ISSN 2277-2685
IJESR/June. 2025/ Vol-15/Issue-3s/483-492
G. Dharani *et. al.*, / International Journal of Engineering & Science Research

Emotion Based Ambiance And Music Recommendation Using Deep Learning

[1]Ms. K Srinidhi Reddy, [2]G. Dharani, [3]P. Divya, [4]M. Harini

[1]Assistant professor, Department of ECE, Bhoj Reddy Engineering College for Women, India

[2,3,4]B.Tech Students, Department of ECE, Bhoj Reddy Engineering College for Women, India

ABSTRACT

In today's hyperactive and fast-paced world, individuals are constantly bombarded with stressors stemming from work obligations, social responsibilities, and the pressure to keep up with a rapidly changing society. This overwhelming pace leaves little time for self-care, particularly when it comes to mental and emotional well-being. As people push themselves to meet endless demands, their psychological health often takes a back seat, leading to long-term negative consequences such as anxiety, burnout, and emotional exhaustion. The growing awareness around mental health has prompted interest in solutions that can seamlessly integrate into daily life without requiring extra effort from already overburdened individuals.

A significant body of research highlights the powerful impact that ambient factors—especially lighting and music—have on human emotions. Warm lighting can evoke feelings of comfort and relaxation, while specific types of music have been shown to reduce stress and promote positive emotional states. Drawing on these insights, our solution proposes an intelligent emotional wellness system that uses a network of cameras and sensors to assess a person's facial expressions and body language in real time. By analyzing these cues, the system determines the individual's current emotional state and automatically adjusts the environment to support their mental well-being—changing lighting schemes, playing mood-appropriate music, and even adjusting scent or room temperature if necessary.

1-INTRODUCTION

The working of our service starts with a webcam attached to any microprocessor/device with enough computation abilities. This web-camera takes image of faces. These images are passed in our trained model. Upon passing it successfully, we get one outcome that defines the current mood of the user. Interpreting the elements correctly, by using deep learning techniques gets very complicated, since the images have features very close to each other.

These co-relatability results in a huge co-variance as a result it is very difficult to separate it. Also, the accuracy of a normal human in identifying a mood of another human is somewhere around (67-73) %. This difficulty could be observed if we give a closer look to the dataset presented within the following classes {"angry", "disgust", "fear", "happy", "sad", "surprise", "neutral"}. Despite of all this system is designed in such a way that it solves the issue of identifying the correct facial mood and give accurate outputs.

All the algorithms and methodologies we used are purely based on Convolution Neural Networks (CNN). There are some specified algorithms in CNN which have been used here. Music and ambiance have a profound impact on human

emotions and well-being, making them integral to enhancing daily experiences. With the rapid growth of streaming platforms and smart environments, there is a growing demand for personalized experiences tailored to an individual's emotional state.

Traditional recommendation systems often rely on user preferences, genres, or collaborative filtering techniques. However, these methods may overlook the nuanced relationship between a person's emotional state and their desired auditory or environmental atmosphere. Emotion-based music and ambiance recommendation systems aim to bridge this gap by leveraging advancements in deep learning and affective computing. These systems analyze emotional cues from various inputs such as facial expressions, speech, physiological signals, or contextual data.

Using these insights, they curate personalized recommendations that align with the user's current mood or desired emotion. Music and ambiance have a profound impact on human emotions and well-being.

And making them integral to enhancing daily experiences. With the rapid growth of streaming platforms and smart environments, there is a growing demand for personalized experiences tailored to an individual's emotional state. Traditional recommendation systems often rely on user preferences, genres, or collaborative filtering techniques. However, these methods may overlook the nuanced relationship between a person's emotional state and their desired auditory or environmental atmosphere.

Emotion-based music and ambiance recommendation systems aim to bridge this gap by leveraging advancements in deep learning and affective computing. These systems analyze emotional cues from various inputs such as facial expressions, speech, physiological signals, or contextual data. Using these insights, they curate personalized recommendations that align with the user's current mood or desired emotional state.

