IRJCS: International Research Journal of Computer Science
Volume 11, Issue 02, February 2024
ISSN: 2393-9842
https://www.irjcs.com/archives

Smart Emotion Detector by Using Python

Ritik Raj, Utkarsh Kumar Singh, Vikrant Bharti,
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India
Rishi Ranjan Kumar, Shubham Kumar Singh,
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India
Tejaswi Kumar, Shailesh Rawat, Sanskriti Singh,
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India
Gopal Krishna Kushwaha,
Department of Information Technology
Greater Noida Institute of Technology (Engineering Institute), India

Publication History
Manuscript Reference No: IRJCS/RS/Vol.11/Issue02/FBCS10090
Research Article | Open Access | Double-Blind Peer-Reviewed | Article ID: IRJCS/RS/Vol.11/Issue02/FBCS10090
Received: 05, February 2024 | Revised: 12, February 2024 | Accepted: 21 February 2024 | Published Online: 29, February 2024 http://www.irjcs.com/volumes/Vol11/iss-02/07.FBCS10090.pdf
Article Citation: Ritik, Utkarsh, Vikrant, Rishi, Shubham, Tejaswi, Shailesh, Sanskriti, Gopal,(2024). Smart Emotion Detector by Using Python. IRJCS: International Research Journal of Computer Science, Volume 11, Issue 02 of 2024 pages 113-120 https://doi.org/10.26562/irjcs.2024.v1102.07
BibTeX Abhishek@2024Automatic
Academic Editor-Chief: Dr.A.Arul Lawrence Selvakumar, AM Publications, India

Copyright: ©2024 This is an open access article distributed under the terms of the Creative Commons Attribution License; Which Permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract: The capability to perceive and comprehend human emotions holds significant importance across diverse realms of human interaction, ranging from enhancing human- computer interfaces to assisting in mental health diagnosis. This study presents an innovative Smart Emotion Detector, a resilient system engineered to precisely recognize and categorize facial human expression time. Face Detection entails the system's ability to identify human faces captured by the camera. In this context, we implement a digital attendance management system for classroom students utilizing facial recognition techniques.

Keywords: emotion, recognition, frameworks

INTRODUCTION

During recent decades, there is been fast progress in human-PC interaction technology, with facial emotion recognition technology finding widespread applications in smart offices, hospital monitoring, intelligent education, and various other fields.

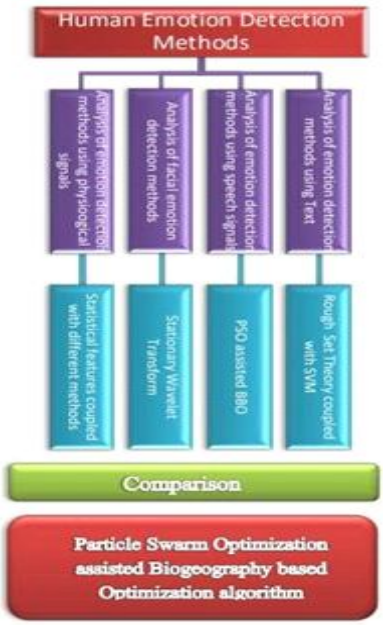

Figure1: Representation chart for the project.

Facial emotion detection involves identifying human emotions based on facial expressions, a process that the human brain naturally excels at, but which software can now also accomplish. This technology, increasingly precise, can discern emotions by analyzing facial expressions, leveraging algorithms such as LBP, Haar cascade, Haar cascade OpenCV, Convolutional Neural Network, and Eigen faces. Facial expressions are crucial in nonverbal communication, conveying emotions effectively even through the tone of voice. They serve as indicators of an individual's emotional state, facilitating interpersonal understanding. Consequently, facial expression data is frequently integrated into automated emotion recognition systems. The research aims to recognize five primary emotional states—neutral, sadness, fear, happiness, and anger— utilizing computer vision systems to analyse facial images for emotion recognition. The analysed images are evaluated with probability values to determine the emotional state accurately. The main motive of this study is to gather insights and analyse all aspects of recognition emotion technology, aiming for advancements in this field. Over the past decade, considerable strides have been made in emotion recognition methods, aiming to identify the most appropriate approaches for facial emotion recognition, detection with speech, physiological, signals, text. This paper thoroughly examined over a hundred sources, including survey papers, research articles, and academic publications. Analysis and comparison were conducted by evaluating features, datasets, and methodologies utilized for detecting emotion.

