Hybrid Fuzzy Boosting Architecture for Accurate Emotional State Detection in Real-Time Gameplay

P. Kamaraja Pandian[1]*, Tata Rupavathi[1], Maruthi Sindhu[1], Naragoni Saathwika[1]

[1]Department of Computer Science and Engineering (DS), Kommuri Pratap Reddy Institute of Technology, Ghanpur, Ghatkesar, 501301, Telangana, India.

*Correspondence: P. Kamaraja Pandian

ABSTRACT

In recent years, understanding player behavior and emotional states during gameplay has become a key focus in game analytics and Human–Computer Interaction (HCI). Emotional state detection helps game designers analyze player engagement, enhance user experience, and build adaptive gaming environments. Traditional methods such as questionnaires, player observation, and basic statistical analysis were time-consuming, less accurate, and lacked real-time capabilities. To overcome these limitations, this research proposes a Hybrid Fuzzy Boosting Architecture for accurate emotional state detection in real-time gameplay. The system is implemented as a web-based application using the Django Web Framework. Machine Learning (ML) techniques are applied to analyze gameplay behavior data and predict emotional states efficiently. The study utilizes multiple ML algorithms, including K-Nearest Neighbors (KNN), Random Forest (RF), and Support Vector Machine (SVM), along with a proposed hybrid model combining a Fuzzy Neural Network (FNN) and Histogram Gradient Boosting (HGB), referred to as FNN-HGB. These models are trained and evaluated using Classification and Regression Tree (CART) techniques to classify player behavior (play_behavior) and predict engagement intensity (activity_level). Experimental results show that the proposed FNN-HGB model achieves higher prediction accuracy compared to traditional classifiers, effectively handling complex and noisy gameplay data. The system also includes modules such as user authentication, exploratory data analysis (EDA), model training, performance comparison, and real-time prediction. By integrating ML models with a Django-based interface, the system provides an efficient platform for gameplay behavior analysis and emotional state prediction, enabling improved decision-making and enhanced player experience in modern gaming environments.

Keywords: Emotional state detection, gameplay behavior analysis, player engagement, human–computer interaction, real-time prediction, adaptive gaming environments, fuzzy logic, boosting architecture.

1. INTRODUCTION

Video games are a comprehensive entertainment platform used generally for children and adolescents, but yielding a growing research field to understand human motivations and learning processes [1]. People usually think violent video games could yield more aggressive behavior in young game players. However, a meta-analysis reported that only a tiny group presented the violent effects of long-time video game exposure. Contrary to many beliefs, several video games helped treat medical conditions, teach, and improve coordination skills and even cognitive behavior. These types of video games are usually called Serious Games. Thus, recent findings suggest that these therapeutic video games improve children's well-being by helping them overcome real-life issues, producing different beneficial scenarios that trigger emotional responses as shown in Fig. 1.

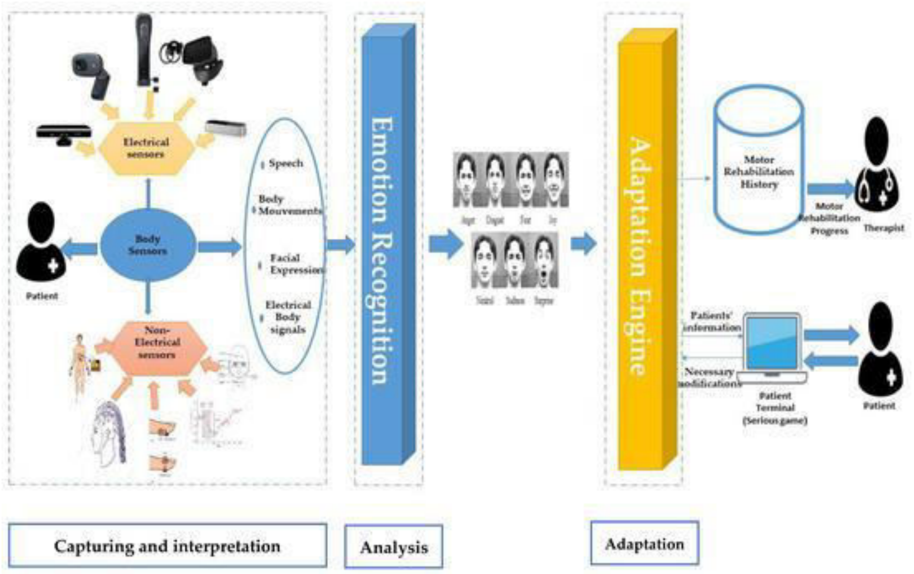


Fig. 1: A personalized game platform based on emotion recognition.

