EmoTunes: Enhancing User Experience with Emotion-Aware Music Recommendations

Saliq Gowhar Khan
Department of Information Technology
National Institute of Karnataka, Surathkal
Mangalore, India
saliqgowhar.211ee250@nitk.edu.in

Harshit Ravindra Gawade
Department of Information Technology
National Institute of Karnataka, Surathkal
Mangalore, India
harshithrg.211ai019@nitk.edu.in

***Abstract*—The rapid growth of music streaming services has revolutionized completely the way people access and consume music. However, the sheer volume and variety of music available pose a significant challenge for users in finding songs that align with their emotional preferences. Traditional music recommendation systems primarily rely on user behavior, such as listening history or explicit feedback, to generate recommendations. While these methods have proven effective to some extent, they often overlook the crucial aspect of emotions, which play a fundamental role in the music-listening experience. In this paper, we propose an emotion-based music recommendation system created by leveraging various deep-learning architectures with the sole aim of bridging the gap between emotions and music recommendations by incorporating users' emotions as a key factor in the process. Emotions are powerful indicators of personal preferences and can significantly impact how individuals perceive and engage with music. By understanding and leveraging these emotional cues, the system can provide tailored recommendations that resonate with users on a deeper level, enhancing user satisfaction and promoting music discovery.**

***Index Terms*—CNN, Facial emotion detection, Haar Cascades, Recommender Systems, Neural network classifiers, Spotify API, Cosine similarity, Tf-IDF, Content-based filtering**

I. Introduction

In the digital age, finding personalized playlists that resonate with our emotions and tastes among the vast array of available music has become a challenge. Traditional music recommendation systems often overlook the crucial role of emotions in our musical choices. By incorporating facial-emotion detection, music recommendation algorithms can establish a stronger emotional connection between the music and the listener. [1] This deeper understanding of users' emotional states enhances the overall music experience, providing comfort, inspiration, and enjoyment.

The ability to identify facial emotions also makes it possible to find new songs. Facial-emotion recognition offers an intuitive and immediate comprehension of users' emotional reactions to music, in contrast to conventional approaches that rely on explicit user feedback. Music recommendation algorithms can broaden users' musical horizons by introducing them to new songs and genres that evoke similar emotional responses. [2] Users are more satisfied and engaged when recognizing and addressing their emotional states. Users are more likely to spend time researching and listening to the recommended music when it corresponds with their emotions, which increases pleasure and overall user retention. [3]

This paper presents a novel approach utilizing deep learning architectures to construct an emotion-based music recommendation system. The proposed system employs Convolutional Neural Networks (CNNs) in conjunction with OpenCV to accurately capture and depict users' real-time mood which is classified into either of these seven distinct classes, namely Angry, Sad, Happy, Neutral, Surprised, Disgusted, and Fear. To facilitate music recommendations, a dataset is curated, by training a neural network model using a dataset, named song moods sourced from Kaggle which comprises of 660 songs with labeled mood information to classify a much larger dataset of 150,000 songs obtained from the Spotify Dataset on Kaggle. The classification process categorizes these songs into four main mood classes: Calm, Energetic, Happy, and Sad and stores them in our dataset with the classified mood label. In order to personalize the music recommendations, the system leverages the Spotify API to access users' Spotify accounts. By analyzing the user's most recently visited playlist, the system obtains valuable insights into their current musical preferences and runs similarity measures such as cosine similarity to finally curate a list of 20 songs for the user based on their current mood as well as their recent music choices.

