Emotion-Sensitive Music Player Leveraging Facial Recognition Technology

SwarnalathaK.S[1] ,ChitteniSai Sanjay[2], C Koushik[3], M Sai Krupananda[4], B M Charan[5]

Professor, Department of Information Science &Engineering, NitteMeenakshi Institute of Technology, Bangalore 560 064, India[1]

Student, Department of Information Science &Engineering, NitteMeenakshi Institute of Technology, Bangalore 560 064, India[2,3,4,5]

Abstract:Facial expression recognition using machine learning and computer vision has become a common practice. Deep learning algorithms are trained on large datasets of labelled facial images, enabling them to identify patterns associated with specific emotions. This has proven effective for facial feature detection and emotion recognition. The rapid advancement of artificial intelligence and machine learning has greatly improved the accuracy of identifying human emotions. Facial recognition technology has emerged as a powerful tool in computer vision, allowing for the identification and analysis of individuals based on their unique facial features. It has applications in attendance tracking, identity validation at ATMs, access control, and even in enabling robots to convey emotions through facial expressions. Additionally, it plays a role in mental illness diagnosis and facilitating human psychological interactions. Algorithms like Face Net, LBPH, SVM, and CNNs are commonly used for image classification and recognition due to their high accuracy. CNNs, in particular, follow a hierarchical model that processes input through interconnected neurons, leading to accurate face expression and recognition results. Overall, facial recognition technology has seen significant advancements thanks to AI and machine learning. It offers a wide range of applications and benefits, but also raises concerns regarding privacy and ethics. With ongoing research and development, facial recognition technology continues to evolve and contribute to various industries.

Keywords:*image, facial, CNN, emotion, VGG16, music, LBPH*

1. INTRODUCTION:

Facial recognition technology has transformed the field of computer vision, revolutionizing the identification and analysis of individuals based on their unique facial features. With advanced algorithms and machine learning, facial recognition systems have gained prominence in various industries. The ability to recognize faces plays a crucial role in identification, and facial recognition technology aims to replicate this ability with remarkable accuracy and efficiency. Recent advancements in processing power, machine learning algorithms, and vast datasets have propelled facial recognition into the mainstream. It has practical applications ranging from security and law enforcement to user authentication and personalized experiences. Face recognition technology is utilized in real-time AI models to search databases and compare faces in scenes or videos. Industries such as attendance tracking, ATM identity validation, and access control benefit from face recognition AI. Robots can also convey emotions through facial expressions, similar to humans. Moreover, it aids in mental illness diagnosis and human psychological interactions. Recognizing facial expressions is a challenging task due to its subjective nature and common challenges in pattern recognition and computer vision. While most advanced algorithms rely on 2D facial data collected in controlled environments, there is a shift towards using 3D facial data for improved recognition results, although it requires more sophisticated equipment and processing methods. Facial expressions are the result of muscle contractions, particularly around the nose, lips, eyelids, eyebrows, and facial skin texture patterns. These expressions typically last a few seconds, with six universal categories: anger, disgust, fear, happiness, sadness, and surprise. Face tracking and recognition of facial expressions using a webcam is a captivating field of computer vision and artificial intelligence. It involves real-time detection, analysis, and recognition of facial features to track movements and identify various expressions. Challenges arise due to variations in position, illumination, occlusion, and individual differences. Human observers adapt quickly to such changes, but automated analysis requires face recognition and normalization of shape and appearance. This paper presents a method utilizing the Supervised Descent Method (SDM) and Rule-Based/SVM classification algorithms for real-time expression recognition in video sequences, achieving excellent results. While implementing a complete system can be complex, libraries like OpenCV and pre-trained models simplify the development process. Intelligent devices facilitating visual interaction between individuals from different cultures rely on facial emotion recognition to interpret feelings. Current research primarily focuses on recognizing emotions based on facial expressions or vocal cues. However, integrating audio and video data simultaneously is seen as a promising approach for future multimedia applications. This paper investigates the effectiveness of combining audio and video for facial expression detection, utilizing deep learning with convolutional neural networks (CNN). The proposed model achieves high accuracy on the Facial Emotion Detection Competition (FERC-2013) and Japanese Female Facial Expression (JAFFE) datasets, surpassing previous models. The study also provides a comprehensive review of facial recognition research, covering technology overview, applications, and feature extraction methods. The findings highlight the potential of integrating audio and visual data for advanced emotion recognition and offer insights into the latest advancements in facial feature extraction.

