IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 4, April 2023

DOI: 10.17148/IARJSET.2023.10431

Real Time Music Recommender System

CH V S V Sastry[1], Janapala Sunil Reddy[2], R Vishnu Sekhar[3]

Student, Department of Computer Science Engineering, Hindustan Institute of Technology and Science, Chennai, India[1,2]

Assistant Professor, Department of Computer Science Engineering, Hindustan Institute of Technology and Science, Chennai, India [3]

Abstract: Music has consistently been a mainstream medium in light of the fact that it can loosen up our pressing factors of life. Nonetheless, what is interesting to an individual could move under his or her various feelings. For instance, the favoured in a pitiful mode is conceivably unique in relation to that in a cheerful way. To focus on this issue, Music information retrieval (MIR) was proposed for perceiving melodic feelings. Previously, although a few investigations have been made on feelings of acknowledgment, their adequacy isn't agreeable. A potential explanation is that the sound highlights removed are not strong enough to segregate the variety of sounds and feelings. Subsequently, in this paper, we propose a successful acknowledgment technique that melds Deep learning (DL), Machine Learning (ML), and Convolutional neural networks (CNN). The significant distinction between the proposed technique and customary sound-based investigations is that the proposed strategy totals the halfway acknowledgment consequences to accomplish better acknowledgment exactness. The trial results on a genuine dataset of CAL500 show that the proposed strategy performs in a way that is better than some other sound-based feeling marking techniques.

Keywords: Convolutional Neural Network, Deep Learning, Music Information Retrieval, Machine Learning.

I. INTRODUCTION

Mood or a person's emotional state can be inferred from their facial expressions. Individuals frequently use their facial expressions to convey their feelings. One of the things that can alter someone's mood is music. In addition to having sensory receptors, lips and eyes can also convey a person's mood. This article explains a computer program called a "emotion-based smart music player" that streamlines the process of making playlists and playing music for users. This project aims to record people's emotions through their facial expressions and play music that fits the circumstance, lifting the user's spirits in the process. Because it can collect and discern the feelings a person is expressing while gradually calming the user's mind, it typically produces a pleasing effect. Webcam interfaces, which are available on computer systems and are made to record human emotions, are used by musicians.

The system snaps a photo of the user when the application launches. The image was captured using a webcam. The phase of rendering starts after saving the previously acquired image. The user's mood may alter after some time, which may or may not change. Images are thus taken after each song or predetermined amount of time. And the following stage receives that image. In some cases, by altering your mood, you can even get over problems like depression and sadness. You can reduce your risk of several health problems and boost your mood by using facial analysis. The two types of image recognition systems are feature-based systems and image-based systems. The original approach takes possibilities from image parts, such as lips, noses, and mouths, and then shapes them to see how they relate to one another. Contrarily, the second system makes advantage of image pixels and is expressed as inbound techniques like principal component analysis and wavelet transform. It is utilized to categorize and identify images. Examples of models in models include happy, sad, angry, and neutral models. Moreover, it features a mixed mood feature. Depending on the user's level, a certain number of songs are given to each pattern model. Current emotional music players take longer than the suggested system.

II. LITERATURE REVIEW

Due to its numerous applications in music retrieval and recommendation, computational modelling of the emotional content of music has seen substantial study recently. Despite significant advancements, this task is still challenging since it is challenging to effectively describe emotion in music. The process of gathering emotional annotations and creating predictive models is complicated by the fundamentally subjective nature of musical emotion as experienced by people. This study tries to authentically convey the emotion of a song, not only subjectivity, rather than presuming that people can agree on the emotion of music.

DOI: 10.17148/IARJSET.2023.10431

We must customize our musical emotion recognition algorithms, which takes time, because musical emotion perception is highly subjective. The quickest way to personalize is explained in this post. Starting with a basic model that was offline trained using a generic user population, the procedure gradually modifies the model to account for music listeners using listener sentiment annotations. Our main goal is to minimize the quantity of user annotations needed for customization.

Three key research questions are investigated in this study, which highlights the strong connection between music and emotions.

Are there variations in how people perceive music (such as emotion, tempo, instrumentation, and complexity) depending on their musical background, experience, and factors like personality or demographics? Are musical qualities derived from music audio signals connected to how people perceive musical qualities like emotion and tempo? Are they connected to the extracted descriptors? We carried out two user studies with separate subject groups to investigate our study question.