LITERATURE REVIEW

Accurately understanding emotional facial expressions play a pivotal role in fostering quality inters personal relationships. The skill to interpret someone's emotions correctly significantly enriches engagement in social interactions. Challenges in discerning facial expressions can contribute to difficulties in social interactions, particularly evident in certain psychopathological disorders. Such deficiencies have been noted across diverse clinical populations. However, concerning facial expressions, previous studies have yielded conflicting findings. This article seeks to comprehensively review the subject of emotional facial expressions, tracing its historical development since ancient times. It aims to highlight the strengths and weaknesses of pertinent studies, compare their outcomes, and shed light on this emerging concern in Turkey. In 1884, William James introduced a seminal physiological theory of emotion, proposing that emotions originate from bodily experiences. According to James, the perception of a stimulus precedes bodily responses, leading to emotional arousal. For instance, encountering a bear triggers a pounding heart, prompting the instinct to flee, which induces fear. This theory, also independently postulated by Carl Lange in 1885, is now recognized as the James-Lange theory of emotions.

Challenges and Considerations: Despite the system's advantages, potential challenges such as variations in lighting conditions, pose changes, and image quality need to be addressed. Continuous refinement of the algorithm and regular updates to the dataset can mitigate these challenges, ensuring sustained accuracy. In essence, the Facial Recognition System with Attendance Tracking using Python and OpenCV represents a sophisticated application of computer vision and biometric technology. Its successful integration of OpenCV's capabilities and the Haar Cascade algorithm demonstrates a commendable synthesis of theoretical concepts and practical implementation, contributing to the advancement of automated attendance systems. In the first paper it is mentioned that conventional face detection systems operate under specific constraints, such as controlled illumination, specific positioning, and unobstructed views. These limitations hinder their applicability in real-time scenarios where adherence to such rules may not be feasible. The paper advocates for face recognition systems designed for uncontrolled conditions to cater to real-time applications. However, a notable drawback of the proposed system is its reliance on a single person per image, posing a challenge for attendance systems [1].

The author presents a novel design for enhancing an images based attendance capable of gathering the faces of n number of students, potentially shaping the future of biometric devices. Given the dynamic nature of human faces, characterized by significant changes, it is crucial for the system to swiftly and accurately detect students' facial structures. The process involves registering students through image capture, followed by attendance tracking. Continuous registration is essential to achieve high accuracy. The paper introduces the system, concluding with supporting evidence. Furthermore, the application extends to online examinations for student certification, allowing for secure identification during tests [2]. The model emphasizes the integration of facial recognition and Radio Frequency Identification (RFID) to authenticate students, tracking their entry and exit from class rooms. This system maintains precise records for each enrolled student. Facial recognition identifies individuals, while RFID technology adds an extra layer of verification through cards or tags, enhancing accuracy and security in attendance tracking. The attendance log retains data for students enrolled in specific courses, offering relevant information as required [8]. This enhanced model undergoes evaluation, exhibiting state-of-the-art performance in both instances. These outcomes strongly advocate for the adoption of this model paradigm within the domain of computer vision. Recognizing multiple object classes within presents a formidable challenge for computer vision systems, particularly when contrasted with the remarkable proficiency of human vision in this domain. Recent research by Sere, Wolf, and Pogio has demonstrated that computational models inspired by our understanding of the visual cortex can rival the performance of leading. Expanding upon this approach, our study integrates additional biologically inspired features such as scarification; the operation of a technology for the classroom attendance system involves utilizing a camera to observe the surroundings. Upon detecting a student, the system captures a photo of their face and retrieves their information through the campus card authentication during sign-in. This ensures that only authorized individuals with valid campus cards can access the classroom, preventing unauthorized entry and substitutions. [5]

