IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Emotion Recognition Entertainer using Deep Learning

[1]Prof. Amar Palwankar, [2]Mr.Arman Nakhwa, [3]Mr.Rushikesh Kadam, [4]Mr.Ved Shirgaonkar, [5]Mr.Sourabh Koravi

[1]Professor [2]Student [3]Student [4]Student [5]Student

Department of Information Technology Finolex Academy of Management and Technology, Ratnagiri, India

Abstract

Human behavior is mainly influenced by emotion. People are significantly impacted by the music and movies they listen and see. Video has emerged as a popular and dominant medium during this pandemic. In order to develop recommendation systems on the basis of user emotions and deliver dynamic content recommendations to users, this paper attempted to present a system design model. We will suggest the list of movies to the user by applying deep neural network technology to capture real-time emotion and combine it with conventional content-based recommendations.

The term "recommender system" refers to a system that may be used to suggest products to a user based on information or criteria such as prior user comments or other user patterns. This study will offer movies or music to the user based on his or her facial expression rather than prior input from the user or any other trend. Therefore, it is important to develop a recommendation system that uses less user data while still functioning well since a user's needs could not be tied to his or her history but rather to the present, which is denoted by the user's expressions.

Keywords: Emotion Recognition, Facial Recognition, Speech Recognition, entertainment Recommender, Multimedia Recommender, Deep Learning, Convolutional Neural Networks.

INTRODUCTION

The science of facial expression recognition software employs algorithms for biometric systems to extract data from human facial expressions, including conclusions about emotions. More precisely, this technology is an emotion analysis system that can recognize a variety of human expressions, including sad, angry, happy, etc.

Nonverbal communication signals are communicated through gestures and facial expressions, which are crucial in interpersonal interactions. Therefore, facial expression recognition is able to give unfiltered, honest emotional reactions as data since it takes and analyses information from an image or video source.

Generally speaking, recommender systems Contains algorithms that try to present consumers relevant things (items being movies to watch, text to read, products to buy or anything else depending on industries) according to their interest.

Automated emotion detection in multimedia entities like music or movies is expanding quickly thanks to technological advancements in digital signal processing and other efficient feature extraction algorithms, and this recommendation system can play a significant role in many potential applications like human-computer interaction systems, music entertainment, and movie recommenders for theatres.

METHODOLOGY

In traditional entertaining websites the user needs to go through the various section in order to get content that they want. The whole process is manual We provide the emotion recognition entertainer which will automatically recommends the songs, movies, news, etc. based on current mood or emotion of the user.

It is like a website so user do not need to install it on personal computer so user can simply access the software using web browser and good internet connection. Web cam permission is also required in according to face recognition.
We also provide users the additional option where users can upload their photograph for mood detection instead of real time mood detection. Whenever due to some circumstances like extremely high and low lighting, their camera not working properly, camera quality issues like unable to detect clear face which results in unable to detect the proper mood or any other problem user simply click on “Choose File” option in order to upload your picture.

OBJECTIVES

Use Deep leaning algorithms to improve emotion prediction.

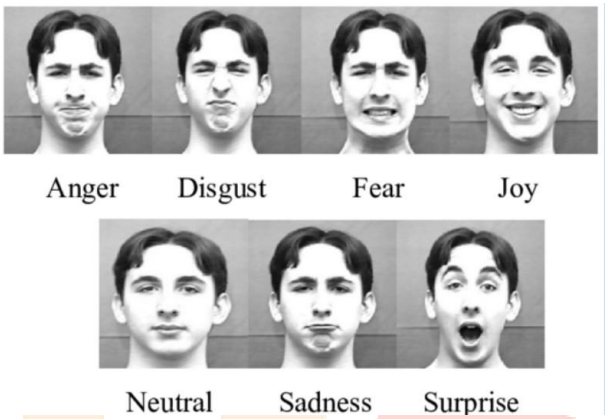
To provide a seamless user interface for multimedia.

Ultimate aim is to provide entertainment source with less user interaction.

SYSTEM

The user's face is treated as an input that was recorded by a camera. The face is the typical site of facial expression, which is the crucial data needed by the system to produce the output.

Providing Entertainment Source recommendations by using voice, emotions, and photos as an input.



