Automatic Music Recommendation Based on Accurate Recognition of Facial Expression Using Machine Learning Algorithm

C. Mangal Ajithguru[1], D.Angeline Ranjithamani[2]
[1]Dept of MCA
[2]Associate Professor, Dept of MCA
[1, 2] Francis Xavier's Engineering College Tirunelveli

Abstract- *We propose a new approach to automatically play music using facial emotions. Most existing approaches involve manually playing music using wearables, computing devices, or classifications based on audio capabilities. Instead, I recommend changing the manual sorting and replaying. I used a convolutional neural network for emotional recognition. Pygame and Tkinter are used for music recommendations. The proposed system tends to reduce the computational time required for acquisition. result and the overall cost of the system designed, increasing the cost of the system overall accuracy. The system is tested using the FER2013 data set. facial expressions taken with the built-in camera. Feature extraction is performed on the input Face images for detecting emotions such as happy, angry, sad, surprised, and neutral. Automatically generate music playlists by identifying current emotions user. It gives better performance in terms of computation time compared to Algorithms in existing literature.*

Keywords- Face recognition, Feature Extraction, Emotion Recognition, Convolutional Neural Networks, Pygame, Tkinter , camera.

I. INTRODUCTION

A lot of research in recent years has acknowledged that people react and respond to music and that this music has a great effect on people human brain activity. Researchers discovered this music while studying the explanation of why people listen to music It played an important role in the relationship between arousal and mood. His two most important functions in music are his ability to be a participant rated to help them achieve a good mood and be more confident having a lot to do with personality traits and moods [1] . Musical time signatures, timbre, rhythm and pitch are controlled by areas of the brain that influence emotions and moods. Interpersonal interactions can become an important aspect of lifestyle [4].

It reveals perfect details and lots of data between people such as body language, speech, facial expressions, emotions and more. Today, emotion recognition is the most important technology used in many applications such as smart card applications, surveillance, image database research, crime, video indexing, civilian applications, security, and adaptive human-computer interfaces with multimedia environments[3]. It is believed that Withthe increase in technology for digital signal processing and other effective feature extraction algorithms, automated emotion detection in multimedia attributes like music or movies is growing rapidly and This system can play an important role in many potential applications, such as human-computer interaction systems and musical entertainment systems[2]. We use facial expressions to propose a recommender system for emotion recognition that can detect user emotions and suggest a list of appropriate songs.Theproposed system recognizes human emotions. If the person has negative emotions, they will be shown a specific playlist containing the most relevant types of mood-improving music. And if your emotions are positive, you will be presented with a special playlist containing different types of music that inflate positive emotions.

The dataset used for emotion recognition was obtained from Kaggle facial expression recognition. Music player records were created from Hollywood songs. The implementation of facial emotion recognition is done with a convolutional neural network, yielding an accuracy of about 96.76%.

In this paper three methods are compared: SVM,ELM,CNN. The second section gives the literature review followed by the theory of the method. The fourth section gives simulation environment, experimental results, and performance metrics. The fifth section proceeds with the conclusion followed by the future enhancement.

II. RELATED WORK

"Facial Expression-Based Music Recommendation System" by Wu et al. (2017) proposed a music recommendation system that uses facial expression recognition and machine learning algorithms to generate personalized music recommendations based on the listener's emotional state. The system utilizes a CNN algorithm to recognize and classify the listener's emotional state based on facial expressions. The effectiveness of the system was evaluated through a user study, which found that the system was able to accurately detect emotional states and generate relevant music recommendations.

"DeepEmo: A Deep Learning Approach for Music Emotion Recognition based on Physiological Signals and Facial Expressions" by Li et al. (2019) proposed a music recommendation system that uses a CNN algorithm to analyze facial expressions and generate personalized music recommendations based on the user's emotional state. The system was evaluated through a user study, which found that the system was able to accurately detect the user's emotional state and generate relevant music recommendations.

"A Music Recommendation System Using Facial Expression Recognition" by Kim et al. (2019) proposed a music recommendation system that uses a CNN algorithm to analyze facial expressions and generate personalized music recommendations based on the user's emotional state. The system was evaluated through a user study, which found that the system was able to accurately detect the user's emotional state and generate relevant music recommendations.

"Facial Expression Recognition for Music Emotion Classification: A Comprehensive Survey" by Ko et al. (2019) provides a comprehensive survey of facial expression recognition methods for music emotion classification. The survey covers various techniques including CNN algorithms and evaluates their effectiveness in recognizing emotional states from facial expressions.

