FACESPOT: A SMART HUMAN FACIAL DETECTION SYSTEM USING SUPERVISED LEARNING

Jaishree Jain[1], Saroj Kushwah[1], Surbhi Saroha[2], Raghuraj Singh[3]
[1]Department of Computer Science & Engineering, Ajay Kumar Garg Engineering College Ghaziabad
[2] School of Computational Sciences & Engineering, Shobhit University, Meerut
[3] Department of Computer Science & Engineering, Galgotias College of Engineering & Technology, Greater Noida, India
jaishree112@gmail.com, sarojkushwahsiem@gmail.com, indiasurbhisaroha12@gmail.com, raghurajnitj4@gmail.com

Abstract— **In current research discussions, facial expression recognition, or FER, receives a lot of attention. A relative examination of face expression recognition systems is presented in the paper. Three steps are typically involved in facial expression recognition: face detection, feature extraction, and expression categorization. A face detection and extraction technique based on Haar cascade features is the main focus of the proposed study. Numerous images, both positive and negative, are used to train the classifier. It is used to extract the features. Convolutional kernels are utilized for conventional Haar feature pictures. Then, using the COHN-KANADE database, the supervised learning technique Fisher face classifier is used to create a facial expression classification system with eight classes (seven basic emotions and neutral). A 65% recognition rate in COHN-KANADE database is achieved.**

Index Terms— **Convolutional Neural Networks (CNNs), OpenCV, Machine Learning, Deep Learning, Emotion Recognition, and Facial Detection**

I. INTRODUCTION

FER plays a significant role in the interpersonal growth of partnerships. By analyzing facial expressions in the social setting, one can predict the intents or circumstances and react accordingly. Human facial expressions can convey emotions and sentiments that regulate and observe one another's conduct. Humans and robots may now communicate thanks to rapidly advancing technology, which allows robots to be controlled by facial expressions. Research on the FER of humans by machines is still in its infancy. The main problems with FER are as follows: variations in expressions, quick processing, and product-specific applications. Due to its wide range of applications, FER has currently emerged as a crucial area of study. Efforts are being made to improve FER accuracy by utilizing a facial expression database that spans a wide age range. In general, there are Disgust, sadness, fear, surprise, happiness, anger, contempt, and neutral are the eight fundamental expressions. Every society and culture uses these expressions. The actions listed below are required for any FER system: (A) Face Recognition (B) Extraction of Features (C) Categorization of Facial Expressions. The authors used a variety of feature extraction techniques to conduct a thorough investigation into FER techniques. A number of face and FER procedures have been proposed. Each of them, though, has unique benefits and drawbacks. A few examples of holistic approaches are Linear Discriminant Analysis (LDA) and the Eigenfaces and Fisher face algorithms, which are constructed using the most recent 2D PCA holistic methods. Numerous authors have examined these techniques, but local descriptors have drawn interest due to their resilience to lighting changes and posed variations. In the sections that follow, some of the alternative strategies are covered in more detail. In the study [1], the authors talked about Automated Facial Expression with salient patch features. Six universal expressions for facial expression recognition have been provided in this research. The authors of [2] talked about FER, which is concerned with aging facial expressions. Here, facial features are extracted using the Gabor and log Gabor filters. When compared to the Gabor filter, the log Gabor filter produces better results. However, the log Gabor filter takes longer to process than the Gabor filter. The precision with which the Gabor filter or log Gabor filter with SVM is implemented classifier outperforms natural human prediction on both synthetic and standard databases. Using two databases, including lifetime and faces gathered from the psychology community, the authors of [3] talked about facial expression recognition, which is related to aging.

They provide thorough explanations of the causes of aging and how they affect FER. The authors of [4] talked about feed forward neural networks in relation to FER. For recognition testing, the Cohn Kanade dataset is utilized. Compared to traditional linear PCA and histogram equalization, it has provided a higher recognition rate. The most recent face recognition algorithms' operation was discussed by the authors in [5]. The results showed that the most recent algorithms outperform the facial recognition algorithms by a factor of ten. a hundred times better than those of 1995 in the year 2002. When it came to face recognition, some of the algorithms could leave human participants behind.

Additionally, even identical twins may be recognized by these algorithms. Principal Component Analysis and Wavelet Packet Decomposition, which allow the use of PCA-based face recognition involving a large number of training images and are faster training than PCA method, were discussed by the authors in [6] in relation to Vytautas Perlibakas (2004). The authors of [7] have presented a classification and face recognition algorithm that relies on linear subspace projection. Using dimensionality and adjacent component analysis (NCA), the subspace is created.

