Journal of Electrical Technology UMY (JET-UMY), Vol. 9, No. 1, June 2025
ISSN 2550-1186 e-ISSN 2580-6823

Development of a Web-Based Music Recommendation System Based on Facial Expression Using a Convolutional Neural Networks Model

Dimas Aditya Putra Wardhana*[1], Fridy Mandita[1], Elvianto Dwi Hartono[1,]
Dwi Hendra Rosli[2], Bima Putra Purnawirawan[2], Muhammad Aldi Syafrillah Hidayat[2]

[1] Department of Robotics and Artificial Intelligence Engineering, Faculty of Engineering,
Universitas 17 Agustus 1945 Surabaya
Jl. Semolowaru No.45, Menur Pumpungan, Kec. Sukolilo, Surabaya, Jawa Timur, Indonesia, (031) 5931800
[2] Department of Information Systems and Technology, Faculty of Engineering,
Universitas 17 Agustus 1945 Surabaya
Jl. Semolowaru No.45, Menur Pumpungan, Kec. Sukolilo, Surabaya, Jawa Timur, Indonesia, (031) 5931800
*Corresponding author, e-mail: dimasapw@untag-sby.ac.id

Abstract – *This research presents the development of a web-based music recommendation system that uses facial expression recognition to match songs with users' emotional states. Real-time facial detection and expression classification are conducted in the browser using two CNN models implemented via the face-api.js library. Each classified expression is mapped to a specific music genre, and relevant songs are retrieved using the SoundCloud API. The system was evaluated through two aspects, accuracy and user satisfaction. Accuracy was measured using a dichotomous questionnaire, with results showing that 91% of users agreed that the recommended songs reflected their current emotions. User satisfaction was also assessed using a similar questionnaire and reached 86%, indicating a high level of comfort and relevance in the user experience. Compared to previous studies that used Likert scales, this study offers a different yet equally effective evaluation approach. The findings suggest that integrating facial expression recognition into music recommendation systems can provide a practical and user-friendly way to support emotional regulation through music.*

Keywords*: CNN, Face-api.js, Facial expression recognition, Music recommendation, Web*

I. Introduction

Music has long played an essential role in everyday life not only as a form of entertainment during leisure time but also as a medium for emotional expression. It is an art form with the ability to influence human emotions since ancient times [1]. In fact, music is also used therapeutically to reduce stress and anxiety [2]. The emotional bond between individuals and music makes song selection a crucial factor in managing emotions. A song, as a structured composition of melody, rhythm, and lyrics, is designed to be sung [3]. With the growth of the music industry, there is now an abundance of songs and genres such as pop, rock, jazz, EDM, indie, and others. This variety often leaves individuals confused when trying to select music that matches their current mood [4]. For example, someone who is feeling sad may struggle to identify songs that truly reflect their emotional state. This is because not everyone is consciously aware of what they need when experiencing certain emotions. As a result, choosing the wrong song can potentially worsen their emotional condition. According to an IFPI report, around 78% of people believe that music helps them relax and cope with stress [5]. Therefore, a solution is needed that can automatically understand the user's emotional condition and assist in selecting music that aligns with their current mood.

Facial expression recognition has become a key

Manuscript received May 2025, revised June 2025

component in emotion-based systems.

Several studies have explored facial expression detection using image processing and artificial intelligence (AI) technologies. Saputra and Nudin developed a web-based music player integrated with facial expression detection using a Convolutional Neural Network (CNN) model. Their system achieved a validation accuracy of 68.69% and an average user satisfaction score of 4.22 out of 5 [6]. However, the system was limited in the number of expressions it could classify, making it less effective in capturing emotional diversity. Sukietra applied CNN to classify music genres using MFCC features from the GTZAN dataset, implementing the system on an Android app integrated with the Spotify API [7]. Although the model reached a validation accuracy of 79%, it focused on genre classification rather than facial expression detection. Ardiansyah et al. developed a desktop-based music recommendation system that utilized facial expression recognition, achieving an average F1-score of 88.5% for four primary emotions (happy, sad, angry, afraid), with 88% user satisfaction based on beta testing [8]. However, this study was not implemented on a web-based platform.