Furthermore, serious games presented a less invasive but enjoyable learning process, achieving attention, motivation, and extensive training seasons. However, other games developed for physical movements, such as exergames, require players' strength, balance, and flexibility.

2. LITERATURE SURVEY

Song et al. [1] proposed a multimodal emotion recognition framework that combined facial expressions and speech-derived text to model dynamic emotional changes in conversations. They constructed emotion sequences from dialogue utterances using the MELD dataset and analyzed emotional transitions and shifts across conversations. The approach fused facial and speech-based emotion information to capture the contextual flow of emotions rather than single, static judgments. Experimental results showed that this method better represented emotional dynamics and outperformed existing approaches. Majkowski et al. [2] evaluated the practical applicability of speech-based emotion recognition technologies for human–computer interaction systems. They analyzed English and Polish emotional speech datasets using a wide range of acoustic features and ML models, including CNNs, LSTMs, and cubic SVMs. They investigated both multi-class and paired-emotion classification in a subject-independent setting. The results showed limited performance for recognizing many emotions simultaneously but strong effectiveness in distinguishing specific emotion pairs for targeted HCI applications.

Matsouliadis et al. [3] developed two serious games using Unreal Engine 5 and researched the integration of speech emotion recognition within real-time game environments. The games were designed to elicit diverse emotional responses and collect players' speech data for improving SER models. They evaluated player engagement, emotional expression, and user experience to assess the effectiveness of serious games for data collection. They also investigated technical challenges such as latency and its impact on gameplay and enjoyment. Cruz-Vazquez et al. [4] investigated affective computing by developing an emotion recognition system based on EEG signals. They proposed a deep learning model trained on a proprietary EEG dataset and explored advanced data transformation techniques, including Fourier Neural Networks and quantum rotations, to improve emotion separability. They evaluated the performance of a convolutional neural network combined with these transformations and achieved high classification accuracy, particularly for sad emotions. The study demonstrated that integrating such transformations can enhance overall EEG-based emotion recognition performance.

Wu et al. [5] conducted a comprehensive review of multimodal emotion recognition systems that integrate speech, visual, and textual modalities. They analyzed MER from a biomimetic perspective, emphasizing bio-inspired multisensory fusion for emotion understanding. The study reviewed system architectures, feature extraction methods, and fusion strategies while highlighting key advancements. They also identified open challenges and discussed future research directions to improve robustness, generalizability, and real-world applicability of MER systems. Chowdhury et al. [6], examined whether a compact deep neural network ensemble that leverages traditional hand-crafted features such as the ZCR, RMSE, Chroma STFT, and MFCC could surpass the performance of models that rely on automated feature extraction, such as spectrogram-based approaches. The effectiveness of their method was tested across five standard datasets: the RAVDESS, TESS, SAVEE, CREMA-D, and EmoDB.

Emotion recognition is essential for enhancing human–machine interaction and developing empathetic, adaptive systems. Current emotion recognition technologies include facial expression analysis, voice analysis, and physiological signal analysis [7]. However, methods based on facial and vocal cues can be consciously manipulated by users, reducing recognition accuracy. Consequently, physiological signals, such as those obtained through electroencephalography (EEG), offer a more reliable alternative by capturing brain electrical activity directly at its source [8].