2-LITERATURE REVIEW

The methodologies to enhance mood of a person and improve have been in existence since a long period of time. There have been various research papers on the same. Various methods have been proved, some being fully automated, while others include human interaction. There were many papers we referred to while our research part. The one from the authors missed to enhance the mood of person, the paper solely consisted of detection of face and identifying the mood of person by the image of face.

They had similar techniques like ours, but it had some of the older techniques for image processing and computation of the final outcome. We made 3-way comparison instead of 2- way comparison made in the paper. it is read by any e-book reader online. The problem we found here was that it didn't work on a human being. Another work consisted ambiance and music changing algorithm but it seemed to have failed in enhancing the mood of a person because the mood was detected only once during a duration of time. the same dataset as we have used, FER 2013 dataset, but they use a different feedback system for music.

They use a manual user prompt, while we use the consecutive emotions of the user as feedback and for music playlist generation we leverage the power of the music application platform unlike the system in the mentioned paper where they completely built the system from scratch hence they don't have the same varied user base and data on each user like music application has. the authors propose a mobile application that can make music

playlists according to the persons mood but the drawback of this system is its inability to scale to large environments and its inability to provide continuous emotion feedback which is solved in our system by keeping each node as a standalone device and connection them via an Realtime Database.

The system named MoodyPlayer, work, describes how the authors managed to detect three different types (Sad, neutral and happy) of mood by various algorithms (like SVM), and manually create a playlist from a selected number of songs. Their model could detect one face. our model detects 7 emotions in total. The mention paper uses a locally hosted database, which is less secured and versatile in compared to Firebase database that persons mood but the drawback of this system is its inability to scale to large environments

Finally, demonstrates about a web-based application, build from the scratch using in JavaScript, MongoDB, Express.js, Angular, Node.js. Again, there is a lot of manual work, consisting uploading of music for the very first time and to change the music, the used has to upload new songs every time else the repeated songs would be played. Some of the other existing systems with similar concept that we came across were, Musicovery SteroMood and MoodFuse.the authors propose a mobile application that can make music playlists according to the persons mood.

3-SOFTWARE REQUIREMENTS

The software and hardware requirements outlined are essential to ensure the smooth development, deployment, and execution of the project. On the hardware side, a system with an Intel i3 processor or higher, at least 4GB of RAM, and a minimum of 500GB storage is recommended to provide the necessary computational power and storage capacity for managing large datasets and applications. On the software front, the project leverages Python as the core programming language, complemented by Flask as the primary web development framework. The use of IDEs like PyCharm or Jupyter Notebook facilitates efficient coding and debugging. Additionally, Anaconda is included as a tool for managing dependencies and environments, ensuring seamless integration of libraries and frameworks. These requirements collectively provide a robust foundation for developing and maintaining a high-performance application.

Data flow

1. The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.
2. The data flow diagram (DFD) is one of the most important modeling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.
3. DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.

System Architecture

1. Load the dataset Data Preprocessing:
2. Handling null values (delete record, replace with zero, and replace with column mean, mode, median)
3. Handling duplicate values (delete duplicate records)
4. Handling Categorical Values (label encoder, one

hot encoding)

