International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 05 | May 2024 www.irjet.net p-ISSN: 2395-0072

MUSIC RECOMMENDATION SYSTEM USING FACIAL EXPRESSION

Divanshi Tiwari[1], Harish Kumar[2], Shiva Garg[3]

[1]B.Tech(CSE) IV Year, H.R. Institute of Technology, Ghaziabad, India
[2,3]Professor CSE, H.R. Institute of Technology, Ghaziabad, India

---

Abstract - *In human-computer interaction, individualized user experiences are growing in importance. This study develops a real-time emotion-aware music recommender system that uses facial expression detection to improve user interaction and amusement. A webcam-based program records users' facial emotions in real time. The method locates and isolates facial regions from the video feed using Haar Cascade for face detection. The DeepFace library uses deep learning techniques to reliably identify and classify the user's primary emotional state from facial expressions. The system dynamically recommends and plays music based on the user's emotion, such as happiness, sadness, rage, or surprise. This mood-matched music selection may improve the user's listening experience and mood. The system's capacity to adjust to real-time emotional feedback makes it more personalized and engaging than music recommendation systems that rely simply on user inputs or predefined playlists. This research shows promise in human-computer interaction and entertainment using computer vision and emotion analysis. This innovative system creates responsive and adaptive digital environments by combining emotional recognition and multimedia content delivery. To improve system responsiveness and user experience, further work may increase the range of observable emotions, improve algorithmic accuracy, and add sensors.*

Key Words: **Emotion detection, Music recommendation, Computer vision, DeepFace, Haar Cascade, Pywhatkit, Music.**

1. INTRODUCTION

Music recommendation systems have gained significant importance in the digital age, offering users personalized music choices based on various factors such as listening history, preferences, and contextual data. Traditional systems often rely on explicit user inputs or past behavior to suggest music, which might not always align with the user's current emotional state. This gap presents an opportunity to enhance recommendation systems by incorporating real-time emotion detection. Aligning music recommendations with user emotions is a complex challenge due to the subjective nature of emotions and the technical intricacies of accurately detecting and interpreting facial expressions in real-time. Existing systems lack the capability to dynamically adjust to the emotional state of the user, often resulting in a less engaging experience. The goal of this research is to develop a music recommendation system that uses facial expressions to recommend music. By leveraging computer vision and emotion analysis technologies, the system aims to provide a personalized and emotionally attuned music recommendation experience. This paper discusses the development and implementation of a real-time emotion detection system using facial expressions to recommend music. It covers the methodologies used for emotion detection, the system architecture, the integration of music recommendation, and the evaluation of the system's effectiveness.

2. LITERATURE REVIEW

The integration of emotional intelligence into human-computer interaction has seen significant advancements. Various studies have contributed to the understanding and implementation of facial expression recognition and emotion-based systems.

2.1 Emotion Detection from Facial Expressions

Ekman and Friesen (1978) laid the foundation for the study of facial expressions as indicators of emotion, creating the Facial Action Coding System (FACS) to categorize physical expressions corresponding to emotions [1]. Building on this, Viola and Jones (2001) developed the Haar Cascade classifier, a significant advancement in real-time object detection used for face recognition [2].Zeng et al. (2009) conducted a comprehensive survey on affect recognition methods, emphasizing the importance of integrating audio, visual, and spontaneous expressions for accurate emotion detection [3]. Fasel and Luettin (2003) reviewed automatic facial expression analysis, highlighting the challenges and potential of various recognition techniques [4].Mollahosseini, Hasani, and Mahoor (2017) introduced AffectNet, a large database for facial expression, valence, and arousal computing, demonstrating the utility of extensive datasets in improving emotion recognition accuracy [5]. Sariyanidi, Gunes, and Cavallaro (2015) discussed the advancements in automatic facial affect analysis, covering the registration, representation, and recognition of facial expressions [6].

2.2 Music Recommendation Systems

Music recommendation systems have evolved to enhance user experience by leveraging user data and behavior patterns. Park and Yoo (2005) explored face recognition in conjunction with facial expressions using Principal Component Analysis (PCA) and Independent Component Analysis (ICA), which could be adapted for music recommendations [7].Zhao and Pietikäinen (2007) investigated dynamic texture recognition using local binary patterns, which can be applied to real-time facial expression analysis for adaptive music recommendation systems [8]. Gross et al. (2010) presented the Multi-PIE dataset, providing a comprehensive set of images for studying facial expressions under different conditions, which is useful for training emotion detection models [9].