In response to these limitations, this study aims to develop a facial expression-based music recommendation system implemented on a web platform. The system performs real-time facial expression detection using two CNN-based models one for face detection and another for expression classification then maps the classified emotions to corresponding music genres. Unlike previous studies that were limited in expression variety, not implemented on the web, or lacked facial detection integration, this study introduces a web-based solution that supports seven expression classes, enabling more varied and emotionally relevant song recommendations for users.

II. Research Methods

II.1. System Design

The facial expression-based music recommendation system was developed as a web-based platform with an architecture illustrated in Fig 1. The system follows an input–process–output workflow and consists of four main components: frontend, internal API, backend, and database. On the frontend, facial expression detection and classification are performed in real-time using a JavaScript library. The frontend communicates with the internal API to send the classified expression data and manage requests for logging data to the backend. The backend is responsible for storing logs in the database and forwarding music genre requests to an external API. The external API utilized in this system is the SoundCloud API, which retrieves a list of songs based on the genre mapped from the user's facial expression. The use of the SoundCloud API plays a significant role in the design of this web-based system. SoundCloud was selected due to its open endpoints that allow for genre-based queries, as well as its seamless integration into web interfaces via an audio player that supports direct music streaming without requiring platform switching [9]. Thus, the system is capable of providing relevant music recommendations that can be played directly through the website.

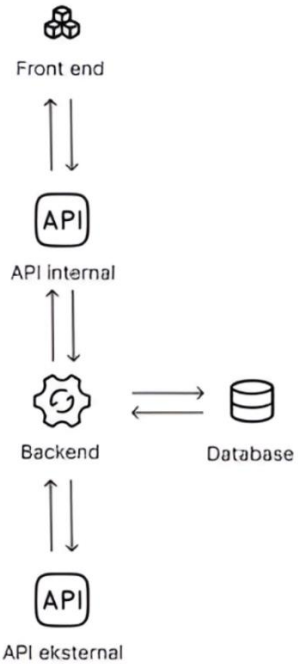

Fig. 1. System architecture as a representation of the workflow between components

Fig. 2 presents a detailed flowchart of the system's operational process. The process begins when a user accesses the website and clicks the capture button on the homepage. A countdown of three seconds is initiated before the system captures an image of the user’s face. The captured image is processed to detect a face. If no face is detected, the system displays a notification and prompts the user to repeat the process. However, if a face is detected, the system proceeds to the expression classification stage to identify the user’s emotion. The classified

emotion is then displayed on the website along with five recommended songs retrieved through the SoundCloud API based on the mapped genre. The user may choose one of the songs, and the system logs the interaction before playing the selected track through the embedded audio player. The session concludes once the song is played, and the system becomes ready for the next use.

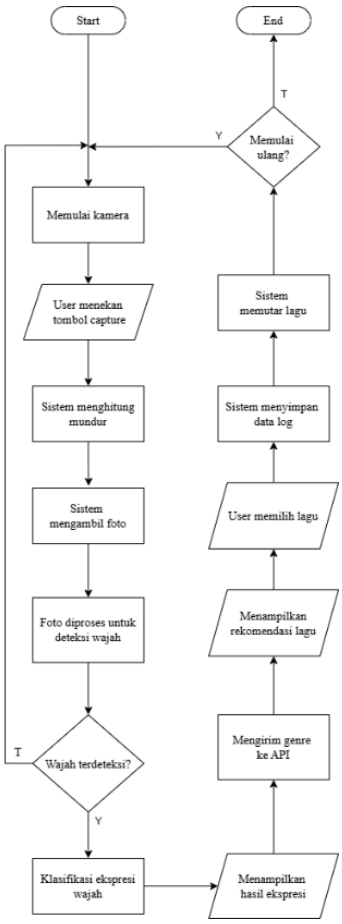

Fig. 2. System flowchart as a description of the processing of input data to output

II.2. Model

This system utilizes two CNN-based models to detect and classify facial expressions. Both models are implemented using the JavaScript library face-api.js [10]. The first model used is TinyFace Detector, which is part of the face-api.js library built on top of TensorFlow.js. TinyFace Detector features a lightweight architecture and fast detection time, making it suitable for web implementation [11]. This model is responsible for detecting the location of the face in the input image before classification is performed.