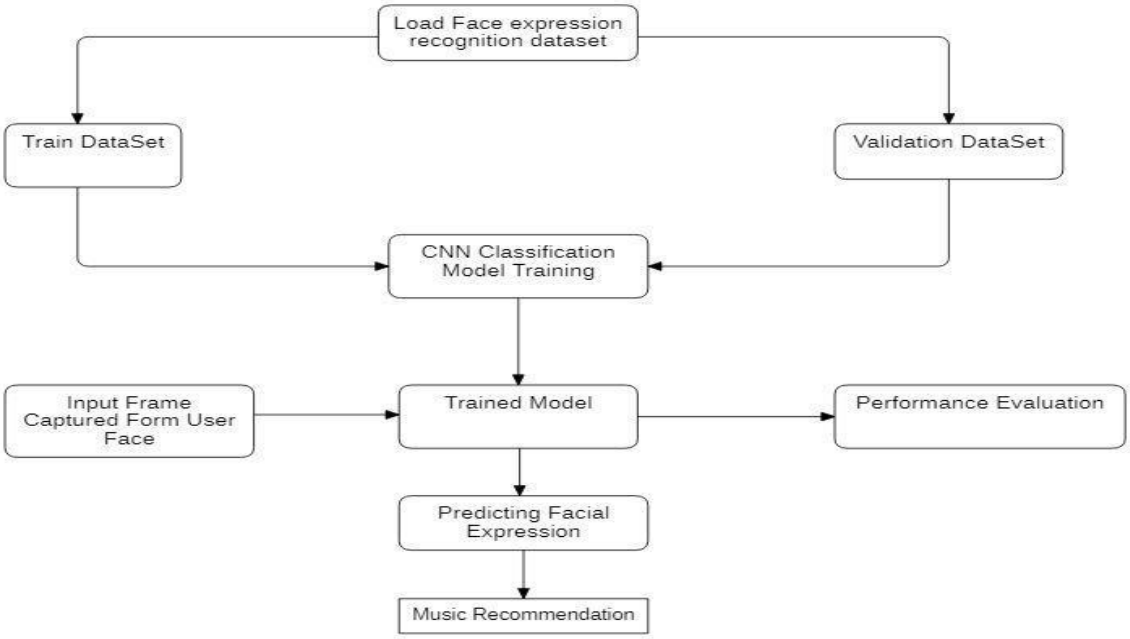

Fig - 2.1 System Architecture

4-MACHINE LEARNING PROJECT TESTING

How to write model tests:

So, to write model tests, we need to cover several issues:

- Check the general logic of the model (not possible in the case of deep neural networks so go to the next step if working with a DL model).
- Control the model performance by manual testing for a random couple of data points.
- Evaluate the accuracy of the ML model.
- Make sure that the achieved loss is acceptable for your task.
- If you get reasonable results, jump to unit tests to check the model performance on the real data.

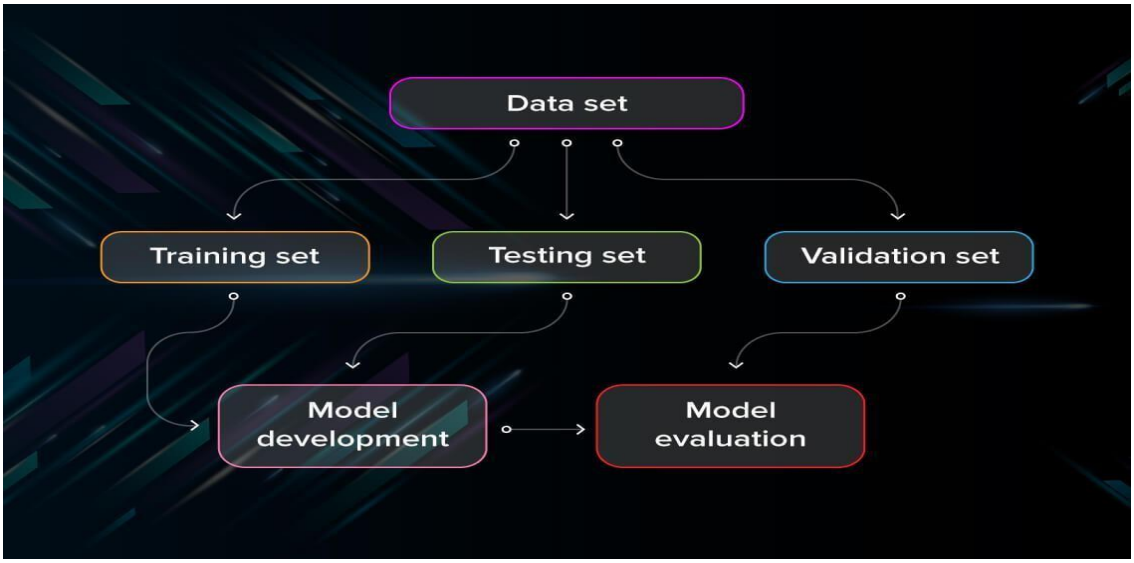

Fig - 3.0 :Testing

First of all, you split the database into three non-overlapping sets. You use a training set to train the model. Then, to evaluate the performance of the model, you use two sets of data:

- **Validation set:** Having only a training set and a testing set is not enough if you do many rounds of hyperparameter-tuning (which is always). And that can result in overfitting. To avoid that, you can select a small validation data set to evaluate a model. Only after you get maximum accuracy on the validation set, you make the testing set come into the game.

