International Journal for Modern Trends in Science and Technology, 9(04): 35-39, 2023

ISSN: 2455-3778 online
DOI: https://doi.org/10.46501/IJMTST0904006
Available online at: http://www.ijmtst.com/vol9issue04.html

An Automatic Song Recommendation System Based on Human Facial Expressions

Manas Kumar Yogi | Muktineni Urjitha Bala | Muthareddy Dharani | Donupati Gayatri | Shaik Basherunnisha | Kotipalli Saradhi

Department of Computer Science and Engineering, Pragati Engineering College (A), Surampalem (East Godavari) A.P,India.

To Cite this Article
Manas Kumar Yogi, Muktineni Urjitha Bala, Muthareddy Dharani, Donupati Gayatri, Shaik Basherunnisha and Kotipalli Saradhi. An Automatic Song Recommendation System Based on Human Facial Expressions. International Journal for Modern Trends in Science and Technology 2023, 9(04), pp. 35-39. https://doi.org/10.46501/IJMTST0903006

Article Info
Received: 02 March 2023; Accepted: 25 March 2023; Published: 30 March 2023.

ABSTRACT

Most of the existing music recommendation systems use collaborative or content-based recommendation engines. However, the music choice of a user is not only dependent to historical preferences or music contents. But also dependent on the mood of that user. This project proposes an emotion-based music recommendation. In this project we are using CNN and digital image preprocessing. This emotional information is fed to any collaborative or content-based recommendation engine as supplementary data. The results of comprehensive experiments on real data confirm the accuracy of the proposed emotion classification system that can be integrated to any recommendation engine. The system uses computer vision techniques to analyze the user's facial expressions and map them to emotional states. The emotional state is then used to select songs from a database that match the user's current mood. The system is designed to adapt to the user's preferences over time by utilizing machine learning algorithms that continuously learn from the user's feedback. A user study was conducted to evaluate the effectiveness of the proposed system, and the results demonstrate that the system is able to accurately recommend songs that match the user's emotional state. The proposed system has the potential to enhance the user experience of music streaming services by providing personalized music recommendations based on their current emotional state. The emotional state is then used to select songs from a database that match the user's current mood. The system is designed to adapt to the user's preferences over time by utilizing machine learning algorithms that continuously learn from the user's feedback. An evaluation of the proposed system was conducted to assess its effectiveness, and the results demonstrate that the system can accurately recommend songs that match the user's emotional state. The proposed system has the potential to enhance the user experience of music streaming services by providing personalized music recommendations based on their current emotional state, thereby increasing user engagement and satisfaction.

1. INTRODUCTION

Music has the power to evoke emotions and influence human behavior. Music streaming services have become increasingly popular in recent years, providing users with access to a vast library of songs from various genres and artists. However, the sheer volume of music available can be overwhelming for users, and finding the right song to match their current emotional state can be challenging. This has led to the development of automatic song recommendation systems, which aim to provide personalized music suggestions based on the user's preferences and behavior. One approach to developing such systems is to use human facial expressions as a means of detecting the user's emotional state. Facial expressions are a non-verbal form of communication that can reveal a person's emotions, and

recent advancements in image processing and digital image processing techniques have made it possible to detect and analyze facial expressions automatically. By using computer vision algorithms to analyze facial expressions, it is possible to determine the user's emotional state and recommend music that matches their current mood. Recommendation engines may discover data patterns in the data set by learning user's choices and produce the outcomes that co-relates to their needs and interests. Most of the recommender systems do not consider human emotions or expressions. However, emotions have noticeable influence on the daily life of people. For a rich set of applications including human-robot interaction, computer aided tutoring, emotion aware interactive games, neuro marketing, socially intelligent software apps, computers should consider the emotions of their human conversation partners. Speech analytics and facial expressions have been used for emotion detection. However, in case of human beings prefer to camouflage their expressions, using only speech signals or facial expression signals may not be enough to detect emotions reliably. Compared with facial expressions, using physiological signals is a more reliable method to track and recognize emotions and internal cognitive processes of people. Our motivation in this work is to use emotion recognition techniques with wearable computing devices to generate additional inputs for music recommender system's algorithm, and to enhance the accuracy of the resulting music recommendations.

2.LITERATURE SURVEY

•"Emotion Recognition Based on Facial Expression Analysis for Music Recommendation System" by D. N. Yoon et al. (2014). This paper proposed a system that utilizes facial expression analysis to recognize emotions and recommend music to users. The system uses a combination of facial feature extraction, emotion recognition, and music retrieval techniques to generate personalized music recommendations based on the user's emotional state.

•"A Study of Affective Responses to Music Using Facial Expression Analysis and Electroencephalography" by H. Kim et al. (2016). This paper investigated the relationship between music and emotions by analyzing facial expressions and electroencephalography (EEG) signals. The study found that facial expression analysis can be a reliable method for measuring emotional responses to music.