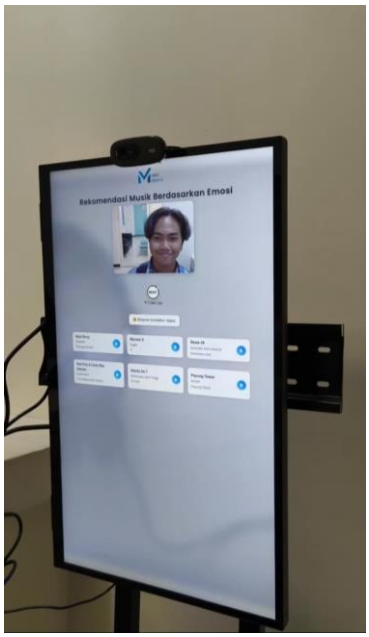

Fig. 3 Implementation of the model on the website

The second model is a CNN-based facial expression recognition model loaded using pre-trained weight manifests provided by face-api.js. Its primary function is to classify the detected face into one of seven emotional categories: happy, sad, angry, fearful, surprised, disgusted, and neutral. These two models were chosen for their lightweight characteristics that support efficient browser-based execution, enabling the system to perform real-time facial expression detection and classification directly within the user's browser. As a result, the system can provide real-time music recommendations based on facial expressions through the web interface.

II.3. Data Collection Technique

This study utilized two types of data collected through different techniques: primary data and secondary data. Primary data refers to information gathered directly by the researcher to evaluate the music recommendation system through user feedback, while secondary data refers to datasets acquired indirectly and used for training the facial expression classification model [12].

The secondary data was obtained from publicly available facial image datasets on the Kaggle platform. This dataset consists of labeled human facial images categorized into seven emotional classes: happy, sad, angry, fearful, surprised, disgusted, and neutral [13]. The dataset was used to train and test the CNN-based expression classification model implemented with the face-api.js library.

Primary data was collected through a quantitative survey using a dichotomous-scale questionnaire to obtain direct feedback from users of the developed system. A dichotomous questionnaire provides only two answer options, namely Yes or No [14]. This instrument was designed to evaluate both the accuracy and the user experience of the facial expression-based music recommendation system. The use of a dichotomous scale was chosen to obtain firm and consistent answers from respondents, as the options provided consist of only two opposing choices. This approach minimizes ambiguity and facilitates easier interpretation of the collected data [15].

III. Results and Discussion

III.1. System Implementation

The system is implemented as a web-based application with a user interface as shown in Figures 4 and 5.

Fig. 4 Main interface display of the system

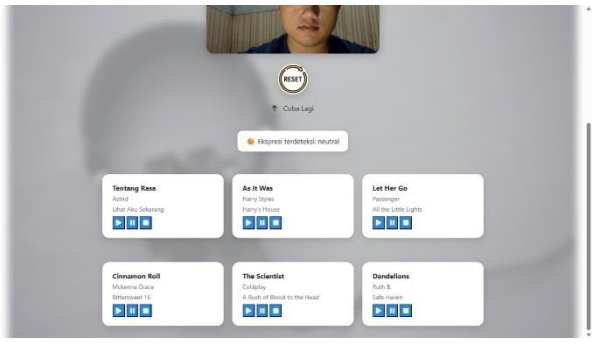

Fig. 5 The interface after capture image will display song recommendations

The main page features a live camera feed to capture the user's face, a capture button, facial expression classification results, and a list of recommended songs that can be played directly via a SoundCloud audio player. Facial detection and expression classification are performed directly in the browser using the face-api.js library, while log storage and song list retrieval processes are managed through the backend and internal APIs.

