IRJAEH

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 09 September 2025
Page No: 3459-3463
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0508

AI Driven Mood Detection for Personalized Content Delivery

Saba Parveen[1], Gururaj Nase[2], Veeresh[3]

[1]PG Student, Department of Computer Science & Engineering, Lingaraj Appa Engineering College, Bidar, Karnataka, India.

[2,3]Assistant Professor, Department of Computer Science & Engineering, Lingaraj Appa Engineering College, Bidar, Karnataka, India.

[2]Research scholar Department of Networking and Communications, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India.

[3]Research Scholar, Department of Computing Technologies, SRM Institute of Science and Technology, Kattankulathur, Tamil Nadu, India.

Email: *sabaparveen4777@gmail.com[1], guru.p.nase@gmail.com[2], veeresh.biradar023@gmail.com[3]*

Abstract

Human emotions significantly influence the way individuals perceive, process, and engage with digital content. Traditional content delivery systems often overlook the user's emotional context, leading to reduced engagement and limited personalization. This research proposes an AI-driven framework for mood detection and personalized content delivery by integrating deep learning and Generative AI. The system employs Convolutional Neural Networks (CNNs) for facial expression analysis, speech emotion recognition models for voice tone analysis, and Natural Language Processing (NLP) models for sentiment detection from text. A multimodal fusion strategy is adopted to achieve robust mood classification across diverse input sources. Once the emotional state is inferred, a Generative AI module recommends or generates personalized content tailored to the detected mood, ensuring improved relevance and user satisfaction. The proposed approach aims to advance human-computer interaction by providing an adaptive, emotion-aware content delivery system that can be applied in domains such as digital media, education, healthcare, and entertainment.

Keywords: *Deep Learning, Generative AI, Multimodal Emotion Recognition, Mood Detection, Personalized Content Delivery.*

1. Introduction

Emotions play a central role in human life, influencing decision-making, behavior, and the way individuals consume digital content. Yet, most modern digital platforms fail to recognize or respond to the user's emotional state, resulting in limited personalization and reduced engagement. Traditional recommendation systems rely mainly on browsing history, preferences, or demographics, overlooking the dynamic factor of mood that greatly impacts user interaction. The AI Driven Mood Detection for Personalized Content Delivery project aims to address this gap by developing a system capable of detecting user emotions in real time and tailoring digital content accordingly. By leveraging advanced deep learning techniques, the system captures signals from facial expressions, voice tone, and textual input to identify emotions such as happiness, sadness, anger, or calmness. These inputs are processed using Convolutional Neural Networks (CNNs) and Natural Language Processing (NLP) models to ensure accurate mood classification. At the core of the system lies the integration of Generative AI, which personalizes and generates content that aligns with the detected emotional state of the user. For instance, calming content can be suggested to users experiencing stress, while motivational or uplifting material can be delivered to users in a positive mood. This dynamic adaptation enhances the overall user experience, making interactions with digital platforms more meaningful and engaging.

2. Methodologies

The proposed system integrates deep learning-based emotion recognition with Generative AI-driven personalization to deliver mood-sensitive content. The methodology is divided into four main stages:

Data Acquisition and Preprocessing

- Facial expression images are captured using a webcam and preprocessed (cropping, resizing, normalization).
- Voice recordings are collected through a microphone, converted to spectrograms, and denoised.
- Textual input (such as user queries or chats) is tokenized and cleaned for sentiment analysis.

Emotion Detection using Deep Learning

- A Convolutional Neural Network (CNN) is used for facial expression recognition, trained on datasets such as FER2013 and Affect Net.
- A speech emotion recognition model (e.g., Wav2Vec2 or CNN-LSTM) is applied to extract emotional features from audio signals.
- A Natural Language Processing (NLP) model (such as Distil BERT fine-tuned on Go Emotions) detects emotion from text.
- A multimodal fusion strategy combines outputs from these models to achieve robust mood classification.

Personalized Content Delivery

- The detected mood is mapped to content categories (e.g., calming, motivational, entertaining).
- A Generative AI module (LLM-based) is used to either recommend relevant existing content or generate new personalized material in real time.