A process created by Ramya Ramanathan and others. [1] An intelligent music player was informed by emotion identification in a proposed article. A fundamental aspect of human nature is emotion. They have the most significant role in life. Human emotions are there for us to share and comprehend. Local music choices made by a user are arranged first according to the mood the album evokes. This is frequently determined by taking into account the song's lyrics. paper is particularly Athavle, M., et al. Online ISSN: 2582-7006 Understanding Human Feelings and How to Preferably Use a Suggested Emotion Recognition System, 3 International Conference on Artificial Intelligence (ICAI-2021) Journal of Informatics Electrical and Electronics Engineering (JIEEE) A2Z Magazine. Also, a quick explanation of playlist creation and emotion classification is provided. The labour-intensive and time-consuming task of manually dividing playlists and annotating songs based on the user's current emotional state was suggested by CH Radhika et al. [8]. In order to automate this process, several methods have been suggested.

Common methods, however those are inefficient, need additional hardware (such EEG structures and sensors), raise the system's overall cost, and are far less precise. This study proposes an algorithm that automatically generates music playlists based on facial expressions, saving time and labour compared to manual creation of these playlists. The methods provided in the work are intended to cut down on the designed system's overall processing time and expense. It also seeks to increase system design correctness. Comparing the system's facial expression recognition engine against user-dependent and user-independent datasets serves to validate it.

Z. Zeng et al. [5] looked into a number of developments in emotion recognition technology. He concentrated on several strategies that may be employed with audio and/or visual recordings of emotional emotions. The audio-visual computer method is thoroughly discussed in this study. The prototype for emotions like happiness, sorrow, fear, wrath, disgust, and surprise has been identified as this impact. The challenges of establishing automatic and spontaneous emotion recognition that is useful for emotion recognition were the main topic of this research. We also noted a number of problems that unimodal emotion recognition neglected.

The concept of automating user-music player interaction was first up by Parul Tambe et al. [7], who suggested learning all of the user's preferences, feelings, and activities in order to enable song selection. The technology captured the user's various facial expressions in attempt to interpret the user's emotions and foretell the music genre.

Collaborative or content-based recommendation engines are used by the majority of current music recommendation systems. Yet, a user's musical preferences are not just determined by their past musical preferences and substance. But it also depends on how the individual is feeling. In this article, we suggest a framework for recommending music based on user moods that learns from signals collected by wearable physiological sensors. Wearable computing systems with physiological photoplethysmography (PPG) and galvanic skin response (GSR) sensors in particular classify human moods. The recommendation algorithms that are collaborative or content-based receive this sentiment data as extra data.

Collaborative or content-based recommendation engines are used by the majority of current music recommendation systems. Yet, a user's musical preferences are not just determined by their past musical preferences and substance. But it also depends on how the individual is feeling. In this article, we suggest a framework for recommending music based on user moods that learns from signals collected by wearable physiological sensors. Wearable computing systems with physiological photoplethysmography (PPG) and galvanic skin response (GSR) sensors in particular classify human moods. The recommendation algorithms that are collaborative or content-based receive this sentiment data as extra data.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified ⁙ Impact Factor 8.066 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 10, Issue 4, April 2023
DOI: 10.17148/IARJSET.2023.10431

III. PROPOSED METHODOLOGY

A type of deep learning algorithms known as convolutional neural networks (CNNs) is particularly effective in processing and recognising images. It has several layers, including fully linked, pooling, and convolutional layers. Convolutional layers, which apply filters to input images to extract characteristics like edges, textures, and forms, are a crucial part of CNNs. A pooling layer is used to minimise the spatial dimension while maintaining the most crucial data after the output of the convolutional layer is delivered to it. One or more fully connected layers used for picture prediction or classification and get the outputs of the pooled layers. A machine learning-based approach for a music recommender system was put out in this paper. The suggested framework can record the client's facial expressions, divide the main places on the face based on appearance, and organise them to record the client's particular feelings. The client is then shown a music that corresponds to their emotion once the emotions have been described.

The following are some of the suggested system's benefits:

1. Very quick computing of characteristics.
2. Effective feature choice.
3. A location- and scale-invariant detector.
4.We scale the features rather than the image itself.
5. High precision.

