South American Publishing

SAP Gamification and Augmented Reality. 2024; 2:38
doi: 10.56294/gr202438

ORIGINAL

A research on a music recommendation system based on facial expressions through deep learning mechanisms

Una investigación sobre un sistema de recomendación musical basado en expresiones faciales mediante mecanismos de aprendizaje profundo

Patakamudi Swathi[1] ✉, Dara Sai Tejaswi[1] ✉, Mohammad Amanulla Khan[1] ✉, Miriyala Saishree[1] ✉, Venu Babu Rachapudi[1] ✉, Dinesh Kumar Anguraj[1] ✉

[1]Koneru Lakshmaiah Education Foundation, Department of CSE. Vaddeswaram, Andhra Pradesh, India.

Cite as
Swathi P, Sai Tejaswi D, Amanulla Khan M, Saishree M, Babu Rachapudi V, Kumar Anguraj D. A research on a music recommendation system based on facial expressions through deep learning mechanisms. SAP Gamification and Augmented Reality. 2024; 2:38. https://doi.org/10.56294/gr202438

Publication History
Submitted: 18-10-2023
Revised: 06-02-2024
Accepted: 28-04-2024
Published: 29-04-2024

Corresponding Author
Patakamudi Swathi ✉

ABSTRACT

In this study, we propose a new music recommendation system (MRS) that combines facial expression recognition technology and deep learning algorithms to respond to the changing music industry environment and provide personalized music recommendations based on the user's emotional state. Our approach includes a thorough study of facial expression recognition, emotion-based music recommendation systems, and deep learning engines, as well as a detailed presentation of the MRS design, system architecture, and deep learning engines used. Through extensive experiments, we evaluate MRS's ability to accurately recognize facial expressions, filter music based on emotional states, and effectively recommend music to users. We analyze the results of follow-up experiments to identify the strengths and limitations of MRS compared to existing approaches, and conduct a comparative study with the latest music recommendation systems based on deep learning and emotion. This comparison highlights the originality and potential of the proposed MRS system to improve user experience and promote the development of artificial intelligence-based music recommendation systems. This study demonstrates the problem of accurately determining a user's emotional state from facial expressions, which requires the integration of facial expression recognition systems, deep learning, and music recommendation systems. Using advanced deep learning techniques and a comprehensive experimental setup, the proposed MRS provides a solution to this problem by facilitating accurate emotional state identification and personalized music recommendations. Overall, MRS represents a powerful and innovative response to the growing demand for accurate and reliable music recommendations and shows significant potential for future collaboration and development of AI-based music recommendation systems.

Keywords: Facial Expression Recognition; Deep Learning Algorithms; Machine Learning Models; Neural Networks; CNN; Ethical Considerations

RESUMEN

En este estudio, proponemos un nuevo sistema de recomendación musical (MRS) que combina tecnología de reconocimiento de la expresión facial y algoritmos de aprendizaje profundo para responder al cambiante entorno de la industria musical y proporcionar recomendaciones musicales personalizadas basadas en el estado emocional del usuario. Nuestro enfoque incluye un estudio exhaustivo del reconocimiento de la expresión facial, los sistemas de recomendación musical basados en emociones y los motores de aprendizaje profundo, así como una presentación detallada del diseño del MRS, la arquitectura del sistema y los motores de aprendizaje profundo utilizados. A través de amplios experimentos, evaluamos la capacidad de MRS para reconocer con precisión las expresiones faciales, filtrar la música en función de los estados emocionales y recomendar eficazmente música a los usuarios. Analizamos los resultados de los experimentos de seguimiento para identificar los puntos fuertes y las limitaciones de MRS en comparación con los enfoques existentes, y realizamos un estudio comparativo con los últimos sistemas de recomendación de música basados en el aprendizaje profundo y la emoción. Esta comparación destaca la originalidad y el potencial del sistema MRS propuesto para mejorar la experiencia del usuario y promover el desarrollo de sistemas de recomendación musical basados en inteligencia artificial. Este estudio demuestra el problema de determinar con precisión el estado emocional de un usuario a partir de expresiones faciales, lo que requiere la integración de sistemas de reconocimiento de expresiones faciales, aprendizaje profundo y sistemas de recomendación de música. Utilizando técnicas avanzadas de aprendizaje profundo y una configuración experimental completa, el MRS propuesto proporciona una solución a este

© 2024; The Authors. This Open access article is distributed under the Creative Commons Attribution 4.0 license (https://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

problema facilitando la identificación precisa del estado emocional y las recomendaciones musicales personalizadas. En general, MRS representa una respuesta potente e innovadora a la creciente demanda de recomendaciones musicales precisas y fiables y muestra un potencial significativo para la colaboración y el desarrollo futuros de sistemas de recomendación musical basados en IA.