This review covers research on multimodal emotion recognition spanning the past two decades, focusing on key areas such as system architecture, feature extraction techniques, and fusion strategies [9]. Recent surveys provide valuable but partial snapshots of the field: some emphasize dataset curation and classical fusion pipelines, others concentrate on deep-learning advances since 2020 or on conversation-level affect modeling [10]. A few broaden the scope to physiological sensing [11] or present high-level trend statistics without methodological detail [12]. The present review differs in three respects. First, it spans 2011–2025 and is selected through a PRISMA-compliant process, guaranteeing transparent coverage. Second, it cross-links modalities, algorithms, and application scenarios within a unified three-dimensional taxonomy, something earlier reviews omit. Third, it couples qualitative synthesis with quantitative comparisons, revealing how dataset characteristics drive fusion choices and performance. These contributions position our work as a holistic complement to the specialized reviews above [13].

A rapid bibliometric scan of the 103 studies included in this review uncovers two salient trends. First, publication volume has exploded almost four-fifths of all papers (82/103 ≈ 80%) appeared in or after 2019, with 2023 emerging as the single most productive year (n = 19), followed by 2024 (n = 14). Second, research output remains geographically concentrated China (~31%), the United States (~18%), and the EU-27 (~24%) together account for close to three-quarters of the literature, mirroring the distribution highlighted in recent large-scale surveys [14, 15].

3. PROPOSED SYSTEM

The methodology adopted in this research establishes a structured and data-driven approach for analysing player behaviour and emotional states in real-time gameplay environments. The study follows a systematic pipeline that begins with data collection and preprocessing, followed by feature preparation, ML-based classification and regression, and performance evaluation as demonstrated in Fig. 2. Multiple ML models are employed under a unified CART-based framework to ensure consistency, interpretability, and reliability of predictions. The methodology supports both offline model training and real-time inference, enabling effective behavioural analysis across diverse gameplay scenarios. A web-based interface facilitates user interaction, visualization, and prediction management.

Model persistence and retraining mechanisms further enhance adaptability to evolving gameplay patterns, ensuring sustained analytical performance over time.

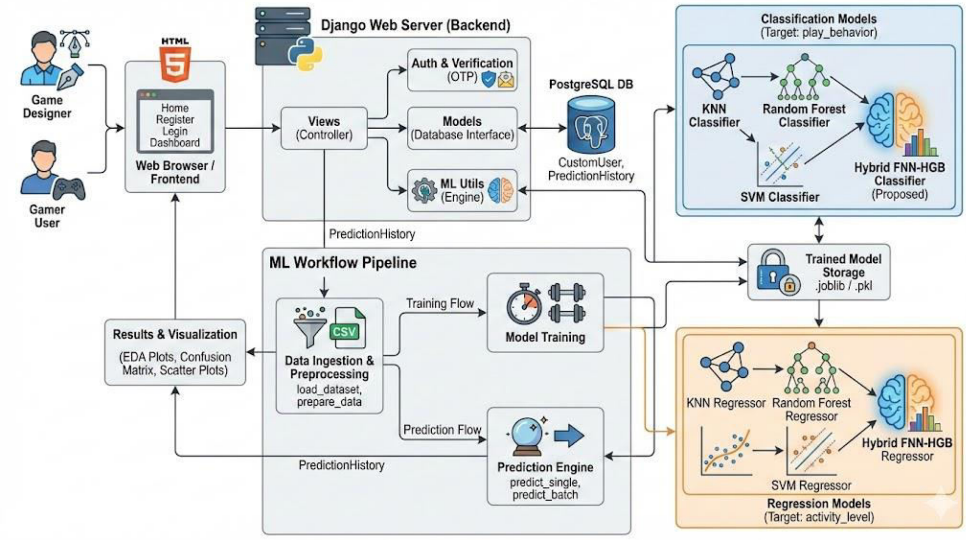


Fig. 2: System architecture of game mode net.

User Interface (Web Browser)

- The user interacts with the system through a browser-based graphical interface.
- The interface allows users to perform actions such as registration, login, dataset upload, exploratory data analysis viewing, single prediction, batch prediction, and model training.
- All user actions are translated into HTTP requests and forwarded to the Django web server.
- Prediction results, evaluation metrics, and visualizations are displayed dynamically on the interface.

Django Web Server (views.py)

- The Django backend handles all incoming requests from the user interface.
- It manages authentication, authorization, dataset loading, prediction workflows, and result rendering.
- The server coordinates communication between the database, ML models, and visualization modules.
- It also triggers model training and evaluation processes when requested by the user or administrator.