II. Literature survey

Shlok Gilda et al. [1] focus on the development of a facial mood detection-based recommender system for music classification. The system utilizes convolutional neural networks (CNNs) to analyze facial expressions and accurately classify the user's mood into four categories: happy, sad, angry, and neutral. The CNN model achieves an impressive accuracy of 90.23% in emotion recognition. The dataset used for training and testing consists of grayscale images of faces from the FER2013 dataset, with a total of 26,217 images categorized into the four chosen emotions. The network architecture includes convolutional layers with ReLU activation, pooling layers for dimensionality reduction, and dense layers for high-level feature extraction and classification. Softmax activation is applied at the output layer to generate a probability distribution for each emotion class. Moving on to the music classification module, the acoustic features of 390 songs are extracted using LibROSA and audio pitch algorithms. These features include spectral, rhythmic, tonal mode, and pitch characteristics. An artificial neural network is trained on these features, in classifying the songs into the four mood classes: exciting and energetic, happy and joyful, sad and melancholy, and calm and relaxed. Recursive Feature Elimination (RFE) is employed to select the most relevant features for accurate mood classification. The overall process involves preprocessing the songs, extracting relevant features, and training the neural network.

Sourav Joshi et al. [2] propose the process of building an emotion classification model using deep learning algorithms. The data for training the model was collected from two datasets obtained from Kaggle, which focused on emotions in text. The selected emotions for training were anger, happiness, love, and sadness. The collected text data underwent pre-processing steps, including cleaning to remove URLs, symbols, and punctuation, as well as tokenization, lowercase conversion, and lemmatization. The resulting tokens were then joined and transformed into feature vectors using the TF-IDF technique.Four deep-learning algorithms were employed for model building: CNN, LSTM, CNN-LSTM, and LSTM-CNN. The CNN model utilized convolutional and pooling layers, while the LSTM model focused on capturing long-term dependencies. The combined CNN-LSTM and LSTM-CNN models incorporated both convolutional and LSTM layers to capture local and global features. After training the models, the LSTM-CNN approach exhibited better performance in emotion classification. It consisted of dropout, LSTM, convolutional, max pooling, dense, and softmax activation layers to generate output probabilities for each emotion class. The user gives text input about their feelings which is then fed into the model from which detected emotion is used as a seed for recommending songs using an API.

Ke Chen et al. [3] focus on content-based music recommendation using a metric learning framework. It consists of two branches: the User Branch and the Audio Branch. The User Branch is responsible for encoding the user's music preferences based on their listening history and demographics. It utilizes a classification problem formulation to classify whether a user likes or dislikes a track. The branch incorporates explicit and implicit feedback from the user's interactions with tracks. The data is converted into lookup embeddings and fed into a network inspired by YoutubeDNN. Hidden layers with leaky-ReLU activation are employed, and a cross-entropy loss is minimized during training. The Audio Branch utilizes a Siamese model to learn Audio embedding (AE) directly from audio data. The User Embeddings (UE) obtained from the User Branch act as an anchor for metric learning. The branch uses log-mel spectrograms as inputs and employs CNNs to extract features. The CNNs share weights and configurations and are followed by a feature vector layer to obtain AE with a dimension of 40. The branch utilizes negative sampling and a max-margin hinge loss to measure pairwise similarity and optimize the model. The trained User Branch is used to generate recommendations, and the Audio Branch is evaluated using hit rate. A dataset with user-liked/disliked tracks is created, and log-mel spectrograms are generated for training. The loss is optimized using stochastic gradient descent. The models are evaluated based on AUC and precision for music recommendation tasks and genre classification.

Jagendra Singh et al. [4] describe a methodology in the paper that involves collaborative filtering-based recommendation techniques in the context of music. These methods rely on the extraction of features from the music objects and utilize classification models such as KNN (K-Nearest Neighbors) and SVD (Singular Value Decomposition) for the recommendation. In the pre-processing and feature extraction stage, features are extracted from the music objects using multidimensional information interpretation. One-hot encoded vectors are generated to represent the items in a feature space. The KNN classification model is applied to construct a model for each user based on their ratings for the items. Additionally, the SVD algorithm is used for matrix factorization to reduce unnecessary parameters and map users and items to a latent factor space. The SVD model is trained and tested with different numbers of latent factors, and the optimal number of factors is chosen based on training RMSE (Root Mean Square Error). Finally, the hybridization approach is adopted to combine the outputs from Factorization Machines and SVD. A deep neural network-based recommendation system is designed to improve the accuracy of predictions. The scores from SVD and Factorization Machines are treated as a classification problem, and the hybrid recommendations for songs are obtained.