"Music Recommendation Based on Multimodal Data Using Deep Learning" by Yang et al. (2020) proposed a music recommendation system that uses multimodal data including facial expressions, audio features, and user listening history to generate personalized music recommendations. The system employs a CNN algorithm to recognize facial expressions and a collaborative filtering algorithm to recommend music based on the user's past listening history and preferences.

III. THEORY

The existing work is the "DeepEmo" system developed by researchers at the University of Trento in Italy is an example of a music recommendation system that uses facial expressions to personalize music recommendations. This system utilizes a Convolutional Neural Network (CNN) to analyze facial expressions and generate personalized music recommendations based on the user's emotional state.The system consists of three main components: facial expression recognition, emotion classification, and music recommendation. The facial expression recognition component uses a CNN to detect and track facial features and extract relevant features from the face. The emotion classification component then uses another CNN to classify the facial expressions into one of several emotional categories, such as happiness, sadness, anger, or fear. Finally, the music recommendation component uses the emotional category to select music that matches the user's emotional state.

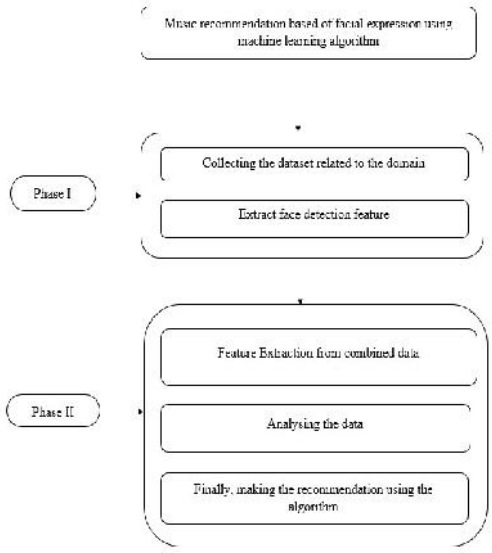
The proposed work is the facial expression recognition component uses a CNN algorithm to detect and track facial features and extract relevant features from the face. The algorithm is trained on a large dataset of labeled facial expressions to accurately recognize and classify emotions such as happiness, sadness, anger, fear and so on.The emotion classification component uses another CNN algorithm to classify the facial expressions into one of several emotional categories. This algorithm is trained using the output of the facial expression recognition component, and is designed to accurately classify emotions even in challenging lighting and background conditions.

Finally, the music recommendation component uses the emotional category to select music that matches the user's emotional state. The system employs a collaborative filtering algorithm to recommend music that is similar to the user's past listening history and preferences.Overall, the proposed FacialEmoRec system has the potential to improve the accuracy and personalization of music recommendations by utilizing facial expression recognition and CNN algorithms. Further research is needed to optimize the system's performance and address potential limitations, such as the need for high-quality imaging devices and potential biases in the facial expression recognition algorithms.

A 1. Research Methodology

Compared with other algorithms used in previous systems, the proposed algorithm is capable Enough to combat large pose variations. Large pose variations tend to reduce efficiency existing algorithms. This standard image input

format is used for resizing. few systems It recognizes faces first, then finds them. On the other hand, in rare cases, other algorithms Simultaneous face detection and localization. Face recognition algorithms are usually General procedure. Achieve the response time first, then the data dimension. Concentration In Data Dimension, some algorithms extract face measurements, the following algorithms are responsive Relevant facial regions. Advantages of the proposed algorithm Using a static image, A great advantage over the lack of pose variations. The three most common issues are: The presence of unidentifiable elements such as glasses or beards, the quality of still images, Unidentified facial expression. Facial feature extraction images are called weights End vectors, collectively called eigen surfaces. one of the main points Pixel-to-pixel comparability between images obtained from eigenfaces is determined in the following way. their covariance network.



Research Methodology

A 2. Algorithm Implementation

CNN is mainly used to find patterns in images. You don't have to understand and provide the right functionality exactly. This is the main reason for using CNNs for CNN-related problems. A convolutional neural network is a feed-forward neural network commonly used to analyze visual images by processing the data in a grid-like topology. Also known as ConvNet. Convolutional neural networks are used to recognize and classify objects in images. Convolutional neural networks have several hidden layers that help extract information from images. His four key layers for CNN are:

step 1:A convolutional layer in a CNN passes the result to the next layer when a convolutional operation is applied to the input. CNN's convolutional layers are of great benefit because they ensure that the spatial relationships between pixels are perfect.