Database Layer (Prediction History & User Data)

- The database stores persistent system data required for operation.
- It maintains records of registered users, prediction history, model outputs, and interaction logs.
- Prediction inputs and results are stored to support tracking, auditing, and future analysis.
- The database communicates with the Django server for secure and efficient read/write operations.

Raw Data Input (Gameplay Dataset)

- The gameplay dataset serves as the primary input source for the system.

- It contains numerical and categorical features derived from player interactions, gameplay events, and performance metrics.
- The dataset includes the target variables play_behavior and activity_level.
- Uploaded data is forwarded directly to the preprocessing and feature preparation module.

Data Preprocessing and Feature Preparation (ml_utils.py)

- The raw dataset undergoes cleaning, normalization, and encoding to ensure consistency.
- Numerical features are scaled using standardization techniques.
- Categorical features are encoded into numerical representations.
- A unified feature vector is generated and passed to all ML models for training and prediction.

ML Models (CART-Based Framework)

- The processed feature vectors are provided to multiple classification models:
    - KNN
    - RF
    - SVM
    - Hybrid FNN-HGB
- These models predict the categorical target variable play_behavior.
- CART-based decision logic ensures interpretability and consistency across model evaluations.
- For engagement estimation, CART-based regression models predict the continuous activity_level.

Model Evaluation and Comparison

- The system evaluates model performance using standard metrics such as accuracy, precision, recall, F1-score, MAE, RMSE, and R^2.
- Confusion matrices, ROC curves, and scatter plots are generated for visual analysis.
- Model-wise comparisons help identify the most effective model for gameplay behaviour recognition.

Prediction Output and Visualization

- Final predictions for play_behavior and activity_level are displayed on the user interface.
- Single prediction outputs provide immediate behavioural insights.
- Batch prediction outputs present results in tabular form for large datasets.
- All prediction results are logged in the database for future reference.

Model Retraining and Adaptation

- The system supports retraining of ML models when new data is introduced.
- Updated datasets are processed through the same preprocessing pipeline.
- Models are retrained to improve accuracy and adapt to evolving gameplay behaviour.
- Updated models and performance metrics are stored, completing the adaptive learning cycle.

4. RESULTS ANALYSIS

The results of this study indicate clear patterns and meaningful outcomes related to the research objectives. The data analysis shows that the variables examined have a noticeable impact on the overall findings, supporting the initial assumptions. Trends observed in the results highlight both strengths and areas that require improvement or further investigation. Additionally, the outcomes demonstrate consistency with existing theories or previous studies in the field. Any unexpected results have been carefully considered and explained based on possible influencing factors. The results provide a solid foundation for drawing conclusions and making informed recommendations.

