A MUSIC RECOMMADATION BASED ON EMOTION DETECTION USING DEEP CONVOLOTION NEURAL NETWORK

MR.A.PANDIAN
Department of ECE
SRM Valliammai Engineering College
Kattankulathur, India
pandiana.ece@srmvalliammai.an.in

SURYA K
Department of ECE
SRM Valliammai Engineering College
Kattankulathur, India
suryakrishnamoorthi17@gmail.com

SURIYAPRAKASH C
Department of ECE
SRM Valliammai Engineering College
Kattankulathur, India
csuriyaprakash02@gmail.com

THINAGARAN T
Department of ECE
SRM Valliammai Engineering College
Kattankulathur, India
thinagarant04@gmail.com

VIGNESH S
Department of ECE
SRM Valliammai Engineering College
Kattankulathur, India
vigneshsivaak@gmail.com

ABSTRACT:
Human emotions are inconsistent and are truly a product of both internal and exterior events taking place in a person's environment. Human emotions have been the subject of extensive study and investment, which has opened the door to many potential uses. The current method includes automatically generating a playlist of music based on genres, artists, etc. Manually organizing audio files into playlists is still another choice. Calculating music similarity and multiple frequency estimates are recent problems. A QBSH (Query by singing and humming) method uses the song's content to determine what it is (tune and rhythm). Nevertheless, the problem with this approach is that it takes time and doesn't always satisfy the consumer. The user's emotion is not taken into account in the current system. A music recommendation system that considers human emotions can be created because they are important in daily activities. A person's emotion can be determined to determine the type of music that would work best. The technology seeks to analyses the data supplied by determining the user's emotion. The classification of the various emotions using a deep learning algorithm is followed by the generation of labels and the playing of appropriate music. The proposed system has produced results with a notable degree of accuracy and it also opens the door for additional study in this field.

Keywords: ***QBSH, Emotions, Deep learning***

INTRODUCTION:

The art form that is recognized to connect with a person's emotions the most is music. It has a special capacity to improve one's mood. A user's listing experience will also be enhanced if a recommendation is made based on his preferences. Recommendations for music have been around for a while. However, in the majority of cases, the recommendation is made after taking into account the user's preferences over time, such as taking into account their previous song preferences, the amount of time they spend listening to music, etc. This study proposes a neural network-based method for song suggestion that analyses a person's mood based on their facial expressions. This strategy is more effective than the ones currently in use and makes it easier for users to search for specific playlists and create them. The way a person is feeling can be inferred from their facial expressions. A face is photographed using a webcam or other camera, and information is then retrieved from the image. Also, the mood of the person is ascertained using this information.

Sharing information or resources among people requires communication. Individuals can communicate information to one another verbally or nonverbally. An individual's facial expressions can be highly helpful in picking up on the subject's mood and conduct. Human emotion is essential in helping people convey their thoughts. The six main types of emotions—sadness, happiness, anger, fear, disgust, and surprise—are the ones they belong to. The form, size, and movement of the mouth, eyes, and eyebrows can all be used to identify certain emotions. The internet-connected devices of today have access to millions of music at any given time. An individual's mood can be changed by music at any time. Our major goal is to create a playlist of music that automatically produced by analyzing these human feelings.

Automatic music classification based on genres, artists, countries, frequency, and other factors is a feature of modern technology. Only a small number of mobile applications, including Spotify, Saavn, Wynk, etc., automatically classify music. Here, users have the option of making their own playlists or listening to those that the software has already created. Find a song is a contemporary application that a user is unable to recognize. In these circumstances, mobile applications like Shazam, track ID, and Sound Hound can be quite helpful. Some applications, such as Music match and YouTube Music, assist with the automatic translation of song lyrics from one language to another as well as with the instantaneous display of song lyrics. It is said that music The world's greatest healing tool, according to legend, is music. As a result, we use that aspect of music to great use for the person experiencing a certain mood.