The authors of [8] talked about the Fisher Linear Discriminant (FLD) feature fusion approach. Two-Dimensional PCA (2DPCA) and Gabor Wavelets are used in the technique to extract features, and FLD is then used to fuse the features. According to a survey of the literature, standard classification methods for facial expression recognition were employed in the 1980s [9]. Neural network classifiers have been used in recent years to extract features such as the mouth, eyes, and so on [10].

This is how the paper is structured. The suggested methodology is examined in Section II, the comparison of the various approaches is highlighted in Section III, and the results and their analysis are explained in Section IV. Section V concludes the paper.

II. LITERATURE REVIEW

Generally speaking, there is now much more regard for experts in the field of facial recognition. As the violation grows more serious, the legal observation is necessary to prove a facial quality, a visual similarity of a casualty's face. Many professionals employ a variety of methods [11]. A reasonable framework exhibition is produced, but ongoing face exploration remains a tedious task. We are discussing two facial recognition models [12]. By suggesting a rectangle element-based classifier and a robust identification calculation that satisfied computation and location execution effectiveness, the continuous face district was found. In this research, the face acknowledgment approach combined with PCA and the identified face region as an acknowledgment input picture One of the clever groups, the multi-layer network, was suggested, and its demonstration was evaluated [13].

This method uses PCA to register the eigen face and sends the preparation images to it as a basic vector as part of the preprocessing calculation of the info face picture [14]. Each image uses the principal vector's load arrangement as an element vector, which simultaneously reduces the image's element. The multi-layer neural organization is then used to do face recognition [15]. Two types of typical strategies are first and second. Location and execution were proposed. One of the clever groups, the face recognition approach using PCA and multilayer network, was developed in this paper employing the recognized face area as an acknowledgement [16]. The input image and its display were assessed. This method communicates the preparation photos with the info face image as a simple vector and uses PCA to register the own face as a pretreatment measurement. Each graphic shows how the load is arranged [17].

III. METHODOLOGY

To recognize a person's face from video frames or still photos using a database of face images. The user's face reveals a great deal about them, and they are frequently more sensitive to environmental changes. Face recognition has seen a lot of studies, but its accuracy rate is lower than that of other biometric information about an individual, such as fingerprints, facial expressions, eyes, palm geometry, retina, etc. Fig. 1 illustrates the face recognition process.

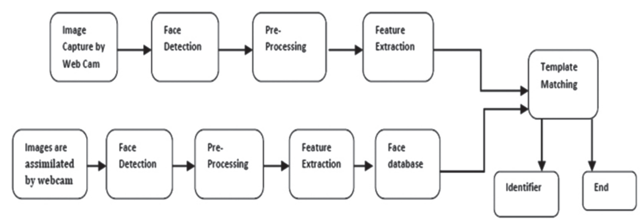


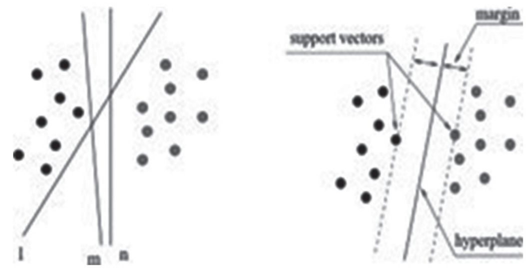
Fig. 1. Face Recognition Process

There are two stages to the system. The training phase is when machine learning models are learned, and the testing phase is when trained models are tested and their performances are assessed. First, the system environment is set up and all required libraries are installed. skLearn for machine learning algorithm instances, CNN (Convolutional Neutral Network) for image classification tasks, NumPy for numerical computation, pandas for dataset creation and data operations, and seaborn for data visualization are the libraries required for the system. Open-source data is employed since data gathering is completed after the environment is created. Since the data is in picture format, it is extracted and used in the following section of the system. Following that, the image is sent to the media channel for facial feature estimation, which identifies the 33 important spots inside a specified frame and assigns a visibility value to each of these points' 3-D coordinates. These 33 crucial points serve as the foundation for a new dataset, which include each critical point's coordinate and visibility as a feature. The data is then cleaned and prepared for a machine learning model using data pre-processing techniques. The data set undergoes normalization as part of the pre-processing step. Making sure that every value in a data collection fall between 0 and 1 is known as normalization. This is primarily because certain machine learning models require data that has been standardized to work effectively and produce positive outcomes. Next, new features are extracted from the dataset's current features using feature engineering. We did this by turning the key points into vectors. These vectors are used to compute the feature measurements, and this one depicts a 3-D face feature. All processed data is then sent to machine learning algorithms that are based on classification for training. First, test data needs to be created or obtained. This data must be brand-new and shouldn't include training-phase duplicates. We attempted to compare different classifiers for face feature classification. Ridge Classifier, Random Forest Classifier, Gradient Boosting Classifier, Logistic Regression Classifier, and KNN Classifier are some of these classifiers. These classification techniques have helped to provide the greatest results with more precision. The dataset was gathered through web scraping. The goal was to gather the photos in the most effective manner possible. Fig.2 displays feature mapping for facial detection.