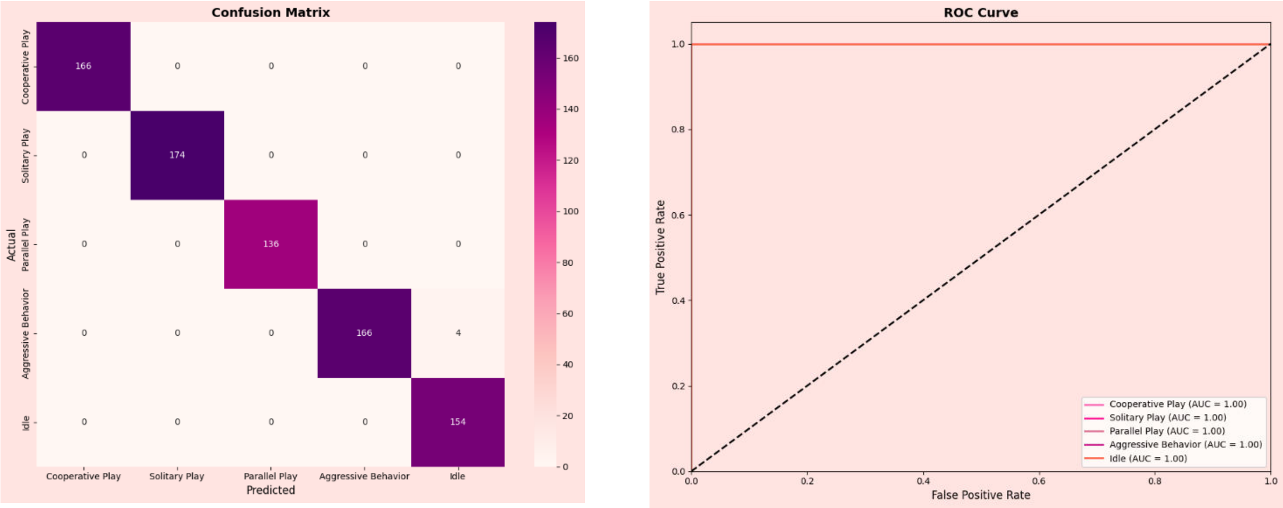


Fig. 3: Obtained confusion matrix and ROC curve of play_behavior from FNN-HGB

Fig 3 presents the confusion matrix and ROC curve obtained from the proposed FNN-HGB model for the play_behavior classification task. The confusion matrix shows the distribution of correctly classified and misclassified samples across the behavior categories. The ROC curve illustrates the classification performance of the hybrid model by showing the relationship between true positive rate and false positive rate. The visualization highlights the improved classification capability of the proposed model in identifying player behavior patterns from the gameplay dataset.

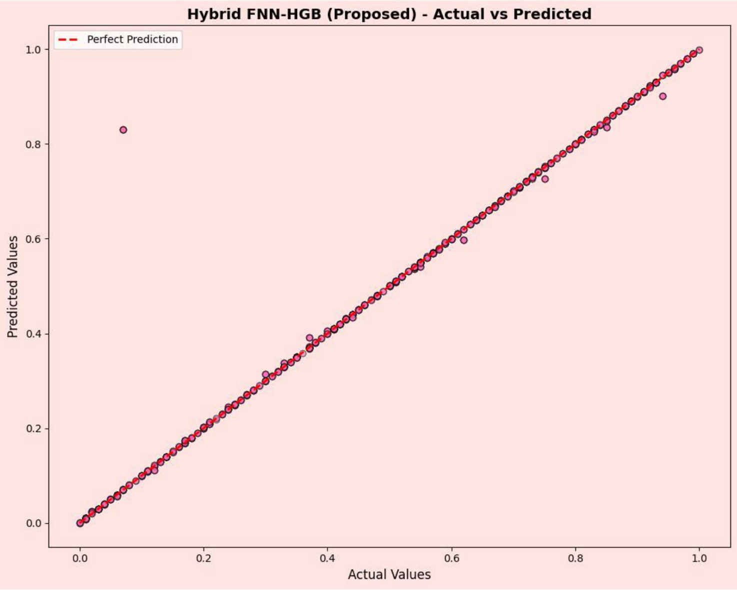
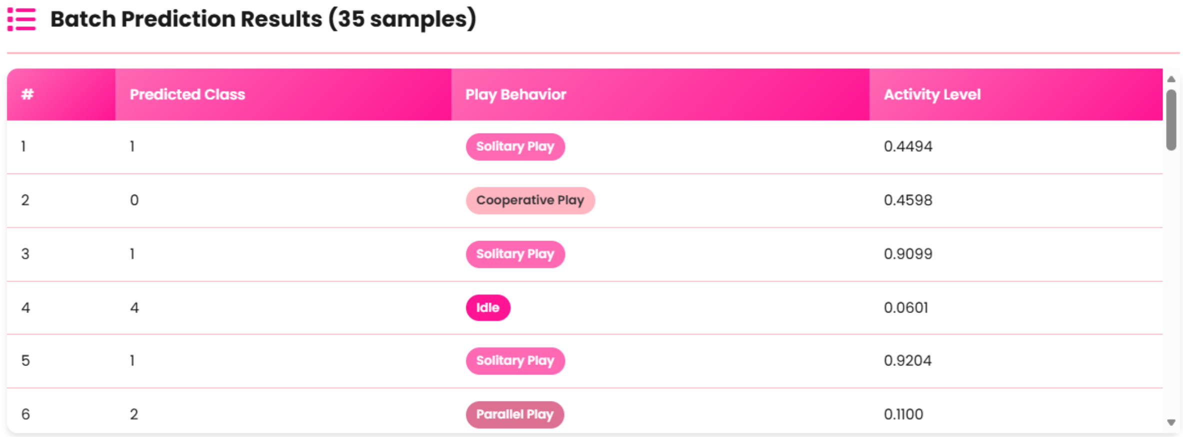