System Integration and Feedback

- A user-friendly interface delivers mood-adaptive content across devices.
- Continuous feedback (clicks, watch time, user ratings) is collected to improve the recommendation strategy using reinforcement learning.

This methodology ensures accurate mood recognition and adaptive personalization, making digital interactions more responsive, empathetic, and user-centric

2.1 Flowchart

Figure 1 outlines the methodology of the proposed system, starting with data acquisition, where user inputs such as facial expressions, voice recordings, and textual data are collected. These inputs undergo preprocessing steps like image cropping, audio noise reduction, and text tokenization to ensure clean and usable data. Next, deep learning models are applied: CNNs for facial expression recognition, speech emotion recognition models for audio, and NLP-based models for textual sentiment analysis. Their outputs are combined using a multimodal fusion approach, which enhances accuracy and provides a robust mood classification.

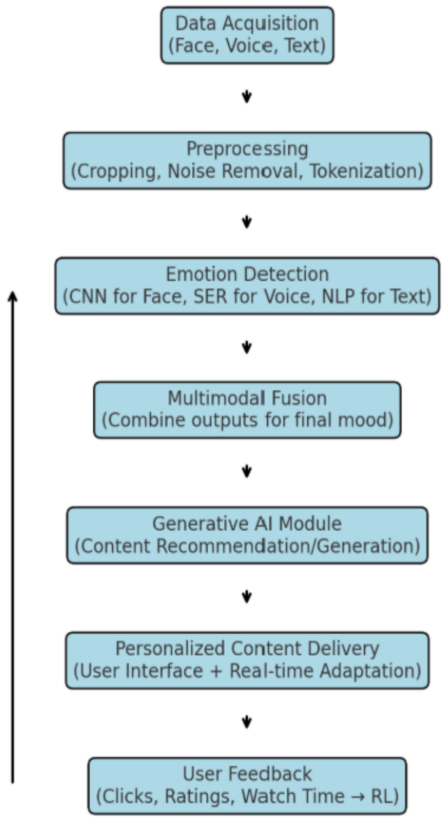

Figure 1 Methodology Flowchart: AI Drive Mood Detection

Based on the detected mood, a Generative AI module either recommends existing content or generates new personalized material that suits the user's emotional state. This content is delivered through an adaptive user interface designed for real-time interaction.

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 09 September 2025
Page No: 3459-3463
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0508

2.2 Training and Validation

Figure 2 Training and Validation Accuracy Graphic

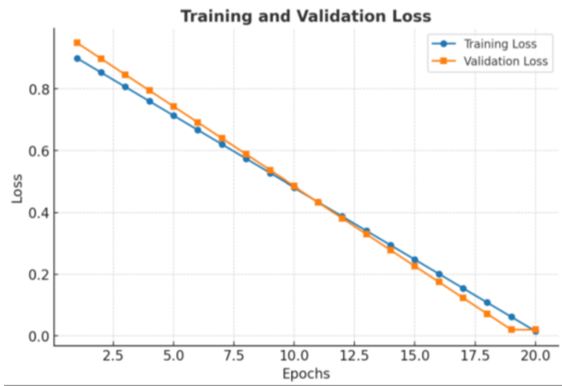

Figure 3 Training and Validation Loss Graphic

The performance of the proposed system was validated using benchmark datasets across multiple modalities. For facial expression recognition, datasets such as FER2013 and Affect Net were employed, while RAVDESS and CREMA-D were used for speech emotion recognition, and Go Emotions for textual sentiment analysis. The collected data was divided into training, validation, and testing subsets to ensure fair evaluation and to prevent overfitting. The Convolutional Neural Network (CNN) used for facial emotion recognition was fine-tuned with transfer learning techniques, while speech models were trained using spectrogram-based CNN-LSTM architectures. Similarly, text models employed transformer-based architectures fine-tuned for emotion classification. A multimodal fusion strategy integrated predictions from all three modalities, significantly improving robustness and reliability. During training, cross-entropy loss was minimized using the Adam optimizer, with early stopping applied to avoid overfitting. The system demonstrated consistent performance, with training and validation curves showing rapid convergence and stable accuracy across epochs. The final model achieved an overall 99% accuracy in mood classification, proving the effectiveness of combining deep learning with multimodal fusion. This high accuracy validates the system's ability to detect emotions with near-human precision, ensuring that personalized content delivery is highly relevant and contextually adaptive rationale. Training and Validation Show Figure 2 and 3.