One of the most significant and intricate systems in the world is the human body. Due to the vast number of nerves that travel throughout our bodies and carry and process information, there is complexity. The human body is now one of the most sophisticated organisms on Earth as a result of this. One may say that this organic information processing mechanism served as an inspiration for neural networks. Many applications have been developed as a result of this research. So, a neural network likewise learns through experience, or training, in the same way that people do.

CLASSIFICATION:

All grouping assignments rely on tagged datasets, which means that everyone must share their understanding of the dataset in order for a neural system to become familiar with the relationship between label and information. It is additionally referred to as supervised learning

- Face detection, identifying all the individual's faces in images, detecting the mood of the individual.
- Object identification in images.
- Gesture recognition in images/videos.
- Sentiment recognition, voice detection.
- Classification of spam in emails, fraudulent (in claims of insurance).

KEY CONCEPT OF NEURAL NETWORK:

The depth of a network is the quantity of node levels that a piece of data must traverse. This characteristic is used to separate single hidden layer networks from deep learning networks. A neural network's depth has a big impact on pattern recognition. A neural network model's early iterations lacked depth. Input, output, and a secret layer in between them made up the majority of it. Yet, something must have at least four layers to be considered a deep learning network (including input and output).

Deep learning network architecture trains each layer by recombining the output features that it receives from the layer before it. More challenging highlights are also seen by neurons as they blend highlights from prior layers as we progress deeper into the neural system.

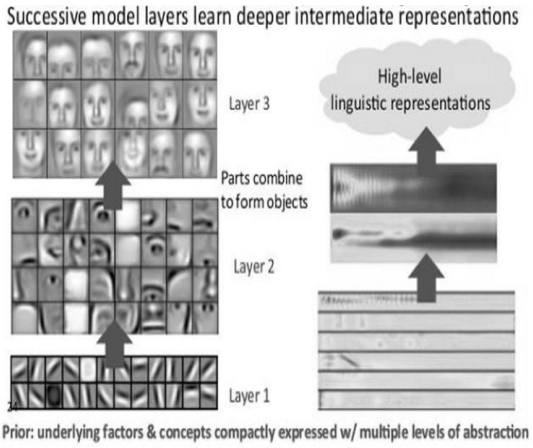


Fig:1.1Feature Hierarchy in Image processing

The feature hierarchy is depicted in the of parameters because to its special property diagram above. It is created as complexity and abstraction levels rise. The deep-learning architecture can manage massive volumes of multi-dimensional data with billions.

In deep learning networks, feature extraction can be carried out automatically. This distinguishes it from many machine-learning algorithms that call for some level of human participation. In this method, we may create a lot of intelligent systems. By recombining the input from the prior samples, features in a deep learning network are automatically trained and learned. With this method, they may also make connections, spot patterns, and deliver the best outcomes.

A specific result or label makes up a deep learning network's output layer. That result is utilized in prediction, for example, we can say that an image is 85% likely to represent a human given the input and the output. The output of emotion analysis based on an image might be a probability value, which is made possible by the architecture's Softmax layer.

EMOTION ANAYSIS:

Emotions significantly influence how someone expresses their sentiments to others. It also has an impact on how people interact with others and live. We can all be said to be controlled by different types of these emotions. These emotions control how we respond to every activity, and depending on the action, they may be either intentional or involuntary. The emotions we are experiencing at the time influence every choice we make and every action we take. It's also the reason why it's advised to avoid making decisions when irritated.

TYPES OF EMOTIONS:

Finding out how, when, and why people react differently to different situations has been a research focus for many psychologists. They have gathered a variety of feelings that people experience. To further clarify how people feel, these emotions have also been categorized.

Basic Emotions: Paul Eckman, a well-known psychologist of the 1970s, identified six basic emotions that he believed were shared by all people, regardless of their cultural backgrounds. He categorized the following emotions: fear, disgust, happiness, sadness,

and rage. He argued that regardless of the circumstance, any person will respond with one of these emotions. He came to the conclusion that all people were drawn to these emotions.

