INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT (IJSREM)
VOLUME: 08 ISSUE: 04 | APRIL - 2024 SJIF RATING: 8.448 ISSN: 2582-3930

Literature Review : Emotion-Based Music Recommendation System with Age Detection

Prof Vijay C P
Department of CSE(AI &ML)
Vidyavardhaka College of Engineering
Mysuru, India

Akash Baskar S
Department of CSE(AI &ML)
Vidyavardhaka College of Engineering
Mysuru, India

J B Akhil Raj
Department of CSE(AI &ML)
Vidyavardhaka College of Engineering
Mysuru, India

K S Praveen
Department of CSE(AI &ML)
Vidyavardhaka College of Engineering
Mysuru, India

Kalpak S
Department of CSE(AI &ML)
Vidyavardhaka College of Engineering
Mysuru, India

***Abstract*— This research introduces an innovative Emotion-Based Music Recommendation System coupled with Age Estimation. Leveraging computer vision techniques, specifically Convolutional Neural Networks (CNNs), the proposed system analyzes facial data, obtained either through a camera or OpenCV, to discern both the emotional state and age of an individual. The synergy of emotion and age information serves as the foundation for personalized music recommendations. To compose playlists and propose songs, a third-party API is used, personalizing the musical experience to the user's emotional state and age group. This integration of facial analysis, deep learning, and music recommendation aims to enhance user engagement and satisfaction by providing a more contextually relevant and emotionally resonant music selection.**

Keywords— Emotion-Based Music Recommendation System, Age Estimation, Convolutional Neural Networks (CNN), OpenCV, Deep Learning, Third-Party API.

INTRODUCTION

The convergence of music recommendation systems with face analysis in the age of digital personalization is a big step toward a more complex and customized user experience. Traditional music recommendation systems frequently rely on user preferences, but this study tries to overcome such constraints by including emotional intelligence and age estimate into the computational decision-making process. Our method decodes the expressions and tiny clues inherent in the user's face by using the capability of Convolutional Neural Networks (CNNs) in facial analysis. This analysis not only determines the user's emotional state, but it also calculates his or her age bracket. Understanding the emotional context is critical in developing a music recommendation system that extends beyond genre or artist preferences. By combining age assessment with emotional cues. We attempt to develop a comprehensive user profile that encompasses the many characteristics of musical taste. The suggested system does more than just play songs; it creates dynamic playlists to ensure that the musical selection matches the user's emotional state and age group. The collaborative synergy of facial recognition technology and a third-party music API ensures a diverse and adaptive music recommendation landscape. This study delves into the uncharted territory of combining facial analysis, deep learning, and music curation to enhance user engagement and satisfaction. Through this interdisciplinary approach, we envision a paradigm shift in personalized music recommendations, ushering in an era where digital platforms not only understand what users like but also how they feel and cater to their evolving preferences across different life stages.