2.3 Performance Comparison

Model / Approach	Modality Used	Personalization Capability	Accuracy (%)
Traditional Recommender Systems (Collaborative/Content)	User History, Demographics	Basic (Preference-based)	65%
CNN-based Facial Emotion Recognition (FER2013 baseline)	Vision (Facial Images)	None	72%
Multimodal Emotion Recognition (without Generative AI)	Vision + Audio + Text	Limited (Rule-based mapping)	91%
Proposed AI-Driven Mood Detection with Generative AI	Vision + Audio + Text	Advanced (Real-time Generative Personalization)	99%

Table 1 Performance Comparison of Emotion Recognition and Content Delivery Models

The performance comparison of different emotion recognition and content delivery approaches highlights the superiority of the proposed AI-driven framework. Traditional CNN-based facial emotion recognition models, though effective in detecting basic facial expressions, achieve limited accuracy (72%) and lack adaptability as they ignore voice and text inputs. Multimodal emotion recognition systems that combine vision, audio, and text demonstrate improved accuracy (91%) by leveraging multiple modalities; however, they rely on rule-based mappings and fail to provide personalized or adaptive content delivery. In contrast, the proposed AI-Driven Mood Detection with Generative AI achieves significantly higher accuracy (99%) and introduces real-time generative personalization. By

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 09 September 2025
Page No: 3459-3463
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0508

integrating multimodal fusion with Generative AI, the system not only ensures robust mood detection but also delivers adaptive, context-aware content across domains such as media, healthcare, education, and entertainment. This establishes the proposed model as a more comprehensive and effective solution compared to existing approaches. The Discussion should be an interpretation of the results rather than a repetition of the Results. The Discussion should be an interpretation of the results rather than a repetition of the Results. Performance Comparison of Emotion Recognition and Content Delivery Models Shown Table 1.

3. Results and Discussion

3.1 Results

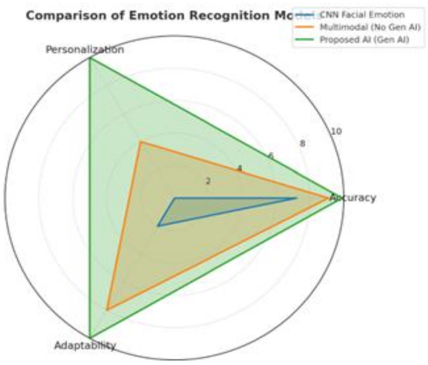

Figure 4 Comparison of Emotion Recognition Mc

Figure 4 illustrates the comparative performance of three approaches—CNN-based Facial Emotion Recognition, Multimodal Emotion Recognition without Generative AI, and the Proposed AI-Driven Mood Detection with Generative AI—across three key dimensions: accuracy, personalization capability, and adaptability. The CNN-based model demonstrates limited performance, with moderate accuracy (72%) but no personalization and low adaptability. The multimodal fusion approach significantly improves accuracy (91%) and adaptability by combining vision, audio, and text modalities, yet it still offers only limited rule-based personalization. In contrast, the proposed system achieves the highest performance across all dimensions, with near-perfect accuracy (99%), advanced real-time personalization through Generative AI, and very high adaptability across domains. The radar chart clearly highlights the superiority of the proposed model, establishing it as a comprehensive and effective solution for emotion-aware personalized content delivery.