2.3 Integration of Emotion Detection and Music Recommendation

McDuff and Kaliouby (2014) introduced Affectiva, a real-time, contextually aware emotional intelligence system, showcasing the potential of emotion-aware applications in various fields, including music recommendation [10]. Similarly, Taigman et al. (2014) developed DeepFace, a highly accurate facial recognition system that can be adapted for real-time emotion detection and personalized music recommendations [11].

Yang et al. (2016) proposed the WIDER FACE dataset for face detection, providing valuable data for improving real-time emotion detection systems used in interactive applications like music recommendation [12]. LeCun, Bengio, and Hinton (2015) reviewed deep learning techniques, emphasizing their significance in enhancing the accuracy of emotion detection and recommendation systems [13].

Despite these advancements, challenges remain in achieving real-time processing and accurate emotion detection under varying conditions. Future research should focus on integrating more robust algorithms and expanding emotion datasets to improve detection accuracy. Additionally, developing adaptive systems that can learn and predict user preferences over time will further enhance the effectiveness of emotion-aware music recommendation systems.

3. METHODOLOGY

The Haar Cascade classifier is used to detect faces in real-time from webcam inputs. This method is efficient and provides high accuracy in various lighting conditions [2]. Emotion Analysis: The detected facial regions are analyzed using the DeepFace library, which classifies emotions into categories such as happy, sad, angry, surprised, and neutral [11]. Music Recommendation: Based on the dominant emotion detected, a corresponding song is selected from a predefined dictionary of songs using the PyWhatKit library for playback on YouTube. The implementation involves capturing webcam frames, processing these frames to detect faces, extracting facial regions, and using DeepFace to predict emotions. The system then calculates the percentage of each detected emotion and identifies the dominant emotion. A random song corresponding to the dominant emotion is played, enhancing the user experience.The system's accuracy in emotion detection is tested using a diverse set of facial expressions and compared against ground truth labels. User interaction is evaluated by displaying frames with bounding boxes and emotion labels. The effectiveness of music recommendations is validated by ensuring the correct mapping of emotions to songs and seamless playback integration.