2. LITERATURE REVIEW:

This research evaluates various algorithms for facial detection and recognition, considering different types of datasets such as white background, shear images, and rotational images. • Face recognition from images or videos is a prominent subject in biometrics research. • Face detection poses challenges in computer vision due to the dynamic nature and visual variability of the human face. Achieving accuracy and speed in identification remains a significant concern. • In the context of video surveillance, several approaches offer comprehensive solutions for image-based face detection and recognition, aiming to improve accuracy and response rates [1].The paper explores various aspects of deep learning, with a specific focus on Convolutional Neural Networks (CNNs). The authors conduct image recognition and detection experiments on the MNIST and CIFAR-10 datasets using only a CPU unit. The accuracy achieved on the MNIST dataset is good, while the accuracy on CIFAR-10 can potentially be improved by training with more epochs and utilizing a GPU unit.CNNs are highlighted as the most commonly used deep neural networks, consisting of convolutional, non-linear, pooling, and fully connected layers. The parameters differ for each layer, with pooling and non-linearity layers being parameter-free. The paper emphasizes the strong performance of CNNs in machine learning problems. The experiments specifically utilize the MNIST dataset for recognizing handwritten digits and the CIFAR-10 dataset for object detection. In the case of the MNIST dataset, the input image is a vector with 784 pixel values. Convolution layers are followed by pooling and fully connected layers, with dropout introduced after each convolutional layer. The pooling layers simplify the output after convolution [2]. The paper discusses the standard K-means algorithm and a more advanced version. In the standard algorithm, k centres are randomly selected, and data objects are assigned to the closest centre based on Euclidean distance. The advanced algorithm introduces two data structures to store cluster labels and distances, allowing data objects to remain in their previous clusters if the distance to the new centre is smaller. Experimental results demonstrate that the improved K-means algorithm achieves faster convergence and improved accuracy compared to the standard K-means. It can produce final cluster results in a shorter time while enhancing the algorithm's overall accuracy [3].The paper presents a framework for real-time emotion recognition using a trained model and the FER2013 dataset. The system incorporates a Haarcascade classifier for face detection and pooling layers for feature extraction. It is capable of recognizing five universal emotions: joyful, sad, neutral, surprise, and angry. The framework is trained with the FER2013 dataset, which consists of around 27,182 images categorizing emotions into five categories. Pre-processing involves pooling layers and activation functions like Elu and Softmax for feature extraction. The trained model is then used for real-time emotion prediction using a webcam. Convolutional Neural Network (CNN) technology is utilized in the model due to its high accuracy in image-related tasks such as detection, identification, recognition, and face recognition. The CNN pre-processes and categorizes input images based on the training data. The model is trained with a learning rate of 0.001 for 450 epochs, using 24,176 images for training and 3,006 images for validation. During testing, live emotion recognition is performed using OpenCV and a Haar-cascade classifier to classify emotions. The proposed model can recognize five emotions: happy, sad, angry, surprise, and neutral [4].This paper presents an efficient face expression identification system that combines deep neural networks and decision trees. The system utilizes the 2-D DCT as a feature extractor to differentiate between neutral and expression images. The proposed technique is evaluated using the CK+ database, containing 123 participants and 70 photos for each expression. The system aim to identify six facial expressions: smile, sadness, disgust, anger, surprise, and fear. A three-node NN-based decision tree achieves remarkable results for the 6-expression FER problem. Additionally, the system incorporates OHL-MLP and DNN nodes to accurately recognize different expressions using features extracted from facial parts [5]. This paper compares two methods for facial landmarks detection and feature extraction: image processing and the Dlib library. A landmarks-based FER system is proposed using face detection, feature extraction from facial regions, and classification with SVM and MLP classifiers. Experimental results show the performance of the Viola-Jones face detector, but limitations exist in multi-face detection and challenging lighting conditions [6].This algorithm analyses segmented image frames using motion vectors to determine facial emotions based on the strongest degree of similarity. It utilizes an AU-coded facial expression database for matching and recognition of facial expressions. Four types of recognition methods are employed, including using emotion space, optical flow, active shape models, and neural networks, to recognize facial expressions [7].

The Facial Action Coding System (FACS), developed by Ekman and Friesen, is a method for analysing facial emotional behaviour and recognizing emotional facial expressions. It relies on human observers to identify subtle variations in facial features. The system utilizes configurable facial models known as Action Units (AUs) or Facial Action Units (FAUs) to represent individual facial actions. FACS consists of 46 action units that correspond to different combinations of facial muscles for each facial expression [8].Deep Learning algorithms, particularly Convolutional Neural Networks (CNNs), have become the main tool in solving various computer vision tasks. One drawback of these algorithms is their reliance on large amounts of data for effective learning. Fortunately, the datasets used in our experiments provide ample data for training, eliminating this concern DL has shown promising performance and is widely used in the field of AI. Our model utilizes CNNs, a specific type of Neural Network architecture that has consistently demonstrated reliable results in image-related tasks. By leveraging CNNs, we are able to build a robust and effective model for our purposes [9].

3. METHODOLOGY:

The flow chart presented in Figure 1 outlines the overall process described in the methodology. The process begins with capturing initial photos using a camera. Each of these images then undergoes face detection using a face detection algorithm, specifically the Haar Cascade classifier, which is known for its ability to identify objects and is applied in this case for detecting faces. After successful face detection, the next step involves face feature extraction. This process aims to extract important facial features, including curved edges and boundaries, which are crucial for training the model accurately. By identifying and extracting these distinctive features, the subsequent steps can make more informed decisions about face recognition and emotion classification. Following feature extraction, a face recognition algorithm is applied to the set of images. This step determines whether a recognized face is present or not by comparing it with the faces that the algorithm has been trained on. By leveraging the training data, the algorithm can accurately recognize familiar faces and differentiate them from unknown individuals. Once a face is detected and recognized, the process proceeds to the emotion classification stage using a Convolutional Neural Network (CNN) model. The choice of a neural network is motivated by its efficiency in classification tasks, particularly when it comes to emotions. The CNN model is trained to classify the detected emotion accurately based on the input image's features. Finally, based on the classified expression, appropriate music is played as a means of representing and detecting emotions. This step connects the classified emotion with the corresponding musical response, creating a sensory experience that aligns with the detected expression.

