

# Performance Evaluation of Complex Face Emotion Recognition Using Convolution Neural Networks Under Diverse Conditions

**Rajashree Jain**

Symbiosis Institute of Computer Studies and Research, Symbiosis International Deemed University, Pune, Maharashtra, India  
rajashreejain@gmail.com

**Milind Talele**

Symbiosis Centre for Research and Innovation, Symbiosis International Deemed University, Pune, Maharashtra, India  
milind1248@gmail.com (corresponding author)

**Taher Muhammad Ali**

Gulf University for Science and Technology, Mubarak Al-Abdullah, Kuwait  
Ali.T@gust.edu.kw

Received: 20 August 2025 | Revised: 3 November 2025 and 27 November 2025 | Accepted: 28 November 2025

Licensed under a CC-BY 4.0 license | Copyright (c) by the authors | DOI: <https://doi.org/10.48084/etasr.14183>

## ABSTRACT

Facial emotion recognition methods are computer vision-based algorithms for identifying facial emotions from an input image or video, with several applications in surveillance systems, healthcare, education, and retail. A complex facial emotion can be visualized as a combination of two or more basic emotions. This study presents a model for recognizing complex facial emotions using a convolutional neural network, trained on the FER2013 dataset and tested on the CEED dataset. This study investigates the performance of the model in variations, such as illumination conditions, demographic diversity, and image characteristics. Systematic and controlled experiments were conducted to validate the performance of the model and calculate the Complex Emotion Scores (CES). The average CES remained at 80% for dark images with an accuracy of approximately 89%. Female images tend to maintain slightly higher CES scores in light conditions than male ones. However, CES remained consistent on image file types, and the CPU utilization was found to be high for certain image formats.

**Keywords**-Face Emotion Recognition (FER); complex face emotion recognition; Convolutional Neural Network (CNN); performance; CFER score

## I. INTRODUCTION

Facial Emotion Recognition (FER) methods use Artificial Intelligence (AI) and Machine Learning (ML) to analyze face expressions and recognize human emotions such as anger, surprise, fear, sadness, happiness, and disgust. FER can improve the user experience and support several applications in healthcare, retail, forensics, and education. According to the Plutchik model [1], combinations of basic emotions lead to more complex facial emotions. Computational systems with various ML and Deep Learning (DL) techniques enable FER. Advances in DL techniques have further enhanced the accuracy and efficiency of these systems, making them increasingly integrated into everyday applications. As these technologies continue to evolve, guidelines are needed to prioritize user trust and ensure responsible use of emotional recognition systems.

Collaboration between interdisciplinary teams is essential to address these challenges, as insights from psychology, ethics, and technology converge to shape the future of emotionally aware systems. [2-5].

Two or more basic emotions can be combined to form complex emotions. The emotion wheel or the physiological model of emotions proposed in [5] has been widely used in FER. Machines can understand basic emotions, but complex emotions are fragile in nature and require a robust setup of mechanisms to identify them. In addition, the lack of open-source datasets also poses challenges in complex FER. Custom CNNs remain popular among FER researchers due to their efficient and automatic feature extraction techniques compared to other pre-trained models, such as ResNet-50 [6].

Although FER and Complex FER (CFER) techniques have evolved significantly, challenges remain in the robustness of the models developed, as image quality, face diversity, and face occlusion pose challenges in model performance. The study of different parameters likely to impact the accuracy of emotion identification has highlighted different light conditions, darkness, brightness, diversity, ethnicity, age, gender, or image type. This study focuses on the impact of parameters such as lighting conditions, age, gender, and race on model performance.

In [7], a review of ML-based CFER techniques focused on challenges such as prediction accuracy, algorithm latency, and constraints such as image resolution and available bandwidth. A new concept for the detection of complex emotions was proposed by combining physiological and psychological concepts with the seven basic emotions. This approach included detecting basic emotions from static images and combining them to infer complex emotional states, thus advancing the field of FER.