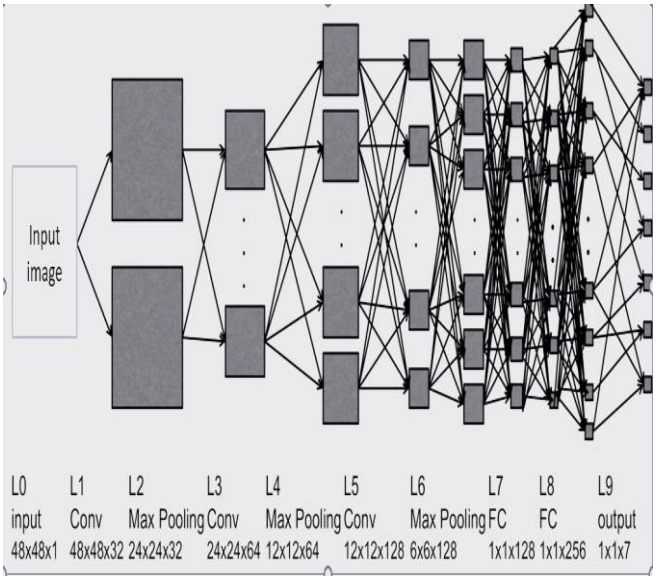

Fig-1 General CNN model

EMOTION DETECTION

Asper our research the paper MUSIC RECOMMENDATION BASED ON FACIAL EXPRESSION BY USING DEEP LEARNING [1], Introduces a comprehensive system comprising distinct modules, including face detection, face prediction, music recommendation, and music playing. In the face detection module, Convolutional Neural Networks (CNNs) and the mediapipe library are employed to identify faces or objects in images or videos. The system captures the user's facial region of interest and forwards it to the face prediction module. The face prediction module utilizes a CNN model that learns features directly from input images and classifies facial expressions into five emotions: happy, angry, sad, surprise, or neutral. This module is implemented using the OS module, numpy, keras, mediapipe, and streamlit packages. The paper concludes that the developed system can capture the user's face, extract relevant features, classify emotions, and generate a music playlist with videos via the YouTube web browser. It claims advantages over existing methods, stating that the system reduces computational time and cost while enhancing accuracy and user experience. Additionally, the paper suggests potential applications in mood alteration and addressing emotional states such as depression and sadness. Overall, the proposed system aims to provide an integrated solution for emotion-based music recommendation and mood enhancement.

The paper Deploying Machine Learning Techniques for Human Emotion Detection [5], presents a real-time emotion detection system for robotic vision, incorporating SRGAN-based image preprocessing, MediaPipe face mesh for 468 facial landmark annotation, FACS-guided selection of 27 key landmarks with subsequent emotional mesh generation, and classification using various machine learning techniques. The evaluation on CK+, JAFFE, and RAF-DB datasets indicates high accuracy and comparative analysis with state-of-the-art methods reveals the proposed approach's superiority, particularly with KNN achieving 97% accuracy, MLP with 94%, SVM with 94%, and LR with 87%, suggesting its potential in diverse robotic vision applications requiring human emotion recognition.

The paper Facial Expression Detection Using CNN with MediaPipe [6] proposes an innovative facial expression detection system utilizing a modified CNN and the MediaPipe framework, encompassing modules for capturing facial images via a web camera, pre-processing with face region detection using MediaPipe, conversion of face images into normalized pixel arrays, and emotion classification with a CNN trained on the FER2013 dataset, achieving a testing accuracy of 66.369%. Additionally, the paper discusses diverse applications in neuromarketing, data-driven animation, sociable robotics, interactive games, and human-computer interaction, while suggesting potential extensions for handling video input sequences, integrating vocal cues, and functioning as a plug-in feature for various applications.

The paper Facial Expression Based Music Recommendation System [8] introduces a novel Facial Expression Based Music Recommendation System, featuring a two-module structure. The facial expression recognition module, employing Python's OpenCV library and an SVM algorithm, demonstrates an impressive accuracy of 90.23% in identifying the user's mood from uploaded images. The music suggestion module, utilizing the recognized emotions, achieves a high classification accuracy of 97.69%, categorizing songs into four distinct mood classes. The proposed system is positioned as a personalized and convenient tool for users to select music aligned with their emotional state, potentially enhancing emotional well-being.

Comparative analysis with existing music recommendation systems emphasizes the proposed system's advantages, boasting superior accuracy, performance, and user satisfaction. Future work suggestions include expanding the range of recognized emotions, incorporating songs from diverse languages and regions, and implementing collaborative filtering to further refine user preferences. This multifaceted approach aims to enhance the system's capabilities and broaden its applicability to cater to diverse user preferences and emotional contexts.

The Paper Human Emotion Detection using Machine Learning Techniques [7] introduces a comprehensive human emotion detection system utilizing machine learning techniques, encompassing face detection through the Viola Jones algorithm, feature extraction employing Local Binary Pattern Histogram (LBPH) and Convolutional Neural Networks (CNN), feature selection via Principal Component Analysis (PCA), and emotion classification using Support Vector Machine (SVM) and K-Nearest Neighbor (KNN). Evaluated on JAFFE and CK+ datasets, the system achieves an accuracy of 93.75% for LBPH and 97.5% for CNN on a dataset of 400 images. Comparative analysis with existing methods demonstrates superior accuracy and F1-score, positioning the system as a promising tool for monitoring emotional health, screening emotion-related physiological and mental diseases, and providing user feedback and support. The paper acknowledges challenges related to occlusion, illumination,

pose, expression variations, and the need for additional data and computational resources.