- **Test set (or holdout set).** Your model might fit the training dataset perfectly well. But where are the guarantees that it will do equally well in real-life? In order to assure that, you select samples for a testing set from your training set — examples that the machine hasn't seen before. It is important to remain unbiased during selection and draw samples at random. Also, you should not use the same set many times to avoid training on your test data. Your test set should be large enough to provide statistically meaningful results and be representative of the data set as a whole.

5-EMOTION BASED AMBIANCE AND MUSIC RECOMMENDATION USING DEEP LEARNING ALGORITHM

Population Survey

A population survey can help identify potential users and their preferences for emotion- based ambiance and music recommendations. Key survey areas include demographic data (age, occupation, location), current music habits, and interest in emotion-driven personalization. Questions can explore preferred use cases (e.g., relaxation, productivity, or entertainment) and willingness to adopt such technology. The survey results can guide the system's features, ensuring it aligns with user needs and expectations.

Existing System

There were many papers we referred to while our research part. The one from the authors in missed to enhance the mood of person; the paper solely consisted of detection of face and identifying the mood of person by the image of face. They had similar techniques like ours, but it had some of the older techniques for image processing and computation of the final outcome. We made 3- way comparison instead of 2- way comparison made in the paper. It consisted a music playing system that matched the mood of an e-book when it is read by any e-book reader online.

The problem we found here was that it didn't work on a human being. Another work, consisted ambiance and music changing algorithm but it seemed to have failed in enhancing the mood of a person because the mood was detected only once during a duration of time. the same dataset as we have used, FER 2013 dataset, but they use a different feedback system for music. They use a manual user prompt, while we use the consecutive emotions of the user as feedback and for music playlist generation we leverage the power of the music application platform unlike the system in the mentioned paper where they completely built the system from scratch hence they don't have the same varied user base and data on each user like music application.

The authors propose a mobile application that can make music playlists according to the persons mood but the drawback of this system is its

inability to scale to large environments and its inability to provide continuous emotion feedback

Proposed System

We used a generalized CNN algorithm and we process the images in real-time, by passing the matrix of the images through a CNN function. The implementations have been validated by using and developing a real-time system, using web-camera. This entire System has been deployed to click pictures and pass into the trained model. For algorithms, we've used python pass into the trained model. For algorithms, we've used python as our coding language. This system uses Firebase as its database and this database is used to store all the values of moods of user over a time and provide the rolling average for the same which is used further.

The proposed system uses deep learning to analyze users' emotions through inputs like facial expressions, voice tones, or text analysis. Based on the detected emotional state, the system recommends personalized music and adjusts the ambiance (e.g., lighting or soundscapes). It integrates with smart devices and platforms to provide seamless, real-time customization. The system continuously learns from user interactions, improving its accuracy and relevance over time. This innovative approach enhances user experience by creating emotionally adaptive environments.

The proposed system leverages deep learning techniques to analyze and interpret users' emotions through various inputs, such as facial expressions, voice modulation, or text-based sentiment analysis. By understanding the emotional state of the user, the system dynamically recommends personalized music and ambiance adjustments, such as altering lighting or soundscapes, to enhance the user's mood. The system integrates with smart devices, IoT platforms, and music libraries, enabling real-time synchronization and seamless functionality. This approach ensures a highly adaptive and responsive experience tailored to the user's emotional needs.

Additionally, the system incorporates continuous learning capabilities, improving its recommendations over time based on user feedback and interaction patterns. By utilizing advanced neural networks and a rich dataset, it ensures high accuracy in emotion detection and relevance in recommendations.

The proposed system is versatile, with applications ranging from personal wellness to professional environments, entertainment, and healthcare. This innovative solution not only enhances the user experience but also promotes emotional well-being by aligning ambiance and music with the user's mood.