The study in [8] investigated the influence of various illumination conditions on the precision of FER systems utilizing the UIBVFED Plus-Light dataset. This study emphasized the necessity for emotion recognition methodologies to evolve in response to different lighting environments, presenting novel challenges for improving accuracy in real-world applications. In [9], it was stated that existing FER techniques need enhancements to effectively address lighting variations, making it a key domain for future research and innovation. In [10], a CNN achieved 95.8% accuracy in real-time FER from multiple faces, even under varying lighting conditions. This marks a significant improvement over existing methods, which often rely on frontal faces and struggle with illumination changes. The system was designed to handle challenges such as head pose, illumination, and age, enhancing its reliability in diverse scenarios. However, this study did not address complex emotions from non-frontal faces or test its performance across different age groups and environments, limiting its applicability in real-world conditions.

The study in [11] recognized the limitations imposed by the restricted sample size and the specific sociocultural contexts of the participants, indicating that its conclusions may not have universal applicability. This study also emphasized the complexity of emotional perceptions, indicating that variables such as the spatial relationships existing between individuals and objects, in conjunction with individual task trajectories, may influence emotional responses. Therefore, it requires more extensive and multifaceted research methodologies for a thorough investigation of emotional perceptions across diverse contexts. Other studies focused on understanding the emotional impact of images through a factor graph model, emphasizing aesthetic effects and emotional prediction on emotion recognition. This indicates the complexities involved in representing the correlations that exist between visual attributes and their associated emotional implications, while simultaneously evaluating traditional methods such as Support Vector Machines (SVM), Random Forest (RF), and Naive Bayes (NB) classification for their tendency to oversimplify

these relationships and ignore significant emotion-related attributes. In addition, the complexity of social networks has been highlighted, where user interactions subtly affect emotional responses to images and the challenge of accurately predicting emotions without considering these social dynamics [12-13].

Authors in [14] focused on the challenges of affective classification of images using facial expressions, highlighting issues such as pose variation, illumination changes, and facial size differences. This study acknowledged that factors such as compression levels and codec formats can affect classification performance, but did not provide a detailed analysis of these aspects. The proposed system addressed these by a multi-view face detector and tracker for pose variation, using fiducial points for feature extraction to reduce registration errors. This system also employed Local Binary Pattern (LBP) descriptors to counteract lighting changes and normalized facial sizes to ensure consistent recognition.

Taking into account the gaps in the literature on CFER under various conditions, this study developed a CFER method using a CNN and evaluated its performance under various conditions, such as lighting, image type, and diversity parameters, such as gender and race.

## II. METHOD USED

The proposed CNN model was trained on the FER2013 [15] dataset and tested on the CEED dataset [16], using a three-fold method:

1. Development of a CFER method.
2. Training the CNN model on FER2013.
3. Testing the CFER model performance under various conditions using the CEED dataset. Illumination variability, role of diversity (such as ethnicity, age, and gender), and the impact of image quality on the efficacy of the CFER model were considered for testing.

### A. Datasets

#### 1) FER2013

This dataset was used for training and testing the CNN model to derive complex emotions from basic ones. This dataset is a widely used open-source resource for evaluating FER models. It consists of 36,321 grayscale images, categorized into seven basic emotions, and is specifically designed for robust training and testing. The image size was kept at 48x48 pixels to reduce computational complexity. The canonical FER2013 split (80% training, 20% testing) ensures comparability with prior studies. Table I shows the number of images used for training and testing the CNN model.

#### 2) Complex Emotion Expression Database (CEED)

CEED [16] is a comprehensive dataset of 480 images portraying both basic and complex emotions, including lovesick, desirous, contemptuous, attracted, flirtatious, brokenhearted, betrayed, jealous, and affectionate emotions. It was developed by Pennsylvania State University researchers and funded by the National Institute of Mental Health (NIMH).