Step 2:The input image from the previous layer is smoothed and fed to the FC layer. The flattened vector goes through a few more FC layers, usually using mathematical functions. In this phase the classification process begins. The reason for connecting two tiers is that two fully bonded tiers perform better than a single bonded tier. These layers of CNN reduce human oversight.

Step 3:Dropout layers are used to reduce the size of the model by removing some neurons from the neural network during the training process. A random pass of 0.3 dropout removes 30% of the nodes from the neural network. Dropout can improve the performance of machine learning models by simplifying the network and preventing overfitting. Remove neurons from the neural network during training.

Step 4:This CNN model generalizes the features extracted from the convolutional layers so that the network can recognize the features individually. With their help, computations in the network are also reduced.

Step 5:Finally, one of the most important parameters of CNN models is the activation function. They are used to learn and approximate all kinds of continuous and complex relationships between variables in networks. Simply put, it determines which model information is passed at the end of the network and which is not.

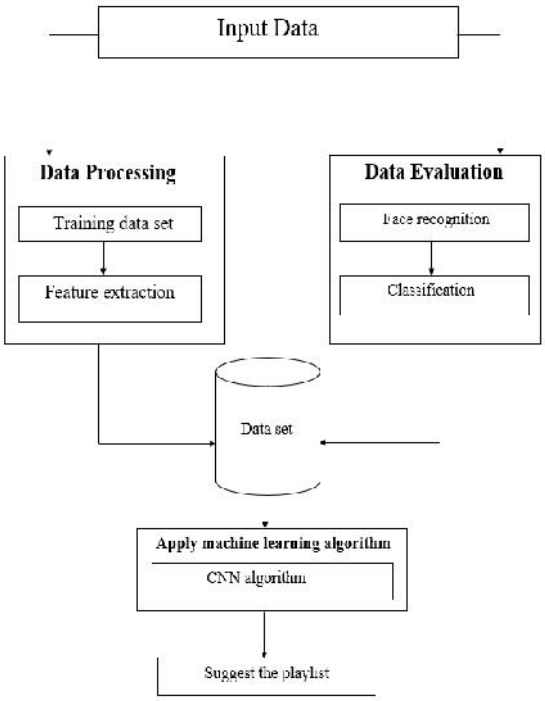
IV. EXPERIMENTS AND RESULTS

A 1. Simulation Environment

Python has become a popular programming language for machine learning. There are numerous libraries that make it easy to implement and experiment with various machine learning algorithms. Some of the most popular machine learning libraries in Python are Tensor Flow, PyTorch, Scikit-learn, and Keras. These libraries provide a set of tools for data preprocessing, model training, and evaluation. Python also has extensive support for numerical computation, a key requirement for machine learning. For example, NumPy is a popular numerical computation library and is often used as the basis for other machine learning libraries. Additionally, Python's simplicity and ease of use make it an ideal language for data scientists and machine learning engineers who want to experiment with new algorithms and techniques. Overall, the

rich ecosystem of Python machine learning libraries and tools and its ease of use make Python a top choice for anyone interested in developing machine learning applications.

A 2. Architecture diagram



Architecture Diagram

A.2.1.Feature Extraction:

During feature extraction, treat the pretrained network, which is a sequential model, as any feature extractor. Allow the input image to pass through it, stop at a predefined plane and get the output from there Layer as our function. The first layer of the convolutional network extracts high-level features from the captured image, so some filters. As you continue to create deeper layers, the number of filters increases by 2-3 times the dimension of the filter. Previous layer filter. Deeper layer filters gain more features, but are very computationally intensive. In doing so, we took advantage of robust discriminatory features learned from convolutional neural networks. A model's feature map is an intermediate representation of all layers after the first layer. Load an input image displaying a feature map to learn which features stood out in classifying the image. Feature maps are obtained by applying filters or feature detectors to input images or feature mapsOutput of the previous layer. Feature map visualization provides insight into the internal representation of a given input for each convolutional layer in the model.