:

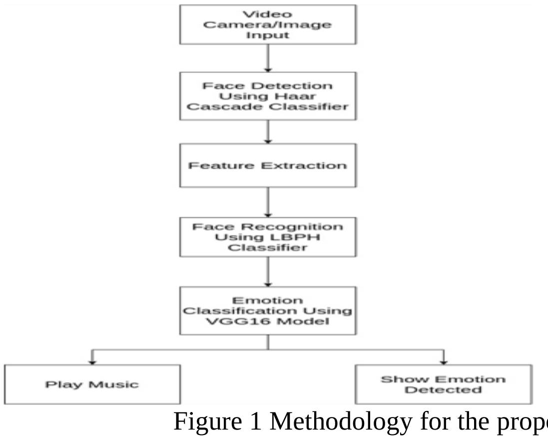

Figure 1 Methodology for the proposed work

3.1 SYSTEM ARCHITECTURE

The system architecture diagram provides an abstract representation of the component architecture within the system. It serves as a helpful tool for understanding the basic interactions between components and the overall functionality of the system. The diagram visually illustrates how the different system components are interconnected and specifies the roles and functions of each component. It offers a high-level overview of the system's structure and provides insights into how the system is designed and organized.The primary purpose of the system architecture diagram is to depict the main system functions and showcase the relationships and dependencies between the various system components. By presenting a visual representation of the system architecture, it enables stakeholders to grasp the overall system design and gain a better understanding of how different components work together to achieve the system's objectives. Figure 2, presented below, represents the specific system architecture for this context. It highlights the stages and components involved in the system, providing a clear illustration of how the system is structured and the flow of information or processes within it. This diagram serves as a valuable reference for understanding the system's design and the various stages or steps involved in its operation.

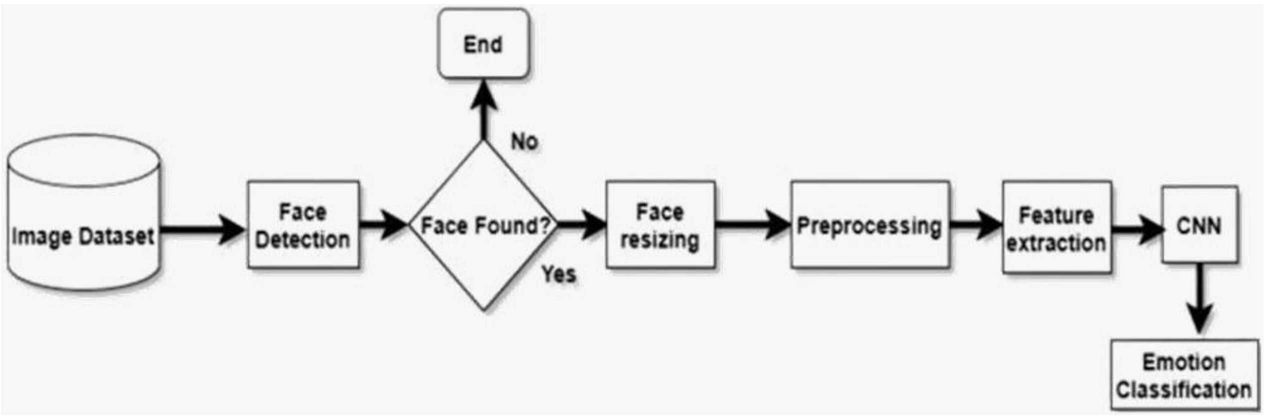

Figure 2 System architecture of the proposed work

- Face Detection: Face detection is a fundamental computer vision task that involves locating and identifying human faces in images or video frames. Various methods and algorithms exist for face detection, with one popular approach being Haar cascades. Haar cascades utilize machine learning techniques to detect faces accurately.
- Face Resizing: Resizing faces can be beneficial for multiple purposes, such as standardizing face sizes for further processing or fitting them into specific proportions. Resizing enables consistency and compatibility in subsequent stages of analysis or application.
- Pre-processing: Pre-processing is a crucial stage in facial recognition tasks within machine learning. It involves transforming and preparing data (in this case, facial photographs) before feeding them into a machine learning model. Pre-processing techniques help enhance the performance and accuracy of facial recognition systems.
- Feature Extraction: Feature extraction is the process of extracting important and discriminative features from facial images. It is a critical stage in recognizing facial features. Machine learning algorithms utilize these features, which represent the unique characteristics of each face, to distinguish and identify different individuals.
- CNN: Convolutional Neural Networks (CNNs) play a crucial role in facial recognition and expression analysis applications. CNNs excel in automatically learning and extracting hierarchical features from raw image inputs. They are well-suited for tasks such as facial recognition, as they can effectively analyze complex visual patterns and capture facial details necessary for accurate recognition and expression analysis.

