Industrial Engineering Journal
ISSN: 0970-2555
Volume : 53, Issue 4, No. 2, April : 2024

EMOTION-DRIVEN AMBIANCE AND MUSIC CONTROL

Prasad Bhosle, Sujeet More, Trinity College of Engineering and Research, Pune, India .
Aboli Mohite, Purva Mene, Prathamesh Abnave, Susmita Mali, Trinity College of Engineering and Research, Pune, India .

Abstract – Our project discusses a proposed solution for tracking a person's emotions using a network of cameras and adjusting lighting, music, and ambiance accordingly to improve their mood. It also mentions the use of computer vision and machine learning for facial expression recognition. Another aspect involves creating a mood-based music player that recommends songs based on the user's mood in real-time, enhancing customer satisfaction. The project highlights the growing importance of music in people's lives and the need for improved music emotion recognition methods using an artificial bee colony algorithm to enhance the structure of neural networks and achieve better recognition results and faster processing.

I. Introduction

In this stressful and struggling world of the 21st century mental health has become a major entity to be dealt with thus people are trying numerous tactics that will help them to analyse and calm their emotions. Like wise to deal with this major issue we have come up with an idea of music player based on emotions. Compared to its predecessors, the main advantage of CNN is that it automatically detects the important features without any human supervision. This is why CNN would be an ideal solution to computer vision and image classification problems. Because their convolutional layers have fewer parameters compared with the fully connected layers of a traditional neural network, CNNs perform more efficiently on image processing tasks. CNNs use a technique known as parameter sharing that makes them much more efficient at handling image data.

The system's operation starts with a webcam connected to a microprocessor or device with sufficient computational capabilities. This webcam captures images of users' faces, and these images are then processed through a trained model. Deep learning techniques are used to interpret these images, which can be challenging due to the high degree of similarity in facial features. This similarity results in significant covariance, making it difficult to distinguish different emotions accurately. Moreover, human accuracy in identifying the mood of another person typically ranges from 67% to 73%. However, despite these challenges, the system is designed to effectively identify facial moods and provide accurate outputs.

II. Existing System

This paragraph highlights the significance of music in human history and its integration into various aspects of daily life. [1] It emphasizes the impact of music on people's emotions and how it's used in different contexts, such as retail stores, exercise, relaxation, and medical care. [1] It discusses the evolution of music storage and retrieval, noting that traditional methods are no longer sufficient for personalized music retrieval, leading to the need for emotion based music classification and retrieval. The paragraph also touches on the challenges of labelling music with emotions and the development of automatic music emotion recognition technology. [1] Various approaches, including machine learning algorithms, have been used to achieve music emotion recognition with recognition rates exceeding 60%. [1]. This passage discusses the significance of deep learning based Music Emotion Recognition and Classification (MER) as an active research field in Music Information Retrieval (MIR).[2] It highlights the challenges in manually retrieving and classifying music in a vast digital library and the importance of automatic retrieval systems, especially in a diverse and multicultural context. [2].

The passage emphasizes the role of music emotions in relation to nationality, age, mood, purpose, and cultural background, and how these factors influence the complexity of MER. Music's connection to human emotions and its therapeutic applications are also highlighted, along with the use of neurotransmitters and hormones to describe the emotional impact of music. [2].

The paragraph underscores the integral role of music in people's lives, both as a source of enjoyment and for medical purposes, such as stress relief. [3] It mentions the advancement of high-end music players with enhanced features, including volume control, modulation, and genre selection. The goal is to create a platform that plays music based on a person's current mood using facial expression recognition in the context of artificial intelligence (AI) and machine learning (ML).[3] The process involves dimensionality reduction for facial expression recognition, starting with face detection and then extracting facial features like the eyes, nose, and mouth to identify emotions. [3] The paragraph also highlights the impact of music on human emotions and cognitive functions, emphasizing its role in enhancing mindfulness and mood. [3] The connection between musical preferences and individual characteristics and the growing use of digital emotion detection in multimedia content like music and movies is noted. The proposed recommender system is described, which can identify user emotions and suggest appropriate songs based on their mood. It uses the Kaggle Look Acknowledgment dataset for emotion detection and Bollywood Hindi songs for the music library. [3] Facial emotion detection is implemented using a Convolutional Neural Network with high accuracy, approximately 95.14%. This system aims to improve users' emotional states through music recommendations based on their detected feelings. [3].

This passage discusses the growing trend of sharing life experiences and opinions through images and videos on social networks, which has led to an explosion of online social data. [4] Analysing these multimedia data, particularly at the emotional level, has gained significant importance, as emotions can influence decisionmaking and have wide ranging applications from marketing to political forecasting. [4] While there has been progress in textual sentiment analysis, emotion analysis of social images poses several challenges. One major challenge is the "affective gap" where finding features to express emotions in images remains a key issue. Previous efforts have typically focused on predicting dominant emotions that are common among viewers, without considering individual and subjective evaluations. [4].

III. Methodology

The system you've described is focused on using facial expression recognition to create music playlists that match the user's emotional state. The main steps and components of the system are as follows:

1) Real-Time Capture: The system captures the users face accurately in real time. This is the initial step to get input for emotion recognition. In this we are using the camera to capture the users face.

