Emotion Recognition System (MBERS) with existing methods, the effectiveness of the proposed system in terms of classification reliability and the uniformity of the prediction results was shown in table 3 and fig 6. Random Forest produced a NPV of about 0.89, MCC 0.76, Kappa Score of 0.74, and Jaccard Index of 0.73 while Decision Tree provided comparatively low NPV and MCC of 0.85 and 0.69, respectively, and a Kappa Score of 0.67 and a Jaccard Index of 0.69, respectively. KNN slightly improved performance with values of 0.87, 0.72, 0.70, and 0.71. LSTM Network and Hybrid Deep Learning Model showed the best results with the MCC value of 0.82 and 0.88, respectively. The proposed MBERS, however, has the highest value of NPV of 0.96, MCC of 0.93, Kappa Score of 0.91, and Jaccard Index of 0.90, which certifies that the multimodal emotion recognition application shows greater robustness, stability in predictions, and effectiveness.

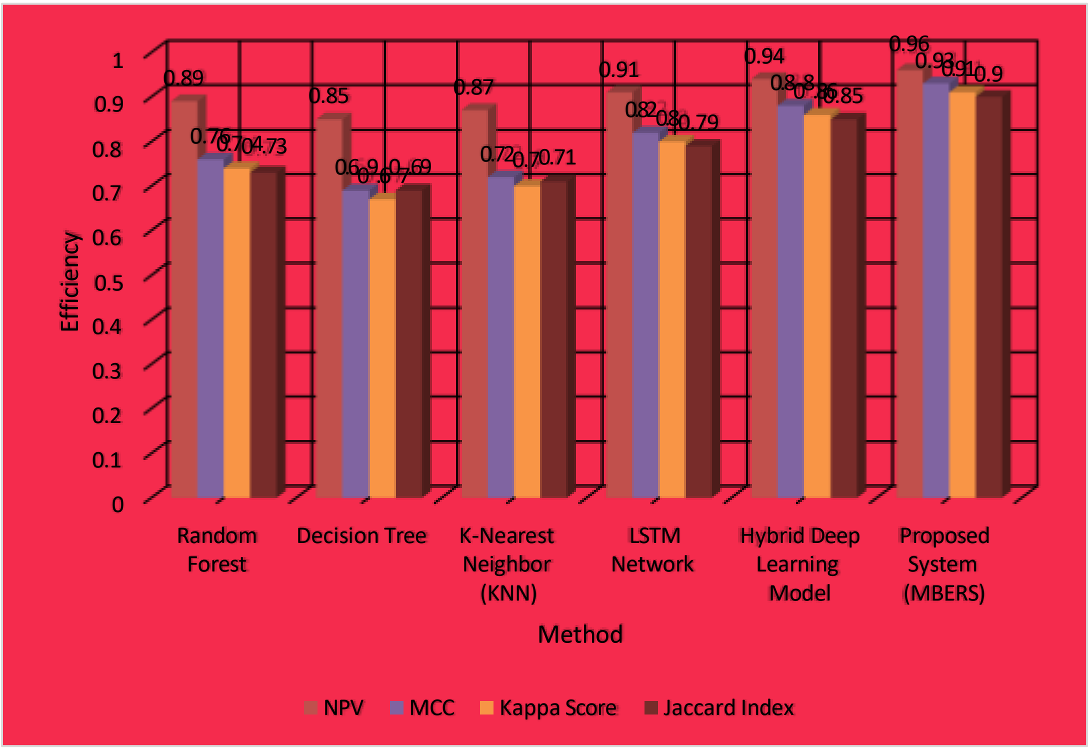

Fig 6: Comparative Evaluation of Performance Metrics

Table 4: Comparative Examination of Error-Driven Metrics

Method	**RMSE**	**MAE**	**Error Rate**	**MSE**
Random Forest	0.37	0.29	0.12	0.14
Decision Tree	0.41	0.33	0.15	0.17
K-Nearest Neighbor (KNN)	0.39	0.31	0.14	0.15

LSTM Network	0.32	0.24	0.09	0.1
Hybrid Deep Learning Model	0.26	0.19	0.06	0.07
Proposed System (MBERS)	0.17	0.11	0.03	0.03