Fig. 2. Feature Mapping for Facial Detection

An algorithm for supervised learning is the support vector machine. Although it has been expanded to be multiclass, it is now a two-class classifier. Regression is another application for it. Vectors of support: The data points nearest to the hyperplane are called support vectors, and they regulate the hyperplane's position and orientation. We can increase the classifier's margin by using these support vectors, and the Hyperplane's position will shift if these support vectors are deleted. In actuality, these are the points that aid in the construction of the SVM. The Support Vectors and the Hyperplane are equally spaced. Because the Hyperplane also shifts if their position does, they are known as support vectors. Accordingly, the Hyperplane is solely dependent on the support vectors and not on any more observations. (Left) Hyperplanes at random. (All right) Fig. 3 displays the optimal separating hyperplane.

In the demonstration of face recognition networks, a semi-supervised learning approach utilizing support vector machines for face recognition is proposed. The recognition mechanism is easy to use and efficient. Therefore, it can be claimed that if distinctive facial traits are added in the right



Fig, 3. (Left) Arbitrary hyperplanes.
(Right) Optimal Separating Hyperplane.

amounts, an image of the original face can be recreated from the native interface. As seen in Fig. 4, each face only displays specific facial traits that might not be seen in the original image.

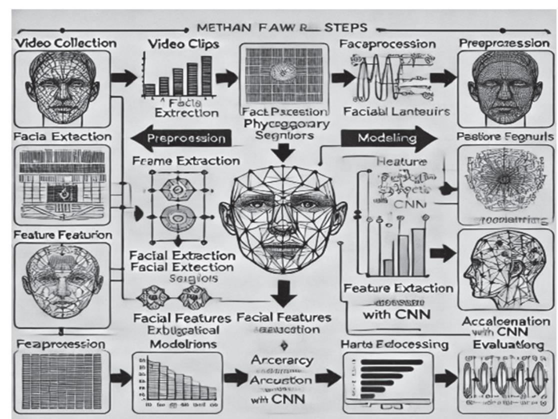


Fig.4. A detailed Data Flow Diagram (DFD) illustrating the process of a Human Facial Recognition System components

Although a little lower, the recognition rate of 95% shows that the model's mechanism can handle these variations; however, further improvement is still necessary to enhance performance across a variety of facial expressions. Additionally, different scenarios were used to test response times. The results demonstrated that, as illustrated in Fig. 5, face recognition performed with 0.5 seconds/frame and standalone weapon detection was able to function at an impressive 0.3 seconds/frame.

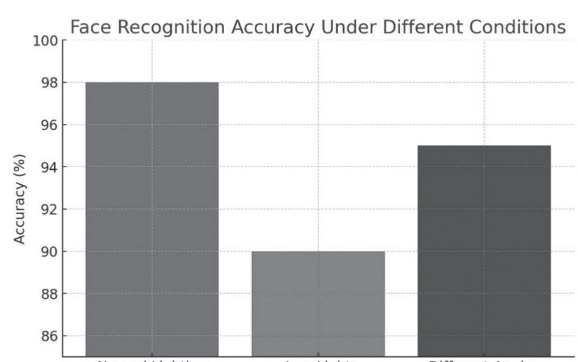


Fig. 5. Graphical Analysis

IV. RESULTS OF FACE RECOGNITION USING CNN

Using a vast database, this technique enables the computer to comprehend the photos by identifying the visual components and identifying any trends. If the similarity is less than the threshold, it may also return "no match." However, the accuracy decreased to 90% in dim light. Less-lit settings that are better suited for infrared cameras or image enhancement techniques may present challenges for the model. Additionally, the devices, IoT systems, and smart cameras were brought about by the angles at which the faces were presented.