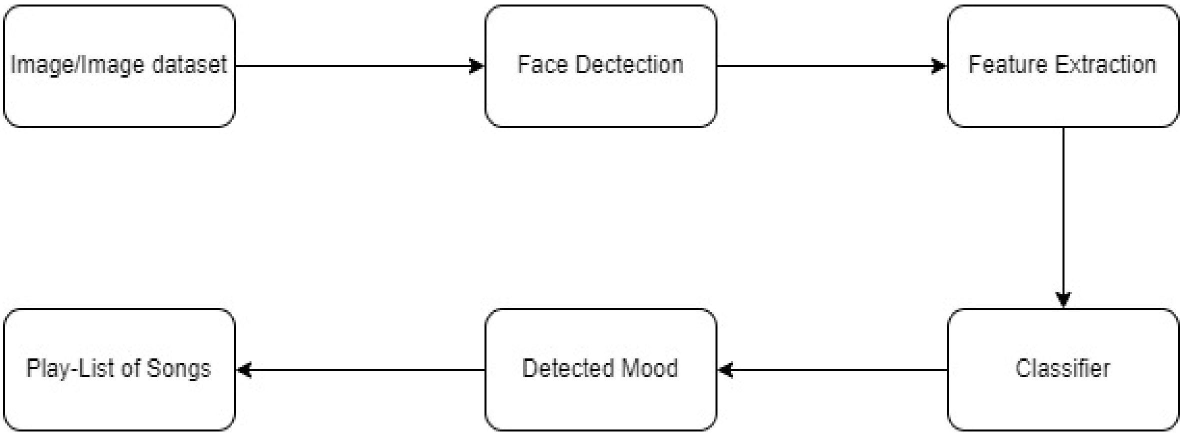

Fig. 1 Flow Chart

Figure 1 depicts the flowchart for the system we intend to create. To identify the user's emotion, the camera first scans their facial expressions, which are then matched to the photographs in the data set. A playlist of songs is afterwards suggested to the person once a classifier has later assessed the person's mood or emotion.

A. Way of Approach:

Dataset Collection: You can collect a dataset of facial expressions and corresponding music genres. This dataset can be used for CNN training and testing.

Pre-processing: Face images can be pre-processed to remove noise and unwanted features and converted to grayscale images. We can also augment the data by introducing random variations in the image, like this: B. Rotation and scaling.

CNN training: A convolutional neural network (CNN) can be trained to classify facial expressions into various emotions, such as happiness, sorrow, and anger, using pre-processed photos. With the aid of an appropriate optimizer like Adam and an appropriate loss function like categorical cross-entropy, a CNN may be trained.

Recommend music: Once taught, the CNN can be used to categorise the various emotions displayed by a user's facial expressions. The algorithm can suggest playlists of music genres that are likely to reflect the user's mood based on the emotions that were detected. For instance, if the system determines that the user is content, it may suggest lively and content music genres like dance and pop. Similar to this, the system might suggest slow and depressing musical genres like jazz and the blues when the user is feeling down.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 4, April 2023

DOI: 10.17148/IARJSET.2023.10431

User feedback: Users may be able to comment on the suggested music using a feedback mechanism included in the system. The system can learn suggestions and advance in the future based on the comments.

Deployment: The system can be deployed as a web or mobile her application that users can access on their devices.

email address or URL your paper, you must type out the address or URL fully in Regular font

A. References

The heading of the References section must not be numbered. All reference items must be in 8 pt font. Please use Regular and Italic styles to distinguish different fields as shown in the References section. Number the reference items consecutively in square brackets (e.g. [1]). When referring to a reference item, please simply use the reference number, as in [2].

Examples of reference items of different categories shown in the References section include:

- example of a book in [1]
- example of a book in a series in [2]
- example of a journal article in [3]
- example of a conference paper in [4]
- example of a patent in [5]
- example of a website in [6]
- example of a web page in [7]
- example of a databook as a manual in [8]
- example of a datasheet in [9]
- example of a master's thesis in [10]
- example of a technical report in [11]
- example of a standard in [12]

IV. RESULT ANALYSIS

In this study, we presented a powerful cascading CNN technique for identifying the eyes in facial photos. Even when the face is obscured, our approach can simultaneously identify and centre the left and right eye positions, and it is unaffected by visible light or infrared imaging. Moreover, the face detector is not dependent on the eye location. We evaluated the technique using more than 5,000 facial photos and discovered that the suggested eye detector is reliable and efficient.