III. Problem Statemnt

To build a system capable of recognizing the facial emotion of the user in real-time, and recommending songs as per the current choice and the mood of the user.

Objectives:

- To detect the facial emotion of the user and classify the mood.
- To analyze the recent playlist of the user on Spotify and figure out similar songs in our dataset.
- To recommend songs to the user as per their recently played songs and their current emotion
- To deploy the model with a well established front end.

IV. Background

A. Haar Cascade Classifier

The Haar cascade algorithm is a robust technique utilized for object detection in images, capable of accurately identifying objects irrespective of their position or scale within the image. Notably, this algorithm exhibits real-time processing capabilities, making it highly efficient for practical applications. Training a Haar cascade detector makes it possible to detect various objects such as cars, bikes, buildings, fruits, and more. The Haar cascade algorithm leverages the concept of cascading windows, systematically examining features within each window to determine if it corresponds to an object of interest. In our paper, we propose the use of Haar Cascades frontal face detector for real-time human face detection using OpenCV.

B. Convolutional Neural Networks

A Convolutional Neural Network (CNN), also known as ConvNet, is a type of special neural network specialized in analyzing data with a grid-like topology, such as images that are nothing but 2D grids of pixel intensities. Convolutional, pooling, and fully connected layers are the three basic layers found in CNN architectures. The convolutional layers are pivotal components of CNNs, employing filters to extract multiple valuable features like edges, textures, and shapes from the input image. Subsequently, the extracted features pass through pooling layers, responsible for downsampling the feature maps, and preserving critical information while reducing spatial dimensions. The resulting outputs then undergo one or more fully connected layers, enabling predictions and image classification. In our paper, we propose the use of CNN for the facial emotion detection of users in real-time. A generalized CNN model is given below in figure 1.

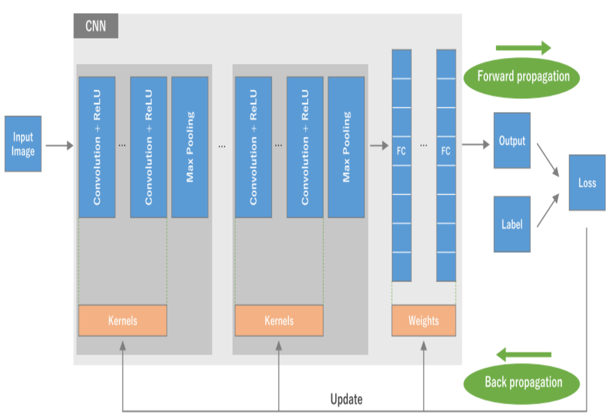

Fig. 1. Generalized CNN architecture [5]

C. Neural Network Classifiers

Neural network classifiers are highly proficient machine learning models renowned for their ability to tackle intricate classification tasks. Drawing inspiration from the intricate workings of the human brain, these classifiers comprise interconnected artificial neurons organized in layered structures. Each neuron receives input, undergoes a mathematical transformation, and generates an output that contributes to the overall decision-making process. By undergoing a training process, neural network classifiers acquire the capacity to discern patterns and generate precise predictions by adjusting the connection strengths between neurons. Typically, this learning is accomplished using algorithms such as backpropagation, which iteratively update the network's parameters based on the disparities between predicted and expected outcomes. The success of neural network classifiers spans diverse domains, including image recognition, natural language processing, and sentiment analysis. Their advancements have revolutionized fields like computer vision, speech recognition, and recommendation systems, driving significant progress.

V. Methodology

The complete methodology is provided below along with the flowchart which is given in Figure 2.