III.2. Accuracy Evaluation

Accuracy evaluation in this study refers to the system's ability to recommend songs that match the user's facial expression. The testing was conducted through a questionnaire distributed to respondents after they interacted with the system. Each respondent was asked to answer a series of questions assessing whether the recommended song appropriately reflected their emotional expression at the time. The responses were dichotomous, with "Yes" indicating a match and "No" indicating a mismatch. The formula used to calculate the accuracy percentage of the system's recommendations is as follows:

$$Accuracy = \frac{score\ obtained}{maximum\ score} \times 100\% \quad (1)$$

Based on the data analysis, 91% of respondents stated that the recommended songs matched their expressions while using the system. This result demonstrates that the facial expression-based music recommendation system developed in this study achieved a high level of appropriateness according to user feedback. It serves as an indicator that the system is capable of recognizing facial expressions and translating them into music recommendations that subjectively align with user emotions.

Compared to the study by Saputra and Nudin [6], which reported a validation accuracy of 68.69% using a CNN-based web system, the accuracy in this research indicates a notable improvement. Although different evaluation approaches were used, the findings highlight that the developed system is more effective in meeting user needs practically. This difference can be attributed to the use of an evaluation method that directly captures user perceptions, making the results more contextual and applicable in real-world usage.

III.3. User Satisfaction Evaluation

In the development of a facial expression-based music recommendation system, evaluating user satisfaction is essential to understand the extent to which the system is accepted and perceived as beneficial by end-users. This aligns with the findings of Ishak et al., who emphasized that user satisfaction evaluation aims to measure the

perceived usefulness of a system [16]. To achieve this, an evaluation was conducted by distributing questionnaires after users interacted directly with the system.

To obtain a measurable overview of user satisfaction, the questionnaire consisted of a series of closed-ended, dichotomous questions, enabling the collection of objective and easily processed data. The formula used to calculate the user satisfaction percentage is as follows:

$$Satisfaction = \frac{score\ obtained}{maximum\ score} \times 100\% \quad (2)$$

Based on the data collected, the user satisfaction level reached 86%. This figure indicates that the majority of users appreciated the experience provided by the system, in terms of comfort, response speed, and the relevance of the song recommendations. These findings reinforce the earlier accuracy evaluation, which demonstrated the system's ability to interpret facial expressions and translate them into music recommendations that users found appropriate.

When compared to the study by Saputra and Nudin [6], which used a Likert-scale evaluation and achieved an average satisfaction score of 4.22 out of 5, this study offers a different yet comparable evaluation approach. While the previous research employed a scaled response format, the current study utilized binary (Yes/No) responses and achieved an overall satisfaction rate of 86%. Although the measurement methods differ, both approaches aim to assess user perceptions of comfort and system relevance. The results show that the system developed in this research also delivers a positive and meaningful experience to users.

IV. Conclusion

This study successfully developed a facial expression-based music recommendation system implemented on a web platform. The system was designed to detect users' facial expressions in real-time using two Convolutional Neural Network (CNN) models based on face-api.js, and map the classified emotions to relevant music genres through integration with the SoundCloud API. With this approach, users can instantly receive music recommendations that align with their emotional state without the need for manual searching. System evaluation was conducted through two key aspects: accuracy and user satisfaction. Based on the questionnaire results, the system achieved 91% accuracy in recommending songs that matched the users' facial expressions. In addition, the satisfaction evaluation showed that 86% of respondents were satisfied with their overall experience using the system, in terms of ease of use, speed, and the relevance of the recommended songs. These findings demonstrate that the system not only performs well in recognizing facial expressions but also provides practical value perceived directly by end-users. Overall, this study proves that integrating facial expression recognition technology with a music recommendation system can deliver an effective and user-centric solution to assist users in selecting songs that reflect their current emotional state. This opens opportunities for further development on other platforms and the exploration of more personalized genres or music preferences.