Mixing Emotions: The "wheel of emotions" is a concept put out by Robert Plutchik, a well-known psychologist. He said that these feelings can occasionally be mixed to create new emotions. He made the comparison to colour, saying that when two colors are combined, a new colour results. Similar to how emotions can combine to create new emotions. These fundamental feelings serve as building blocks. For instance, hate can be produced when anger and disgust are united.

Overview of one Emotion analysis:

Affective computing includes the emotional analysis as a core component. It's possible that we refer to it as emotion AI or artificial emotional intelligence. To determine human emotion, data must be gathered from faces, body language, or speech. It relates to the field of cognitive computing. A variety of human-computer interactions that involve emotion analysis allow a computer to recognize and design a response that is appropriate for the user.

In order to create a model that can recognize diverse emotions, deep learning techniques are applied in this situation. In psychology, the term "affect" is used to describe a patient's emotional state. It has been crucial to conduct numerous studies in this area of emotional analysis and develop a wide range of applications.

We constructed a prototype of one such application, the "Emotion based Music Recommendation System," by doing emotion analysis using convolutional neural networks.

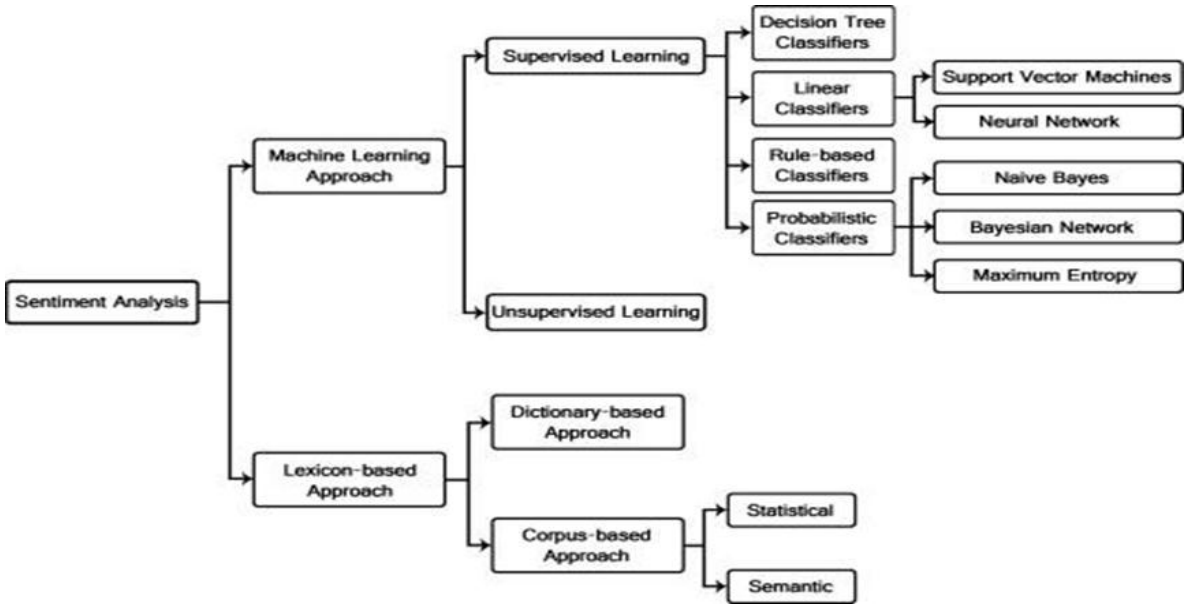


Fig 1.2 Different ways of performing sentiment analysis.

PROPOSED SYSTEM:

The creation of a playlist based on the person's mood is the main subject of this essay. Several images of the user are collected at that precise moment using a camera with the user's consent. These photos go through a rigorous testing and training process to determine the person's mood.