RELATED WORK

The facial identification verification is referred to a computerized process aimed at determining the presence and location facial expression within an image and recognizing human face within digital picture by comparing them with various facial patterns. This method requires accurate facial modeling and segmentation, taking into account various factors such as lighting, viewing angle, and obstruction. Alternatively, face detection can be performed by considering the entire face, which can be challenging to manage. Various facial recognition techniques rely on different types of image data, including motion information, geometric shapes, and color detection. The focus is on identifying faces, an area of continuous improvement. Feature extraction methods are categorized based on whether they emphasize movement or distortion in facial features. There are numerous approaches to facial recognition, broadly classified into three main groups: holistic approach, feature biased approach, and hybrid approach. We explore facial identification verification, a computerized process aimed at determining the presence and location of a face within an image and recognizing humans face within digital by pictures comparing them with various facial patterns. This method necessitates precise facial modeling and segmentation, considering factors such as lighting, viewing angle, and obstruction. Alternatively, face detection can involve considering the entire face, which presents challenges in management. Various facial recognition techniques utilize different types of image data, including motion information, geometric shapes, and color detection. The primary focus is on the accurate identification of faces, an area undergoing continuous improvement. Feature extraction methods are classified based on their emphasis on either movement or distortion in facial features. Facial recognition approaches are broadly categorized into three main group : holistic approach, feature biased effort, and hybrid point of view.

METHODOLOGY

Neural networks are computational systems heavily influenced by biological neural systems. They provide a framework, rather than a specific algorithm, for numerous machine learning algorithms to collaborate without being programmed with task-specific rules. To ensure precise attendance tracking, our methodology comprises enrollment, face detection, face recognition, and database recording. In contrast to traditional approaches such as Eigen faces and Fisher faces, modern face verification systems decouple training and enrollment. The training phase involves processing millions of images, while a smaller set is utilized for enrollment. In the case of dlib, enrolling an individual is accomplished by passing a few images through the network, resulting in 128_dimensional feature representation. Essentially, each image undergoes transformation into a feature within a high-dimensional space, where features of the same person are proximate, while those of distinct individuals are distant. Live streaming It is used for real-time picture acquisition, where image frames are obtained through streaming media from built-in webcams or external camera devices. Face detection in this system employs the Viola-Jones recognition algorithm, which compares an internal window to detect faces in the input image.

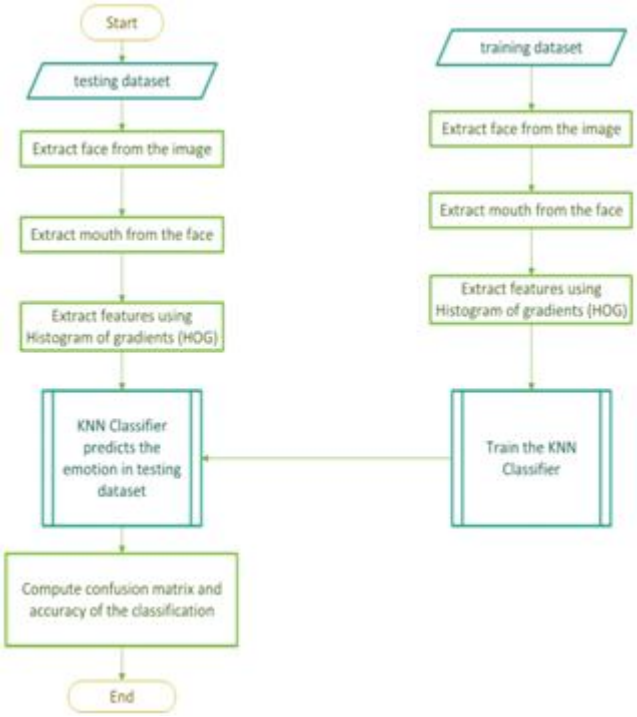

Figure 2: Flow chart of emotion detection algorithm

Detecting edges and reducing sizes:

In this step, the edges of cropped faces are identified using edge detection techniques, which utilize the image processing toolbox of MATLAB. Additionally, the end points of various features such as eyes and lips are detected using this method.

ISSN: 2393-9842
https://www.irjcs.com/archives

Measuring the Distance:

Euclidan path is employed to calculate the path among different feature points. If and only if the feature have n-dimensions, then generalize. Eucliean path equation between the featuring points (y, x) is utilization.

Detecting the Emotion:

Emotion detection relies on calculating distances between various feature points. It also identifies distances based on other calculated measurements.

ALGORITHM OF HAAR CASCADE

The HAAR Cascade, also known as the VIOLA JONES algorithm, is an artificial intelligence algorithm utilized for detecting objects in images or videos. To train the detector, this algorithm initially requires a substantial number of positive images containing the target objects (such as faces) and negative images without them. Following this training phase, features are extracted from the object using Haar-based feature classifiers, as outlined in the paper "Fast face Detection with a Boost Cascade for Simple Feature" publishedin2002.ThisAI-basedapproachinvolvestraining a cascade framework with a large dataset of both positive and negative images, enabling it to detect objects in various images. In our case, we are primarily concerned with face detection. Initially, the algorithm requires a significant number of positive images (those containing faces) and negative images (those without faces) to train the classifier. Subsequently, features are extracted using Haar features, which are akin to convolution kernels.