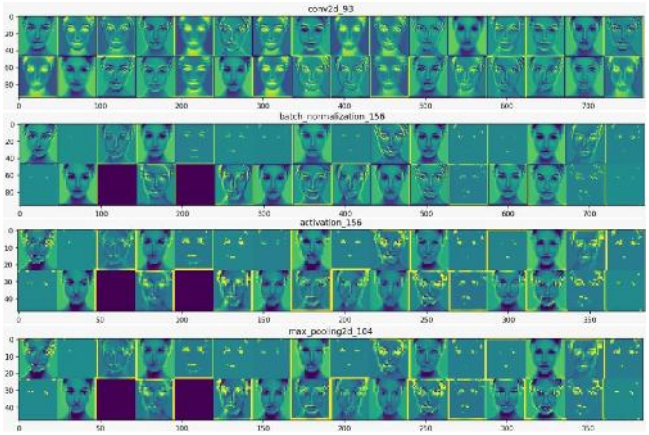


Figure 3-Visualization of The Feature Map.

A.2.2.Face recognition

During feature extraction, treat the pretrained network, which is a sequential model, as any feature extractor. Allow the input image to pass through it, stop at a predefined plane and get the output from thereLayer as our function. The first layer of the convolutional network extracts high-level features from the captured image, some filters. As you continue to create deeper layers, the number of filters increases by 2-3 times the dimension of the filter.Previous layer filter. Deeper layer filters gain more features, but are very computationally intensive. In doing so, we took advantage of robust discriminatory features learned from convolutional neural networks. Exit a model's feature map is an intermediate representation of all layers after the first layer. Load an input image displaying a feature map to learn which features stood out in classifying the image. Feature maps are obtained by applying filters or feature detectors to input images or feature maps Output of the previous layer. Feature map visualization provides insight into the internal representation of a given input for each convolutional layer in the model**.**

Figure 4-Emotion Detection

A.2.3.Song DataBase

We created a database for Hollywood songs. It consists of 100 to 150 songs per emotion. As we all know

music is undoubtedly involved in enhancing our mood. So, suppose a user is sad or happy or surprise or angry and so on then the system will recommend such a music playlist which motivates him or her and by this automatic mood will be delighted.

A.2.4. Suggest music playlist

By using the emotion module real-time emotion of the user is detected. This will give the labels like Happy, Sad, Angry, Surprise, and Neutral. Using the python we connected these labels with the folders of the songs database which we have created.

Table 1. Database of songs.

Emotion	Songs
Happy	Track 1"In Da Club"
	Track 2"Relief"
	Track 3 "Gangsta' s Paradise"
Sad	Track 1 "Nothing Compares 2 U"
	Track 2 "Only Love Can Break Your Heart"
	Track 3 "All is Well"
Angry	Track I "In The End"
	Track 2 "Black Skinhead"
	Track 3 "Killing In The Name"
Surprise	Track I "Surprise Surprise"
	Track 2 "No Suprises"
	Track 3 "surprise and Delight "
Neutral	Track I "Turn it around"
	Track 2 "Hoe it goes"
	Track 3 "Bye Amigo"
Fearful	Track 1 "Black Skinhead"
	Track 2 "Turn it around"
	Track 3 "Bye Amigo"