Once the images are captured, they go through various stages like

1. Preprocessing
2. Segmentation
3. Feature extraction
4. Emotion classification
5. Web service integration

At each stage these images go through various filtering process to extract the exact mood of the individual. Features like mouth, eyebrows and eyes are extracted based on which various calculations are done to find the emotion of the person.

A deep learning algorithm called Convolution Neural Network (CNN) is applied to classify the various emotions. Each emotion is associated

with a value and when the extracted valueof the image falls within the range of the defined values of each emotion, person is said to be in that particular emotional state.

A person's emotion can be determined to determine the type of music that would work best. The technology seeks to analyses the data supplied by determining the user's emotion. The proposed system has produced results with a high degree of accuracy and opens up new opportunities for this field of study.

METHODOLOGY:

GATHERING DATASET:

Webcams or mobile cameras are used to classify emotions in real-time datasets. Once the user has given permission, images are taken. For the purpose of classifying the emotions, the collected photos are compared with FER 2013 datasets. The FER 2013 datasets are made up of 37887 images in grey scale with 7 different emotions: 0 represents anger, 1 disgust, 2 fear, 3 happiness, 4 sadness, 5 surprise, and 6 neutrality.

FACIAL DETECTION AND RECOGNITION:

The method of locating a human face within an image is known as facial detection. A face detector must be able to identify each of the five faces in an image if there are five of them. This is accomplished by using Haar's cascade. It frequently works by looking for human eyes, or a valley region.

Haar Cascade:

If the region displays any features or objects, this method outputs one; otherwise, it outputs zero. An individual can search any area without having to identify specific objects and features by changing the search window**.**

Rather than just scaling the image, the classifier is designed so that it tends to scale effectively, making it more effective at finding objects of different sizes. So, the output method needs to be finished several times at different scales in order to find an object of a mysterious size in the image.

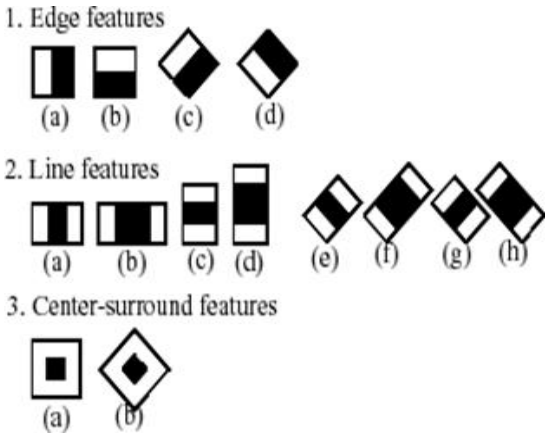


Fig 1.2- Haar-Cascade Classifier

Facial recognition is used to identify one of the five individuals. It accomplishes more than just recognising faces. It makes use of biometric technology to take a picture of a person's face and compare it to other photographs that are already in a database. It will be able to recognise the individual whose face is being displayed. For this, LBPH face recognizer is employed. The face is extracted, cropped, and resized by the LBPH face recognizer. These edited photos are then transformed to grey scale.

ALGORITHM:

1: CNN Algorithm:

The Convolution Neutral Network, as it is known, was created primarily to accurately identify patterns and features inside a picture. CNN also has the advantage of being able to identify the key elements of an image without the aid of humans. The feature extraction portion and the categorization portion make up a CNN. With the help of convolution and pooling layers, features are extracted.

MODULES:

1. Image Capturing
2. Pre-Processing
3. Segmentation
4. Feature extraction

5. Emotion Classification

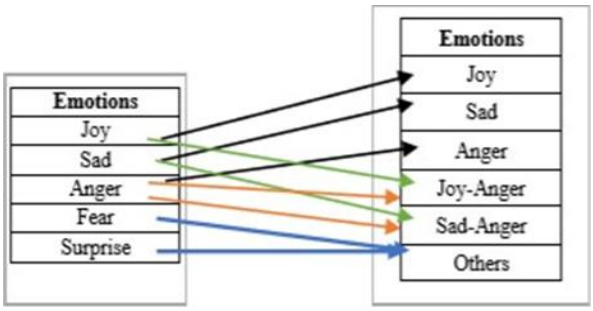


Fig 1.4 Basic and Combined Emotions

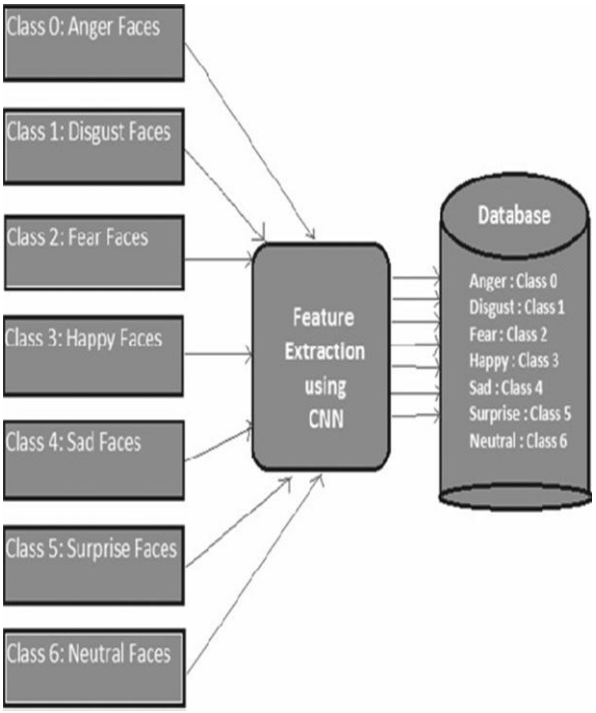
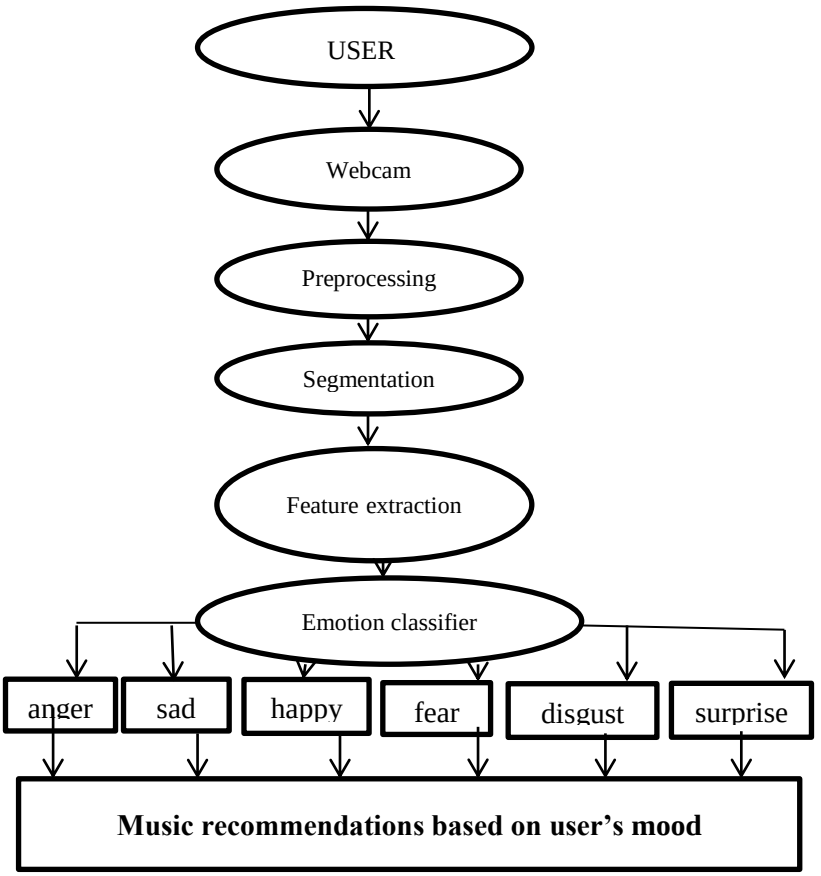