It has a diverse collection of emotional expressions performed by eight trained actors aged 18–27, evenly split by gender (4 males, 4 females) and race (4 white and 4 black). The actors used acting techniques to produce authentic expressions, which were independently validated by approximately 800 participants. The dataset is split into 243 images for basic emotions and 237 for complex emotions, with individual contributions ranging from 15 to 88 images per actor. CEED offers high-quality, rigorously reviewed images, making it a valuable resource for research on emotion perception, processing, and developmental sensitivity to complex emotional expressions. Table II describes the CEED dataset. This dataset was used to test the proposed CNN model for complex emotions.

TABLE I. FER2013 BASIC EMOTION CATEGORIES

Category	Training images	Testing images	Total images
surprise	3236	828	4064
sad	4969	1170	6139
neutral	5013	1247	6260
happy	7195	1856	9051
fear	4134	1049	5183
disgust	467	142	609
angry	4024	991	5015
Total	29038	7283	36321

TABLE II. DETAILS OF THE CEED DATASET

Category	Subcategory	Total images
Emotion	Affectionate	36
	Attracted	19
	Betrayed	20
	Brokenhearted	36
	Contemptuous	19
	Desirous	46
	Flirtatious	22
	Jealous	9
	Lovesick	30
	Total Complex	237
Gender	Male	240
	Female	240
Age Group	18–20	265
	21–23	15
	24–27	200
Race	White	241
	Black	239
Actors	Total	8 actors (4 male, 4 female)

### B. CNN Model

The CNN model includes three convolutional layers (32, 64, 128 filters, kernel size 3×3), ReLU activations, max pooling layers, dropout (0.5), and two fully connected layers. The Adam optimizer was used with a learning rate of 0.001, a batch size of 64, and 50 epochs. Categorical cross-entropy was used as the loss function. The Complex Emotion Score (CES) is calculated as [3]:

$$CES = \sum_{i=1}^n c W_n \quad (1)$$

where  $W_n$  is the weight of basic emotions,  $n$  is between 1 and 7, and  $c$  is a multiplication factor.

The two maximum dominant basic emotions and their respective weights  $W_{max1}$  and  $W_{max2}$  were prioritized based on an average weight value as:

$$W_a = Average (W_{max1} + W_{max2}) \quad (2)$$

and

$$\begin{cases} c = 0, & \text{if } W_n < W_a \\ c = 1, & \text{if } W_n > W_a \end{cases} \quad (3)$$

The threshold condition is grounded in experimental findings and supported by the prior literature on emotion modeling. In our experiments, the two strongest contributing basic emotions consistently exhibited higher weights compared to the remaining emotions. Emotions with lower weights were disregarded ( $c = 0$ ) as their influence was negligible in forming complex emotions. In contrast, weights above the average threshold were treated as dominant ( $c = 1$ ), ensuring that only significant contributors factored into the final CES. This aligns with Plutchik's circumplex model [1], which emphasizes that dominant emotions play the primary role in shaping complex affective states.

### C. Training the CNN Model

The model was trained as shown in Figure 1. The input face image was fed into the CNN model for feature extraction and classification. The training phase extracts facial features through convolution and pooling. The extracted features from the pooling layer are converted into a 1D vector for processing in the fully connected layer and classifying underlying emotions using a fully connected layer. Variations in the dataset were introduced by changing lighting conditions to analyse their impact on the model's performance and accuracy for complex emotion identification. The dataset also considers factors such as diversity, age, ethnicity, and image type to observe their influence on the model's performance. These parameters significantly affect the model's accuracy. Plutchik's emotion wheel provides a foundation for understanding both basic and complex emotional states. While this model outlines primary emotions such as joy, fear, anger, and surprise, it also highlights the formation of complex emotions through their interaction. In this study, approximately 40% of complex emotions were experimentally identified, with selected examples discussed. Complex emotions arise when two or more primary emotions blend; for instance, awe emerges from fear and surprise. Additional examples include love (joy+trust), guilt (joy+fear), optimism (anticipation+joy), and contempt (disgust+anger).