Facial emotions like happy, sad, anger, etc. can be correlated with particular movies or musical styles. Here are a few instances:

Emotions	Movies	Songs
Happy	Comedy.	Pop
Sad	Tragedy	Soft
Anger	War	Rock

MODEL

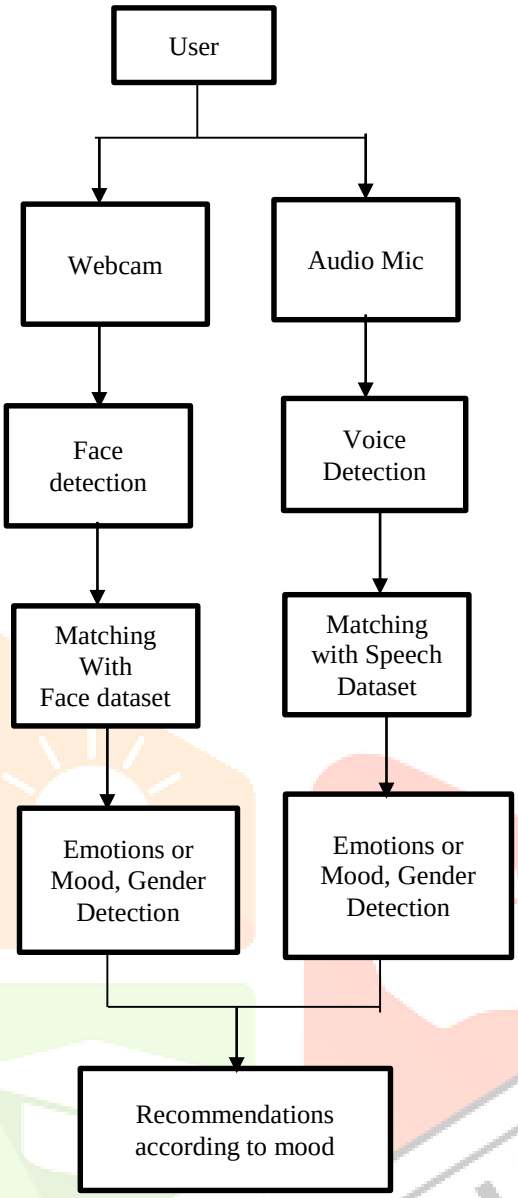
Random Forest

Supervised machine learning algorithms like random forest are frequently employed in classification and regression issues. On various samples, it constructs decision trees and uses their average for classification and majority vote for regression.

The Random Forest Algorithm's ability to handle data sets with continuous variables, as in regression, and categorical variables, as in classification, is one of its most crucial qualities. For categorization issues, it delivers superior outcomes.

Feature	Models	Accuracy	Inference	Go/No-go
Emotions	Deep Learning	1.00	Overfitting	No-go
	Knn	0.77	Result not too accurate	No-go
	Random Forests	0.83	Almost correct predictions	Go
Gender	Logistic Regression	0.98	Overfitting	No-go
	Deep Learning	0.99	Overfitting	No-go
	Random Forests	0.97	Almost accurate predictions	Go

ARCHITECTURAL DIAGRAM



PHASES

1. Capturing Phase: In capturing phase system capture the speech or face of the user by using Audio mic and Webcam Respectively.

2. Detection Phase: after capturing it detects the face is clearly visible or not. Phase in this phase different analysis done in order to extract features from face or speech. Then the features are detected and classified according to classes.

3. Matching Phase: the features extracted from detection phase matched with the datasets used which contains different types of faces and voices in different moods like angary, sad.

4. Emotions /Mood, Gender Detection Phase: Emotions are detected in this phase. With the help of voice, we are also able to detect gender along with mood of the user.

5. Recommendation Phase: According to mood of the user Recommendations are fetched and displayed to user.

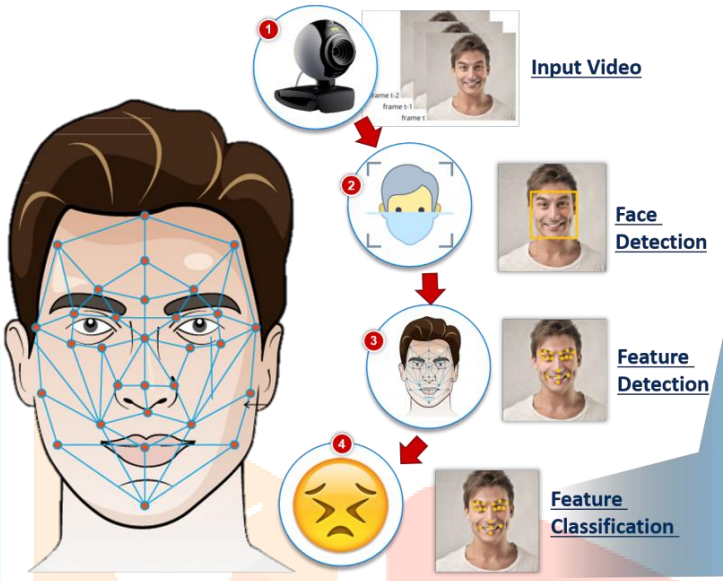
PROCESSES

1. Face Detection

Here, we'll use classification. A data set is classified into classes through the process of classification. This system will use this technique to divide the different data sets into different categories of emotions.