3.2 FACE DETECTION USING HAAR CASCADE CLASSIFIER:

Face detection using Haar cascade classifier is a popular method in computer vision. It involves training a classifier using positive and negative samples of faces, where the positive samples contain annotated faces and the negative samples contain non-face images. The classifier then scans the input image using a sliding window approach, applying Haar-like features to identify potential face regions based on differences in intensity. These potential regions are then evaluated using a cascade of classifiers, progressively filtering out false positives, resulting in the detection of human faces with high accuracy and efficiency. Below figure shows the Face detection using Haar cascade classifier, the below figure 3 shows the working of Haar cascade classifier.

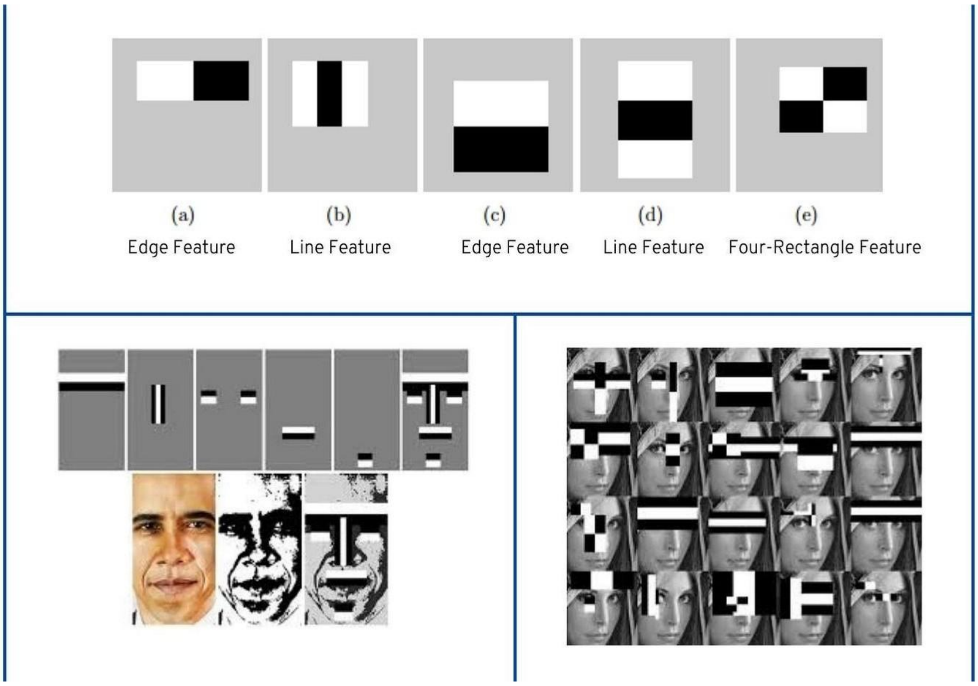

Figure 3 working of Haar Cascade classifier

3.3 FACE RECOGNITION USING LBPH

Face recognition using Local Binary Patterns Histograms (LBPH) is a technique that extracts local texture features from facial images. It divides the face into smaller regions and analyzes the texture patterns within each region. LBPH encodes the relationship between neighboring pixels and generates a histogram representation for each region. These histograms are then compared with the histograms of known individuals during the recognition phase. By capturing unique texture patterns, LBPH enables accurate and robust face recognition, even in varying lighting conditions and facial expressions. The below figure shows the working of LBPH.

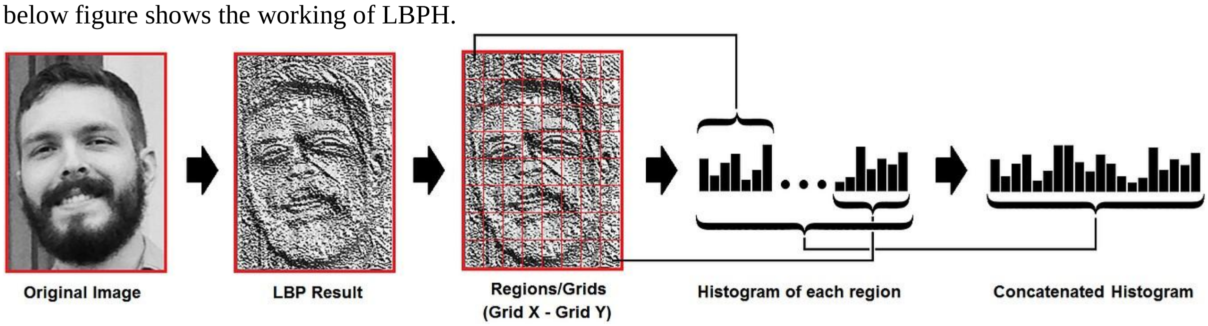

Figure 4 LBPH process

Training Phase in LBPH

- Face Image Collection: Collect diverse face images to create a training dataset covering lighting variations, expressions, and poses.

- Face Detection: Detect and localize faces in the collected images using techniques like Haar cascades or deep learning.
- Image Pre-processing: Normalize, resize, and convert images to grayscale for consistency.
- Feature Extraction: Compute Local Binary Patterns (LBPs) to capture local texture information and generate binary patterns.
- Face Encoding: Encode LBPs into histograms to represent the distribution of patterns and unique facial features.