Acknowledgements

The authors would like to express their gratitude to Universitas 17 Agustus 1945 Surabaya for its support in funding this research and to the Department of Robotics and Artificial Intelligence Engineering for providing the necessary facilities during the research and development process.

References

[1] N. N. Hanifah and D. C. Aprilia, "Analisis Penggunaan Bahasa Pada Lirik Lagu 'BUKTI' Karya Virgoun Sebagai Bentuk Ekspresi Dan Komunikasi Emosional," *J. Ilmu Bhs. Dan Pendidik. Guru Sekol. Dasar*, vol. 1, Jan. 2025.

[2] Lussy Putri Khadijah, "Efektivitas Terapi Musik Untuk Menurunkan Tingkat Stres Dan Kecemasan," *Detect. J. Inov. Ris. Ilmu Kesehat.*, vol. 1, no. 3, pp. 91–98, Jul. 2023, doi: 10.55606/detector.v1i3.2101.

[3] A. A. Reki and Y. Yensharti, "Analisis Struktur Lagu Biar Menjadi Kenangan Ciptaan Ahmad Dhani," *J. Sendratasik*, vol. 9, no. 3, p. 15, Sep. 2020, doi: 10.24036/jsu.v9i1.109300.

[4] N. Ula, C. Setianingsih, and R. A. Nugrahaeni, "Sistem Rekomendasi Lagu Dengan Metode Content-Based Filtering Berbasis Website," *E-Proceeding Eng.*, vol. 8, p. 12193, Dec. 2021.

[5] C. Hatton, "IFPI's global study finds we're listening to more music in more ways than ever," IFPI. Accessed: Jun. 26, 2025. [Online]. Available: https://www.ifpi.org/ifpis-global-study-finds-were-listening-to-more-music-in-more-ways-than-ever/

[6] F. A. Saputra and S. R. Nudin, "Pengembangan Sistem Rekomendasi Pada Pemutar Musik menggunakan Face Emotion Detection dan Resnet Berbasis Website," *J. Manaj. Inform.*, vol. 2, May 2024.

[7] D. Sukietra, "Aplikasi Rekomendasi Musik Berdasarkan Klasifikasi Genre Menggunakan Convolutional Neural Networks," *KALBISIANA J. Sains Bisnis Dan Teknol.*, vol. 10, no. 1, pp. 91–98, Mar. 2024, doi: 10.53008/kalbisiana.v10i1.3527.
[8] A. Ardiansyah, R. A. Ramadhani, and D. Swanjaya, "Pengembangan Sistem Rekomendasi Lagu Berdasarkan Ekspresi Wajah Menggunakan YOLO v8," *Semin. Nas. Teknol. Sains STAINS*, vol. 4, Jan. 2025.
[9] "API - Panduan - Pengembang SoundCloud." Accessed: Jul. 26, 2025. [Online]. Available: https://developers.soundcloud.com/docs#authentication
[10] V. Mühler, *justadudewhohacks/face-api.js*. (Jun. 28, 2025). TypeScript. Accessed: Jun. 29, 2025. [Online]. Available: https://github.com/justadudewhohacks/face-api.js
[11] S. Iftekar, "The 'face-api.js' Library for Accurate Face Recognition in Web- Applications and Possible use Cases with Accuracy Metrics," *Int. J. Comput. Appl.*, vol. 186, 2024.
[12] Y. Indrasari, "Efesiensi Saluran Distribusi Pemasaran Kopi Rakyat Di Desa Gending Waluh Kecamatansempol (Ijen) Bondowoso," *J. Manaj. Pemasar.*, vol. 14, no. 1, pp. 44–50, Mar. 2020, doi: 10.9744/pemasaran.14.1.44-50.
[13] "FER-2013." Accessed: Jul. 24, 2025. [Online]. Available: https://www.kaggle.com/datasets/msambare/fer2013
[14] D. R. Kusuma and F. P. Artharina, "Analisis Implementasi Projek Penguatan Profil Pelajar Pancasila Sekolah Dasar Negeri 01 Penyangkringan Kecamatan Weleri Kabupaten Kendal," *Didakt. J. Ilm. PGSD FKIP Univ. Mandiri*, vol. 10, no. 04, Dec. 2024.
[15] R. F. Utami, S. Prasetyo, and D. Z. Nurizin, "Validitas dan Reliabilitas Kuesioner Chinese Positive Youth Development Scales (CPYDS) Mengukur Keterampilan Hidup Pelajar SMP di Babakan Madang Kabupaten Bogor 2019," *J. Biostat. Kependud. Dan Inform. Kesehat.*, vol. 2, no. 3, Jul. 2022, doi: DOI: 10.7454/bikfokes.v2i3.1029.
[16] S. Ishak, M. H. Koniyo, and N. Pakaya, "Analisis Pengaruh Kualitas Sistem Informasi dan Kualitas Informasi Terhadap Kepuasan Pengguna Sistem Informasi Skripsi dan Kerja Praktek (SISKP) Jurusan Teknik Informatika Universitas Negeri Gorontalo," *Diffus. J. Syst. Inf. Technol.*, vol. 2, Jan. 2022.