Palabras clave: Reconocimiento de Expresiones Faciales; Algoritmos de Aprendizaje Profundo; Modelos de Aprendizaje Automático; Redes Neuronales; CNN; Consideraciones Éticas.

INTRODUCTION

The deep connection between music and human emotions is widely recognized across cultural boundaries and historical periods, highlighting music's remarkable ability to evoke a variety of emotional responses in people. In recent years, significant progress has been made in bringing technology and music closer together, with particular attention being paid to integrating facial emotion recognition technology into music recommendation systems. This innovative approach aims to revolutionize music streaming. We analyze listeners' facial expressions in real time and make recommendations. In this paper, we aim to explore the end-to-end development and implementation of a facial emotion-based music recommendation system to provide a comprehensive understanding of its capabilities and identify potential applications in the music industry and beyond. The main goal of this study is to go beyond simple research. We aim to untangle the complex layers of a facial emotion-based music recommendation system, explore its technical complexities, identify its potential benefits, and overcome the challenges associated with its implementation. As highly accurate facial recognition technology becomes more common and music streaming platforms become more prevalent, the combination of these technologies promises to be a transformative force that redefines the music experience by tailoring music recommendations to individual tastes. By providing users with personalized and emotionally resonant musical experiences, these systems are poised to redefine the structure of musical interaction and introduce new levels of personalization and emotional engagement. The goal of this multifaceted research is to trace the evolutionary trajectory of facial emotion-based music recommendation systems, analyze their technical foundations, and provide a holistic understanding of the system, highlighting its potential impact in various domains. Through a thorough review of existing literature, a thorough investigation of machine learning and deep learning algorithms, and a comprehensive empirical study to evaluate system performance, this paper aims to provide valuable information about the practical feasibility of these systems. We will also carefully consider ethical considerations related to data privacy, security protocols, and informed consent to ensure ethical and responsible implementation of facial emotion recognition technology. Therefore, this study aims not only to advance the discussion on facial emotionbased music recommendation systems, but also to promote a deeper understanding of ethical nuances and broader application of facial emotion recognition technology. In conclusion, this article serves as a beacon to illuminate the uncharted territory of facial emotion-based music recommendation systems, going beyond traditional research and revealing the transformative potential inherent in this technological convergence. By unraveling the tangled web of these systems, explaining their technological underpinnings, and highlighting their impact on various fields, this study attempts to chart a course toward a future in which musical activity moves beyond traditional paradigms. Through a thorough review of existing research, rigorous empirical evaluation, and a strong commitment to ethical integrity, this article aims to advance the discourse and contribute to a better understanding of the profound interactions between technology, music, and human emotions.

METHOD

The proposed system aims to change the interaction between users and music players by introducing an innovative method based on facial expressions. The main goal is to evaluate the user's emotional state using a camera and several well-defined components. Starting with accurate face capture, the user's emotions are predicted based on facial expressions using CNN (Convolutional Neural Network), which is excellent at image analysis. It then uses this emotional data to create personalized playlists tailored to the user's mood, dynamically targeting different emotions such as happiness, sadness, neutral, or surprise. This process includes real-time capture to accurately recognize facial expressions, face recognition using CNN to analyze facial features, emotion recognition to extract emotional signals from facial images, and music recommendation to recommend songs that match the emotions detected by the user. Included. It passes through four main modules: including mapping. They are sorted by mood category in the system's music database.