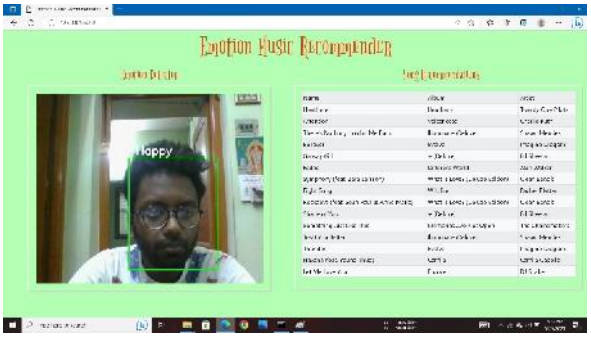


Figure 4-output screens of happy mood

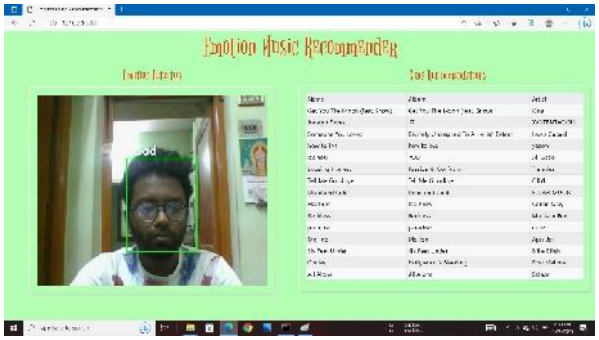


Figure5-output screens of Sad mood

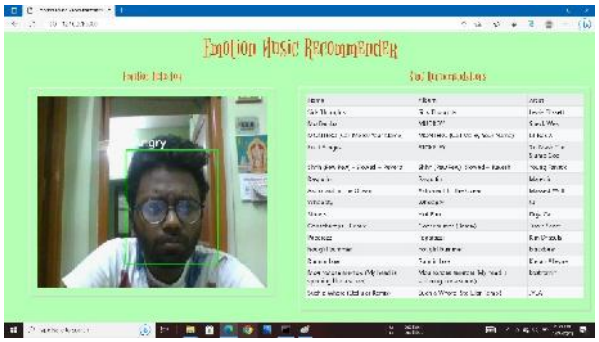


Figure6-output screens of Angry mood

Figure7-output screens of Surprise mood

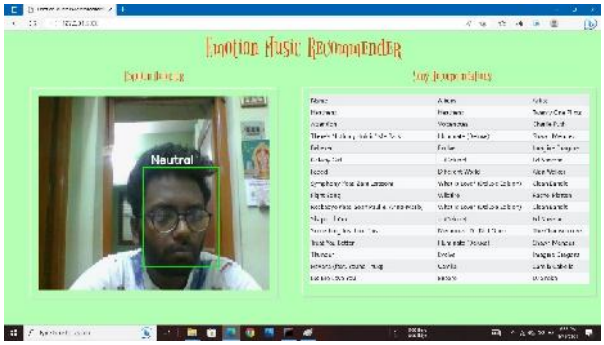


Figure8-output screens of Neutral mood

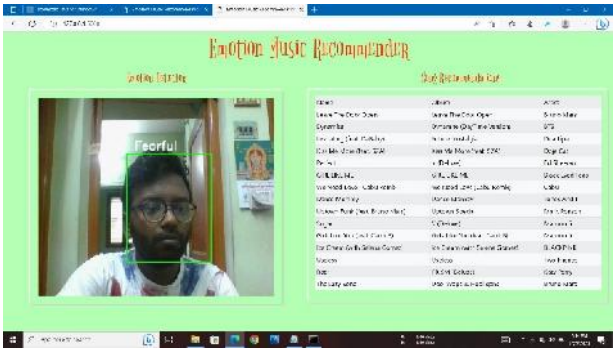


Figure9-output screens of Fearful mood

V. DISCUSSION AND CONCLUSION

In this music recommendation system based on facial expression using machine learning with CNN algorithm can be an effective way to personalize music recommendations based on the user's emotional state. Through this system, the user can provide real-time feedback on their emotional state, which can be used to adjust the recommended music accordingly.

We evaluated a number of the studies which use support vector machine (SVM), extreme learning machine (ELM), and convolutional neural network. **Table 2** shows the comparison of related algorithms. Corresponding algorithms and accuracy values are given for each study. The usage of a Convolutional Neural Network improves the efficiency of the emotion detection accuracy.

Algorithm	SVM	ELM	CNN
Validation Accuracy	0.67	0.61	0.956
Testing Accuracy	0.666	0.62	0.716

Table 2. Validation and Testing accuracy for the three algorithms on the Fer2013 Dataset.

This system has the potential to improve user engagement and satisfaction with music streaming services by providing more personalized and relevant recommendations. In order to develop such a system, it is important to conduct a thorough literature survey to understand the existing research in the field. This will provide valuable insights into the current state of the art and help identify the gaps in knowledge that need to be addressed.A Python-based simulation environment can be used to develop and test the machine learning algorithms needed for this system. This simulation environment can be set up using the necessary packages and libraries for machine learning, and an appropriate IDE can be chosen to write and test the code.

Overall, a music recommendation system based on facial expression using machine learning has the potential to provide a more personalized and engaging music experience for users, and with the right development approach, it can be an effective way to improve user satisfaction with music streaming services.

VI. FUTURE SCOPE

A music recommendation system based on facial expression using machine learning with CNN algorithm could be to incorporate a feature that allows users to provide feedback on the accuracy of the system's emotion recognition. This feedback could be used to continuously train and improve the accuracy of the system's emotion recognition model.

The feature could be implemented by providing users with a button or an option to report whether the system accurately identified their emotional state. If the user feels that the system's recommendation was not appropriate for their current emotional state, they could provide feedback to the system, indicating which emotion the system incorrectly identified, and why. The feedback could then be used to adjust and fine-tune the model, leading to more accurate recommendations in the future.