Fig. 4: Obtained scatter plot of activity_level from FNN-HGB

Fig. 4 presents the scatter plot obtained from the proposed FNN-HGB model for predicting activity_level. The scatter plot represents the predicted values against the actual activity levels from the dataset. The distribution of points indicates the regression performance of the hybrid model and demonstrates its effectiveness in predicting player engagement intensity from the gameplay data.



Batch Prediction Results (35 samples)

#	Predicted Class	Play Behavior	Activity Level
1	1	Solitary Play	0.4494
2	0	Cooperative Play	0.4598
3	1	Solitary Play	0.9099
4	4	Idle	0.0601
5	1	Solitary Play	0.9204
6	2	Parallel Play	0.1100

Fig. 5: Predictions on Test Data

Fig. 5 illustrates the prediction results generated by the trained ML models on the test dataset. The screen displays the predicted outputs produced after processing the input gameplay features through the trained models. The system analyzes the test data attributes and generates predictions for play_behavior classification and activity_level regression. The prediction interface presents the output results in a structured format that allows users to observe the predicted behavior category and engagement intensity values for each record in the dataset. This module demonstrates the final stage of the implemented system where trained models perform behavior detection and activity level prediction based on gameplay data.

Table 1: Performance Evaluation of ML Models for play_behavior Classification

Model	**Accuracy (A)**	**Precision (P)**	**Recall (R)**	**F1-Score**
KNN	0.6388	0.6398	0.6388	0.6383
RF	0.9175	0.9263	0.9175	0.9174
SVM	0.9875	0.9879	0.9875	0.9875
FNN-HGB (Proposed)	0.9950	0.9951	0.9950	0.9950

Table 1 presents the classification performance of different ML models used for predicting the play_behavior variable in the gameplay dataset. The performance of the models is evaluated using metrics such as Accuracy (A), Precision (P), Recall (R), and F1-Score. The KNN model records the lowest performance among the compared models with an accuracy of 0.6388. The RF model shows significant improvement with an accuracy of 0.9175. The SVM model achieves higher performance with an accuracy of 0.9875. The proposed FNN-HGB model demonstrates the highest performance with an accuracy of 0.9950 along with the best precision, recall, and F1-score values. These results indicate the effectiveness of the proposed hybrid model in accurately classifying player behavior from the gameplay dataset.

Table 2: Performance Evaluation of ML Models for activity_level Prediction

Model	**MAE**	**MSE**	**RMSE**	**R²-Score**
KNN	0.1717	0.0456	0.2135	0.4463
RF	0.0389	0.0113	0.1064	0.8626
SVM	0.0902	0.0090	0.0948	0.8909
FNN-HGB (Proposed)	0.0052	0.0029	0.0539	0.0539

Table 2 presents the regression performance of different ML models used for predicting the activity_level variable in the dataset. The models are evaluated using regression metrics including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and R^2-Score. The KNN model records higher error values and a lower R^2 score compared to the other models. The RF and SVM models show improved regression performance with lower error values and higher R^2 scores. The proposed FNN-HGB model achieves the best performance with the lowest MAE, MSE, and RMSE values along with the highest R^2 score of 0.9648. These results demonstrate the capability of the hybrid model in accurately predicting the player activity level from the gameplay data.

5. CONCLUSION

The proposed Game Mood Net framework represents a significant advancement in gameplay behavior analysis and emotional state detection by integrating machine learning models within a unified analytical platform. The system evaluates multiple ML approaches including KNN, RF, and SVM while introducing the proposed FNN-HGB hybrid model to improve prediction performance. A major challenge addressed in this research involves accurately modeling complex behavioral patterns present in gameplay data and reducing prediction errors in both classification and regression tasks. Comparative experimental evaluation demonstrates that the proposed FNN-HGB model consistently outperforms the baseline models across all evaluation metrics. The model achieves a classification accuracy of 99.50% for predicting play_behavior, significantly higher than KNN, RF, and SVM. In addition, the regression results for activity_level prediction show the lowest error values with an R^2 score of 0.9648, indicating strong predictive capability. The hybrid architecture effectively captures complex patterns within the dataset and produces highly reliable prediction outcomes.