A. Dataset preparation and EDA

The dataset used for facial emotion detection is the FER 2013 dataset sourced from Kaggle. It contains images in greyscale which correspond to seven different classes of emotions namely angry, sad, happy, neutral, fear, disgust, and surprise. The images have a dimension of 48 by 48 and do not require any pre-processing. It consists of a total of 35,887 images. The labeled images per emotion distribution include 4,897 for angry, 547 for disgust, 5,016 for fear, 8,029 for happy, 4,660 for sad, 3,372 for surprise, and 6,890 for neutral. The data is available both in CSV as well as image format. As for our model, we have utilized the image dataset.

In terms of song recommendations, two datasets were employed. The first dataset, named "Spotify music data to identify moods," encompasses 661 songs. Each song is labeled with a specific mood, such as sadness, happiness, calmness, or energy. Additionally, it includes 17 other features that describe the songs in terms of metrics like valency, acoustics, danceability, and loudness, among others. The second dataset, "Spotify songs," consists of 150,000 songs and shares the same features as the previous dataset, excluding the mood label.

As far as preprocessing of the Spotify dataset is concerned, we have dropped the unnecessary columns like the year of release, the date of release, etc and we have also normalized the features like loudness which have values more than 1 so as to get all the features in the same range of values.

B. Facial emotion recognition

The image of the user is obtained in real-time using the Python framework OpenCV, and the face of the user is obtained using the frontal face Haar Cascade classifier. The emotion of the person is detected using a CNN model trained using the FER2013 dataset. The model is designed for classifying facial emotions into seven categories: anger, sad, calm, happy, neutral, fear, surprise, and disgust.

The model begins with a sequential structure, which allows for building the network layer by layer. It starts with a series of convolutional layers, which are responsible for learning spatial features from the input images. The first convolutional layer consists of 64 filters with a size of 3x3, preserving the input shape of grayscale images (48x48 pixels). The subsequent layers include a second convolutional layer with 128 filters of size 5x5 and a third and fourth convolutional layer, both with 512 filters of size 3x3. These layers are accompanied by batch normalization, which normalizes the inputs, and ReLU activation functions, introducing non-linearity to the network. Max pooling layers with a pool size of 2x2 follow each convolutional layer, reducing the spatial dimensions. To prevent overfitting and enhance model generalization, dropout layers with a rate of 0.25 are inserted after each max pooling layer. After the final convolutional layer, the feature maps are flattened into a 1-dimensional vector.

The subsequent fully connected layers help capture higher-level representations from the flattened features. The model includes a fully connected layer with 256 neurons and another with 512 neurons. Both fully connected layers are followed by batch normalization, ReLU activation, and dropout with a rate of 0.25. The model concludes with an output layer comprising the number of classes, which is seven in this case. The activation function used for the output layer is softmax, which generates probabilities for each class, representing the confidence of the model's prediction for each emotion. During training, the Adam optimizer is utilized with a learning rate of 0.0001. The categorical cross-entropy loss function is employed to measure the dissimilarity between the predicted and actual emotion labels. The model is compiled with the accuracy metric to monitor its performance.

In summary, this model consists of several convolutional layers to extract meaningful features from facial images, followed by fully connected layers for higher-level representation learning. The dropout and batch normalization layers aid in preventing overfitting, and the softmax output layer provides the probability distribution over the seven emotion classes.

C. Song mood classification

Utilizing the Spotify Dataset consisting of 661 songs, encompassing essential features such as loudness, popularity, artist name, release year, valence, tempo, and more, we constructed a robust model. This model was subsequently employed to classify a significantly larger dataset of 1.5 lakh songs into four distinct categories: happy, sad, energetic, and calm. By categorizing these songs, we effectively expanded our song pool for recommendation purposes. This approach enables us to provide a more diverse and comprehensive range of song recommendations to users, accommodating their varying preferences and moods. The song features are scaled using techniques like Min-Max scaling to ensure that all features are normalized within a specific range, thereby removing any potential bias and allowing for fair comparison and accurate modeling. After scaling, the mood labels associated with the songs are encoded into numerical representations. This encoding process converts the categorical mood labels into numerical values, enabling machine learning models to interpret the labeled data effectively. With the pre-processed and encoded data, a model is constructed.