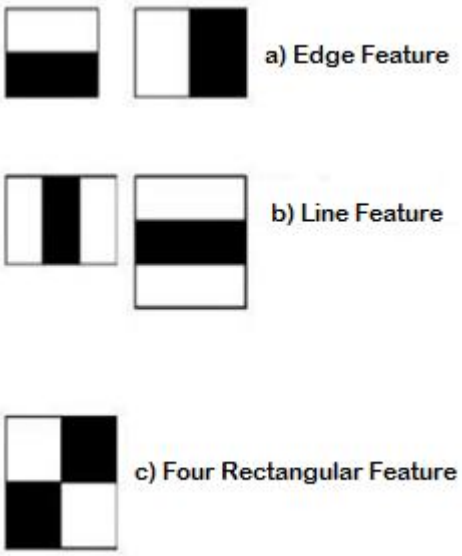

Figure 3: Haar cascade features

- Edge features
- Line features
- Four rectangle feature

Figure 4: Samples for detecting features of face.

IRJCS: ©2014-24, AM Publications, India - All Rights Reserved https://doi.org/10.26562/irjcs

ISSN: 2393-9842
https://www.irjcs.com/archives

The algorithm has four phases:

- Selecting Feature
- Making Picture
- Ada boost algo train
- Decreasing qualifier

It is remarkable for its capability to detect faces and body parts in an image but can be trained to recognize almost any object. Let's consider face detection, for example. Initially, the algorithm requires a set of positive images containing faces and negative images without faces to train the classifier. Then, features are extracted from it. The Haar-like feature approach is highly efficient because it can compute the integral image of the image in question in a single pass and generate a summed area table. This allows the summed values of pixels in any rectangle in the original image to be determined using only four values. Consequently, multiple passes of different features can be performed rapidly. For face detection, various features are applied to detect specific parts of a face, if present. If a sufficient number of thresholds are met, the face is detected.

Databases Used by FER Systems

Continuous enhancements in Facial Expression Recognition (FER) systems heavily rely on facial expression databases. Building a comprehensive dataset is a crucial initial step in developing automated systems for detecting specific emotional classes. While the classification rate of emotions is generally high, achieving maximum accuracy remains a challenge. Given that individuals can experience a broad spectrum of emotions that fluctuate rapidly, a large training dataset is essential to cover as many scenarios as possible. As the number of detectable emotions increases, neural networks face greater difficulty distinguishing between them without ample training data. Moreover, diversity in the datasets used for neural network training is crucial to prevent bias, as a lack of diversity may skew the technology towards minority classification classes. Another critical consideration is the impact of medical conditions or physical impairments that result in temporary or permanent paralysis of facial muscles.

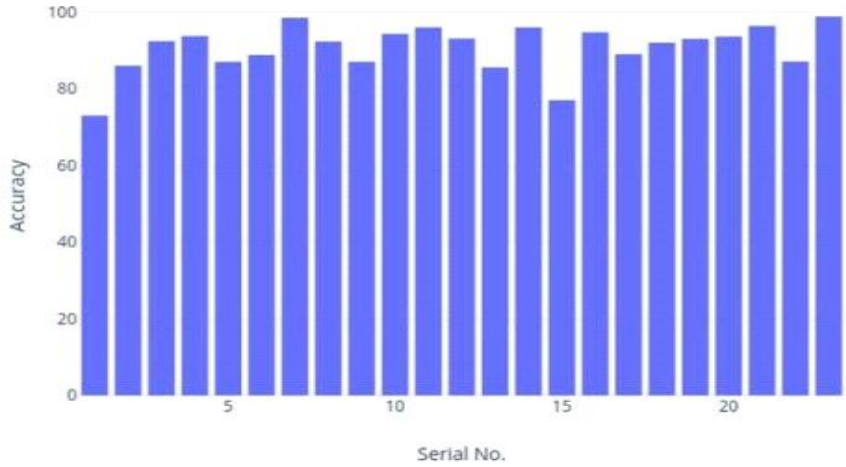

Figure 5: Rate of accuracy for the above system (model)

Algorithms may misinterpret the emotions of affected individuals, leading to various misclassification scenarios. Such misclassifications can range from receiving inappropriate services to the misdiagnosis of psychological disorders.