Block Diagram

The block diagram of the project emotion based ambiance and music recommendation using deep learning algorithm

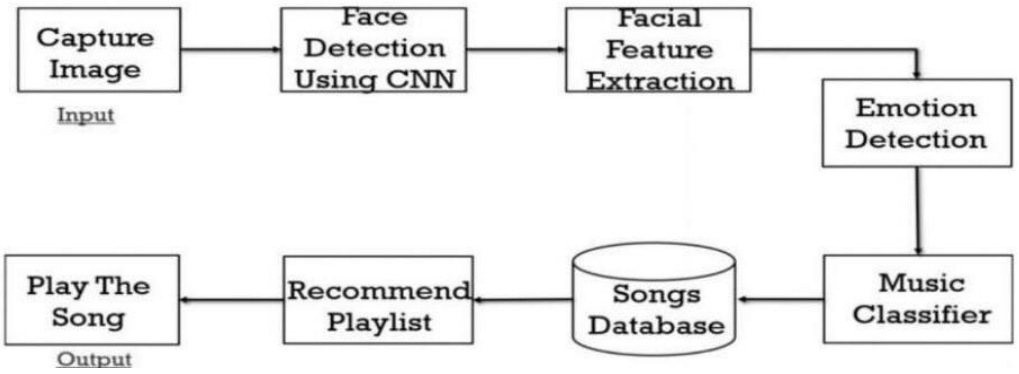

Fig - 4.0 :Block Diagram

This block diagram illustrates the process of an Emotion-Based Ambiance and Music Recommendation System using deep learning. Here's the step-by-step explanation:

1. Capture Image (Input)

- The system begins by capturing an image of the user using a camera or other image- capturing device.
- This input serves as the basis for analyzing the user's facial expressions.

2. Face Detection Using CNN

- The captured image undergoes face detection, where a Convolutional Neural Network (CNN) is used to locate and identify the user's face in the image.

6-ADVANTAGES, DISADVANTAGES AND APPLICATIONS

Advantages

1. Personalized Experience

   The system adapts to the user's emotions, providing a highly personalized ambiance and music selection tailored to their mood and preferences.

2. Enhanced User Satisfaction

   By recommending music and ambiance that align with the user's current emotional state, the system improves user satisfaction and engagement.

Disadvantages

1. Lighting and Background Dependence:

   For the system to work accurately, it needs the face to be well-lit and the background to be clutter-free. Poor lighting or a noisy background can make it difficult for the algorithm to identify facial features accurately. For example, shadows or uneven lighting might cause errors in detecting expressions.

2. Dataset Limitations:

   The system is trained on a limited dataset that only covers four basic emotions (likely happiness, sadness, anger, and surprise). This restricts its ability to recognize more nuanced or complex emotions, such as envy, embarrassment, or contentment, which might be necessary in real-world applications.

Applications

1. Entertainment Industry

   Enhance user experience in streaming platforms like Spotify or Netflix by providing mood-based music or content recommendations.

2. Healthcare and Therapy

   Assist in mental health therapy by using mood-specific ambiance and music to help patients relax, reduce stress, or uplift their mood.

7-RESULT

The phase-1 output of the project Emotion based ambiance and music recommendation using deep learning.



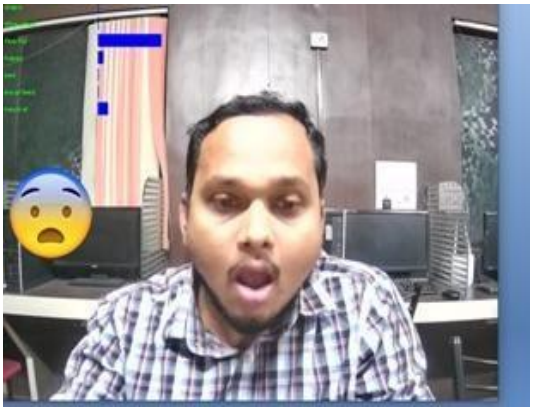

Fig - 6.0 :Outputs

Once the emotion is detected, the system dynamically adjusts the environment and music recommendations to match the user's emotional state. For instance, the ambiance is modified by controlling IoT-enabled devices like smart lights.

Simultaneously, the system selects music tracks from a curated database or online streaming platforms, offering personalized suggestions in real-time. The integration of real-time feedback allows users to rate the accuracy of the detected emotion and the appropriateness of the recommendations, which helps refine the system over time. This seamless blend of emotion detection, IoT integration, and adaptive learning ensures a personalized and immersive user experience.