The model is a sequential model built using the Keras API from the TensorFlow library. An input layer, one or more hidden layers, and an output layer make up this structure. The input layer receives the song features as input, with a dimension of 10. There is one hidden layer with 8 neurons and ReLU activation. The output layer, consisting of 4 neurons, produces the final predictions using the SoftMax activation function. The model is trained using categorical cross-entropy loss and the Adam optimizer, which adapts the learning rate dynamically. Evaluation of the model's performance is done using accuracy, measuring the proportion of correctly predicted mood labels. The model is trained for 300 epochs with a batch size of 200. Once the model is trained and validated, it can be deployed to predict the mood of songs based on their features. Given the feature values of an unseen song, the trained model processes the input and generates a predicted mood label. The model's performance is assessed using a number of criteria, including accuracy, precision, recall, and F1 score. These metrics provide insights into the model's effectiveness in accurately classifying songs based on their mood. Furthermore, the results and analysis derived from the model's predictions are examined to gain deeper insights into the relationship between song features and mood labels. Visualization techniques such as plots, graphs, or heatmaps may be employed to illustrate these relationships and facilitate a better understanding of the classification outcomes.

D. Song recommendation system

To create a music recommendation system, we follow a multi-step process. After classifying the mood of the 1.5 lac songs in the previous step, we reduce the dataset based on the user's mood preference. For example, if the user indicates that they are feeling happy, we narrow down the dataset to include only songs that are happy, calm, or energetic. This reduction in dataset size allows us to focus on songs that align with the user's current mood.

Next, we request the user to provide their Spotify details, and specifically, their most recent playlist. By analyzing this playlist, we apply TF-IDF vectorization and cosine similarity techniques. TF-IDF vectorization converts the textual representation of songs into numerical vectors, capturing their unique features. Cosine similarity measures the similarity between the user's playlist and the songs in the reduced dataset. Using the calculated similarity scores, we identify the most similar 20 songs from the reduced dataset that closely match the user's playlist. These songs are then recommended to the user as they are likely to resonate well with their current music taste and mood. Additionally, we go a step further by generating a playlist directly into the user's Spotify account. This enables the user to easily access and enjoy the recommended songs in their preferred music streaming platform.

Overall, the music recommendation system combines mood classification, dataset reduction based on user preference, TF-IDF vectorization, and cosine similarity to provide personalized song recommendations tailored to the user's mood and musical preferences. The generated playlist in the user's Spotify account enhances the user experience by seamlessly integrating the recommended songs into their existing music library.

E. Deployment

The recommender system is deployed using Flask framework in Python which is used to integrate the front end and the back end. The front end is created using HTML and CSS and the backend consists of the facial emotion recognizor and the music recommender. The user is asked to sign into their Spotify account and then the user's image is captured and 20 songs are recommended and displayed on the webpage along with their Spotify links and cover photos of their respective albums.