TABLE I
Face Recognition Results Under Various Test Scenarios

Test Scenario	Face Recognition Accuracy (%)
Normal Lighting	88% - 98%
Low Lighting	90%
Dynamic Angle	95%

TABLE II
System Response Time under Different Test Cases

Test Cases	Response Time (Sec)
Weapon Detection Only	0.3 sec
Face Recognition Only	0.5 sec
Full System (Both Tests)	0.8 sec

When incorporated into a second system, the system as a whole processed data in 0.8 seconds, guaranteeing prompt alerts for unknown people or possible dangers. The outcomes, as shown in Tables I and II, indicate the system's effectiveness for real-time monitoring, making it a practical option for visually impaired people in dynamic and potentially dangerous situations.

V. CONCLUSION AND FUTURE SCOPE

Using OpenCV and Python, this study effectively created a real-time emotion recognition system that incorporates deep learning approaches to improve speed and accuracy. We can be certain that everything is centered around automation and technology as we enter the Fourth Industrial Revolution. The industry that needs technical advancements is changing, as we can see. In addition to altering how businesses function, these advancements also altered how we live our daily lives. Strong data security measures to guarantee that no criminals can access private data to support forensic analysis. In this project, we have tried a variety of approaches and every one of those face recognition techniques has been successful. Face recognition is the foundation of the face recognition system. The unknown individuals can be identified using this system. This saves time and money and is easy to use. using the algorithm for machine learning. Anyone can customize the system to meet their needs, and it is more dependable and straightforward.

Future work includes multimodal data integration, combining speech emotion recognition, body language, and physiological signals for the more accurate analysis of emotions. Advanced deep learning models such as Vision Transformers (ViTs) and Efficient Net can further increase accuracy. The system will be optimized for low-power edge devices, which enables real-time applications in mobile

Acknowledgment
I gratefully acknowledge the contributions of all the participants who have agreed to participate and have spent time providing their data. It has only been through such cooperation that the present study has been possible to complete, as the findings could only be confirmed through this effort.

I would like to extend my heartfelt gratitude to my institution, whose resources, infrastructure, and academic environment have provided a strong foundation for this research. I am particularly thankful to the esteemed faculty members and my fellow peers for their insightful discussions, constructive suggestions, and continuous encouragement. Their valuable perspectives have greatly contributed to refining and enhancing various aspects of this work.
I am extremely thankful to my family and friends, who have constantly supported, understood, and motivated me throughout the process. The unshakeable belief they had in me made me more powerful during all the challenges of this research endeavor.

REFERENCES

P. Ekman and W. V. Friesen," Facial Action Coding System (FACS): A Technique for the Measurement of Facial Movement," Consulting Psychologists Press, 1978.

A. Mollahosseini, B. Hasani, and M. H. Mahoor," AffectNet: A Database for Facial Expression, Valence, and Arousal Computing in the Wild," *IEEE Transactions on Affective Computing*, vol. 10, no. 1, pp. 18–31, 2016.

B. Schuller, S. Steidl, A. Batliner, et al.," The INTERSPEECH 2018 Computational Paralinguistics Challenge: Atypical & Self-Assessed Af- fect, Crying & Heart Beats," *Proceedings of INTERSPEECH 2018*, pp. 122–126, 2018.

G. Chanel, J. J. M. Kierkels, M. Soleymani, and T. Pun," Short-term emotion assessment in a recall paradigm," *International Journal of Human-Computer Studies*, vol. 67, no. 8, pp. 607–627, 2009.

G. Bradski," The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

D. E. King," Dlib-ml: A Machine Learning Toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

P. Viola and M. Jones," Rapid Object Detection Using a Boosted Cascade of Simple Features," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern*

Recognition (CVPR), vol. 1, pp. 511–518, 2001.

K. Simonyan and A. Zisserman," Very Deep Convolutional Networks for Large-Scale Image Recognition," *International Conference on Learning Representations (ICLR)*, 2015.

F. Chollet," Xception: Deep Learning with Depthwise Separable Convo- lutions," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1251–1258, 2017.

M. Abadi, P. Barham, J. Chen, et al.," TensorFlow: A System for Large- Scale Machine Learning," *Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation (OSDI)*, pp. 265–283, 2016.

C. M. Bishop," Pattern Recognition and Machine Learning," Springer, 2006.

S. Datta and A. Sharma," Ethical Concerns in Facial Recognition Technology," *Journal of Ethics and Technology*, vol. 29, no. 3, pp. 195–204, 2017.

