Innovation and Integrative Research Center Journal
ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-4 | April-2026 | Page 292-301

Mood Melody: Real-Time Emotion-Based Music Recommendation System Using Facial Expression Recognition

[1]Mr. Subham Kumar Das, [2]Mr. Ayush Sahoo, [3]Ms. Mausamita Barai,
[4]Dr. Poonam Mishra, [5]Dr. Kranti Kumar Dewangan
[1,2,3]B. Tech Computer Science & Engineering,
[4]Associate Professor, [5]Head, Department of Computer Science & Engineering
[1,2,3,4,5]Amity University Chhattisgarh, Raipur, India- 493225

Abstract

Mood Melody is an intelligent system designed to detect human emotions in real time using facial expression analysis and provide personalized music recommendations. The system captures live video through a webcam, processes facial features using computer vision techniques, and classifies emotions such as happy, sad, angry, neutral, romantic, party, and stress. Based on the detected emotion, the system retrieves and plays relevant music using the Spotify API.

The methodology integrates OpenCV for image processing and deep learning models for emotion classification. The results demonstrate effective real-time emotion detection and dynamic music recommendation, improving user experience. The system highlights the potential of artificial intelligence in creating adaptive and emotion-aware applications.

Keywords: *Emotion Recognition, Music Recommendation System, Computer Vision, Deep Learning, OpenCV, Spotify API, Human-Computer Interaction*

1. Introduction

Music has a deep and well-established connection with human emotions, influencing mood, behavior, and even cognitive states. People often rely on music to relax, energize, focus, or cope with stress, making it an integral part of daily life. However, traditional music recommendation systems primarily depend on past listening history, user preferences, or manual input [1][2]. While these approaches are useful, they often fail to capture the user’s current emotional state, which can change dynamically throughout the day. As a result, the recommendations may not always align with what the user truly needs in a given moment.

With the rapid advancement of artificial intelligence and machine learning, it has become possible to design systems that go beyond static preferences and respond to real-time human emotions. Technologies such as TensorFlow and OpenCV enable systems to analyze visual data, interpret facial expressions, and classify emotions with considerable accuracy. This opens up new possibilities for creating intelligent applications that can adapt content dynamically based on how users feel, rather than relying solely on historical data.

Mood Melody addresses this limitation by integrating facial emotion recognition with music recommendation into a single cohesive system. The application captures facial expressions through a webcam, processes them in real time, and identifies the user's emotional state using deep learning techniques. Based on this detected emotion, the system automatically recommends and plays suitable music, eliminating the need for manual selection. This creates a seamless and intuitive user experience, where the system proactively responds to the user's mood.

The primary focus of this research is to develop an efficient, real-time system that enhances interaction between users and digital music platforms. By combining artificial intelligence with user-centric design, Mood Melody demonstrates how technology can evolve from passive tools into adaptive systems that understand and respond to human emotions, ultimately delivering a more personalized and engaging experience[3].

2. Literature Review

Traditional music recommendation systems are largely built on collaborative filtering and content-based filtering techniques. These methods analyze user behavior, preferences, and listening history to suggest relevant songs. While they are effective in identifying long-term interests, they lack the ability to adapt to a user's immediate emotional state. As a result, the recommendations may not always align with how the user feels at a particular moment, limiting their effectiveness in dynamic, real-time scenarios.

In recent years, research has shifted toward emotion-aware recommendation systems that aim to bridge this gap. These systems utilize facial recognition and computer vision techniques to detect emotions directly from user expressions. Tools like OpenCV are commonly used for face detection and image preprocessing, while machine learning models classify emotions based on extracted features. This approach enables systems to respond to real-time emotional changes, making recommendations more relevant and personalized.

The introduction of deep learning, especially Convolutional Neural Networks (CNNs), has significantly enhanced the accuracy of emotion detection. Frameworks such as TensorFlow and Keras allow models to learn complex patterns from facial data, improving their ability to recognize subtle emotional cues[4]. These advancements have made emotion-based systems more reliable and suitable for real-time applications.