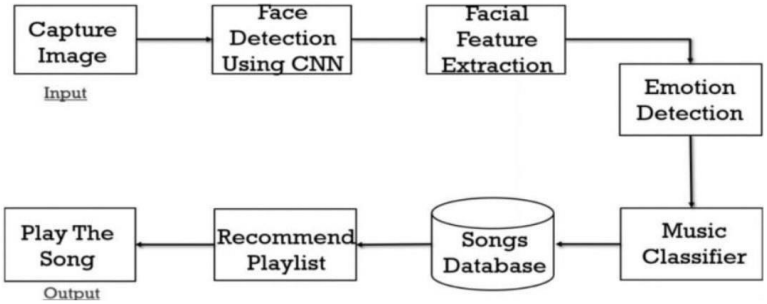

Figure 1. Proposed architecture

https://doi.org/10.56294/gr202438

This methodology is based on the use of convolutional neural networks (CNNs), which require a robust dataset to build and train the model. The FER2013 dataset selected by Kaggle includes training (24 176 images) and testing (6 043 images) datasets, each representing five emotions: happy, sad, angry, surprise, and neutral, with a resolution of 48x48 pixels. Contains a grayscale image of the face. The creation of the FER-2013 dataset included a comprehensive Google image search, and efforts were made to address imbalances in emotional expressions by applying the SoftMax approach with a weighted training loss. The emotion detection module includes important components such as face recognition using the Haar cascade for accurate face identification and feature extraction. Here, the pre-trained CNN acts as a random feature extractor to extract high-level features from the input image of a given layer. Extraction and preparation: Emotion detection.