Music recommender system we created produced very good results. The model which was trained over a large training and testing datasets were very successful in recognising the emotion of the person and suggested music depending on the emotion detected. The below image represents the epoch graph plotted for training accuracy against the validation accuracy

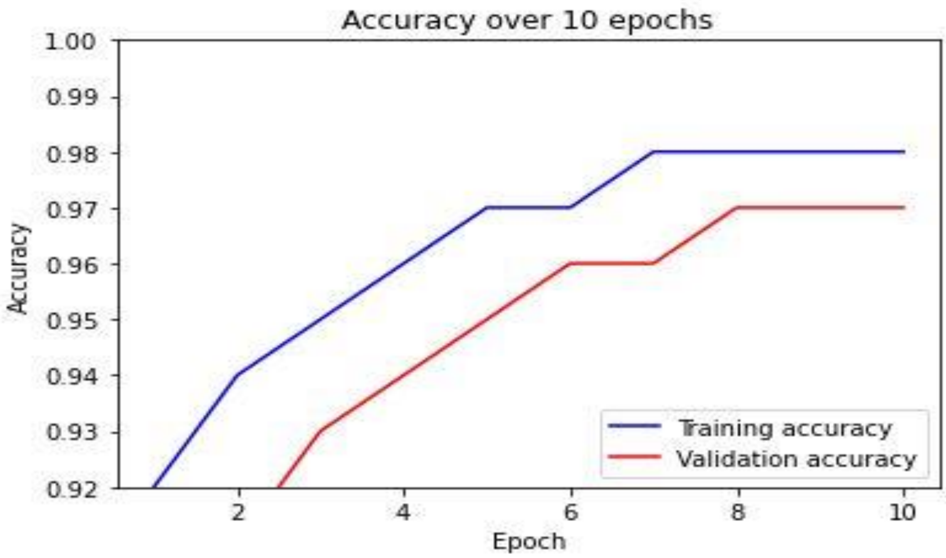

Fig. 2 Training vs Validation accuracy

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁑ Impact Factor 8.066 ⁑ Peer-reviewed / Refereed journal ⁑ Vol. 10, Issue 4, April 2023

DOI: 10.17148/IARJSET.2023.10431

CNN based expression-based music recommendation algorithms have recently produced promising results. This method's goal is to match the listener's emotional state with music by analysing their facial expressions. Several studies have shown that CNN-based models can accurately recognize facial expressions, which can be used to recommend music. However, this approach still has some limitations that should be addressed in future work.

Fig. 3 Playlist recommended for surprise expression

The requirement for sizable data sets on musical preferences and related facial expressions is one restriction. Such data requires time-consuming and expensive collection. Moreover, individual variations in musical preferences and facial expressions may have an impact on the system's accuracy.

The possibility of bias in datasets and models is another drawback. The system might not accurately represent everyone's tastes if the dataset that was used to train the model is not sufficiently diverse. Furthermore, suggestions may maintain any biases included in the model if it was trained on biased data. By gathering bigger and more varied datasets, creating models that can adjust to individual variances, and putting techniques in place to lessen data and model bias, future work in this field will overcome these constraints. You can concentrate on your work. In order to enhance the precision and customization of the system, it may also be intriguing to take into account the utilisation of additional forms of data, such as physiological signals or user input.

V. CONCLUSION

In conclusion, this project In this study, we presented a powerful cascading CNN technique for identifying the eyes in facial photos. Even when the face is obscured, our approach can simultaneously identify and centre the left and right eye positions, and it is unaffected by visible light or infrared imaging. Moreover, the face detector is not dependent on the eye location. We evaluated the technique using more than 5,000 facial photos and discovered that the suggested eye detector is reliable and efficient. CNN-based expression-based music recommendation algorithms have recently produced promising results. This method's goal is to match the listener's emotional state with music by analysing their facial expressions. Several studies have shown that CNN-based models can accurately recognize facial expressions, which can be used to recommend music.

However, this approach still has some limitations that should be addressed in future work. The requirement for sizable data sets on musical preferences and related facial expressions is one restriction. Such data requires time-consuming and expensive collection. Moreover, individual variations in musical preferences and facial expressions may have an impact on the system's accuracy. The possibility of bias in datasets and models is another drawback. The system might not accurately represent everyone's tastes if the dataset that was used to train the model is not sufficiently diverse.

Furthermore, suggestions may maintain any biases included in the model if it was trained on biased data. By gathering bigger and more varied datasets, creating models that can adjust to individual variances, and putting techniques in place to lessen data and model bias, future work in this field will overcome these constraints In order to enhance the precision and customization of the system, it may also be intriguing to take into account the utilization of additional forms of data, such as physiological signals or user input.