Despite these improvements, several challenges remain. Many existing studies rely on limited or non-diverse datasets, which can reduce model generalization in real-world conditions. Additionally, most systems focus primarily on emotion detection without incorporating deeper personalization based on user preferences or behavior. Practical implementation also presents challenges, such as handling varying lighting conditions, ensuring real-time performance, and maintaining user privacy.

ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-4 | April-2026 | Page 292-301

Mood Melody builds upon these existing approaches by combining real-time emotion detection with live music streaming through the Spotify API. By integrating accurate emotion recognition with dynamic content delivery, the system addresses both technical and usability gaps. It not only detects user emotions in real time but also translates them into meaningful and personalized music recommendations, offering a more complete and practical solution compared to traditional systems[5].

3. Methodology

The Mood Melody system follows a structured and sequential workflow that transforms raw visual input into meaningful music recommendations. This pipeline consists of data acquisition, preprocessing, emotion detection, emotion mapping, and output delivery, ensuring smooth and real-time system performance.

3.1 Data Acquisition

The process begins with capturing real-time video input through the user's webcam. Using computer vision techniques provided by OpenCV, the system continuously reads video frames and detects the presence of a human face. Once a face is identified, the relevant facial region is extracted from each frame. This step ensures that only useful visual data is passed forward for analysis, reducing noise and improving efficiency.

3.2 Preprocessing

Before feeding the extracted facial images into the model, preprocessing is applied to standardize the input. The images are converted into grayscale to remove unnecessary color information while preserving essential facial features. They are then resized to match the input dimensions required by the model and normalized to scale pixel values into a consistent range. These steps improve computational efficiency and help the model perform more accurate and stable predictions[6].

3.3 Emotion Detection

The pre-processed images are passed into a trained Convolutional Neural Network (CNN) model developed using TensorFlow and Keras. The model analyzes facial features such as eyes, eyebrows, and mouth patterns to classify emotions. It identifies categories including happy, sad, angry, neutral, romantic, party, and stress. This step is the core intelligence of the system, enabling it to interpret human emotions accurately.

3.4 Emotion Mapping

Once an emotion is detected, it is mapped to a predefined music category. This mapping is based on logical associations between emotional states and musical moods. For example, a happy emotion may correspond to energetic or upbeat music, while a sad emotion may be linked to calm or soothing tracks. This step bridges the gap between emotion recognition and music recommendation.

3.5 Music Recommendation

After mapping the emotion, the system communicates with the Spotify API to fetch relevant songs or playlists. The API enables real-time retrieval of music data, including track details and playback options. This ensures that recommendations are dynamic, up-to-date, and aligned with the user's current emotional state[7].

3.6 Output Display

Finally, the system presents the results through the user interface. The detected emotion is displayed in real time, along with the recommended songs. Music playback is initiated automatically, providing a seamless and interactive experience. The user can also control playback through options such as play, pause, and skip, ensuring both automation and flexibility in interaction.

4. System Architecture

The Mood Melody system is built using a modular architecture that separates responsibilities into distinct layers, allowing each component to function independently while maintaining seamless communication with others. This design improves flexibility, scalability, and ease of maintenance, ensuring that updates or enhancements in one layer do not disrupt the entire system.

The frontend layer acts as the user-facing component, responsible for displaying the webcam feed, detected emotion, and music playback controls. Built using web technologies or frameworks like Streamlit, it captures real-time input from the user and presents outputs in an intuitive and responsive interface. This layer ensures smooth interaction by continuously sending captured frames to the backend and updating results dynamically.

The backend layer serves as the core processing unit of the system. It receives input data from the frontend and performs operations such as face detection, preprocessing, and communication with external services. Using libraries like OpenCV, the backend extracts facial features and prepares the data for analysis. It also handles communication with external APIs, ensuring efficient data flow and real-time response.