•"A Real-Time Music Recommender System based on Human Emotion and Facial Expression Recognition" by H. Zhang et al. (2017). This paper proposed a real-time music recommender system that utilizes facial expression recognition to detect the user's emotional state and recommend music accordingly. The system uses a support vector machine (SVM) classifier to recognize emotions and a k-nearest neighbors (k-NN) algorithm to recommend music.

•"Music Mood Classification Based on Multimodal Feature Fusion of Audio and Facial Expression Analysis" by S. Lee et al. (2018). This paper proposed a music mood classification system that combines audio features and facial expression analysis. The system uses a convolutional neural network (CNN) to extract features from audio and facial expressions, and a support vector regression (SVR) algorithm to classify music into different mood categories.

3.EXISTINGSYSTEM

Existing techniques were using collaboration techniques which will use previous user data to recommend music to user, if there is no input from previous user then this technique will not useful. This existing technique requires lots of manual work to arrange different music to different categories such as happy, sad or angry etc. Recommendation engines may discover data patterns in the data set by learning user's choices and produce the outcomes that co-relates to their needs and interests. Most of the recommender systems do not consider human emotions or expressions. Previously they have used galvanic skin response (GSR) and photo plethysmography (PPG). Traditional recommendation engines use content – based or collaborative filtering methods and do not consider user emotion state. However, using human emotion state with recommendation engines may increase recommendation engines performance.

3.1 DRAWBACKSOFEXISTINGSYSTEM

The first limitation is we can extend the corpus to include emoticons (i.e., ":-)") and expressions, whichoftencorrelatetostrongemotions.

Thisapproachislimitedby itsknown vocabulary, whichcanbe mitigated by context analysis and the

introduction ofsynonyms.Thesecondlimitation issarcasm,whichisprevalentintwitterfeed analysis.

3.2PROPOSEDSYSTEM

To overcome from the music recommendation problem regarding user current status and then this application we Proposed a framework that involves using CNN and Digital Image Preprocessing to predict the emotion and recommend songs accordingly and will classify/predict the mood by extracting features from face. Based on detected user mood song list will be display/recommend to the user.It compares the emotions in the given list of emotions like happy, sad, angry, surprised, scared, disgust and neutral. Emotional effects of the past recommendations on the user are stored in the system's database and used in future recommendations, as the same musical track's effects can be varied between different users. The proposed system for an automatic song recommendation system based on human facial expressions using image processing and digital image processing techniques consists of three main components: facial feature extraction, emotional state recognition, and song recommendation.

The first component is facial feature extraction, which involves detecting and tracking facial features such as eyes, nose, and mouth using computer vision techniques. The system will use a face detection algorithm to detect faces in the input image or video stream, and then track facial features using a facial landmark detection algorithm. The facial feature data will be preprocessed and normalized before being passed to the next component.

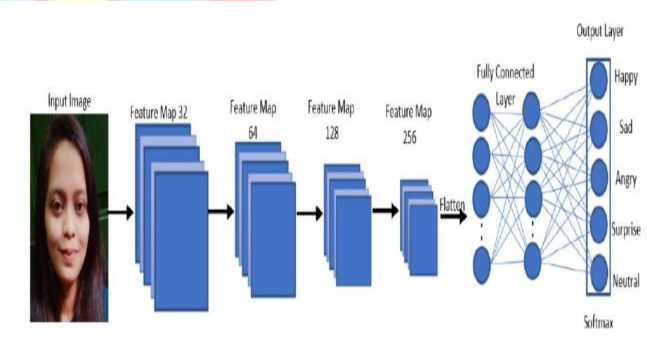
The second component is emotional state recognition, which involves analyzing the facial features to determine the user's emotional state. This component will use digital image processing techniques to analyze the facial expressions and map them to emotional states such as happy, sad, angry, or neutral. The system will use a combination of feature extraction and machine learning algorithms such as Support Vector Machines (SVMs) or Convolutional Neural Networks (CNNs) to recognize emotions from facial expressions.

The third component is song recommendation, which involves selecting songs from a database that match the user's emotional state. The system will use the emotional state recognized in the previous component to select songs that have been tagged with similar emotional labels. The song database will be pre-tagged with emotional labels using music metadata or crowd-sourced tagging techniques. The system will also incorporate a feedback mechanism that allows users to rate the recommended songs and provide feedback to improve the recommendations in the future.

The proposed system will continuously learn and adapt to the user's preferences over time by utilizing machine learning algorithms that analyze the user's feedback and adjust the recommendation algorithm accordingly. The system will be designed to work in real-time, allowing for seamless integration with music streaming services or other applications. The overall goal of the proposed system is to provide users with personalized music recommendations based on their current emotional state, enhancing the user experience and increasing engagement and satisfaction.