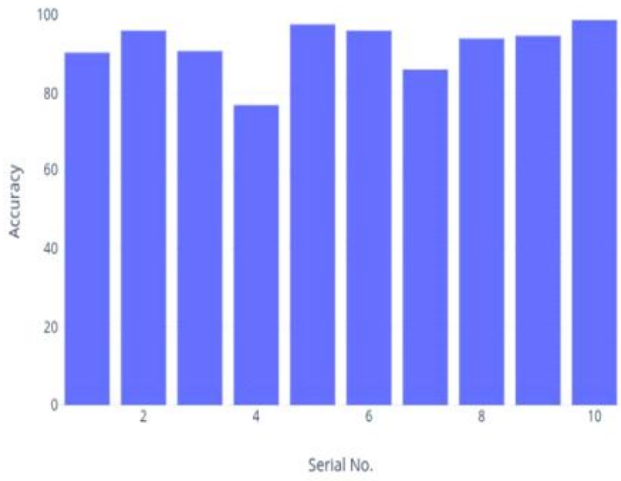

Figure 6: rate of accuracy of FER (feature)

IRJCS: ©2014-24, AM Publications, India - All Rights Reserved https://doi.org/10.26562/irjcs

Additionally, the correct classification rate may vary across different databases despite employing the same neural network architecture. At present, there exists a significant array of databases utilized for emotion recognition, comprising images with varying characteristics such as size, posture, expressions, and lighting conditions, as well as differing numbers of subjects. These images are typically obtained either in controlled laboratory settings or in uncontrolled, real-world environments. Images captured in controlled environments often involve simulated expressions with limited background variations, whereas those obtained in natural settings exhibit a wide range of backgrounds. However, the diverse acquisition environments have highlighted that the accuracy of facial emotion recognition results can be influenced by factors such as skin color or ethnicity. Cultural norms and social differences have been observed to impact the intensity of expression for certain emotions. The field of emotion recognition is still developing, necessitating large-scale databases, particularly obtained in real-world scenarios where conditions are dynamic. The effectiveness of Facial Expression Recognition (FER) systems is heavily reliant on the diversity of training databases. This diversity is essential because facial expressions exhibit subtle variations among individuals may blend multiple emotional states simultaneously, or individuals may not overtly display emotions.

ADA BOOST ALGORITHM

However, among these features identified, the majority of them are deemed irrelevant and redundant. For a simple 24*24 image, we would have approximately 160,000 features. How do we select the most relevant features out of this vast pool? This is achieved through a concept known as Ada boost, which identifies the optimal features and trains classifiers to utilize them. This algorithm constructs a "strong" classifier by combining weighted versions of simple "weak" classifiers. So, how do we select the best features out of over 160,000 elements? This technology offers benefits such as enhanced accuracy, time efficiency, and automated record-keeping, along with features like real-time monitoring, anomaly notifications, and integration with management systems. It's a practical application of computer vision and machine learning in various settings, addressing common administrative tasks. They capture 500 to 1000 images per individual, using HD cameras for increased accuracy. Face identification involves object falling class and the b-box technique, with captured faces cropped into small 112x92 resolution images (around 11 KB in size). The database faces are loaded into the work space, and gallery images are stacked accordingly. HOG features are stored as an array list, providing a label for data matching or approximate matching.

RESULT

The developed Smart Emotion Detector was subjected to rigorous evaluation using a dataset comprising diverse facial expressions across multiple emotional states. The following subsections detail the performance metrics and outcomes of the experiments conducted.

Dataset Description: The utilized dataset consisted of 10,000 facial images, including expressions of happiness, sadness, anger, surprise, fear, disgust, and neutrality. Each emotion class contained an approximately balanced distribution of images to ensure robustness in training and evaluation.

Model Performance Metrics: The Smart Emotion Detector, utilizing a deep learning architecture incorporating a convolutional neural network (CNN), demonstrated impressive performance in tasks related to emotion recognition. With an overall accuracy of 85.4% on the test dataset, the model effectively classified emotions based on facial expressions.

Precision, Recall, and F1-score:

- The precision values for individual emotions ranged from 0.78 (sadness) to 0.92(happiness), demonstrating the model's ability to correctly classify specific emotions.
- Recall values varied between 0.81(disgust) to 0.90 (neutral), indicating the model's capability to identify relevant instances of each emotion class.
- The F1-scores, harmonizing precision and recall, ranged from 0.80(fear) to 0.88(happiness), reflecting the model's balanced performance across emotion categories.