Fig 1.5 Proposed CNN Architecture

USE CASE DIAGRAM:

They are typically employed to visualise the various operations carried out by the application. Also, they show the many users who have access to carry out these functions. Because they place a strong emphasis on the actions taken by the users (or actors), use-case diagrams fall within the category of behaviour diagrams. The various responsibilities that the application performs are

1. Getting permission for webcam
2. Image pre-processing
3. Image segmentation
4. Feature extraction
5. Emotion classification
6. Music recommendation based on the mood



FEATURE EXTRACTION:

The traits that make a face different while displaying various emotions must be recognized in order to determine a person's expression. The lips, eyebrows, and eyes are these features. We can determine a person's emotion based on the way these characteristics move.

For instance, we can tell someone is astonished if their mouth is open and their eyes are bigger than usual.

The following facial features are extracted

Eye: Because of its white eye and iris, human eyes have sharp vertical edges. In this way, the Sobel cover is attached to the image, and the vertical facilitate of the eyes may be determined by the even projection of vertical edges.

Eyebrow: The locations for the eyebrows are two rectangular portions in the edge image that are located above the eye regions. For

further improvement, the edge images of these two locations are obtained. Since the Sobel technique can distinguish more edges than the Roberts strategy, it was now used to acquire the edge picture. The images here are then increased to fill in the spaces. The generated images are used to fine-tune the locations of the eyebrows.

Mouth: The left, right, top, and bottom-most points are taken to determine the mouth's centroid.

EMOTION CLASSIFICATION:

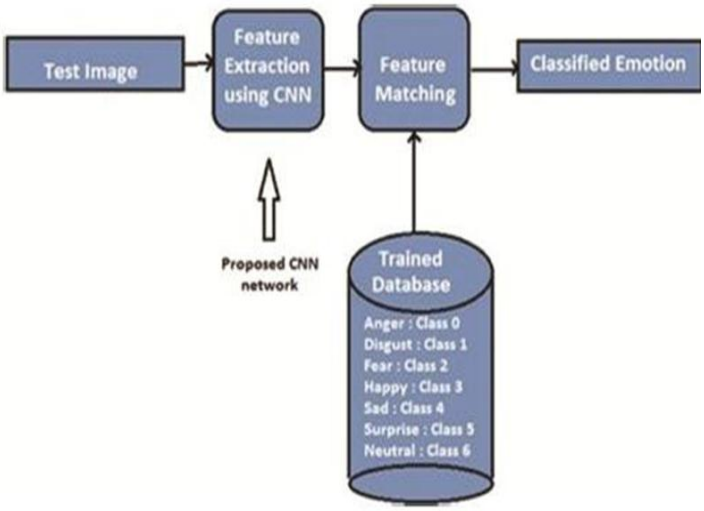


Fig 1.6 **Proposed testing architecture**

RESULT:

OUTPUT:

Fig 1.7 Surprised emotion

Fig 1.8 Fearful emotion

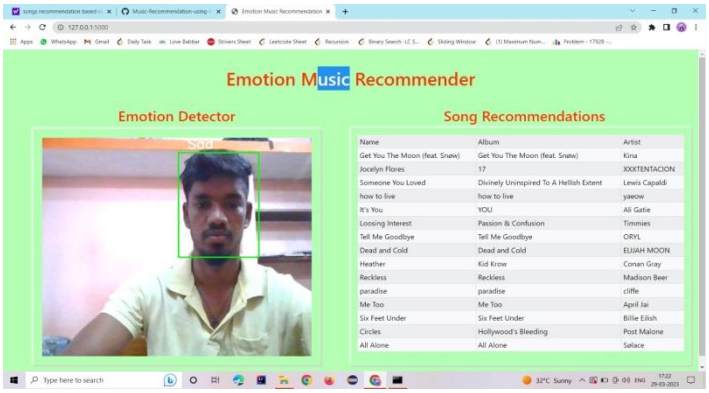


Fig 1.9 SAD Emotion.