P. Singh, M. Kansal, R. Singh, S. Kumar, and C. Sen," A Hybrid Approach based on Haar Cascade, Softmax, and CNN for Human Face Recognition: A HYBRID APPROACH FOR HUMAN FACE RECOGNITION," *Journal of Scientific & Industrial Research (JSIR)*, vol. 83, no. 4, pp. 414–423, 2024.

A. Bisaria, D. Chauhan, H. Varshney, S. Kumar, U. Rastogi, and V. Sharma," Human Detection to Mitigate Excessive Electrical Energy Usage in Communal Spaces," in *Proceedings of the 2024 International Conference on Intelligent Systems for Cybersecurity (ISCS)*, 2024.

S. Singh, S. Kumar, and B. K. Tripathi," A Comprehensive Analy- sis of Quaternion Deep Neural Networks: Architectures, Applications, Challenges, and Future Scope," *Archives of Computational Methods in Engineering*, pp. 1–28, 2024.

J. Jain, R. Tyagi and R. Verma," A Review on Distinctive Hybrid Cloud Mode", GLIMPSE -Journal. of Computer Science • Vol. 10, No. 2, Jul-Dec 2019.

J. Jain, N.Sagar and J. Chauhan, "Brain Tumor Detection Model Based CNN and Threshold Segmentation", GLIMPSE -Journal of Computer Science •Vol.2(2), JULY-DECEMBER 2023, pp.25-30.

ABOUT THE AUTHORS

Dr. Jaishree Jain, Associate Professor in Computer Science & Engineering Department, Ajay Kumar Garg Engineering College, Ghaziabad, affiliated to AKTU, Lucknow, Uttar Pradesh, INDIA. She has 14 years teaching experience in CSE & IT Departments that include Chandigarh University, Punjab and college of AKTU, University, Lucknow since July 2010. She had done her B.Tech. in Information Technology from Uttar Pradesh Technical University, Lucknow. She had done her M.Tech. in Software Engineering from MNNIT, Allahabad. She had qualified GATE two times in 2007 & 2008 with 94.86 percentile. She had done her Ph.D. from TMU, Moradabad in 2023. She had Published more than 50 research papers in SCI/SCOPUS/ UGC-Care/ UGC/National and International Journals, Conferences and 5 Book chapters published in International Book.

Dr. Saroj Kushwah currently working as an Assistant Professor in the Computer Science and Engineering Department at AKGEC, Ghaziabad, UP, India. She received her Ph.D degree in "Machine Learning" in 2023. She received her Bachelor's degree in Information Technology in 2010-11, and her Master's in Computer Science and Engineering in 2013-14. He has 4.5 years of academic and research experience. Her research interests include Data Science, Big Data Analytics, Artificial Intelligence, Machine Learning, Cyber Security, Wireless Sensor Network, and Algorithm Design. She has attended several International and National conference and workshop as well as published many research papers and book Chapters in various reputed international Journals and Conferences so far.

Dr. Surbhi Saroha is an accomplished Associate Professor in the School of Computational Sciences & Engineering at Shobhit University, Meerut, bringing nearly a decade of academic and professional experience. She has previously served in esteemed institutions such as Meerut College, Radha Govind Group of Institutions, Subharti University, and Vidya Knowledge Park, and held the position of Head of the Computer Science Department at Karan Public School. . She has taught a wide range of subjects including Artificial Intelligence, Cloud Computing, Java, DBMS, Network Security, and more. Her research includes multiple Scopus-indexed publications focusing on human–computer agents, emotional analysis, IoT, and soft computing. She has actively participated in academic events, serving as session chair at conferences like ICCCIE-2022 and ICOCS-2022, and as a panel/advisory board member for ICEMSS-2025. Dr. Saroha has authored several academic books, such as Java Programming (A Mini Guide) and Introduction to Artificial Intelligence. She also holds patents related to intelligent cloud-based robotics and smart highway systems using machine learning.

Raghuraj Singh is working as an Assistant Professor at the Department of Computer Science & Engineering, Galgotias College of Engineering & Technology, Greater Noida. He has done B.Tech and M.Tech in CSE & Pursuing Ph.D from B.R. Ambedkar National Institute of Technology, Jalandhar (Punjab), India. His area of research includes Software Fault Prediction, Software Engineering, Machine Learning and Deep Learning. He has 20 years of Teaching Experience in various technical Universities/ Institutions. He has published several paper in SCI and Scopus Journal. He is Mentor and guiding research investigations for under-graduate, post-graduate engineering students in the areas including, Software Fault Prediction, Software Engineering, Machine Learning, Deep Learning, Artificial Intelligence, Soft Computing, and Image Processing. He has expertise in Salesforce and AWS.