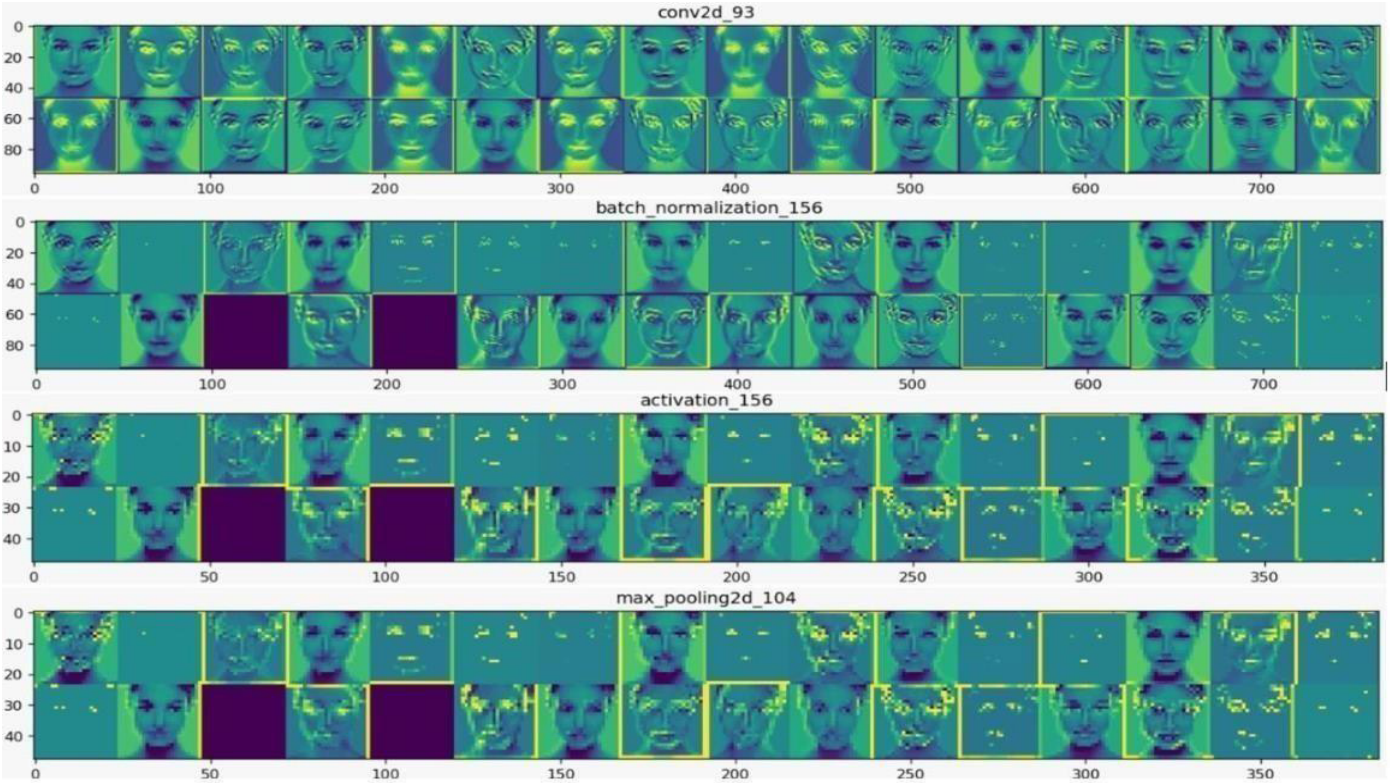

Figure 2. Feature evolution in CNN layers

The Convolutional Neural Network (CNN) architecture plays a key role in the emotion detection module, which processes input images using filters and generates feature maps using Rectified Linear Unit (ReLU) activation functions. These filters identify various elements of the image, such as edges, vertical and horizontal lines, and curves. The CNN then analyzes these feature maps to predict emotion, including pooling for translation invariance. This process of extracting features from different CNN layers reveals the internal representation of specific input data. We also demonstrate the effectiveness of our system in accurately labeling emotions based on user images in real time.

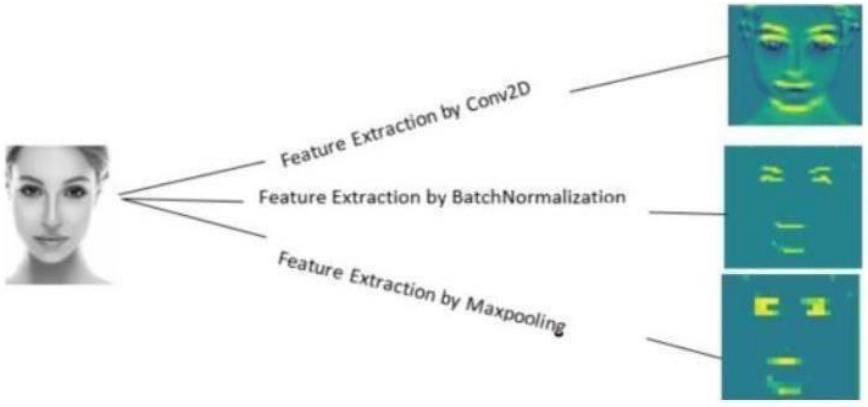

Figure 3. Labelling of real-time user emotions

In the music recommendation module, the system uses a diverse database of Hindi and Bollywood songs to effectively tailor playlists to the user's emotional state. This database, which contains 100 to 150 songs in each emotional category, confirms the

https://doi.org/10.56294/gr202438

profound impact music has on mood.

After real-time emotion detection, the emotions are categorized into individual labels such as Happy, Sad, Angry, Surprised, and Neutral, and linked to the corresponding folders in the song database using the Python operating system listdir. This seamless integration allows you to dynamically and automatically recommend music playlists based on the user's emotional state. In conclusion, this methodology seamlessly combines facial emotion recognition