Recognition Phase:

- Face Detection: Detect and locate faces in the input image using the same face detection techniques.
- Feature Extraction: Compute LBPs for the detected face region, capturing local texture patterns.
- Face Comparison: Compare the LBP histogram of the input face with those in the training dataset using distance metrics.
- Recognition Decision: Assign the input face the label corresponding to the closest match. Apply a threshold to determine reliable identification, classifying as unknown if below the threshold.

3.4 EMOTION CLASSIFICATION

Emotion classification involves recognizing emotions from input data like images using machine learning or deep learning models. For image-based emotion classification, facial features are extracted, such as landmarks or texture patterns. We use a CNN architecture, specifically the VGG16 model, known for extracting hierarchical features. The pre-trained VGG16 model is fine-tuned using the FER2013 dataset, which contains labelled facial images for various expressions as shown in the below figure 5. The CNN model learns to map extracted features to emotions during training. Once trained, the model predicts emotions for new images. Our approach combines CNNs, VGG16 model, and FER2013 dataset for accurate facial expression classification. Emotion classification using VGG-16 with its architecture involves the following steps:

- Architecture Overview: VGG-16 is a deep convolutional neural network (CNN) architecture that consists of 16 layers, including convolutional layers, max-pooling layers, and fully connected layers. It is known for its ability to extract hierarchical features from input images.
- Pre-processing: Input images for emotion classification are pre-processed by resizing them to a fixed size and normalizing the pixel values. This ensures consistency and facilitates efficient processing by the VGG-16 model.
- Feature Extraction: The VGG-16 model is utilized as a feature extractor by removing the fully connected layers. The input image is passed through the convolutional layers, which apply a series of filters to extract relevant features at different spatial resolutions.
- Classification Layers: The extracted features from the convolutional layers are flattened and fed into the fully connected layers of the VGG-16 architecture. These layers learn to map the extracted features to the specific emotions present in the input image.
- Softmax Activation: The final layer of the VGG-16 model uses the softmax activation function to produce probability scores for each emotion class. The scores indicate the likelihood of the input image belonging to each emotion category.
- Training: The VGG-16 model is trained using a labelled dataset of images with their corresponding emotion labels. During training, the model adjusts its internal parameters to minimize the difference between predicted and actual emotion labels.
- Evaluation: The trained VGG-16 model is evaluated on a separate dataset to assess its performance in accurately classifying emotions. Evaluation metrics such as accuracy, precision, recall, and F1 score are commonly used to measure the model's effectiveness.
- Inference: After training and evaluation, the VGG-16 model can be used for emotion classification on new, unseen images. The input image is fed into the model, and the model predicts the most probable emotion based on the learned patterns and features.

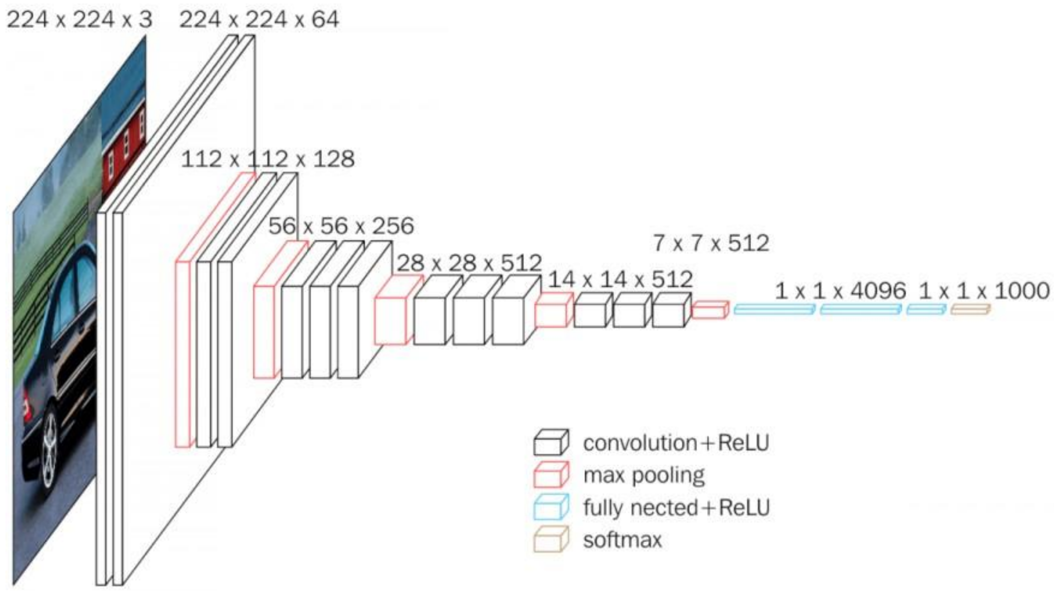

Figure 6 VGG 16 Architecture

4. RESULTS:

The process begins by capturing images using a camera as input. Subsequently, a face detection algorithm, specifically the Haar Cascade classifier, is applied to detect faces in the images. Following face detection, the next step involves face feature extraction, which aims to capture important facial features, including curved edges and boundaries. Accurate feature extraction is crucial for effective model training. After feature extraction, a face recognition algorithm is applied to determine if the detected faces are recognized. This step relies on the training of the algorithm to identify known faces. The following figures provide visual representation of the recognition results. Figure 7 confirms successful recognition of the detected face, whereas Figure 2 indicates that the detected face is unrecognized or unknown.