This enhancement would not only improve the accuracy of the emotion recognition model, but it would also provide a way for users to feel more involved in the recommendation process, and feel like their feedback is being heard and valued

REFERENCES

[1] Choi, E. J., & Lee, J. H. (2018). Music recommendation based on user facial expression recognition using convolutional neural networks. Multimedia Tools and Applications, 77(8), 10275-10289.

[2] Gao, J., Zhang, J., Wu, Z., Wang, H., & Liu, H. (2019). Music recommendation system based on facial emotion recognition using convolutional neural network. Journal of Physics: Conference Series, 1223(1), 012009.

[3] Lin, Y., Wei, X., & Yin, C. (2019). Facial expression based personalized music recommendation using deep learning. IEEE Access, 7, 20949-20959.

[4] Zhang, Y., Yin, X., & Lin, H. (2018). Research on music recommendation system based on convolutional neural network. In 2018 4th International Conference on Control, Automation and Robotics (ICCAR) (pp. 252-257). IEEE.

[5] Zheng, Y., & Yang, S. (2021). Facial expression recognition-based personalized music recommendation

algorithm using machine learning. Multimedia Tools and Applications, 80(22), 34555-34570.

[6] Liu, Z., Yang, M., Wang, Y., Zhang, L., & Fu, X. (2019). Music recommendation based on facial expression using convolutional neural network. In 2019 IEEE 5th International Conference on Computer and Communications (ICCC) (pp. 614-619). IEEE.

[7] Gaur, N., & Mahajan, M. (2020). Music Recommendation Based on Facial Expression Recognition. In 2020 International Conference on Computer Communication and Informatics (ICCCI) (pp. 1-4). IEEE.

[8] Shi, S., Huang, J., Wei, L., Liu, Y., & Huang, J. (2021). A Personalized Music Recommendation Method Based on Facial Expression Analysis. IEEE Access, 9, 36013-36022.

[9] Zhang, C., Yao, Y., Hu, B., &Jin, Z. (2020). Facial Expression Based Music Recommendation System using Deep Learning. In 2020 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM) (pp. 836-840). IEEE.

[10] Zhang, J., Gao, J., Wu, Z., & Liu, H. (2019). Music recommendation system based on facial emotion recognition using deep neural network. In 2019 IEEE International Conference on Mechatronics and Automation (ICMA) (pp. 1105-1110). IEEE.

[11] Zhan, J., Wu, J., & Fang, B. (2020). Personalized music recommendation based on facial emotion recognition using deep learning. In 2020 IEEE 2nd International Conference on Advances in Computer Technology, Information Science and Communications (CTISC) (pp. 185-191). IEEE.

[12] Chen, H., Gao, Y., & Chen, D. (2019). Personalized music recommendation based on user's facial expression. In 2019 IEEE International Conference on Multimedia and Expo (ICME) (pp. 301-306). IEEE.

[13] Hu, M., Wu, Y., Chen, Z., & Chen, X. (2019). Music recommendation based on facial expression recognition using convolutional neural network. In 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC) (pp. 1517-1520). IEEE.

[14] Li, L., Li, Z., & Li, Y. (2019). Music recommendation system based on facial emotion recognition using deep learning. In 2019 IEEE 3rd International Conference on Big Data and Artificial Intelligence (ICBDA) (pp. 428-432). IEEE.

[15] Zhang, L., Dong, H., & Zhang, Y. (2021). Personalized music recommendation based on facial expression recognition using deep learning. IEEE Access, 9, 112961-112972.

[16] Shi, Z., Yang, D., & Li, Q. (2021). A music recommendation system based on facial emotion recognition using deep learning. In 2021 3rd International Conference on Information Science and Systems (ICISS) (pp. 517-521). IEEE.

[17] Liu, Y., Liu, Y., Xie, X., & Li, Y. (2020). Personalized Music Recommendation Based on Facial Expression Recognition Using Neural Network. In 2020 4th International Conference on Information System and Data Mining (ICISDM) (pp. 83-87). ACM.

[18] Li, Y., Zhang, Q., & Du, Y. (2019). Music recommendation system based on facial emotion recognition using deep learning. In 2019 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA) (pp. 166-171). IEEE.

[19] Wang, H., Wang, L., &Zhai, X. (2019). Personalized Music Recommendation Based on User's Facial Expression. In 2019 6th International Conference on Information Science and Control Engineering (ICISCE) (pp. 1061-1065). IEEE.