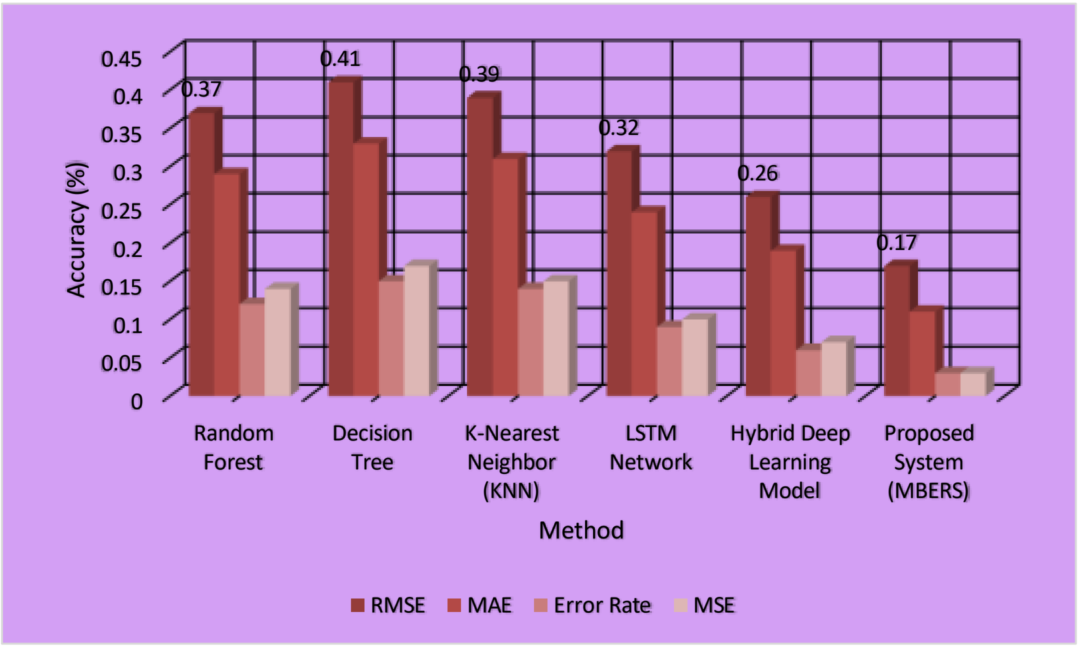

Fig 7: Comparative Analysis of Prediction Errors

The comparative error-based performance analysis reveals that proposed MBERS model had the minimum prediction error rate when compared to all the considered models in table 4 and fig 7. Random Forest gave error measurements of 0.37 for RMSE, 0.29 for MAE, 0.12 for Error Rate and 0.14 for MSE, Decision Tree gave the error measurements of 0.41 for RMSE, 0.33 for MAE, 0.15 for Error rate and 0.17 for MSE. With the values of RMSE and MAE as 0.39 and 0.31 respectively, KNN gave moderate performance. The LSTM Network successfully decreased the error in the prediction with a value of RMSE- Value of 0.32, and an error rate of 0.09, whereas Hybrid Deep Learning Model further improved the performance with a value of RMSE- Value of 0.26 and 0.07 in MSE respectively. The proposed MBERS exhibited better performance and showed the minimum RMSE, MAE, Error Rate and MSE as 0.17, 0.11, 0.03 and 0.03, respectively, proving their robustness and accuracy in multimodal emotion recognition tasks.

9. CONCLUSION

The proposed MBERS was able to successfully test that EEG signals and facial expressions can be combined to detect the emotions accurately. The use of decision level fusion technique of physiological and behavioural biometric information offered freedom to the conventional single modal approaches. The introduction of preprocessing techniques like normalization and the median filter resulted in better quality of the signal whereas the feature extraction technique namely GF, PCA, and STFT enhanced the representation of emotional features. The classification by the tools NN and SVM gave reliable emotions classification (happiness, sadness, anger, fear, surprise, disgust, neutrality).The superiority of the proposed MBERS over the existing methods was verified using experimental analysis. The results showed that the framework achieved an accuracy of 97.1%, precision of 0.95, recall of 0.94, TNR of 0.98 and reduced FPR of 0.02, which showed that the framework was robust and reliable. Emotion interpretation was greatly enhanced and false classifications were significantly decreased with the multimodal fusion strategy. The suggested framework has promising application potential in the healthcare monitoring, adaptive learning systems, intelligent human computer interaction, as well as emotion-aware technologies.