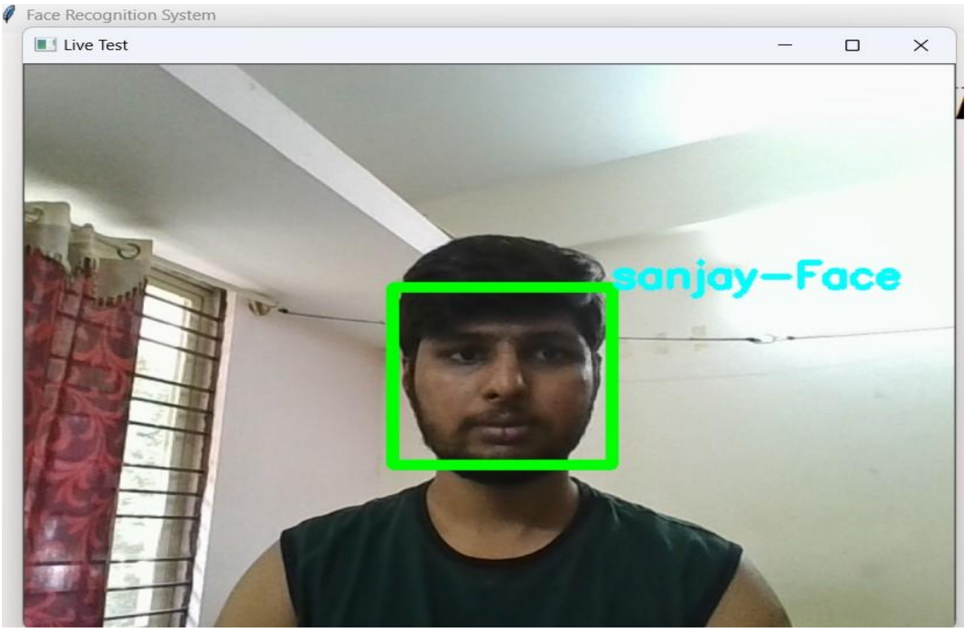

Figure 7 Detected face

Figure 8 Unrecognized face

Our system follows a multi-step process to detect and interpret emotions based on facial expressions. Firstly, the system detects and recognizes the face of an individual. Once the face is successfully detected, a Convolutional Neural Network (CNN) model, a powerful type of neural network, is employed for the classification of the expressed emotion.Neural networks, such as CNNs, have demonstrated remarkable efficiency in tasks related to emotion classification. Leveraging the capabilities of the CNN model, our system accurately classifies the emotion expressed by the individual. This classification enables us to understand the emotional state of the person.Upon determining the classified emotion, our system takes it a step further by playing corresponding music. By associating specific music with different emotions, our system provides audio cues that aid in the detection and interpretation of emotions.To evaluate the performance of our approach, we present the results in Figure 9. The proposed method utilizes the VGG16 model, which is renowned for its effectiveness in image recognition tasks. In our case, we employ the VGG16 model to detect various facial expressions, including happiness, surprise, anger, sadness and neutral. By showcasing the results of our approach, we highlight the successful utilization of the VGG16 model in accurately identifying and classifying different facial expressions, leading to a better understanding of the individual's emotions. We present a real-world scenario that demonstrates the prevalence of a neutral sentiment. At first glance, the dominant sentiment observed may appear to be neutral. However, with the aid of the live probability list displayed alongside the webcam feed, a more nuanced understanding emerges.Upon closer examination, it becomes evident that the second most prominent sentiment, as indicated by the probabilities in the live list, is happiness. This additional information provides valuable insights into the emotional state of the individual being observed. By incorporating the live probability list alongside the webcam feed, our system enables a more comprehensive analysis of the observed sentiments, allowing for a more accurate interpretation of the individual's emotional state in real-time.

Figure 9 Emotion identification

PERFORMANCE MEASURES

The F1 score is a valuable metric when dealing with imbalanced datasets or when the costs of false positives and false negatives vary. While accuracy is useful for symmetric datasets, F1 provides a better evaluation in such cases. Precision measures the confidence in correctly identifying positive instances, while recall focuses on capturing all the true positive instances. Figure 10 represents the accuracy curve, showing how accuracy improves with each epoch during training. Figure 11 displays the loss curve, indicating the decreasing loss as the number of epoch's increases. The achieved overall accuracy is 95. Additionally, figure 12 presents the classification report, including precision, recall, F1-score, and support for each class.

TRAINING ACCURACY:

The training accuracy graph is a visual representation of how the training accuracy of a machine learning model evolves during the training process. It is created by plotting the training accuracy values against the number of epochs or iterations. The x-axis of the graph represents the number of epochs or iterations, which refers to the number of times the model has iterated through the entire training dataset during the training process. The y-axis represents the training accuracy, indicating the percentage of correctly predicted labels on the training data.