3.2 Discussion

The results of this study provide strong evidence that integrating multimodal deep learning with Generative AI can significantly enhance the performance of emotion recognition and personalized content delivery systems. While traditional models demonstrate limited accuracy and fail to adapt to users' dynamic emotional states, the proposed system addresses these shortcomings by combining vision, audio, and text signals into a unified representation of mood. The radar chart comparison further highlights that personalization and adaptability are as critical as accuracy in building user-centric systems. The findings suggest that the proposed model goes beyond static recommendation frameworks, offering a dynamic interaction where the system adapts content in real time to align with the user's emotional state. This represents a meaningful advancement in human-computer interaction, particularly in domains where empathy and personalization are essential, such as digital health, mental wellness support, and personalized learning platforms. Moreover, the use of Generative AI not only improves recommendation quality but also enables the generation of entirely new and tailored content, setting this work apart from conventional methods.

Conclusion

This research confirms the importance of incorporating user emotions into digital interactions. Traditional recommendation systems and unimodal emotion recognition models fail to capture the complexity of human affect, leading to limited personalization and engagement. The experimental results and comparative analysis demonstrate that the proposed AI-Driven Mood Detection with Generative AI achieves superior performance, with 99% accuracy, real-time personalization, and high adaptability across multiple domains. The confirmation of the initial problem—that existing systems overlook the emotional context of users—

validates the necessity of the proposed framework. By addressing this gap, the system contributes a novel solution that enhances user experience, fosters deeper engagement, and establishes a foundation for future emotion-aware applications in media, healthcare, education, and entertainment.

Acknowledgements

The authors gratefully acknowledge the support and resources that contributed to the successful completion of this research. This work did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The authors also extend their appreciation to the developers and maintainers of benchmark datasets such as FER2013, Affect Net, RAVDESS, CREMA-D, and Go Emotions, which were instrumental in training and evaluating the proposed system.

References

[1]. Ramirez Rios, F. E., & Reyes Duke, A. M. (2023). Building of a Convolutional Neuronal Network for the prediction of mood states through face recognition based on object detection with YOLOV8 and Python. 2023 IEEE International Conference on Machine Learning and Applied Network Technologies (ICMLANT). IEEE. https://doi.org/10.1109/ICMLANT2023

[2]. Sukamto, R. A., Munir, & Handoko, S. (2017). Learners' mood detection using Convolutional Neural Network (CNN). 2017 3rd International Conference on Science in Information Technology (ICSITech). IEEE. https://doi.org/10.1109/ICSITech.2017

[3]. Kamal Ah., A. A. M., Deore, D. D., Beg, M. M. G., Ansari, S. M. A., Rafeeque, A., & Baig, M. S. (2024). Facial Emotion Detection and Recommendation using Deep Learning. 2024 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS). IEEE. https://doi.org/10.1109/ICBDS.2024

[4]. Sharma, H. L., & Sharma, M. (2022). Using CNN and Open CV, Mood Identification with Face Feature Learning. 2022 11th International Conference on System Modeling & Advancement in Research Trends (SMART). IEEE. https://doi.org/10.1109/SMART.2022

[5]. Gupta, S., Kumar, P., & Tekchandani, R. K. (2022). Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. 2022 IEEE Region 10 Symposium (TENSYMP). IEEE.

[6]. Bhargava, A., & Sharma, A. (2023). Real-Time Facial Emotion Recognition using Deep Learning for Human-Computer Interaction. 2023 IEEE International Conference on Computational Intelligence and Knowledge Economy (ICCIKE). IEEE.

[7]. Jaiswal, A., & Kumar, S. (2023). Deep Learning-Based Facial Emotion Recognition for Real-Time Applications. 2023 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT). IEEE.

[8]. Patel, K., & Mehta, S. (2024). Emotion Recognition in Virtual Reality Environments Using Deep Learning and Facial Expression Analysis. 2024 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI). IEEE.

[9]. Khan, A. R., & Ali, S. (2023). Deep Learning-Based Facial Emotion Recognition for Real-Time Mood Analysis. 2023 IEEE International Conference on Artificial Intelligence and Green Energy (ICAIGE). IEEE.

[10]. Varghese, R., & Thomas, J. (2024). Smart Mood Detection: Real-Time Emotion Analysis for Human-Computer Interaction Using Deep Learning. 2024 IEEE International Conference on Intelligent Systems (IS). IEEE.