and music recommendation to provide a user-centric, dynamic music experience. From dataset selection and model training to real-time emotion detection and playlist recommendations, every step is carefully designed for accuracy and personalization using cutting-edge technologies such as convolutional neural networks that revolutionize music recommendations based on facial expressions.

RESULTS

The 'facial expression-based music recommendation system using a deep learning engine' aims to enrich the music listening experience by improving user satisfaction by providing personalized music recommendations based on real-time facial expression analysis. Applications for this system include developing advanced algorithms using deep learning techniques, seamless integration with webcams for facial expression detection, and creating an intuitive user interface to improve interaction with music recommendations.

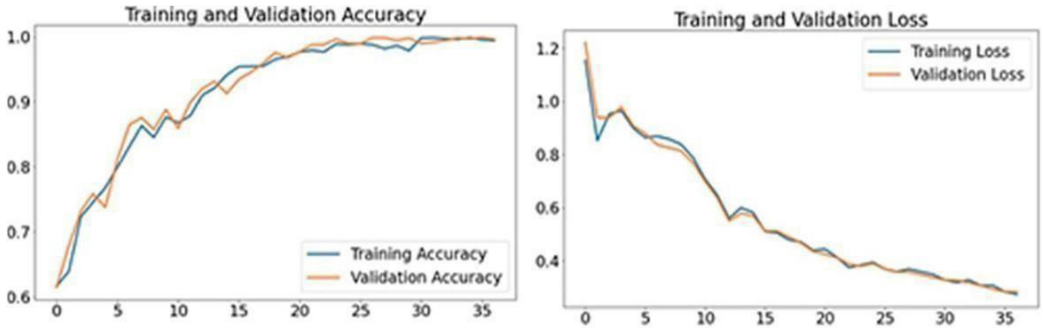

Figure 4. Training and testing validations

The music system described above demonstrates remarkable accuracy, ensuring that recommended playlists exactly match the user's emotional state. It accurately detects user emotions in real time using sophisticated deep learning algorithms and facial expression recognition technology. By accurately detecting emotional signals in facial expressions, the system can classify them into different mood categories such as happy, sad, angry, surprised, and neutral. This high level of accuracy allows the system to generate personalized music playlists tailored to each user's emotional state with exceptional reliability and efficiency.