Authors' information

Dimas Aditya Putra Wardhana received a Bachelor of Applied Engineering degree from Intitut Teknologi Sepuluh Nopember, and a Master of Applied Engineering from Politeknik Elektronika Negeri Surabaya. He is currently pursuing a Doctoral degree in Electrical Engineering at Universiti Tun Hussein Onn Malaysia (*UTHM*). He works as a Lecturer and Head of Study Program at the Department of Robotics and Artificial Intelligence Engineering, Faculty of Engineering, Universitas 17 Agustus 1945 Surabaya. His research interests include robotics, electric vehicles, and industrial automation.

Fridy Mandita received a Bachelor of Computer Science degree from Universitas Gadjah Mada Yogyakarta, and a Master of Science from King Mongkut's University of Technology North Bangkok. He is currently pursuing a Doctoral degree in Computer Science at Gadjah Mada Yogyakarta. He works as a Lecturer at Universitas 17 Agustus 1945 Surabaya. His research interests include machine learning, computer vision, natural language processing, and project management.

Elvianto Dwi Hartono received a Bachelor of Engineering degree from Universitas 17 Agustus 1945 Surabaya, Master of Engineering from Institut Sains dan Teknologi Nasional Jakarta, Master of Computer Science from STMIK Nusa Mandiri Jakarta, Master of Management from STIE Mahardhika Surabaya. He is currently pursuing a Doctoral degree in Agricultural Industrial Technology at Universitas Brawijaya Malang. He works as a Lecturer Universitas 17 Agustus 1945 Surabaya. His research interests include computer vision, internet of things, and robotika.

Bima Putra Purnawirawan is currently pursuing a Bachelor's degree in Information Systems and Technology at Universitas 17 Agustus 1945 Surabaya. He graduated from SMK Negeri 2 Bojonegoro. His research interests in Information Technology.

Dwi Hendra Rosli is currently pursuing a Bachelor's degree in Information Systems and Technology at Universitas 17 Agustus 1945 Surabaya. He graduated from SMA Negeri 6 Surabaya. His research interests in data analytics.

Muhammad Aldi Syafrillah Hidayat is currently pursuing a Bachelor's degree in Information Systems and Technology at Universitas 17 Agustus 1945 Surabaya. He graduated from SMA Al Islam Krian, Sidoarjo. He has a strong interest in programming as well as data and system analytics, and he is passionate about exploring how technology can be leveraged to solve real-world problems efficiently.