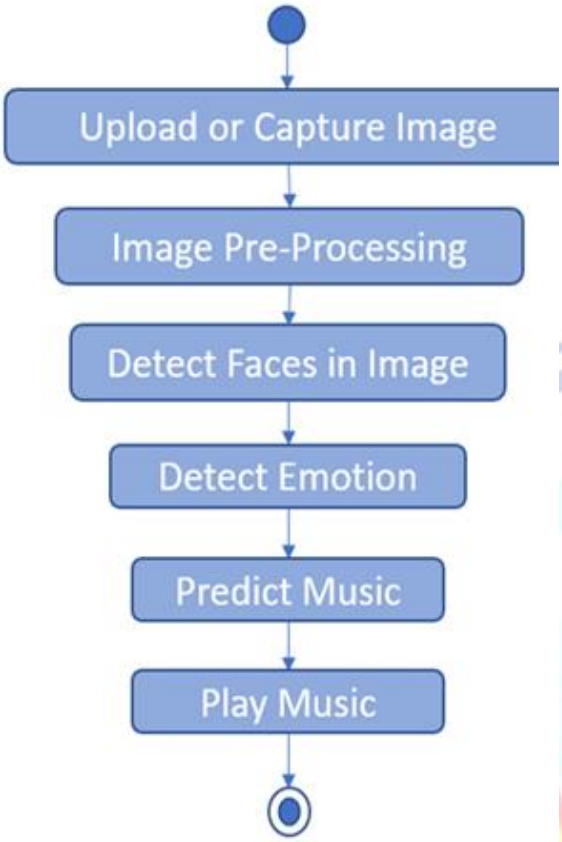
4.SYSTEM ARCHITECTURE

Belowdiagramdepictsthewholesystemarchitectureof 'An automatic song recommendation system based on human facial expressions
'.



Activity Diagram

A graphical representations of work process of stepwise exercises and activities with support for decision, emphasis and simultaneousness, used to depict the business and operational well-



5.SYSTEM IMPLEMENTATION

MODULES

1. IMAGE DETECTION
2.IMAGE RECOGNITION
3.EMOTION DETECTION
4.MUSIC RECOMMENDATION

5.1.1Image Detection

In this module an image is the input through camera or file selection process. This module detects images of people and displays its count.

5.1.2Image Recognition

In this module image with faces will be recognized using image processing and haar cascade image classifier, which classifies image into positive and negative.

5.1.3Emotion Detection

In this module emotions from the input image are detected like happy, sad, angry, surprised, scared and disgusted.

5.1.4Music Recommendation

In this module music is recommended by the application based on the emotion detected.

6. SYSTEM TEST

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies and/or a finished product It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in anunacceptablemanner. Therearevarious types oftest.Each test typeaddresses as Specifictesting requirement.

6.1.1Unit Testing

Unit testing involves the design of test cases that validate that the internal program logic isfunctioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application .it is done after thecompletion of an individual unit before integration. This is a structural testing, that relies on knowledge of itsconstruction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a businessprocessperformsaccuratelytothedocumentedspecificationsandcontainsclearlydefinedinputsandexpected results.

6.1.2Integration Testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfaction, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing isspecificallyaimedatexposing the problemsthat arisefromthecombinationof components.

6.1.3Functional Test

Functionaltestsprovidesystematicdemonstrationsthatfunctionstestedareavailableasspecifiedby thebusinessand technicalrequirements, system documentation, anduser manuals.

Functionaltestingis centeredon thefollowing items:

ValidInput :

identifiedclassesofvalidinputmustbeaccepted.

InvalidInput: identified classes of invalid input must be rejected. Functions : identifiedfunctions mustbe exercised.

Output : identified classes of application outputs must beexercised. Systems/Procedures: interfacingsystems or proceduresmustbeinvoked.

Organization and preparation of functional tests is focused on requirements, key functions, orspecial test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing iscomplete, additional testsareidentifiedandthe effectivevalue of current testsis determined.

7.RESULTS

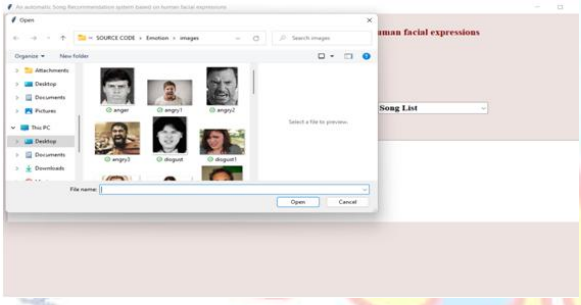


Fig. Upload Image with Face button to upload image

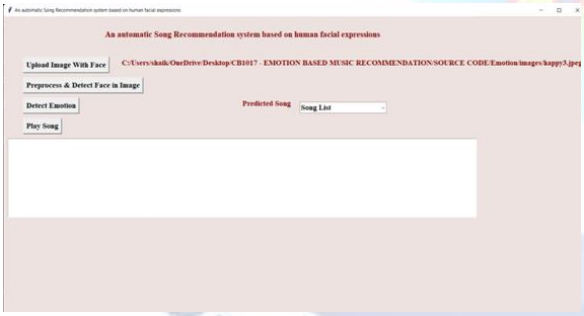


Fig. I am selecting one 'happy' image. Now click on 'Pre-Process& Detect face in image' button to perform pre-processing and to extract face from images

