SMART EMOTION BASED MUSIC PLAYER USING FACIAL RECOGNITION

Deepali Patil
Information technology
Shree L R Tiwari College Of Engineering (Mumbai university)
Mumbai, India

Mohit Yadav
Information technology
Shree L R Tiwari College Of Engineering (Mumbai university)
Mumbai, India

Gopal Gehlot
Information technology
Shree L R Tiwari College Of Engineering (Mumbai university)
Mumbai, India

Deepak Tiwari
Information technology
Shree L R Tiwari College Of Engineering (Mumbai university)
Mumbai, India

Himanshu Barai
Information technology
Shree L R Tiwari College of Engineering (Mumbai university)
Mumbai, India

***Abstract* —Songs, as a medium of expression, have invariably been a well-liked option to depict and perceive human emotions. Reliable feeling primarily based classification systems will go in helping us and parsing the meaning. However, analysis within the field of emotion-based music classification has not yielded optimum results. In this paper, we pretend to a cross-platform music player. EMP, that recommends music based on real-time mood of the user. Our music player contains three modules: Emotion Module, Music Classification Module and Recommendation Module. The Emotion Module takes an image of the user's face as an input and makes use of deep learning algorithms to identify their mood with an accuracy of 90.23%. The Music Classification Module makes use of audio features to achieve a remarkable result of 97.69% while classifying songs into 4 different mood classes. The Recommendation Module suggests songs to the user by mapping their emotions to the mood type of the song, taking into consideration the preferences of the user.**

***Index Terms*---Emotion Recognition, Artificial neural networks, Music Player**

I. Introduction

Current analysis within the field of music scientific discipline has shown that music induces a transparent emotional response in its listeners [1]. Musical preferences are incontestable to be extremely correlative with temperament traits and moods. The meter, timber, rhythm and pitch of music area unit managed in areas of the brain that subsume emotions and mood [2]. beyond question, a user's emotive response to a music fragment depends on an oversized set of external factors, like gender, not age [3], culture [4], personal preferences, feeling and context [5] (e.g. time of day or location). However, these external variables put aside, human's area unit ready to systematically categorize songs as being happy, sad, angry or relaxed. Acknowledging the barrier, we tend to focus our efforts on audio feature extraction and analysis of contemporary Yankee and British English songs so as to map those options to four basic moods. Automatic music classification victimization some mood classes yields promising results. Facial expressions area unit the foremost ancient and natural approach of transference emotions, moods and feelings. The aim of this paper, is to categorize facial expressions into four totally different emotional classes, viz. happy, sad, angry and neutral. The main objective of this paper is to style a cheap music player that mechanically generates a sentiment aware of listing and support the emotion of the user. The feeling module determines the feeling of the user. Relevant and significant audio data from a song is extracted by the music classification module. This technique provides significantly higher accuracy and performance than existing systems.

II. RELATED WORK

Various methodologies are projected to classify the behavior to land emotion of the user. Maseetal. Focused on using movements of facial muscles while Tianetal. attempted to recognize Actions Units (AU) developed by Ekman and Friesen in 1978 using permanent and transient facial features. With evolving methodologies, the employment of Convolutional Neural Networks (CNNs) for feeling recognition has become progressively well-liked. Music has additionally been classified using lyrical analysis [6]. Whereas this tokenized technique is comparatively easier to implement, on its own it's not appropriate to classify songs accurately. Another obvious concern with this technique is that the barrier that restricts classification to one language. Another technique for music mood classification is using the acoustic options like tempo, pitch and rhythm to spot the sentiment sent by the song. This technique involves extracting a group of options and victimization those feature vectors to find patterns that characterize to a specific mood [7],

III. FLAWS IN CURRENT SYSTEM

In this section, we tend to study the usage of convolutional neural networks (CNNs) within the context of feeling recognition. CNNs area unit far-famed to simulate the human brain once analyzing visuals; but, given the procedure needs and quality of a CNN, optimizing a network for efficient computation is important. Thus, a CNN is enforced to construct a procedure model that successfully classifies feeling into four moods, namely, happy, sad, angry and neutral.