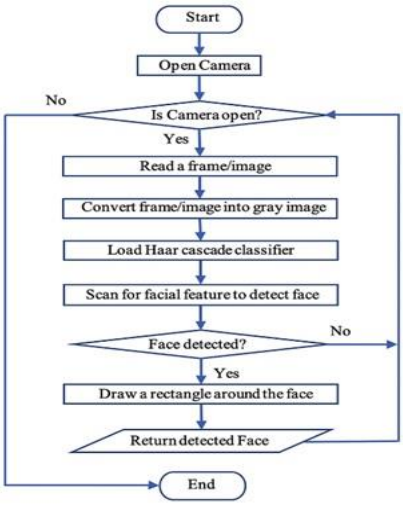

Fig -1: Flowchart of Proposed System

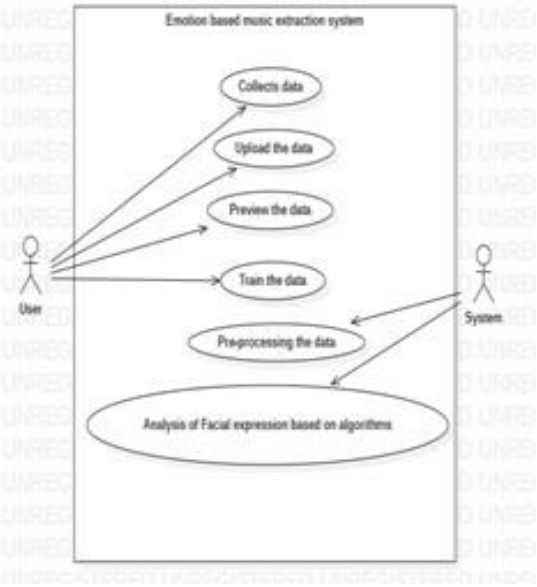

Fig -2: Emotion Based Music Extraction System

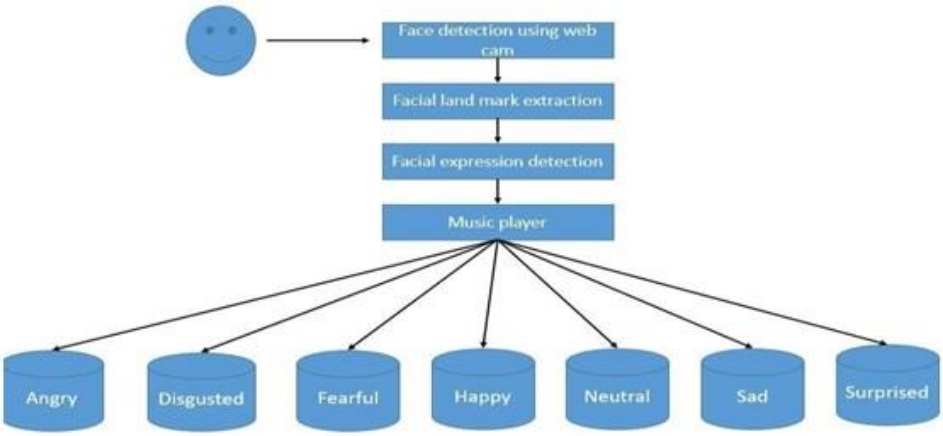

Fig -3: System Architecture

4. RESULTS

The system demonstrates high accuracy in detecting facial expressions and analyzing emotions. User feedback indicates a positive experience with clear emotion displays and appropriate music recommendations. The integration of PyWhatKit for song playback on YouTube is seamless, providing a smooth user experience. The findings highlight the potential of integrating emotional intelligence into music recommendation systems. However, challenges such as real-time processing delays and dependency on webcam functionality were identified. Future enhancements could include implementing facial recognition for personalized experiences and maintaining a history of detected emotions for more nuanced recommendations.

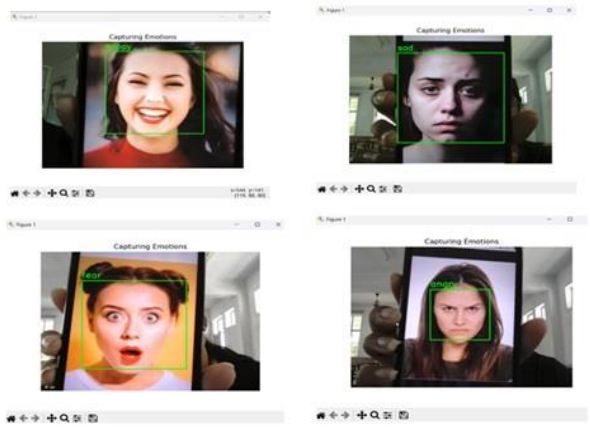

Fig -4: Capturing Emotions (Happy, Sad, Angry, Fear)