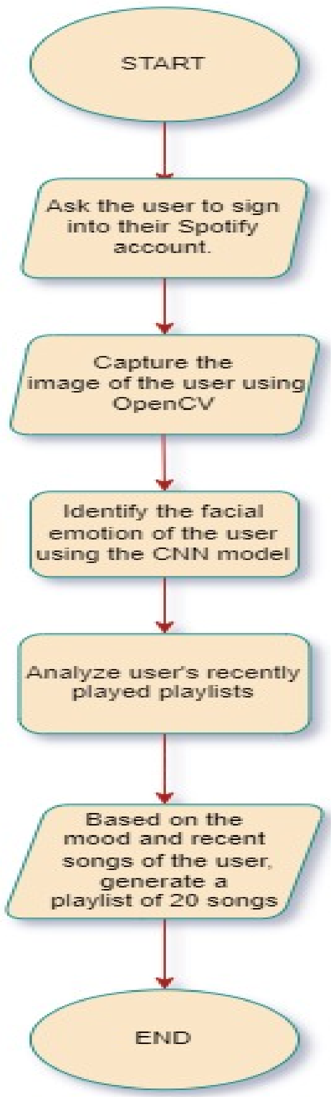

Fig. 2. Flowchart for complete methodology

VI. Results

1) We have performed exploratory data analysis over the Spotify dataset using various visualization techniques and gained the following insights:
   - Factors such as the year of release, loudness, and energy exhibited a positive correlation with popularity. In contrast, features like acousticness and instrumentalness demonstrated an inverse relationship with popularity. It is interesting to note that loudness, energy is associated with happy songs while acousticness, instrumentalness are associated with sad songs.
   - It is observed that the top 10 most popular tracks each year predominantly consist of happy songs. This suggests a strong trend where upbeat and positive songs tend to resonate more with the general audience, leading to higher popularity. The data indicate that listeners are more inclined towards joyful and uplifting music, as reflected by the consistent presence of happy songs in the top rankings year after year.
   - Contrary to the observation regarding the overall popular tracks, our dataset reveals an interesting pattern when considering the artists with the most tracks. It is found that many of these artists have their highest-ranking songs on the popularity charts associated with sad or emotionally poignant themes. This suggests that while the general audience might gravitate towards upbeat songs for overall popularity, certain artists have built a strong following and recognition through their ability to evoke deep emotions with their sad songs.
   - An intriguing observation is that while there are artists who exclusively produce happy songs, there are no artists who exclusively focus on producing sad songs.

2) The accuracy obtained for the facial emotion detection model is around 84% for the training set and 78% for the cross-validation set. Although most of the predictions come out to be either happy or neutral. This is because the number of images corresponding to these emotions is more in number and also because it's very hard for a person to actually look sad or angry as it mostly comes out to be a neutral look only.

3) To analyze the results of the mood classification model for songs, we have used 10-fold cross-validation over the neural network which gave a mean baseline accuracy of 80% and a standard deviation of 4.17% from the mean accuracy. The songs are classified into four categories namely Sad, Calm, Energetic, and Happy.

4) The output of the recommender system is given below in Figure 4.

VII. Conclusion and Future works

In conclusion, this research paper has presented a comprehensive study on facial emotion detection using Convolutional

```
Predicted Emotion:  Neutral
Recommended songs are as under:
                                                        name
73394                                        On Your Way Down
25725                                 A Day In The Life - Remix
5359             Down In the DM (feat. Nicki Minaj) - Remix
27017    Tokyo Drift (Fast & Furious) - From "The Fast ...
74160    The Way Life Goes (feat. Nicki Minaj & Oh Wond...
25425             With A Little Help From My Friends - Remix
36303    Waka Waka (This Time for Africa) [The Official...
38387                   All Night Longer REMIX (feat. B.o.B)
93889    UP! (feat. Chris Brown, Pleasure P) (R&B Remix...
16700                   In Your Eyes (feat. Kenny G) - Remix
9285                                  50K Remix (feat. T.I.)
73421                                          Down the Road
58222                                          8 Ball - Remix
155245                  Plain Jane REMIX (feat. Nicki Minaj)
67720             There But for the Grace of God Go I - Remix
46485                   I'm a Flirt Remix (feat. T.I. & T-Pain)
8440                               I Lay My Love on You - Remix
114251                       When The Lights Go Out - US Remix
73422                                          Down on the Farm
38937      Whole Lotta Choppas (Remix) [feat. Nicki Minaj]
```