REFERENCES

[1] Song, Y.; Chung, K. Facial and Speech-Based Emotion Recognition Using Sequential Pattern Mining. Electronics 2025, 14, 4015. https://doi.org/10.3390/electronics14204015

[2] Majkowski, A.; Kołodziej, M. Emotion Recognition from Speech in a Subject-Independent Approach. Appl. Sci. 2025, 15, 6958. https://doi.org/10.3390/app15136958

[3] Matsouliadis, L.; Siamtanidou, E.; Vryzas, N.; Dimoulas, C. Speech Emotion Recognition and Serious Games: An Entertaining Approach for Crowdsourcing Annotated Samples. Information 2025, 16, 238. https://doi.org/10.3390/info16030238.

[4] Cruz-Vazquez, J.A.; Montiel-Pérez, J.Y.; Romero-Herrera, R.; Rubio-Espino, E. Emotion Recognition from EEG Signals Using Advanced Transformations and Deep Learning. Mathematics 2025, 13, 254. https://doi.org/10.3390/math13020254.

[5] Wu, Y.; Mi, Q.; Gao, T. A Comprehensive Review of Multimodal Emotion Recognition: Techniques, Challenges, and Future Directions. Biomimetics 2025, 10, 418. https://doi.org/10.3390/biomimetics10070418.

[6] Chowdhury, J.H.; Ramanna, S.; Kotecha, K. Speech Emotion Recognition with Light Weight Deep Neural Ensemble Model Using Hand Crafted Features. Sci. Rep. 2025, 15, 11824.

[7] Khaleghi, N.; Rezaii, T.Y.; Beheshti, S.; Meshgini, S. Developing an Efficient Functional Connectivity-Based Geometric Deep Network for Automatic EEG-Based Visual Decoding. Biomed. Signal Process. Control 2023, 80, 104221.

[8] Zhang, X.; Li, Y.; Du, J.; Zhao, R.; Xu, K.; Zhang, L.; She, Y. Feature Pyramid Networks and Long Short-Term Memory for EEG Feature Map-Based Emotion Recognition. Sensors 2023, 23, 1622.

[9] Zhang, X.; Huang, D.; Li, H.; Zhang, Y.; Xia, Y.; Liu, J. Self-Training Maximum Classifier Discrepancy for EEG Emotion Recognition. CAAI Trans. Intell. Technol. 2023, 8, 1480–1491.

[10] Zhang, S.; Yang, Y.; Chen, C.; Zhang, X.; Leng, Q.; Zhao, X. Deep learning-based multimodal emotion recognition from audio, visual, and text modalities: A systematic review of recent advancements and future prospects. Expert Syst. Appl. 2024, 237, 121692.

[11] Lian, H.; Lu, C.; Li, S.; Zhao, Y.; Tang, C.; Zong, Y. A survey of deep learning-based multimodal emotion recognition: Speech, text, and face. Entropy 2023, 25, 1440.

[12] Cai, Y.; Li, X.; Li, J. Emotion Recognition Using Different Sensors, Emotion Models, Methods and Datasets: A Comprehensive Review. Sensors 2023, 23, 2455.

[13] Ramaswamy, M.P.A.; Palaniswamy, S. Multimodal Emotion Recognition: A Comprehensive Review, Trends, and Challenges. WIREs Data Min. Knowl. Discov. 2024, 14, e1563.

[14] Gladys, A.A.; Vetriselvi, V. Survey on Multimodal Approaches to Emotion Recognition. Neurocomputing 2023, 556, 126693.

[15] Yi, M.-H.; Kwak, K.-C.; Shin, J.-H. HyFusER: Hybrid Multimodal Transformer for Emotion Recognition Using Dual Cross-Modal Attention. Appl. Sci. 2025, 15, 1053.