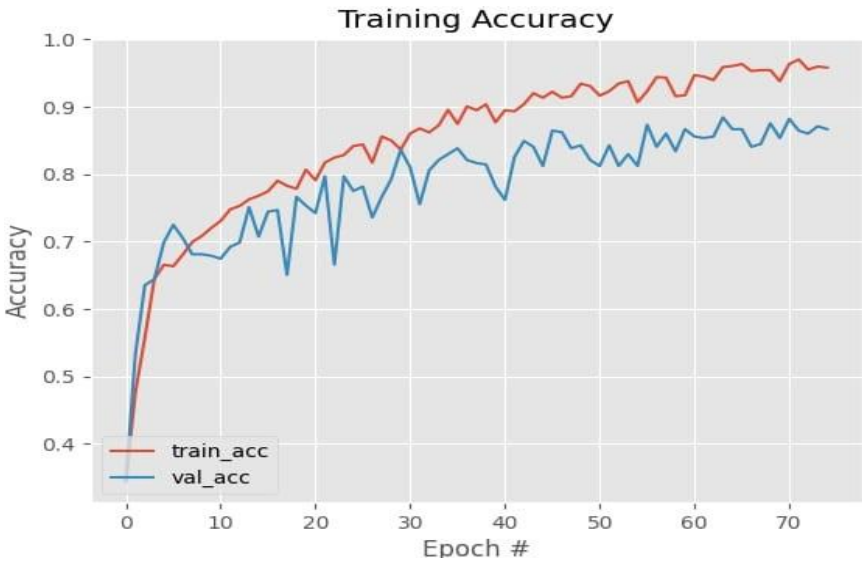

Figure 10 Training accuracy

TRAINING LOSS:

The training loss graph is a visual representation of how the loss or error of a machine learning model changes during the training process. It is created by plotting the training loss values against the number of epochs or iterations. The x-axis of the graph represents the number of epochs or iterations, indicating the number of times the model has iterated through the entire training dataset. The y-axis represents the training loss, which quantifies the dissimilarity between the predicted outputs and the actual labels.

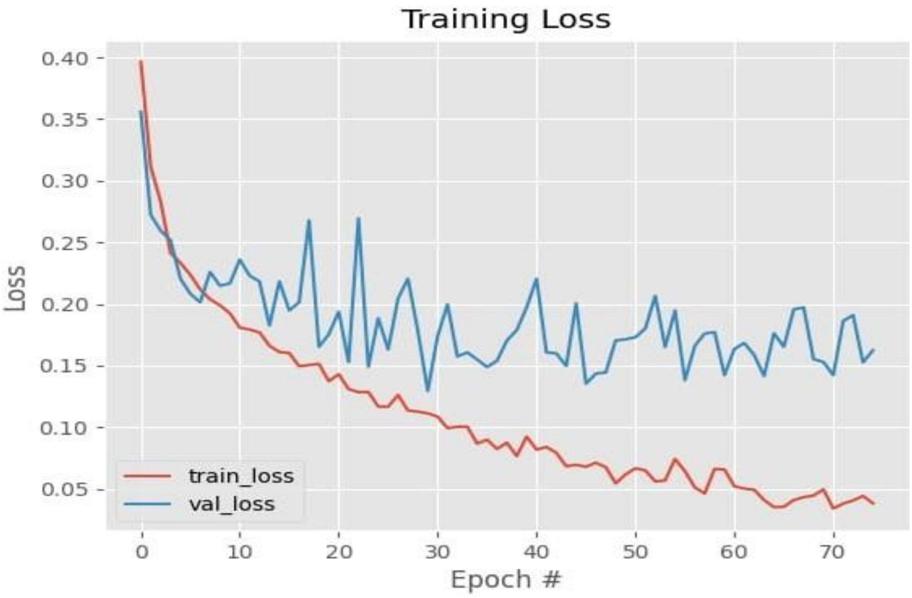

Figure 11 Training loss

The classification of facial expressions report is accompanied by a detailed description below, outlining the key aspects of the analysis. The report is based on the FER2013 dataset, which can be accessed at the following link: https://www.kaggle.com/datasets/msambare/fer2013.

	precision	recall	f1-score	support
angry	0.95	0.09	0.14	67
disgust	0.83	0.80	0.65	95
fear	0.85	0.97	0.90	76
happy	0.89	0.45	0.60	53
neutral	1.00	0.02	0.05	42
sad	0.93	0.05	0.50	43
surprise	0.94	0.25	0.30	53
macro avg	0.92	0.28	0.35	733
weighted avg	0.94	0.22	0.65	733

Figure 12 classification report

CONCLUSION:

The combination of the VGG16 model for facial recognition and the LBPH algorithm for expression-based music player presents an effective system that utilizes computer vision techniques to enhance the user experience. Deep learning, including applications in finance and medicine, has gained worldwide popularity. By using live video as input, this system accurately detects human emotions, which is crucial for communication, interactions, behavioural research, and medical rehabilitation. The non-invasive techniques using facial images offer quick and efficient emotion detection. The integration of neural networks in this experiment achieved a remarkable 95% accuracy in real-time emotion recognition. This integration of the VGG16 model and the LBPH algorithm opens up exciting possibilities for a facial recognition and expression-based music player, delivering personalized and emotionally-aware music experiences.