The Paper Music Recommendation Based on Face Emotion Recognition [9] introduces a music recommendation system predicated on the user's facial emotion, employing a Convolutional Neural Network (CNN) for emotion detection. The system, comprising real-time capture, face recognition, emotion detection, and music recommendation modules, utilizes the FER2013 dataset to train the CNN model and a custom Bollywood Hindi songs database for playlist generation. Face detection, music playback, and GUI development are facilitated through OpenCV, Pygame, and Tkinter libraries respectively. In its conclusion, the paper assesses the system's performance, emphasizing its high accuracy, minimal computational time, and cost-effectiveness in comparison to alternative methods like SVM and ELM. The paper explores potential applications in music therapy and mood enhancement, envisioning broader implications for emotional well-being. Future enhancements are suggested, including the incorporation of more emotions, improvements in light conditions, and heightened camera resolution, demonstrating the paper's commitment to refining the system's capabilities and addressing potential limitations. This work adds valuable insights to the literature on emotion-driven music recommendation systems.

The paper Mood based Music Recommendation System [10] introduces a mood-based music recommendation system, comprised of two integral modules: facial expression recognition and music recommendation. The facial expression recognition component leverages a MobileNet model implemented with Keras to effectively identify seven distinct emotions from the user's facial expressions. Subsequently, the system suggests playlists tailored to match the detected mood, drawing from a Firebase database. An added feature allows users to manually select an emoji, enabling personalized playlist generation based on their preferred mood.

In its conclusion, the paper asserts the system's proficiency in accurately detecting human moods and enhancing user experience by recommending devotional, motivational, and patriotic songs, particularly targeted for negative emotions. Despite reporting an accuracy of approximately 75%, which is comparatively lower than certain existing systems.

AGE DETECTION

The paper Facial Age Estimation Using Machine Learning Techniques: An Overview [2] employs a comprehensive methodology encompassing a survey of diverse automatic age estimation models, including both handcrafted and deep learning-based approaches, with a thorough comparative analysis of their advantages and disadvantages. The authors meticulously examine benchmark datasets used for training and testing age estimation models, delving into their characteristics and associated challenges. Additionally, the paper reviews recent age estimation methods utilizing various techniques such as feature extraction, regression, classification, and generative adversarial networks, providing a detailed account of their performance and evaluation metrics. The authors, in concluding their findings, emphasize the challenging yet captivating nature of age estimation as a machine learning problem, underscoring obstacles related to data scarcity, ageing pattern variability, and external factors like illumination and facial modifications. They advocate for future research directions, including the exploration of new data sources, enhanced data augmentation techniques, and the integration of domain knowledge and human feedback into models. Furthermore, the authors stress the current methods' suboptimal accuracy and advocate for the development of more comprehensive and equitable evaluation metrics and benchmarks.

The Paper Rank consistent ordinal regression for neural networks with application to age estimation [4] proposes a novel method known as the Consistent Rank Logits (CORAL) framework for addressing ordinal regression problems. In this framework, the ordinal regression problem, which involves predicting ordered categories, is transformed into a series of binary classification tasks. Specifically, for a problem with K ranks, CORAL formulates K-1 binary classification problems. To achieve rank consistency among the binary classifiers, the framework minimizes a weighted cross-entropy loss function. Rank consistency ensures that the predicted probabilities are decreasing, and the binary labels exhibit a rank-monotonic pattern. The paper extends the application of the CORAL framework to age estimation from face images by incorporating it into the ResNet-34 Convolutional Neural Network (CNN) architecture. Comparative evaluations are conducted against existing approaches, including the ordinal regression CNN (OR-CNN) and the standard cross-entropy classification CNN (CE-CNN).In the conclusions, the paper highlights the substantial improvements in predictive performance achieved by the CORAL framework across three independent datasets compared to OR-CNN and CE-CNN.