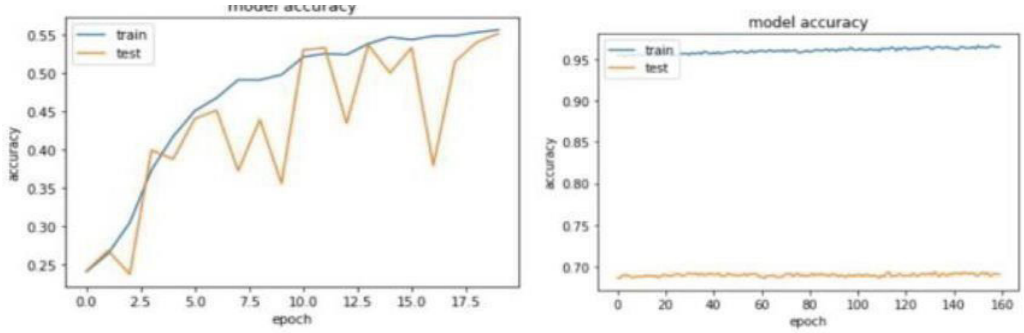

Figure 5. Model accuracy rates

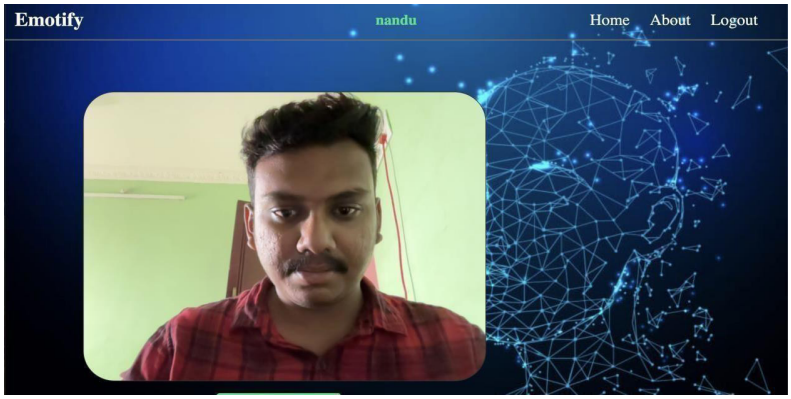

Figure 6. Detecting face emotion

https://doi.org/10.56294/gr202438

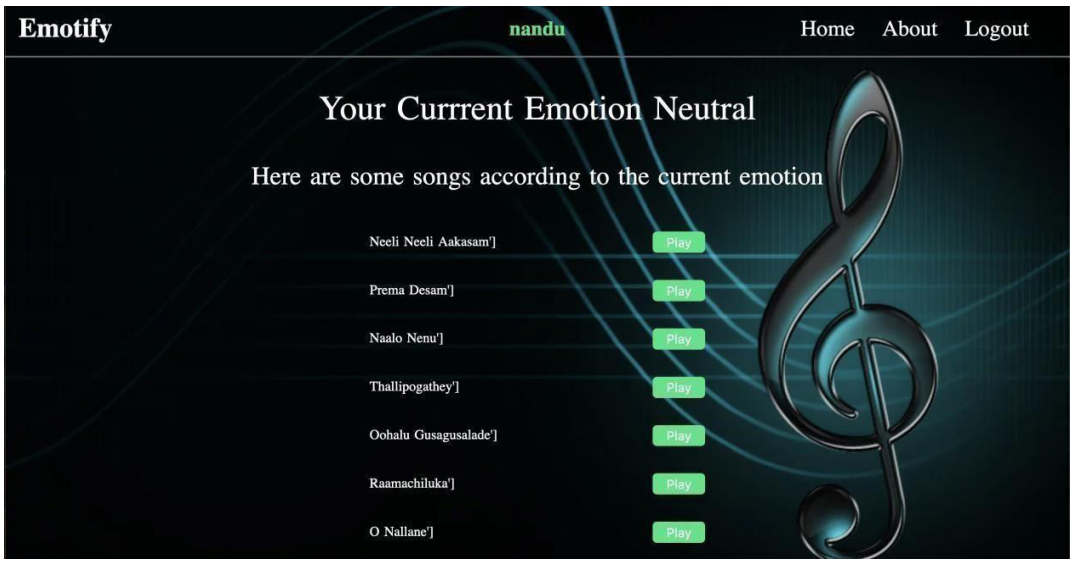

Figure 7. Presenting the outcome of facial emotion detection

DISCUSSION

Building on the foundation laid by the previous discussion, several directions for future work and research are identified. First, increasing the diversity and depth of the song database can significantly improve the performance of music recommendation systems. Including a wider range of genres and languages beyond Bollywood and Hindi songs allows us to reach a more diverse user base catering to different musical preferences and emotional states. Additionally, incorporating user feedback mechanisms will further enhance the recommendation process, allowing the system to adapt and evolve based on individual preferences and reactions to suggested playlists. Additionally, learning advanced machine learning techniques and emotion detection algorithms improves accuracy and efficiency, ultimately improving overall system performance and user experience. Additionally, research on the integration of real-time user feedback and physiological signals such as the heart Speed variability provides greater insight into the user's emotional state, enabling more accurate and personalized music recommendations. Finally, considering ethical implications and ensuring user privacy is paramount in future developments, requiring ongoing efforts to maintain data protection standards and transparency of data practices. Through continued research and innovation, music recommendation systems can evolve into sophisticated and essential tools for increasing user engagement and satisfaction with digital music consumption.

Literature

We conducted several literature reviews to understand existing research in this field. Research on emotion-based music recommendation systems has attracted the attention of researchers, engineers, and scientists around the world, as evidenced by the extensive literature on the topic.[6,3] This comprehensive review examines existing methodologies, highlights their strengths and weaknesses, and provides important implications for future innovation in the field. Understanding the nuances of current research, theoretical developments, and methodological developments in emotion-based music recommendation systems is essential for developing new solutions that resonate with user experiences. Integrating facial expressions into powerful indicators of human emotion is emerging as a central theme for these systems. The system uses non-verbal signals to personalize music recommendations in real time based on your emotional state. Ayush Guidel and others advance the discourse by emphasizing that a person's state of mind and emotional state can be observed through facial expressions.[7] Their advanced system uses convolutional neural networks for face recognition to recognize basic emotions such as happiness, sadness, anger, excitement, surprise, disgust, fear, and neutral. Ramya Ramanathan,et al.,propose an intelligent music player that includes emotion,recognition by initially grouping local music selected by the,user based on the emotion conveyed by the album,,taking song lyrics into account.[1] Various emotion recognition technologies are being applied to various fields such as medicine, education, entertainment, and the music industry.[3] The potential use of emotion-based systems in patient care, teaching strategies, and live performance demonstrates the broad societal impact of these technologies. Methodological contributions to emotion-based music recommendation systems go beyond facial expression recognition and include the integration of physiological signals.[3,12] Future research directions may include exploring innovative approaches to address current issues, improve algorithms, and further improve the overall user experience.[13,14,15,16] Although technological advancements have advanced the field, ethical considerations and user privacy remain paramount concerns. The integration of facial recognition technology requires critical thinking about data security, informed consent, and responsible data use. Additionally, ongoing efforts to improve the accuracy of algorithms and address existing limitations highlight the desire to develop the field ethically and responsibly.[17,18] Beyond the music industry, the versatility of emotion recognition technology is expected to have applications in a variety of fields, from medicine to entertainment.[6,7] The potential use of emotion-based systems in patient care, teaching strategies, and live performance demonstrates the versatility and broad societal impact of these technologies. The reviewed literature highlights methodological contributions to emotion-based music recommendation systems, ranging from the use of facial expression recognition to the integration of physiological signals.[19,20,21,22] Future research directions may include exploring