Future Work

A future study with the aim to improve emotion analysis using deep learning architectures and the inclusion of more physiological sensors is proposed. The proposed system can be readily extended to multi-lingual and multicultural emotion recognition systems. Optimization for real-time deployment in smart wearable devices and IoT devices under portable or on-site conditions can further enhance its practical application, adaptability, and accessibility.

Funding: "This research received no external funding"

Conflicts of Interest: "The authors declare no conflict of interest."

REFERENCES

1. K. M. Rukhiran, N. Phaokla and P. Netinant, "Adoption of environmental information chatbot services based on the Internet of Educational things in smart schools: Structural equation modeling approach", Sustainability, vol. 14, no. 23, pp. 15621, Nov. 2022.
2. D. Shah and V. Haradi, "IoT based biometrics implementation on Raspberry Pi", Procedia Comput. Sci., vol. 79, pp. 328-336, 2016.
3. M. Rukhiran and P. Netinant, "IoT architecture based on information flow diagram for vermiculture smart farming kit", TEM J., vol. 9, no. 4, pp. 1330-1337, Nov. 2020.
4. A. George, A. Ravindran, M. Mendieta and H. Tabkhi, "MEZ: An adaptive messaging system for latency-sensitive multi-camera machine vision at the IoT edge", IEEE Access, vol. 9, pp. 21457-21473, 2021.
5. S. Taheri and J.-S. Yuan, "A cross-layer biometric recognition system for mobile IoT devices", Electronics, vol. 7, no. 2, pp. 26, Feb. 2018.
6. P. Bellmann, P. Thiam and F. Schwenker, "Person identification based on physiological signals: Conditions and risks", Proc. ICPRAM, pp. 373-380, 2020.
7. A. Phokajang and P. Netinant, "Developing software architecture for a smart school digital framework", Proc. ICSIM, pp. 22-27, 2021.
8. R. A. Hamid, U. A. Mokhtar, M. M. Yusof and A. Warland, "Electronic records management in schools: The case study of school examination analysis system", J. Pengurusan, vol. 57, pp. 142-154, May 2019.
9. A. Al Abdulwahid, N. Clarke, I. Stengel, S. Furnell and C. Reich, "Continuous and transparent multimodal authentication: Reviewing the state of the art", Cluster Comput., vol. 19, no. 1, pp. 455-474, Mar. 2016.
10. S. Ayeswarya and J. Norman, "A survey on different continuous authentication systems", Int. J. Biometrics, vol. 11, no. 1, pp. 67, 2019.
11. P. Arias-Cabarcos, C. Krupitzer and C. Becker, "A survey on adaptive authentication", ACM Comput. Surv., vol. 52, no. 4, pp. 1-30, Sep. 2019.
12. M. Singh, R. Singh and A. Ross, "A comprehensive overview of biometric fusion", Inf. Fusion, vol. 52, pp. 187-205, Dec. 2019.
13. I. C. Stylios, O. Thanou, I. Androulidakis and E. Zaitseva, "A review of continuous authentication using behavioral biometrics", Proc. SouthEast Eur. Design Autom. Comput. Eng. Comput. Netw. Social Media Conf. SEEDA-CECNSM, pp. 72-79, 2016.
14. G. Fenu, M. Marras and L. Boratto, "A multi-biometric system for continuous student authentication in e-learning platforms", Pattern Recognit. Lett., vol. 113, pp. 83-92, Oct. 2018.
15. S. Srivastava and P. S. Sudhish, "Continuous multi-biometric user authentication fusion of face recognition and keystoke dynamics", Proc. IEEE Region 10 Humanitarian Technol. Conf. (R-HTC), pp. 1-7, Dec. 2016.
16. W. Meng, D. S. Wong, S. Furnell and J. Zhou, "Surveying the development of biometric user authentication on mobile phones", IEEE Commun. Surveys Tuts., vol. 17, no. 3, pp. 1268-1293, 3rd Quart. 2015.
17. I. Masi, Y. Wu, T. Hassner and P. Natarajan, "Deep face recognition: A survey", Proc. 31st SIBGRAPI Conf. Graph. Patterns Images (SIBGRAPI), pp. 471-478, Oct. 2018.