The AI model layer is responsible for performing facial emotion recognition. This layer utilizes deep learning models developed with TensorFlow and Keras. The model analyzes processed facial data and classifies emotions into predefined categories. This layer forms the intelligence of the system, enabling accurate interpretation of user emotions[8].

The API integration layer connects the system to external music services such as the Spotify API. It handles authentication, sends requests for music data, and retrieves relevant songs or

Innovation and Integrative Research Center Journal
ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-4 | April-2026 | Page 292-301

playlists based on detected emotions. This layer ensures that music recommendations are dynamic, up-to-date, and aligned with user mood.

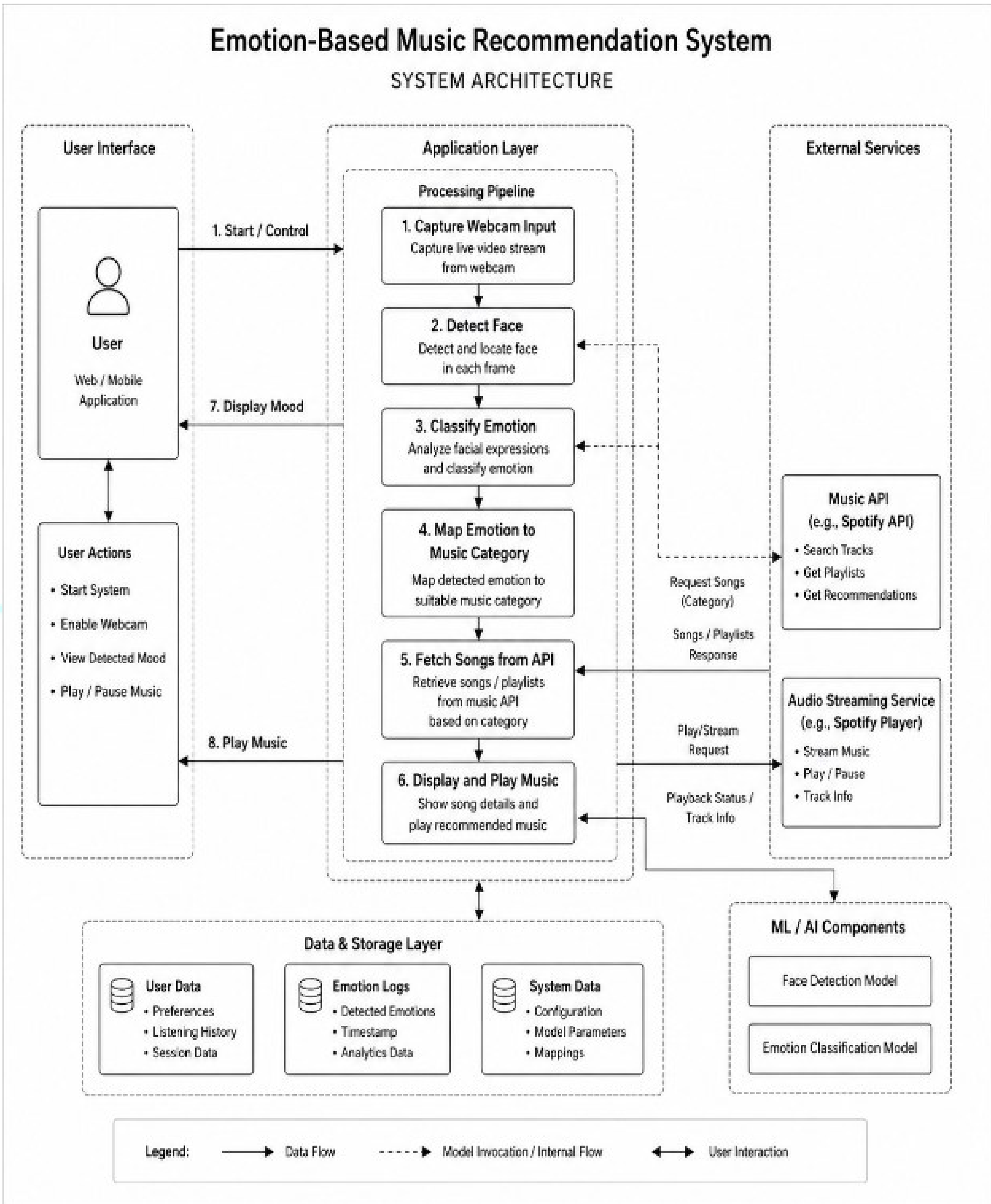

Figure 1: System Architecture

Overall, the architecture ensures smooth interaction between all components, creating a continuous workflow from input capture to output delivery. By maintaining clear separation of concerns and efficient communication between layers, the system achieves real-time processing, reliability, and scalability [3].

5. Results and Evaluation

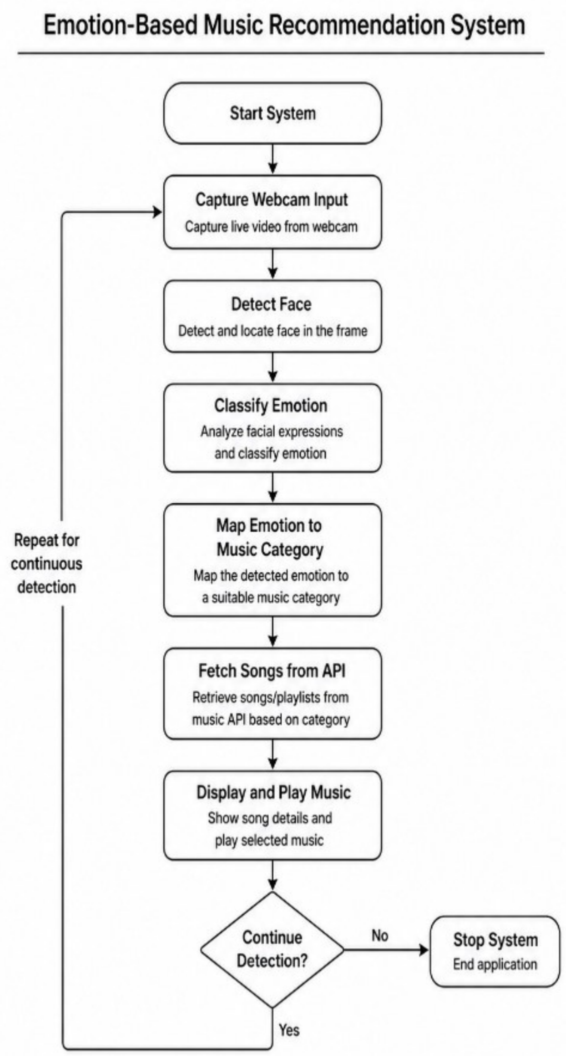

Figure 2: Flowchart and use case diagram

The Mood Melody system demonstrates effective performance in real-time emotion detection and music recommendation under standard operating conditions. By utilizing computer vision techniques through OpenCV and deep learning models built with TensorFlow and Keras, the system is able to accurately detect facial expressions and classify emotions with satisfactory reliability[4]. The Convolutional Neural Network (CNN) model performs particularly well in well-lit environments, where facial features are clearly visible, leading to higher prediction accuracy.

Innovation and Integrative Research Center Journal
ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-4 | April-2026 | Page 292-301

One of the key strengths observed during evaluation is the system's minimal response time. The processing pipeline—from capturing video frames to displaying results—operates efficiently, ensuring a smooth and seamless user experience. Real-time feedback allows users to see immediate changes in detected emotions and corresponding music recommendations, enhancing interactivity and engagement.

The integration with the Spotify API further strengthens the system's practical usability. It enables quick retrieval of songs and playlists based on detected emotions, ensuring that recommendations are dynamic and relevant. The predefined mapping between emotions and music categories proves effective in delivering context-aware suggestions that align with user mood.