Fig. 3. Final output of the recommender system

Neural Networks (CNN) and its application in building an emotion-based song recommendation system. The proposed approach has successfully demonstrated the effectiveness of CNN in accurately detecting facial emotions, paving the way for personalized music recommendations based on the user's mood. By leveraging audio features to classify songs as happy or sad, the recommendation system has shown promising results in providing tailored song suggestions to enhance user experience and engagement. The findings of this research highlight the potential of integrating facial emotion detection with music recommendation systems. However, there are still areas for improvement and future works. One significant aspect is the inclusion of demographic factors such as age, gender, and location-specific recommendations. We aim to transition our face detection approach from haarcascades to facenet. Haarcascades, although widely used, can exhibit shortcomings in accurately detecting faces under challenging conditions such as variations in pose, lighting, and occlusions. By leveraging the capabilities of facenet, which has demonstrated superior performance in handling these challenges, we anticipate more robust and precise face detection results for our future work. By considering these factors, the recommendation system can further enhance personalization and relevance, ensuring that the suggested songs align with individual preferences.

VIII. Individual Contributions

1) Saliq Gowhar
   - Worked on facial emotion detection and its real-time analysis.
   - Worked on song mood classification.
   - Worked on building the recommender system.
2) Harshit Gawade
   - Worked on facial emotion detection and its real-time analysis.
   - Performed EDA, pre-processing, and multiple feature-oriented techniques.
   - Worked on building the recommender system and integrating Spotify API.
   - Deployment of the model.

References

[1] S. Gilda, H. Zafar, C. Soni and K. Waghurdekar, "Smart music player integrating facial emotion recognition and music mood recommendation," 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2017, pp. 154-158, doi: 10.1109/WiSPNET.2017.8299738.

[2] S. Joshi, T. Jain and N. Nair, "Emotion Based Music Recommendation System Using LSTM - CNN Architecture," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 01-06, doi: 10.1109/ICCCNT51525.2021.9579813.

[3] K. Chen, B. Liang, X. Ma and M. Gu, "Learning Audio Embeddings with User Listening Data for Content-Based Music Recommendation," ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 2021, pp. 3015-3019, doi: 10.1109/ICASSP39728.2021.9414458.

[4] J. Singh, "Collaborative Filtering based Hybrid Music Recommendation System," 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS), Thoothukudi, India, 2020, pp. 186-190, doi: 10.1109/ICISS49785.2020.9315913.

[5] Yamashita, R., Nishio, M., Do, R.K.G. et al. Convolutional neural networks: an overview and application in radiology. Insights Imaging 9, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9

[6] N. V. Durga Malleswari, K. Gayatri, K. Y. Sai Kumar, N. Likhita, P. K and D. Bhattacharyya, "Music Recommendation System using Hybrid Approach," 2023 Second International Conference on Electronics and Renewable Systems (ICEARS), Tuticorin, India, 2023, pp. 1560-1564, doi: 10.1109/ICEARS56392.2023.10085059.

[7] D. Bogdanov, M. Haro, F. Fuhrmann, A. Xambo, E. Gomez and P. Herrera, "A content-based system for music recommendation and visualization of user preferences working on semantic notions," 2011 9th International Workshop on Content-Based Multimedia Indexing (CBMI), Madrid, Spain, 2011, pp. 249-252, doi: 10.1109/CBMI.2011.5972554.

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

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

[10] Peter Burkert, Felix Trier, Muhammad Zeshan Afzal, Andreas Dengel, and Marcus Liwicki, "DeXpression: Deep Convolutional Neural Network for Expression Recognition.

[11] Gil Levi and Tal Hassner, "Emotion Recognition in the Wild via Convolutional Neural Networks and Mapped Binary Patterns."

[12] A. Kołakowska, A. Landowska, M. Szwoch, W. Szwoch, and M. R. Wrióbel, "Human-Computer systems interaction: back-grounds and applications," 3, ch. Emotion Recognition and Its Applications, Cham: Springer International Publishing, 2014, pp. 51–62.