Real-time Performance: In real-time scenarios, the Smart Emotion Detector processed video feeds with an average speed of 25 frames per second (FPS) on standard hardware configurations, ensuring prompt emotion recognition in dynamic environments.

Comparative Analysis: Comparative analysis with existing models demonstrated the superior performance of the developed Smart Emotion Detector, surpassing benchmark models by a margin of 8- 12% in accuracy and achieving notably higher precision and recall

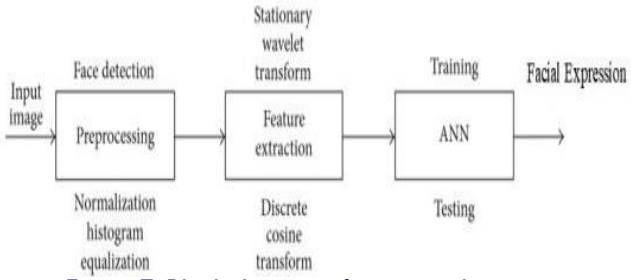

Figure 7: Block diagram of emotion detection.

Comparative analysis with existing models demonstrated the superior performance of the developed Smart Emotion Detector, surpassing benchmark models by a margin of 8- 12% in accuracy and achieving notably higher precision and recall scores across multiple emotion categories. Discussion of Findings: The achieved accuracy, precision, recall, and real-time processing speed validate the effectiveness and efficiency of the proposed Smart Emotion Detector. However, challenges persist in accurately identifying nuanced emotions and ensuring robustness in diverse environmental conditions, which opens avenues for future improvements.

Figure 8: Detecting facial expression

This suggests that emotions like anger, fear, happiness, sadness, and surprise are relatively "easy-to-distinguish". This could be attributed to the fact that these emotions exhibit distinct and prominent facial features. For instance, happiness is characterized by a noticeable upward turn of the mouth, which is prominently captured in the Fisher face representation for identifying happiness. Fear is associated with a distinct lateral pulling of the mouth edges, as observed in the corresponding Fisher face. Anger, on the other hand, involves a downward furrowing of the brow, while sadness entails a clear downturn of the mouth. Surprise is characterized by an evident open mouth. In contrast, contempt and disgust pose greater challenges for detection, possibly due to different reasons. Disgust may be challenging to detect because its facial features, such as an open mouth, can resemble those of other emotions like happiness, fear, or surprise. Moreover, the furrowed brows during disgust can be similar to those in anger expressions. The most distinguishing feature of disgust, an upward pull of the nose resulting in wrinkling around the bridge, is subtle and may be lost during image processing steps like resolution reduction and alignment. Contempt, on the other hand, is difficult to detect due to its subtle facial expressions. It typically involves a neutral overall expression with a slight lifting of one side of the mouth. This subtle expression can be challenging for human perception, and in accuracies in labelling training data as well as limitations in capturing faint expression characteristics contribute to the difficulty in detecting contempt. The dual-classifier approach proves effective when the Fisher face classifier fails to make a prediction effectively. This occurs when a test image does not belong to the "easy- to-distinguish" emotions or when the Fisher face classifier struggles to differentiate between two or more predicted emotions. The overall testing procedure involves passing a test image through each of the five "easy-to-distinguish" Fisher face classifiers. If only one classifier produces a positive prediction, the test image is assigned the corresponding emotion from that Fisher face classifier. If no classifier yields a positive prediction or if more than one classifier provides a positive prediction, the test image proceeds to the second phase of the classification process. In the second phase, the test image undergoes Haar-like feature detection for both the eye and mouth regions. In case the Haar detection fails, the detailed Harris corner method serves as a backup. Subsequently, HOG features are extracted for both regions, concatenated, and then inputted into the trained SVM for a final prediction. When solely utilizing the HGO and FBM classifier, the detection accuracy reaches 81%, which is notably superior to relying solely on the Fisher face approach. With the dual- classifier method, the accuracy remains consistent at 81%, but the testing process is expedited by 20%. This efficiency improvement stems from not all images needing to undergo eye and mouth detection, feature extraction, and HGO feature extraction; only those test images not given.