innovative approaches to solve current problems, improve algorithms, and improve the overall user experience.[12] In conclusion, the literature on emotion-based music recommendation systems highlights the importance of facial expressions and physiological signals in understanding human emotions. Advanced technologies, machine learning algorithms, and wearable devices have laid the foundation for innovative systems that can personalize music recommendations in real time based on emotional states.[1,3,6] Facial expressions, which are hidden indicators of human emotions, become an important element of these systems.[6] In the realm of interpersonal communication, nonverbal signals, including hand gestures, facial expressions, and tone of voice, are powerful means of conveying emotions.[23,24] Facial expressions, in particular, make a valuable contribution to emotion-based systems as they provide real-time insight into an individual's current mental state. Prima et al. Noting the time-consuming task of manually creating playlists, they proposed a music player that could select songs based on the user's mood.[6] Using the Viola-Jones algorithm for face detection and expression extraction, the system uses support vector machines (SVMs) to classify emotions across basic categories such as anger, joy, surprise, sadness, and disgust. Yusuf Yaslan et al. present an innovative approach that integrates physiological signals obtained from wearable computing devices for emotion-based music recommendation.[3] The system uses GSR (galvanic skin response) and PPG (photoplethysmography) sensors to study the user's emotional state. This methodology increases the granularity of emotion recognition beyond traditional facial expression analysis by predicting arousal and valence based on multi-channel physiological signals.

CONCLUSIONS

The study 'Facial expression-based music recommendation system using deep learning engine' provides deep insight into the relationship between emotions, facial expressions, and the potential impact of technology. Through a thorough literature review, we explore different methodologies for emotion-based music recommendation systems and highlight the need for intelligent systems that dynamically generate playlists based on facial expressions. Inspired by previous research and using advanced techniques such as the Viola-Jones algorithm and convolutional neural networks, the proposed system aims to revolutionize music consumption by personalizing playlists in real time. Feasibility studies ensure that legal, ethical, and cultural considerations are met, and stated goals emphasize ongoing research and user involvement. Essentially, this system pioneers the field of emotion-aware music recommendation by enriching the combination of technology and human emotions.

FINANCING

The authors did not receive financing for the development of this research.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION

Conceptualization: Patakamudi Swathi, Dara Sai Tejaswi, Mohammad Amanulla Khan, Miriyala Saishree.
Data curation: Dara Sai Tejaswi.
Formal analysis: Patakamudi Swathi.
Research: Mohammad Amanulla Khan.
Methodology: Miriyala Saishree.
Project management: Dara Sai Tejaswi.
Resources: Mohammad Amanulla Khan.
Software: Miriyala Saishree.
Supervision: Venu Babu Rachapudi, Dinesh Kumar Anguraj.
Validation: Venu Babu Rachapudi, Dinesh Kumar Anguraj.
Display: Patakamudi Swathi, Dara Sai Tejaswi, Mohammad Amanulla Khan, Miriyala Saishree.
Drafting - original draft: Patakamudi Swathi, Dara Sai Tejaswi, Mohammad Amanulla Khan, Miriyala Saishree.
Writing - proofreading and editing: Patakamudi Swathi.

BIBLIOGRAPHIC REFERENCES

1. Ramya Ramanathan, Radha Kumaran, Ram Rohan R, Rajat Gupta, and Vishalakshi. Prabhu, an intelligent music player based on emo- tion recognition, 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2017. https://doi.org/10.1109/CSITSS.2017.8447743

2. Shlok Gilda, Husain Zafar, Chintan Soni, Kshitija Waghurdekar, Smart music player integrating facial emotion recognition and music mood recommendation, Department of Computer Engineering, Pune Institute of Computer Technology, Pune, India, (IEEE),2017. https://doi.org/10.1109/WiSPNET.2017.8299738

3. Deger Ayata, Yusuf Yaslan, and Mustafa E. Kamasak, Emotion-based music recommendation system using wearable physiologica.

4. Ahlam Alrihail, Alaa Alsaedi, Kholood Albalawi, Liyakathunisa Syed, Music recommender system for users based on emotion

detection through facial features, Department of Computer Science Taibah University, (DeSE), 2019. https://doi.org/10.1109/DeSE.2019.00188

5. Research Prediction Competition, Challenges in representation learning: facial expression recognition challenges, Learn facial expres- sion from an image, (KAGGLE).

6. Preema J.S, Rajashree, Sahana M, Savitri H, Review on facial expression-based music player, International Journal of Engineering Re- search & Technology (IJERT), ISSN2278- 0181, Volume 6, Issue 15, 2018.

7. AYUSH Guidel, Birat Sapkota, Krishna Sapkota, Music recommendation by facial analysis, February 17, 2020.

8. CH. sadhvika, Gutta.Abigna, P. Srinivas reddy, Emotion-based music recommendation system, Sreenidhi Institute of Science and Technology, Yamnampet, Hyderabad; International Journal of Emerging Technologies and Innovative Research (JETIR) Volume 7, Is-sue 4, April 2020.