However, the evaluation also reveals certain limitations. The system's performance may decline under challenging conditions such as poor lighting, partial face visibility, or occlusions. In such scenarios, face detection and emotion classification accuracy can be affected, leading to less reliable outputs. These findings indicate areas for future improvement, particularly in enhancing robustness and adaptability in real-world environments[8].

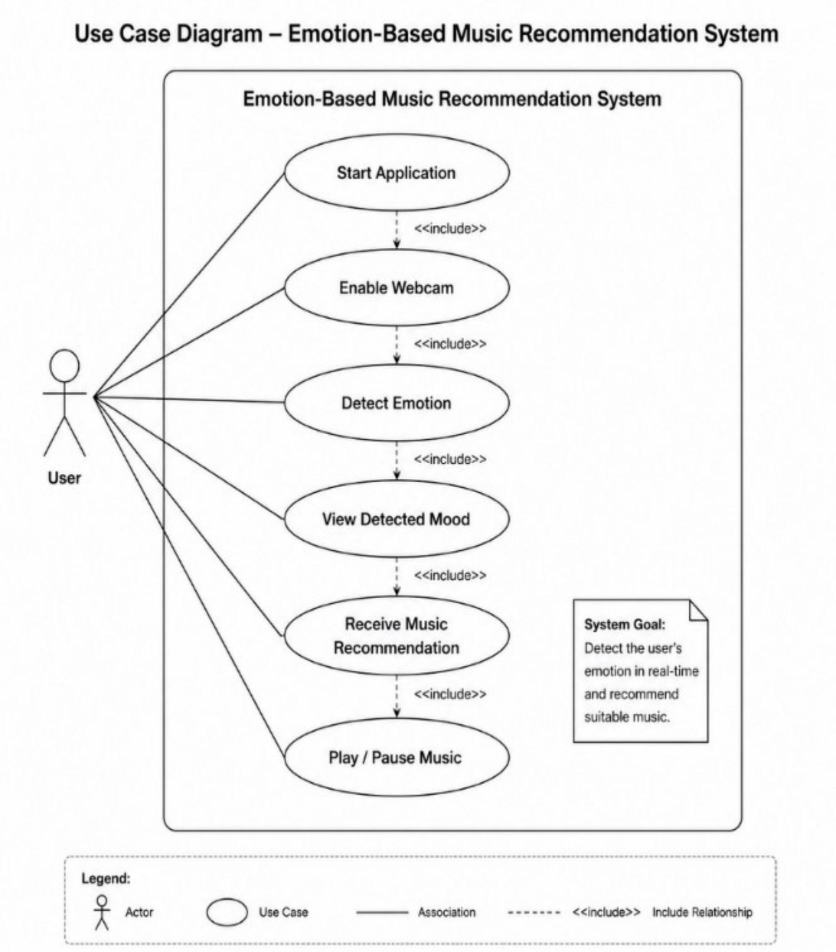

Figure 3: User case diagram

Overall, the results confirm that Mood Melody is capable of delivering efficient, real-time emotion-based music recommendations, while also highlighting opportunities for further optimization and refinement.

6. Discussion

The Mood Melody project demonstrates the practical value of integrating artificial intelligence into real-world applications, particularly in enhancing user experience through personalization. By combining facial emotion recognition with music recommendation, the system moves beyond traditional approaches that rely solely on historical data or manual input. Instead, it introduces a more intuitive interaction model where the system adapts dynamically to the user's current emotional state. Technologies such as TensorFlow, Keras, and OpenCV enable this real-time adaptability, making the system responsive and user-centric.

The discussion also brings attention to several challenges that must be addressed for large-scale deployment. One major concern is dataset bias, where the training data may not represent diverse facial features, expressions, or demographics, potentially affecting model accuracy. Environmental factors such as lighting conditions, camera quality, and background noise can also impact system performance, especially in real-world scenarios. Additionally, privacy and data security remain critical considerations, as the system processes sensitive visual data from users.