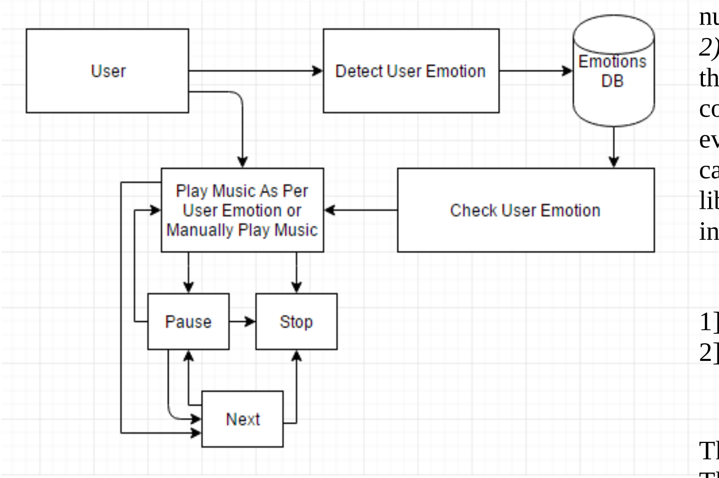

Fig.3.1 Overview of system

IV. Emotion Module

In this section, we study the usage of convolutional neural networks (CNNs) in the context of emotion recognition .CNNs are known to simulate the human brain when analysing visuals; however, given the computational requirements and complexity of a CNN, optimizing a network for efficient computation is necessary. Thus, a CNN is implemented to construct a computational model which successfully classifies emotion into 4 moods, namely, happy, sad, angry and neutral, with an accuracy of 90.23%.

A. Dataset Description

The dataset we used for training the model is from a Kaggle Facial Expression Recognition Challenge. The data consists of 48×48 pixel grayscale images of faces. Each of the faces are organized into one of the 4 emotion classes: angry, happy, sad, and neutral. For this research, we have made use of 4 emotions: angry, happy, sad and neutral. There is a total of 26,217 images corresponding to these emotions. The breakdown of the images is as follows: happy with 8,989 samples, sad with 6,077 samples, neutral with 6,198 samples, angry with 4,953 samples.

TABLE I.

		Datasets per Emotions			
		Angry	Happy	Sad	Calm
Actual class	Angry	981	34	13	21
	Happy	27	1497	16	33
	Sad	41	46	1218	111
	Neutral	29	56	85	1036

B. Model Description

A multi-layered convolutional neural network is programmed to evaluate the features of the user image. The convolutional neural network contains an input layer, and an output layer. These layers are linearly stacked in sequence.

1). Input Layer: The input layer has fixed and predetermined dimensions. So, for pre-processing the image, we used OpenCV for face detection in the image before feeding the image into the layer. Pre-trained filters from Haar Cascades along with Adaboost are used to quickly find and crop the face. The cropped face is then converted into grayscale and resized to 48-by-48 pixels. This step greatly reduces the dimensions from (3, 48, 48) (RGB) to (1 , 48, 48) (grayscale) which can be easily fed into the input layer as a numpy array.

2).Output Layer: The output is represented as song based on the detected each emotion class. Models with various combinations of hyper-parameters were trained and evaluated utilizing a 4 GiB DDR3 NVIDIA 840M graphics card using the NVIDIA CUDAR Deep Neural Network library (cuDNN). This greatly reduced training time and increased efficiency in tuning the model.

V. Future Work

1] Adding various emotion in future such as fear, disgust.
2] Training on large dataset.

VI. Conclusion

The results obtained on top of area unit terribly promising. The high accuracy and fast interval of the appliance makes it appropriate for many sensible functions. The music classification module specially, performs considerably well; it achieves high accuracy within the “angry” class, whereas additionally playing appreciably well within the “happy” and “sad” classes. Thus, EMP reduces user efforts for generating playlists with efficiency mapping the user’s feeling to the proper song category with associate degree overall accuracy of ninety 97.69%, it achieves optimistic results for the four moods studied.

We additionally acknowledge the area for improvement. it's been fascinating to research however the system performs once all four basic emotions area taken into consideration; extra songs from totally different languages and regions may also be supplemental to create the advice system additional strong. User preferences will be collected to boost the general system using cooperative filtering. We tend to conceive to address these problems in our future work

References

[1] Swathi Swaminathan and E. Glenn Schellenberg, “Current emotion research in music psychology,” *Emotion Review*, vol. 7, no. 2 , pp. 189–197, Apr. 2015.

[2] “How music changes your mood”, Examined Existence. [Online].Available: http://examinedexistence.com/how-music-changes-yourmood/. Accessed: Jan. 13, 2017.

[3] Kyogu Lee and Minsu Cho, “Mood Classification from Musical Audio Using User Group-dependent Models.”

[4] Daniel Wolff, Tillman Weyde, and Andrew MacFarlane, “Culture-aware Music Recommendation.”

[5] Mirim Lee and Jun-Dong Cho, “Logmusic: context-based social music recommendation service on mobile device,” Ubicomp’14 Adjunct, Seattle, WA, USA, Sep. 13–17, 2014.

[6] D. Gossi and M. H. Gunes, “Lyric-based music recommendation,” in *Studies in computational intelligence*. Springer Nature, pp. 301–310, 2016.

[7] Bo Shao, Dingding Wang, Tao Li, and Mitsunori Ogihara, “Music recommendation based on acoustic features and user access patterns,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 17, no. 8, Nov. 2009.

[8] Ying-li Tian, T. Kanade, and J. Cohn, “Recognizing lower. Face action units for facial expression analysis,” in *Proceedings of the 4th IEEE International Conference on Automatic Face and Gesture Recognition (FG’00)*, Mar. 2000, pp. 484–490.