We require a large data collection since the more datasets used to train the classifier, the more accurate it will become. You may develop these datasets yourself or obtain them from the internet. The requirement is that each subfolder should be labelled with the name of the emotion that it is related to.

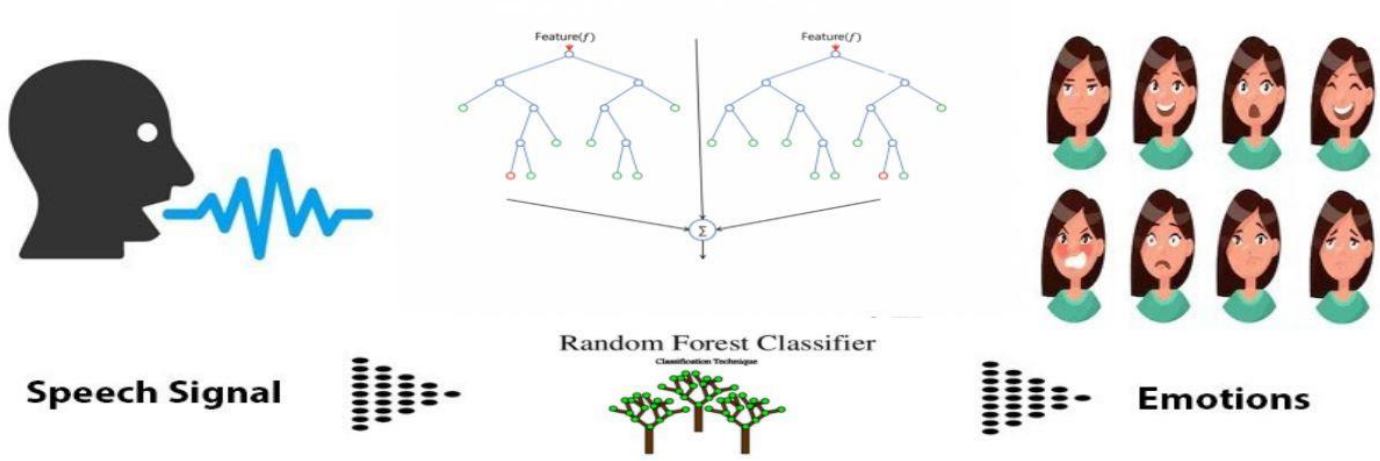
Since the training dataset has labels assigned to them, we are using supervised learning in this case. The training dataset is then compared to the testing dataset using algorithms (classification in this case), and the outcome is the labels associated with the testing dataset.



2. Speech Detection

Convolutional Neural Network (CNN) is applied as advanced deep neural networks to classify each word from our data set as a multi-class classification task. The proposed deep neural network returned 97.06% as word classification accuracy with a speech sample. CNN is used to train and test our data. Speech dataset is basically divided into 2 parts male and female so it also helps the gender of the user with the help of their speech.

Since the training dataset has labels assigned to them, we are using supervised learning in this case. The training dataset is then compared to the testing dataset using algorithms (classification in this case), and the outcome is the labels associated with the testing dataset.



3. Emotion detection

Now that we have a dataset of faces with labels corresponding to different emotions like happy, anger, and sad. Using the VGG- 16 (16-layer convolutional Neural Network), a Convolution Neural Network (CNN) for image classification, the dataset is now being converted into vectors. An input layer, an output layer, and several hidden layers make up the CNN neural network type.

Since Logistic Regression has the best accuracy and comparatively low error rates, it is used as the classification model. This will categories the test image into the category it belongs to in this case, Emotions.

FUTURE ASPECTS

Future improvements may include the addition of support for compatibility with other emotions in the facial expression recognition-based entertainment recommendation system.

Additionally, it may be linked to streaming services like Netflix, Amazon Prime, Spotify, and others, enhancing the system's library.

Facial or Speech recognition can be used for authentication purposes.

CONCLUSION

In this research study, the recommender system essentially merges three distinct recommender systems, one for movies one for music and one for news. by exploiting the human emotion expressed through facial expression using face recognition and classification algorithm to obtain the emotion. Then, using a tool that retrieves the necessary data for the multimedia item, which may be either music or a movie, online automation is performed using the genre that corresponds to the emotion.

REFERENCES

1. RAVDESS Dataset: "The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS)" by Livingstone & Russo is licensed under CC BY-NA-SC 4.0
2. [TESS Dataset: Pichora-Fuller, M. Kathleen; Dupuis, Kate, 2020, and "Toronto emotional speech set (TESS)", Scholars Portal Dataverse, V1
3. Leo Pauly, Deepa Sankar, “A Novel Online Product Recommendation System Based on Face Recognition and Emotion Detection”, (ICCICCT, 2015)
4. S.Nithya Roopa, Research on Face Expression Recognition, (IJITEE, 2019).
5. OpenCV: Cascade Classifier