DOI: 10.17148/IARJSET.2023.10431

REFERENCES

[1]. R. Ramanathan, R. Kumaran, R. Ram Rohan, R. Gupta and V. Prabhu, "An Intelligent Music Player Based on Emotion Recognition," 2017 2nd International Conference on Computational Systems and Information Technology for Sustainable Solution (CSITSS), Bengaluru, India, 2017, pp. 1-5, doi: 10.1109/CSITSS.2017.8447743.

[2]. H. Immanuel James, J. James Anto Arnold, J. Maria Masilla Ruban, M. Tamilarasan, R. Saranya" EMOTION BASED MUSIC RECOMMENDATION SYSTEM": pISSN: 2395-0072 IRJET 2019.

[3]. Hafeez Kabani, Sharik Khan , Omar Khan , Shabana Tadvi"Emotion Based Music Player" International Journal of Engineering Research and General Science Volume 3, Issue 1, January- February , 2015

[4]. Shlok Gikla, Husain Zafar, Chuntan Soni, Kshitija Waghurdekar"SMART MUSIC

[5]. INTEGRATING AND MUSIC MOOD RECOMMENDATION"2017 International

[6]. Conference on Wireless Communications, Signal Processing and Networking(WiSppNET)

[7]. Zeng Z, Pantic M, Roisman GI and Huang TS 2008 A survey of affect recognition methods Audio, visual, and spontaneous expressions IEEE transactions on pattern analysis and machine intelligence 31 39-58

[8]. T.-H. Wang and J.-J.J. Lien, "Facial Expression Recognition System Based on Rigid and Non-Rigid Motion Separation and 3D Pose Estimation" J. Pattern Recognition , vol. 42, no. 5,pp. 962-977, 2009

[9]. Tambe, Parul, Yash Bagadia, Taheri Khalil and Noor UlAin Shaikh. "Advanced Music Player with Integrated Face Recognition Mechanism." (2015).

[10]. 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.

[11]. F Mahmud, M T Khatun, S T Zuhori, S Afroge, M Aktar dan B Pal, 2015 Face recognition using Principal Component Analysis and Linear Discriminant Analysis Electrical Engineering and Information Communication Technology (ICEEICT), International Conference on, pp. 1-4.

[12]. Menezes, P., Barreto, J.C. and Dias, J. Face tracking based on Haar-like features and eigenfaces. 5th IFAC Symposium on Intelligent Autonomous Vehicles, Lisbon, Portugal, July 5- 7, 2004

[13]. Byeong-jun Han, Seungmin Rho, Roger B. Dannenberg and Eenjun Hwang, ―SMERS: music emotion Recognition using support vector regression, 10thISMIR, 2009.

[14]. Patil, J. V., & Bailke, P. (2016). Real time facial expression recognition using RealSense camera and ANN. 2016 International Conference on Inventive Computation Technologies (ICICT).

[15]. Puja, Er. Rachna Rajput 2016 Feature Extraction in Face Recognition using SVMLBP Detection Technique International Journal of Innovative Research in Computer and Communication Engineering Vol. 4.

[16]. Jyoti Rani and Kanwal Garg 2014 Emotion Detection Using Facial Expressions A Review International Journal of Advanced Researching Computer Science and Software Engineering 4.

[17]. Vivek JD, Gokilavani A, Kavitha S, Lakshmanan S and Karthik S2017 A novel emotion recognition- based mind and soulrelaxing systemIn2017 International Conference on Innovations in Information, Embedded and Communication Systems1-5 IEEE.

[18]. Viola, P., and Jones, M. Rapid object detection using a boosted cascade of simple features Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1, pp. 511518 IEEE, 2001 (2001)

[19]. Lucey P, Cohn JF, Kanade T, Saragih J, Ambadar Z and Matthews I2010 the extended Cohn Kanade dataset (ck+) A complete dataset for action unit and emotion-specified expressionIn2010 IEEE computer society conference on computer vision and pattern recognition-workshops94-101 IEEE.

[20]. Anwar S, Milanova M, Bigazzi A, Bocchi L and Guazzini A 2016 Real time intention recognition In IECON 2016-42nd Annual Conference of the IEEE Industrial Electronics Society1021-1024 IEEE.

[21]. 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.

[22]. M. Vinny, P. Singh (2020) Review on the Artificial Brain Technology: BlueBrain. Journal of Informatics Electrical and Electronics Engi-neering,1(1), 3, 1-11.

[23]. Katleho L Masita, Ali N Hasan, "Deep Learning in Object Detection: a Review", IEEE Open Access Journal, vol 7, 2020.