18. Z. Ming, M. Visani, M. M. Luqman and J.-C. Burie, "A survey on anti-spoofing methods for facial recognition with RGB cameras of generic consumer devices", J. Imag., vol. 6, no. 12, pp. 139, Dec. 2020.

19. X. Ben, Y. Ren, J. Zhang, S.-J. Wang, K. Kpalma, W. Meng, et al., "Video-based facial micro-expression analysis: A survey of datasets features and algorithms", IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 9, pp. 5826-5846, Sep. 2022.

20. G. Castaneda and T. M. Khoshgoftaar, "A survey of 2D face databases", Proc. IEEE Int. Conf. Inf. Reuse Integr., pp. 219-224, Aug. 2015.

21. M. Chihaoui, A. Elkefi, W. Bellil and C. Ben Amar, "A survey of 2D face recognition techniques", Computers, vol. 5, no. 4, pp. 21, Sep. 2016

22. F. Jiang, Y. Liu, H. Si, J. Meng and Q. Li, "Cross-scenario unknown-aware face anti-spoofing with evidential semantic consistency learning", IEEE Trans. Inf. Forensics Security, vol. 19, pp. 3093-3108, 2024.

23. B. Subramanian, J. Kim, M. Maray, and A. Paul, ‘‘Digital twin model: A real-time emotion recognition system for personalized healthcare,’’ IEEE Access, vol. 10, pp. 81155–81165, 2022.

24. A. Santra, V. Rai, D. Das, and S. Kundu, ‘‘Facial expression recognition using convolutional neural network,’’ Int. J. Res. Appl. Sci. Eng. Technol., vol. 10, no. 5, pp. 1081–1092, 2022.

25. G. S. Monisha, G. S. Yogashree, R. Baghyalaksmi, and P. Haritha, ‘‘Enhanced automatic recognition of human emotions using machine learning techniques,’’ Proc. Comput. Sci., vol. 218,

26. pp. 375–382, Jan. 2023. [Online]. Available: https://www.sciencedirect. com/science/article/pii/S1877050923000200

27. X. Jin, Z. Lai, W. Sun, and Z. Jin, ‘‘Facial expression recognition based on depth fusion and discriminative association learning,’’ Neural Process. Lett., vol. 54, pp. 2025–2047, Jan. 2022.

28. C. Zhang, G. Zhou, D. Ma, R. Wang, J. Xiao, and D. Zhao, ‘‘A deep learning-enabled human-cyber-physical fusion method towards human–robot collaborative assembly,’’ Robot. Comput.-Integr. Manuf., vol. 83, Oct. 2023, Art. no. 102571. [Online]. Available: https://www. sciencedirect.com/science/article/pii/S0736584523000479

29. P. Arnau-González, M. Arevalillo-Herráez, and N. Ramzan, ‘‘Fusing highly dimensional energy and connectivity features to identify affective states from EEG signals,’’ Neurocomputing, vol. 244, pp. 81–89, Jun. 2017. [Online]. Available: http://www.sciencedirect. com/science/article/pii/S0925231217305222

30. F. Mazza, M. P. Da Silva, and P. L. Callet, ‘‘Investigating electrophysiology for measuring emotions triggered by audio stimuli,’’ in Proc. Asilomar Conf. Signals, Syst. Comput., Nov. 2013, pp. 1350–1364.

31. H. F. García, M. A. Álvarez, and Á. A. Orozco, ‘‘Dynamic facial landmarking selection for emotion recognition using Gaussian processes,’’ J. Multimodal User Interfaces, vol. 11, no. 4, pp. 327–340, Dec. 2017.

32. N. Chanthaphan, K. Uchimura, T. Satonaka, and T. Makioka, ‘‘Novel facial feature extraction technique for facial emotion recognition system by using depth sensor,’’ Int. J. Innov. Comput. Inf. Control, vol. 12, pp. 2067–2087, Dec. 2016.Z. Zhang, L. Cui, X. Liu, and T. Zhu, ‘‘Emotion detection using Kinect 3D facial points,’’ in Proc. IEEE/WIC/ACM Int. Conf. Web Intell. (WI), Oct. 2016, pp. 407–410.