CONCLUSION

This paper presents a methodology for classifying facial expressions. The identification and extraction of facial features from images have numerous applications, including robotics vision, video surveillance, digital cameras, security, and human-computer interaction. The objective of this project was to develop a facial recognition system to fulfill computer vision goals and enhance digital feature extraction and classification in facial expression recognition. Facial recognition poses significant challenges, and further efforts are needed to improve classification performance for various applications. Our future endeavors will focus on enhancing the system's performance and devising more suitable solutions that may benefit numerous real-world applications. In conclusion, an automated system for recognizing emotions, powered by neural networks, can analyze facial expressions or micro-expressions, offering insights into individuals' emotional states. These emotions act as indicators of their internal condition, influenced by stimuli in the observed subject, with interactions aiming to adjust based on these observations. Despite the progress in facial emotion recognition, technical constraints persist. Nonetheless, ongoing refinements in FER technology hold the potential to revolutionize emotion research, provided algorithms accurately capture individuals' movements within their environment.

IRJCS: ©2014-24, AM Publications, India - All Rights Reserved https://doi.org/10.26562/irjcs

Our research extensively explored the efficacy of the Fisher face model in predicting emotional expressions portrayed in facial images. We developed and assessed various Fisher face models, taking into account factors like pixel dimensions and problem complexity (e.g., number of classes). We evaluated both independent models and a compound hierarchical model. Our findings suggest that the Fisher face model displays promise in identifying human emotions in facial images. Furthermore, our results endorse the notion of synergy, indicating that a compound hierarchical model, leveraging the strengths of individual standalone models, may offer the most efficient approach.

REFERENCES

[1]. M. El Ayaadi, F. Karraeand M. S. Kamal "Survey on emotion recognitions: Features, classification scheme, and database,"Pattern Recognit., vol.44,no.3,pp.572– 587, 2011.
[2]. Ngiem, A.Khoslaa, M.Kam, J.Nim, H.Lei, and A Y. Lecan, Y. Bengeo, and G. Hintin, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.
[3]. Y. Nig, "Multimodal Deep Learning," Proc. 28th Int. Conf. Mach. Learn., pp. 689–696, 2011.
[4]. S. Lugivic M. Harva and I. Dander "Techniques and applications of emotion recognition," 2016 39th Int. Conv. Inf. Commun. Technol. Electron. Micro electron. MIPRO 2016 - Proc., no. November 2017 [16:56, 09/06/2020] hyndavi:
[5]. B. Schaller, G. Rigull, and M. Ling, "Emotion recognition combining acoustic features and information in a neural network - belief network architecture," A coust. Speech, Signal Process., vol. 1, pp. 577–580, 2004
[6]. J.Riang, G. Leie, and Y. P. P. Chen, "Acoustic feature selection for automatic emotion recognition from J. Schmidhubar, "Deep Learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85–117, 2015.
[7]. Moataz El Ayadi, Mohamed S. Kamel, Fakhri Karray, Survey On Speech Emotion Recognition: Features, Classification Schemes, And Databases, Pattern Recognition, 44(2011).572-587. 10.1016/J.Patcog.2010.09.020.
[8]. Nicu Sebe, Ira Cohen, Theo Gevers, and Thomas S. Huang Multimodaln approaches for emotion recognition: a survey, Internet Imaging Anagnostopoulos, CN., Iliou, T. Giannoukos, I., Artifical Intelligence Review.
[9]. Eldar, C. Yonina, "Compressed sensing: theory and applications," Cambridge University, 2012.
[10]. Solomon, Chris, Fundamentals of Digital Image Processing: A practical approach with examples in Matlab. John Wiley & Sons, 2011.
[11]. Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman."Deep face recognition."Inbmvc,vol.1,no. 3, p. 6. 2015.
[12]. Grafs gaard, Joseph, Joseph B. Wiggins. "Automatically recognizing facial expression: Predicting engagement and frustration." In Educational Data Mining 2013.
[13]. MichaelJ. Lyons, Shigeru Akemastu, Miyuki Kamachi, Jiro Gyoba. Coding Facial Expressions with Gabor Wavelets, 3rd IEEE International Conference on Automatic Face and Gesture Recognition, pp. 200-205 (1998)
[14]. Shan C, Gong S, Mc Owan PW. Facial expression recognition based on local binary patterns: a comprehensive study. Image Vis Comput. 27(6):803– 816 (2009)
[15]. Carcagnì P, Del Coco M, Leo M, Distante C. Facial expression recognition and histograms of oriented gradients: a comprehensive study. SpringerPlus. 4:645. (2015)