To overcome these challenges, future improvements should focus on enhancing model robustness and generalization by using more diverse and comprehensive datasets. Expanding the range of detectable emotions can make the system more nuanced and context-aware. Incorporating user feedback mechanisms can further refine recommendations, allowing the system to learn and adapt over time. Moreover, strengthening privacy measures and ensuring secure data handling will be essential for building user trust and enabling broader adoption.

Overall, the project highlights both the potential and the limitations of emotion-aware systems, providing valuable insights for future research and development in intelligent and adaptive applications [5].

7. Conclusion

Mood Melody successfully demonstrates the practical implementation of an emotion-aware music recommendation system by effectively integrating facial emotion recognition with music streaming capabilities. By leveraging technologies such as OpenCV for real-time video processing and TensorFlow with Keras for emotion classification, the system is able to interpret user emotions and respond accordingly. The integration with the Spotify API enables seamless retrieval and playback of music, creating a fully automated and personalized experience.

Innovation and Integrative Research Center Journal
ISSN: 2584-1491 | www.iircj.org
Volume-4 | Issue-4 | April-2026 | Page 292-301

The project highlights how artificial intelligence can enhance human-computer interaction by making systems more adaptive, intuitive, and responsive to user needs. Instead of relying on static inputs or past behavior, Mood Melody dynamically aligns digital content with real-time emotional states, offering a more engaging and meaningful interaction [7].

Overall, the system validates the potential of AI-driven solutions in transforming everyday applications. It not only meets its intended objectives but also lays the groundwork for future advancements in emotion-aware computing, where systems can better understand and respond to human emotions in increasingly sophisticated ways.

8. Future Work

The Mood Melody system can be further enhanced by expanding its capabilities and improving its adaptability for real-world deployment. One important direction is the integration with mobile platforms, which would make the system more accessible and convenient for users. A mobile-based application can leverage built-in cameras and sensors to provide continuous emotion-aware recommendations anytime and anywhere.

Advancements in deep learning can also be explored to improve the accuracy and robustness of emotion detection. While current models built using TensorFlow and Keras perform well, adopting more advanced architectures can enhance the system's ability to detect subtle and complex emotional expressions.

Another significant improvement is the incorporation of multi-modal emotion detection, where facial recognition using OpenCV is combined with voice and text-based sentiment analysis. This approach can provide a more comprehensive understanding of user emotions, especially in cases where facial cues alone are insufficient.

Personalization can be further refined by integrating user listening history and preferences into the recommendation process. By analyzing behavioral data along with real-time emotions, the system can deliver more accurate and tailored music suggestions. Integration with platforms like the Spotify API can support this by utilizing user-specific data such as playlists and liked tracks.

Finally, strengthening privacy and security mechanisms is essential for future development. Implementing secure data handling practices, encryption techniques, and strict authentication protocols will help protect user information and build trust. These enhancements will make Mood Melody more scalable, reliable, and suitable for widespread adoption.

References

[1] Ekman, P., & Friesen, W. V. (1978). Facial action coding system (FACS): A technique for the measurement of facial movement. Consulting Psychologists Press.

[2] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. https://www.deeplearningbook.org/

[3] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems. https://doi.org/10.1145/3065386

[4] Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. IEEE CVPR. https://doi.org/10.1109/CVPR.2001.990517

[5] Mollahosseini, A., Hasani, B., & Mahoor, M. H. (2017). AffectNet: A database for facial expression, valence, and arousal computing in the wild. IEEE Transactions on Affective Computing, 10(1), 18–31. https://doi.org/10.1109/TAFFC.2017.2740923

[6] Ko, B. C. (2018). A brief review of facial emotion recognition based on visual information. Sensors, 18(2), 401. https://doi.org/10.3390/s18020401

[7] Li, S., & Deng, W. (2020). Deep facial expression recognition: A survey. IEEE Transactions on Affective Computing, 13(3), 1195–1215. https://doi.org/10.1109/TAFFC.2020.2981446

[8] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. IEEE CVPR. https://doi.org/10.1109/CVPR.2016.90