2) Facial Recognition: The captured user's face is then used as input for the facial recognition module. This step involves using a convolutional neural network (CNN) that is trained to analyse the user's facial features. Convolutional Neural Networks (CNNs) are designed to automatically and adaptively learn spatial hierarchies of features from input images through convolutional layers. This enables them to capture intricate patterns and representations, making them highly effective for tasks like image classification, object detection, and segmentation.

3) Emotion Detection: In this phase, the features extracted from the user's facial image are used to identify emotions. The system then provides music recommendations based on the detected emotions. Feature extraction most important part of classification and emotion recognition. After the preprocessing of an image the facial features with high expression intensity are extracted like eyebrows, forehead wrinkles, nose, jawline, mouth corner. The facial feature extraction method id carried out by Local Binary pattern algorithm. The local

binary pattern technique operates by pointing the pixels of an image and comparing it with the neighborhoods pixels by using binary number.

4) Music Recommendation: This module recommends songs based on the user's emotions and the desired music ambiance.The module plays the songs on the basis of emotions like happy , sad , ngry , etc.

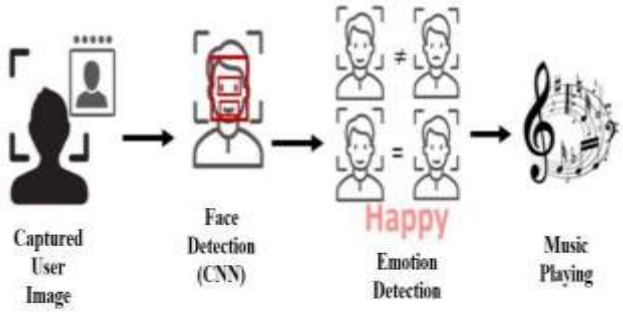

Fig.1. System Architecture

CNN Algorithm

A Convolutional Neural Network (CNN) is a type of deep learning algorithm that is particularly well-suited for image recognition and processing tasks. It is made up of multiple layers, including convolutional layers, pooling layers, and fully connected layers. The convolutional layers are the key component of a CNN, where filters are applied to the input image to extract features such as edges, textures, and shapes. The output of the convolutional layers is then passed through pooling layers, which are used to down-sample the feature maps, reducing the spatial dimensions while retaining the most important information. The output of the pooling layers is then passed through one or more fully connected layers, which are used to make a prediction or classify the image. Convolutional neural networks are based on neuroscience findings. They are made of layers of artificial neurons called nodes. These nodes are functions that calculate the weighted sum of the inputs and return an activation map. This is the convolution part of the neural network. Each node in a layer is defined by its weight values. When you give a layer some data, like an image, it takes the pixel values and picks out some of the visual features. When you're working with data in a CNN, each layer returns activation maps. These maps point out important features in the data set. If you gave the CNN an image, it'll point out features based on pixel values, like colours, and give you an activation function. Usually with images, a CNN will initially find the edges of the picture. Then this slight definition of the image will get passed to the next layer. Then that layer will start detecting things like corners and colours groups. Then that image definition will get passed to the next layer and the cycle continues until a prediction is made.

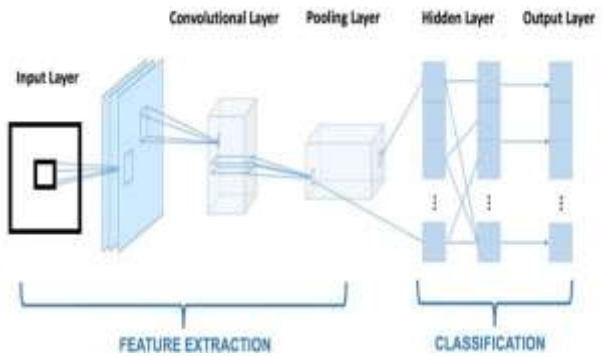

Fig.2. CNN Architecture

IV. Conclusion

Emophony is a groundbreaking system that seamlessly integrates with numerous music platforms, making it effortless for users to discover music tailored to their current emotional state. Users receive a curated list of songs based on their emotions, eliminating the need for manual track searches. Emophony efficiently directs users to gaana.com, where they can instantly play their chosen music with a simple click. The system utilizes affective computing methods to accurately predict and cater to users' changing emotional states. It offers a diverse selection of music, enriching the listening experience by aligning songs with users' emotions. Moreover, the paper introduces a Music Information Retrieval (MIR) network designed to classify music based on emotional characteristics.

V. References

1) Huang, C.; Zhang, Q. Research on Music Emotion Recognition Model of Deep Learning Based on Musical Stage Effect. Sci. Program. 2021.
2) Shikhar C. Maheshwari, Amit H. Choksi and Kaiwalya J. Patil. Emotion Based Ambiance And Music Regulation using Deep Learning, 2020.
3) Mohak Kedia, Vaibhav Bhagat, Yash Khannade, Nayan Laturiya. Emotion Based Ambiance And Music Regulation Dept. Of Computer Science & Engineering, 2023.
4) Sicheng Zhao, Hongxun Yao, Member, IEEE, Yue Gao, Senior Member, IEEE, Guiguang Ding, and Tat-Seng Chua. Predicting Personalized Image Emotion Perception In Social Network. 2018.
5) Han, X.; Chen, F.; Ban, J. Music Emotion Recognition Based on a Neural Network with an Inception-GRU Residual Structure. Electronics 2023.