Our system provides an accuracy of around 74%, when any image or video is passed through the system model. There are various figures describing the same. The image is of same person, in various moods. To check if system works when gender is changed, we have tested image provided in. And to check its working on various people, we passed on a picture with various people, and the results were shown. Now, we provide the training results of the model through which all the images and videos are passed. The accuracy graph is shown. The loss graph is shown, the accuracy reaches up to 74, over a period of time and at the same time the loss decreases gradually. We compare our results with another work used a simple architecture in compared to use.

To prove how our results, outperform the results from that paper , accuracy and loss comparison graphs are shown. These graphs clearly show that how the accuracy of our model gets better overtime and loss is also decreased over time. Hence, our model has clearly performed better than the one demonstrated in their work, by a scale of 5-8%.

Hence after getting all appropriate results of face and mood, this score is collected back and then used to send to over Music API.

The code for Music API is designed in such a way that it collects the songs from one year ago, previously played and very recently played. Hence, a playlist is created in music application, which is totally personalized to the interest of user. An auto bot login through the music application login page, which automatically inputs E-mail, password and hence, as the personalized playlist is created, the song starts playing.

In same way, the lights are controlled, the score from Firebase helps to detect various phases of mood and according to that, the value is passed and the ambiance of room is changed. Hence, the system is complete and works perfectly.The loss graph is shown, the accuracy reaches up to 74, over a period of time and at the same time the loss decreases gradually.There are various figures describing the same. The image is of same person, in various moods.

8-CONCLUSION

By leveraging the power of Machine learning to train on the FER 2013 database to generate a model to detect emotions, and then converting them to numerical scores as a function of confidence score and weights relating to emotions, which are then converted to energy scores to make a playlist on music application and then playing them using a Selenium based bot, to make an all in one system which works with zero human interaction and tries to improve the atmosphere in the room relative to the overall mood. The entire above process can be packaged on a single board computer to run at a single command and then continues running with zero human interaction.

REFERENCES

1. FER 2013 Dataset, Available [Online] https://www.kaggle.com/c/challenges-in-representation-learning-facialexpression-recognition-challenge/data.

2. Ji Chen, Jin Wang, Lin Wang, Method for driving virtual facial expressions by automatically detecting facial expressions of a face image, US7751599B2.

3. Ioana Bacivarov, Galway (IE); Peter Corcoran, Claregalway(IE),"(12) United States Patent," vol. 2, jul 16, 2013.

4. S. W. Ryan, (12) System and Method for Programming a Weighing Scale Usinga Key Signal to Enter a Programming Mode, vol. 1, no. 12. 2009.

5. M. Based and O. N. A. Mood, "(12) United States Patent,"vol. 2, no. 12, 2013.

6. S. Gilda, H. Zafar, C. Soni, and K. Waghurdekar, "Smartmusic player integrating facial emotion recognition and music mood recommendation," Proc. 2017 Int. Conf. Wireless Communication. Signal Process. Networking, WiSPNET 2017, vol.2018-Janua,
pp. 154– 158, 2018.

7. M. Gupta, U. Pradesh, S. Singhal, U. Pradesh, M. Pandey, and U. Pradesh, "Real time emotions recognition and analysis-based music player," vol. 5, no. 2, pp. 1462– 1465,2019.

8. A. R., A. Vollal, P. B., S. Yadav, and R. M., "MoodyPlayer:A Mood based Music Player," Int. J. Comput. Appl., vol.141, no. 4, pp. 21–25, 2016.

9. A. Gupte, A. Naganarayanan, and M. Krishnan, "EmotionBased Music Player - XBeats," Int. J. Adv. Eng. Res. Sci.,vol. 3, no. 9, pp. 217–223, 2016.

10. O. Arriaga, M. Valdenegro-Toro, and P. G.Plöger, "Real-time convolutional neural networks for emotion and gender classification, "ESANN 2019 - Proceedings, 27th Eur. Symp. Artificial. Neural Networks, Compute. Intelligent Machine Learning., pp.221–226, 2019.