FUTURE SCOPE:

Advancements in deep learning models and algorithms can enhance facial recognition and emotion analysis, focusing on real-time performance and resource-constrained devices. Future systems can integrate multi-modal analysis, combining facial recognition with other biometric modalities for a comprehensive understanding of user emotions. Facial recognition and expression-based music players have potential applications in therapy, personalized music recommendations, and immersive AR/VR experiences, while addressing ethical and privacy concerns.

REFERENCES:

1. FaizanAhmad,AaimaNajam,ZeeshanAhmeds. ”Image-based Face Detection and Recog- nition.”,IJCSI,2018.)

2. Rahul Chauhan,Kamal Kumar Ghanshala.” Convolutional Neural Network (CNN) for Im- age Detection and Recognition .”, IEEE, 2019.

3. Shi Na,LiuXumin,”Research on K Means Algorithm-Improved K Means Algorithm”, IEEE 2019.

4. Shreya Mishra, Shubham Kumar, ShubhamGautam, JaspreetKour, ”Real Time Expres- sion Detection of Multiple Faces Using Deep Learning”, 2021 Inter- national Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), pp.537-542, 2021.

5. M. R. N. AimiRuzainie, Y. Xiao, T. Uno, L. Ma, K. Khorasani, ”A New Facial Expression Recognition System Using Decision Tree and Deep Neural Networks”, 2021 IEEE 6th International Conference on Signal and Image Processing (ICSIP), pp.49-54, 2021.

6. S. Moore and R. Bowden, ”Local binary patterns for multi-view facial expression recogni- tion”, Computer Vision Image Unders, pp. 541-558, 2011.

7. Pushpaja V. Saudagare, D.S. Chaudhari. ”Facial Expression Recognition us- ing Neural Network .”,International Journal of Soft Computing and Engineering (IJSCE),2019

8. MurugappanM. ,MutawaA.”Facial geometric feature extraction based emotional expres- sion classification using machine learning algorithms”,Feb 18th, 2021.

9. BabakulovBekzodMamatkulovich,MannonovSarabekShuhrato’g’li,BobomurotovJafar-bekJasurjonovich.”SPECIAL DEEP CNN DESIGN FOR FACIAL EXPRESSION CLAS- SIFICATION WITH A SMALL AMOUNT OF DATAs”,ISSN: 2776-0979, Volume 4, Issue 3, Mar., 2023.

10. TheresaKu¨ntzler,T.Tim A.Ho¨fling, Georg W.Alpers,”Automatic Facial Expression Recognition in Standardized and Non-standardized Emotional Expressions”2021.

11. ShrutiJaiswal ,G. C. Nandi, “Robust real-time emotion detection system using CNN architecture,” Springer 2019.

12. Kumar, A.; Kaur, A.; Kumar, M. Review Paper on Face Detection Techniques. IJERT 2020, 8, 32–33.5th International Conference on Electronic Design (ICED) 2020 Journal of Physics: Conference Series 1755 (2021) 012006 IOP Publishing doi:10.1088/1742-6596/1755/1/012006Shiju Varghese. Using standard library packages. In Go Recipes, pages 103–127. Springer.

13. AkritiJaiswal, A. KrishnamaRaju, Suman Deb, “Facial Emotion Detection Using Deep Learning” 2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020

14. LUZ SANTAMARIA GRANADOS , MARIO MUNOZ ORGANERO GUSTAVO RAMIREZ GONZALEZ, ENAS ABDULHAY AND N. ARUNKUMAR “Using Deep Con- volutional Neural Network for Emotion Detection on a Physiological Signals Dataset (AMI- GOS),” 2018 IEEE Access.

15. M. Karnan,R. Vinodini ”Face detection and recognition system based on hybrid statistical, machine learning and nature-based computing”,International Journal of Biometrics, 2022

16. N. Kumar and S. Behal, ”An Improved LBP Blockwise Method for Face Recognition”, International Journal of Natural Computing Research (IJNCR), vol. 7, no. 4, pp. 45-55, 2018.

17. C. Marechal et al., Survey on AI-Based Multimodal Methods for Emotion Detection , in High-Performance Modelling and Simulation for Big Data Applications: Selected Results of the COST Action IC1406 cHiPSet, J. Ko-lodziej et H. Gonzalez-V´elez, E´d. Cham: Springer International Publishing, 2019, p. 307-324.

18. A. Mollahosseini, B. Hasani, et M. H. Mahoor, AffectNet: A Database for Facial Expres- sion, Valence, and Arousal Computing in the Wild, IEEE Trans. Affect. Computer, vol. 10, no 1, p. 18-31, janv. 2019, doi: 10.1109/TAFFC.2017.2740923.

19. D. K. Jain, P. Shamsolmoali, et P. Sehdev, Extended deep neural network for fa- cial emotion recognition , Pattern Recognit. Lett., vol. 120, p. 69-74, avr. 2019, doi: 10.1016/j.patrec.2019.01.008.