Fig 1.10 HAPPY Emotion

CONCLUSION:

A general model is put into place to recommend music based on user emotions because music has the ability to emote users' feelings. Human emotions are crucial in helping people convey their thoughts. The system's primary objective is to recognize changes in the user's emotional state and play music in accordance with their preferences by listening to a variety of music tracks. The device employs CNN algorithm to categories emotions based on changes in the size, shape,

and movement of the eyebrows, eyes, and mouth. They fit into one of the six main emotional categories—sadness, happiness, anger, fear, disgust, and surprise—and a playlist is created for them based on those categories. The primary benefit of adopting CNN over The strength of SVM is its capacity to identify the key elements of an image without the aid of humans. Also, it is discovered that SVM's prediction accuracy is lower than CNN's accuracy. The suggested system has produced results with a high degree of accuracy. It is challenging to achieve 100% accuracy since human emotions are inconsistent and truly arise as a result of both internal and exterior events taking place in a person's environment. But, a flawless emotion-based music recommendation system can be created with a superior algorithm and more research. A web camera is used as a test subject for the proposed system. The overall cost of carrying out this job is incredibly low. Average expected time for the system's various modules.

REFERENCES:

- "Face Detection and Facial Expression Recognition System," by Anagha S. Dhavalikar and Dr. R. K. Kulkarni. 2018 International Conference on Electronics and Communication System (ICECS - 2014). (ICECS -2014).
- "Emotional Recognition from Facial Expression Analysis using Bezier Curve Fitting," by Yong-Hwan Lee, Woori Han, and Youngseop Kim. The 16th annual conference on network-based information systems will be held in 2019.
- Using Animated Mood Pictures in Music Recommendation, Arto Lehtiniemi and Jukka Holm, 2020 16th International Conference on Information Visualization.
- F. Abdat, C. Maaoui and A. Ruskin, "Human computer interaction utilizing emotion recognition from facial expression", 2021 UKSim 5th European Conference on Computer
- Anagha S. Dhavalikar and Dr. R. K. Kulkarni "Face Detection and Facial Expression Recognition System "Institute of Electrical and Electronics Engineers (IEEE 2014) (IEEE 2014)
- Byeong-jun Seungmin and Han "Music emotion classification and context-based music recommendation," by Rho, Sunghoon Jun, and Eenjun Hwang. Springer Science + Business Media, LLC 2009
- "Music Recommendation System Based on User's Sentiments Extracted from Social Networks," 2015 IEEE International Conference on Consumer Electronics, by Renata Lopes Rosa, Demosthenes Zegarra Rodrguez, and Graca Bressan
- Mitsunori Ogihara and Yajie Hu 12th International Society for Music Information Retrieval Conference, "NEXTONE PLAYER: A MUSIC RECOMMENDATION SYSTEM BASED ON USER BEHAVIOR"
- "Music Recommendation System Using Emotion Triggering Low-level Features" IEEE Transactions on Consumer Electronics 2012 by Kyoungro Yoon, Senior Member, Jonghyung Lee, and Min-Uk Kim
- "Emotion-based Music Recommendation by Association Discovery from Film Music," by Fang-Fei Kuo1, Meng-Fen Chiang2, Man-Kwan Shan2, and Suh-Yin Lee, published in 2005.
- "Predicting the Probability Density Function of Music Emotion Using EmotionSpace Mapping," by Yu-Hao Chin, Jia-Ching Wang, Senior Member, Jud-Chiang Wang, and Yi-Hsuan Yang, Member, IEEE. IEEE 2018
- "Creating Emotional Machines: Detecting Picture Feelings via Deep Neural Networks," by Hye-Rin Kim, Yeong-Seok Kim, SeonJoo Kim, and In-Kwon Le. IEEE TRANSACTIONS ON MULTIMEDIA 2018