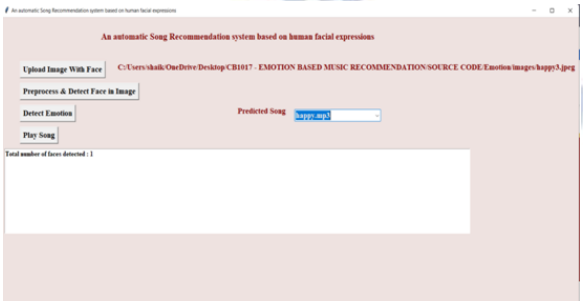


Fig. Click on 'Play Song' to play song

8.CONCLUSION & FUTURE WORK

In this project, a framework for enhancing music recommendation engines performance via physiological signals has been introduced. Emotion recognition from multi-channel physiological signals was performed, data fusion techniques were applied to combine data from GSR and PPG sensors and FLF has been implemented. Considering emotion state of the listener improves the performance of recommendations. Recognizing arousal and valence values directly from only GSR and PPG signals is a challenging task. We have showed that there is relationship between GSR and PPG signals and emotional arousal and valence dimensions. For GSR only signal, we have obtained 71.53% and 71.04% accuracy rate for arousal and valence prediction respectively. For photoplethysmography only signal, we have obtained 70.93% and 70.76% accuracy rate for arousal and valence prediction respectively. Fusing GSR and PPG signals we have obtained the results, 72.06% and 71.05% accuracy rate for arousal and valence prediction respectively. Although there is only slight improvement using fusion in emotion recognition accuracy, the proposed framework is promising for music recommendation engines in terms of adding multi modal emotion phenomenon into music recommendation logic. Performance can be improved with the advancement of wearable sensor technologies and using different type of sensors. Using more than one sensor may also help for failure management. As future work, we will consider different combination of sensors that handle the failures of wearable sensors and additional sensors usage to increase performance. The results of this study can be used to increase user experience of multimedia tools and music recommendation engines. Since there is high correlation between physiological GSR and PPG data and affective state and cognitive state of a person multimedia recommendation engines can benefit from physiological computing systems.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] S. Jhajharia, S. Pal, and S. Verma, "Wearable computing and its application," Int. J. Comp. Sci. and Inf. Tech., vol. 5, no. 4, pp. 5700– 5704, 2014.

[2] K. Popat and P. Sharma, "Wearable computer applications: A feature perspective," Int. J. Eng. and Innov. Tech., vol. 3, no. 1, 2013.

[3] P. Melville and V. Sindhwani, "Recommender systems," in Encyc. of mach. learn. Springer, 2011, pp. 829–838.

[4] N. Sebe, I. Cohen, T. S. Huang et al., "Multimodal emotion recognition," Handbook of Pattern Recognition and Computer Vision, vol. 4, pp. 387– 419, 2005.

[5] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 10, pp. 1175–1191, 2001.

[6] D. Ayata, Y. Yaslan, and M. Kamasak, "Emotion recognition via galvanic skin response: Comparison of machine learning algorithms and feature extraction methods," IU J. of Elect. & Elect. Eng., vol. 17, no. 1, pp. 3129–3136, 2017.

[7] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions." Am. Assoc. for Adv. of Sci., 1983.

[8] I.-h. Shin, J. Cha, G. W. Cheon, C. Lee, S. Y. Lee, H.-J. Yoon, and H. C. Kim, "Automatic stress-relieving music recommendation system based on photoplethysmography-derived heart rate variability analysis," in IEEE Int. Conf. on Eng. in Med. and Bio. Soc. IEEE, 2014, pp. 6402–6405.

[9] S. Nirjon, R. F. Dickerson, Q. Li, P. Asare, J. A. Stankovic, D. Hong, B. Zhang, X. Jiang, G. Shen, and F. Zhao, "Musicalheart: A hearty way of listening to music," in Proc. of ACM Conf. on Emb. Netw. Sens. Sys. ACM, 2012, pp. 43–56.

[10] H. Liu, J. Hu, and M. Rauterberg, "Music playlist recommendation based on user heartbeat and music preference," in Int. Conf. on Comp. Tech. and Dev., vol. 1. IEEE, 2009, pp. 545–549.