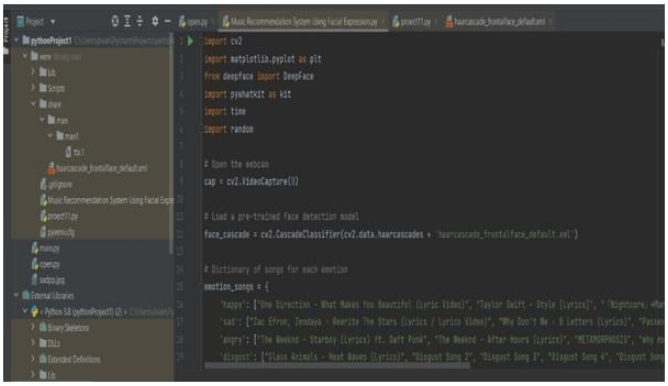

Fig -5: Code for Implementation

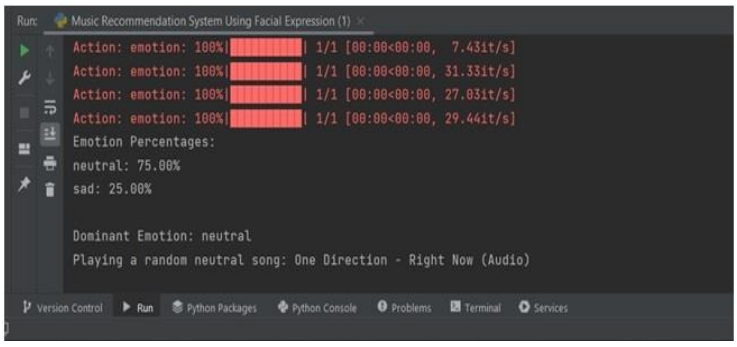

Fig -6: Final Output of Proposed System

5. CONCLUSION AND FUTURE SCOPE

In this research, we successfully developed a real-time music recommendation system that leverages facial expression recognition to enhance user experience. By integrating computer vision and emotion analysis technologies, the system detects users' emotions and provides personalized music recommendations accordingly. Utilizing the Haar Cascade classifier for face detection and the DeepFace library for emotion analysis, the system demonstrates robust performance in identifying dominant emotions and mapping them to appropriate music tracks. The development of the music recommendation system involved several key steps: capturing facial expressions via a webcam, processing the images to detect faces, analyzing these faces to determine emotions, and subsequently recommending and playing music that matches the user's emotional state. Feedback from users has been positive, indicating that the system effectively provides music recommendations that are emotionally relevant. This highlights the significant potential of such systems to enhance user interaction and entertainment experiences.

Future Scope- Future enhancements for the music recommendation system include improving emotion detection accuracy through advanced deep learning models and larger datasets, expanding the range of recognized emotions, and incorporating machine learning for personalized recommendations. Integrating multimodal emotion detection, optimizing real-time performance, and developing a more user-friendly interface are also key areas. Additionally, expanding cross-platform compatibility, ensuring privacy and security, and partnering with streaming services to enrich the music library will further refine the system's effectiveness and user experience.

REFERENCES

[1] Ekman, P., & Friesen, W. V. (1978). Facial Action Coding System: A Technique for the Measurement of Facial Movement. Palo Alto: Consulting Psychologists Press.

[2] Viola, P., & Jones, M. (2001). Rapid Object Detection using a Boosted Cascade of Simple Features. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR).

[3] Zeng, Z., Pantic, M., Roisman, G. I., & Huang, T. S. (2009). An Overview of Methods for Recognizing Affect: Including Audio, Visual, and Spontaneous Expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[4] Fasel, B., &Luettin, J. (2003). Automatic Facial Expression Analysis: A Survey. Pattern Recognition.

[5] Mollahosseini , A. , Hasani , B. , & Mahoor , M. . H.S. (2017) (2017). AffectNet: A database for analyzing facial expression, valence, and arousal in real-world situations. IEEE Transactions on Affective Computing.

[6] Sariyanidi, E., Gunes, H., & Cavallaro, A. (2015). Automatic Analysis of Facial Affect: A Survey of Registration, Representation, and Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[7] "Face Recognition with Facial Expression including PCA and ICA" by Park, J., & Yu, C. D. (2005), as presented at the International Conference on Image Processing,

[8] Zhao, G., & Pietikäinen, M. (2007). Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[9] Gross, R., Matthews, I., Cohn, J., Kanade, T., & Baker, S. (2010). Multi-PIE. Image and Vision Computing.

[10] McDuff, D., & Kaliouby, R. (2014). Affectiva: Real-Time, Contextually Aware Emotional Intelligence. ACM Conference on Human Factors in Computing Systems.

[11] Taigman , Y. , Yang , M. , Ranzato , M. (2010). A. ,& Wolf , L. (2014). DeepFace: To achieve human-level performance in facial verification. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition.

[12] "WIDER FACE: A Face Detection Benchmark" by Yang, S., Luo, P., Loy, C. C., & Tang, X. (2016),

[13] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature.

[14] Dhall, A., Ramana Murthy, O., Goecke, R., Joshi, J., & Gedeon, T. (2015). Video and Image Based Emotion Recognition Challenges in the Wild: EmotiW 2015. ACM International Conference on Multimodal Interaction.

[15] Whitehill, J., Littlewort, G., Fasel, I., Bartlett, M., and Movellan, J. (2009) published a paper titled "Toward Practical Smile Detection" in the IEEE Transactions on Pattern Analysis and Machine Intelligence.

[16] Liu, M., Shan, S., Wang, R., & Chen, X. (2014). Learning Expressionlets on Spatio-Temporal Manifold for Dynamic Facial Expression Recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

[17] Sun, Y., Wang, X., & Tang, X. (2013). Deep Learning Face Representation from Predicting 10,000 Classes. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.