9. Vincent Tabora, Face detection using OpenCV with Haar Cascade Classifiers, Becominghuman.ai,2019.

10. Zhuwei Qin, Fuxun Yu, Chenchen Liu, Xiang Chen. How convolutional neural networks see the world - A survey of convolutional neural network visualization methods.

11. Mathematical Foundations of Computing, May 2018.

12. Ahmed Hamdy AlDeeb, Emotion- Based Music Player Emotion Detection from Live Camera, ResearchGate, June 2019.

13. Frans Norden and Filip von Reis Marlevi, A Comparative Analysis of Machine Learning Algorithms in Binary Facial Expression Recogni- tion, TRITA-EECS-EX2019:143.

14. Singh A, Sharma R, Pandey MS, Asthana S, Gitanjali, Vishwakarma A. Facial Expression Based Music Recommendation System Using Deep Learning. En: Namasudra S, Trivedi MC, Crespo RG, Lorenz P, editores. Data Science and Network Engineering, Singapore: Springer Nature; 2024, p. 31-40. https://doi.org/10.1007/978-981-99-6755-1_3.

15. Bakariya B, Mohbey KK, Singh A, Singh H, Raju P, Rajpoot R. An Efficient Model for Facial Expression Recognition with Music Recommendation. Natl Acad Sci Lett 2023. https://doi.org/10.1007/s40009-023-01346-4.

16. Athavle M, Mudale D, Shrivastav U, Gupta M. Music Recommendation Based on Face Emotion Recognition. Journal of Informatics Electrical and Electronics Engineering (JIEEE) 2021;2:1-11. https://doi.org/10.54060/JIEEE/002.02.018.

17. S SS, S S, T P. An Improved Music Recommendation System for Facial Recognition and Mood Detection. ITM Web Conf 2023;56:01004. https://doi.org/10.1051/itmconf/20235601004.

18. Parasar D, Sahi I, Jain S, Thampuran A. Music Recommendation System Based on Emotion Detection. En: Pandit M, Gaur MK, Rana PS, Tiwari A, editores. Artificial Intelligence and Sustainable Computing, Singapore: Springer Nature; 2022, p. 29-43. https://doi.org/10.1007/978-981-19-1653-3_3.

19. Pallavi Reddy R, Abhinaya B, Sahithi A. A Deep Learning Technique to Recommend Music Based on Facial and Speech Emotions. En: Choudrie J, Mahalle PN, Perumal T, Joshi A, editores. ICT for Intelligent Systems, Singapore: Springer Nature; 2023, p. 25-40. https://doi.org/10.1007/978-981-99-3982-4_3.

20. Sanapala D, Muthalagu R, Pawar PM. Hybrid Deep Face Music Recommendation Using Emotions. En: Roy S, Sinwar D, Dey N, Perumal T, Tavares JMRS, editores. Innovations in Computational Intelligence and Computer Vision, Singapore: Springer Nature; 2023, p. 35-48. https://doi.org/10.1007/978-981-99-2602-2_4.

21. Prasad VM, Swetha GN, Ali KMR. Creation of A Music Recommendation System using Facial Expression Recognition with MATLAB. International Journal of Intelligent Systems and Applications in Engineering 2024;12:320-7.

22. Annam ST, Bodapati JD, Konda R. Emotion-Aware Music Recommendations: A Transfer Learning Approach Using Facial Expressions. En: Tiwari S, Trivedi MC, Kolhe ML, Singh BK, editores. Advances in Data and Information Sciences, Singapore: Springer Nature; 2024, p. 1-11. https://doi.org/10.1007/978-981-99-6906-7_1.

23. Sharath P, Kumar GS, Vishnu BKS. Music Recommendation System Using Facial Emotions. Advances in Science and

Technology 2023;124:44-52. https://doi.org/10.4028/p-4s4w34.

24. Mittal S, Ranjan A, Roy B, Rathore V. Mus-Emo: An Automated Facial Emotion-Based Music Recommendation System Using Convolutional Neural Network. En: Dhar S, Mukhopadhyay SC, Sur SN, Liu C-M, editores. Advances in Communication, Devices and Networking, Singapore: Springer; 2022, p. 267-76. https://doi.org/10.1007/978-981-16-2911-2_29