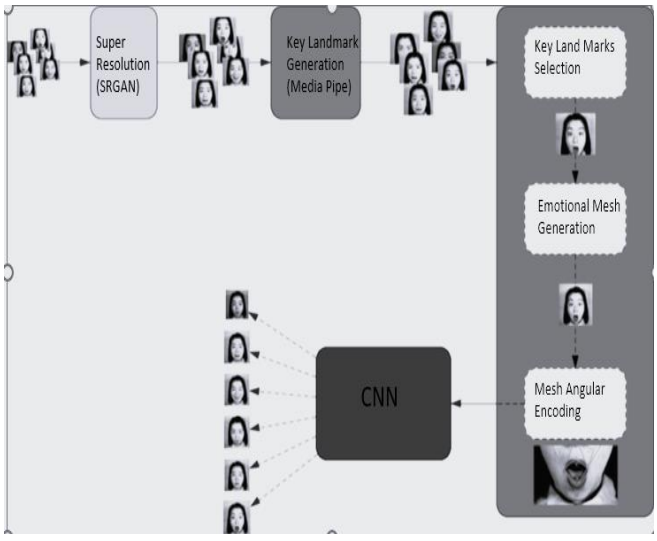

Fig-2 Emotion recognition pre-processing steps

Reference	Dataset	Accuracy (%)
[1]	Fer-2013	80.50
[2]	Imdb-wiki, Hoip	-
[3]	Kaggle-Age detection	-
[4]	Morph, Cacd, Afad	-
[5]	Ck+, Jafee, Raf-db	96
[6]	Fer-2013	66.4
[7]	Lumed, Swell, Yaad	97.5
[8]	Fer-2013	85-90
[9]	Fer-2013	97
[10]	Fer-2013, Mma	75

Fig-3 Datasets used along with the accuracies

CONCLUSION

In this comprehensive literature survey, various machine learning techniques are explored for emotion and age detection. Emotion detection systems, employing CNNs and diverse frameworks, showcase high accuracies in classifying facial expressions, offering potential applications in music recommendation, mood alteration, and emotional well-being. The papers emphasize advantages such as reduced computational time and cost, superior accuracy, and enhanced user experience. Additionally, they propose future enhancements, ranging from expanding recognized emotions to refining system capabilities. On the other hand, age detection models are reviewed, with a focus on addressing challenges like data scarcity and illumination variations. The CORAL framework is introduced for ordinal regression in age estimation, demonstrating significant improvements in predictive performance. Overall, this literature survey provides valuable insights into the evolving landscape of emotion and age detection, highlighting achievements, challenges, and future research directions.

REFERENCES

[1] S. Sunitha, V. Jyothi, P. Ramya, S. Priyanka, "MUSIC RECOMMENDATION BASED ON FACIAL EXPRESSION BY USING DEEP LEARNING“, International Research Journal of Modernization in Engineering Technology and Science(irjmets),Jan-2023.

[2] Khaled ELKarazle, Valliappan Raman and Patrick Then, "Facial Age Estimation Using Machine Learning Techniques: An Overview“, Multidisciplinary Digital Publishing Institue,2022.

[3] Darab Qasimi, Earlham Richmond, "AGE DETECTION USING NEURAL NETWORKS AND MACHINE LEARNING ",ACM Conference,2017.

[4] Wenzhi Caoa, Vahid Mirjalilib, Sebastian Raschkaa, “Rank consistent ordinal regression for neural networks with application to age estimation ”, Pattern Recognition Letters(patrec),2020.

[5] Ali I. Siam , Naglaa F. Soliman , Abeer D. Algarni, Fathi E. Abd El-Samie, and Ahmed Sedik , "Deploying Machine Learning Techniques for Human Emotion Detection ",Computational Intelligence and Neuroscience(Hindawi),2022.

[6] Mrs.V.Lavanya, Mr.K.Azarudeen, Dr.G.Vinoth Chakkaravarthy, Mrs.CB.Selvalakshmi, "Facial Expression Detection Using CNN with MediaPipe ", European Chemical Bulletin, 2023.

[7] Punidha A, Inba. S, Pavithra. K.S., Ameer Shathali. M, Athibarasakthi, "Human Emotion Detection using Machine Learning Techniques ", SSRN Electronic journal,2021.

[8] Abutalib K, Amandeep Gautam, Aryan Khan, Manish Tiwari, Aditya Dayal Tyagi, "Facial Expression Based Music Recommendation System", International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), April-2023.

[9] Madhuri Athavle, Deepali Mudale, Upasana Shrivastav, Megha Gupta, "Music Recommendation Based on Face Emotion Recognition",Journal of Informatics Electrical and Electronics Engineering(ISSN), 2021.

[10] Ankita Mahadik, Shambhavi Milgir, Vijaya Bharathi Jagan, Vaishali Kavathekar, Janvi Patel, “Mood based Music Recommendation System”, International Journal of Engineering Research & Technology (IJERT),2021