### D. Using the Trained Model to Validate CFER Under Various Conditions

This was the major focus of this work. Variations in the dataset were introduced by changing lighting conditions to analyze their impact on the model's performance and accuracy for complex emotion identification. Additionally, different image types were utilized to study performance variations, as shown in Figure 3. The dataset also considers factors such as diversity, age, ethnicity, and image type to observe their influence on the model's performance. These parameters significantly affect the model's accuracy.

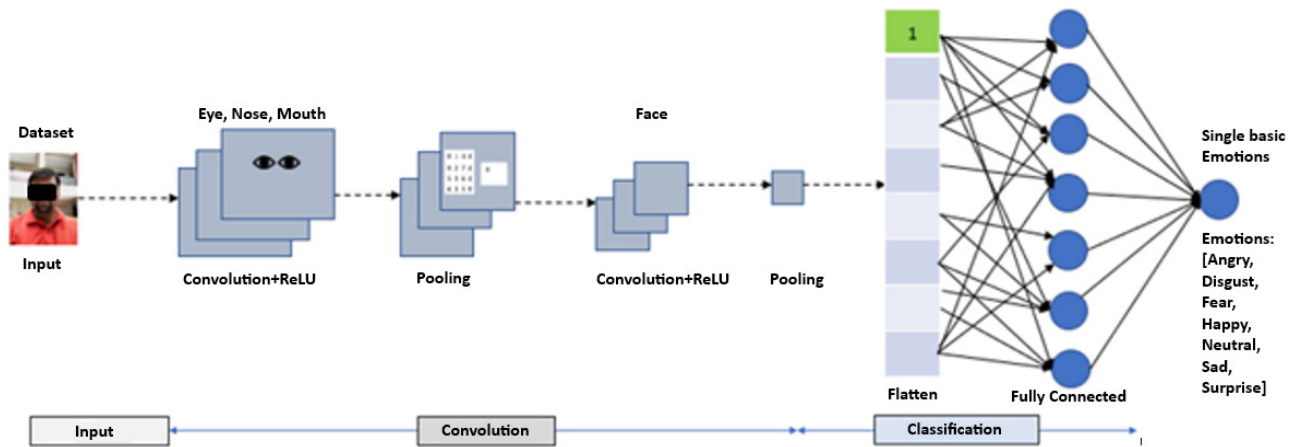


Fig. 1. CNN architecture of the CFER model.

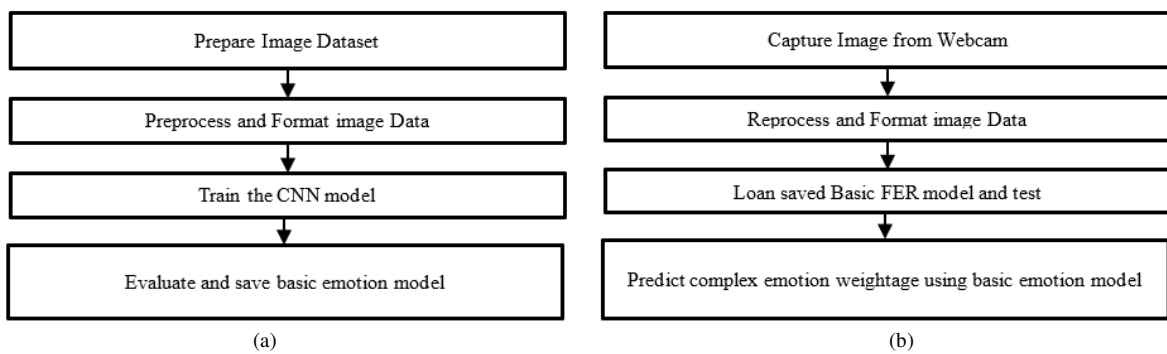


Fig. 2. Block diagram of the CFER process: (a) Training, (b) Testing.

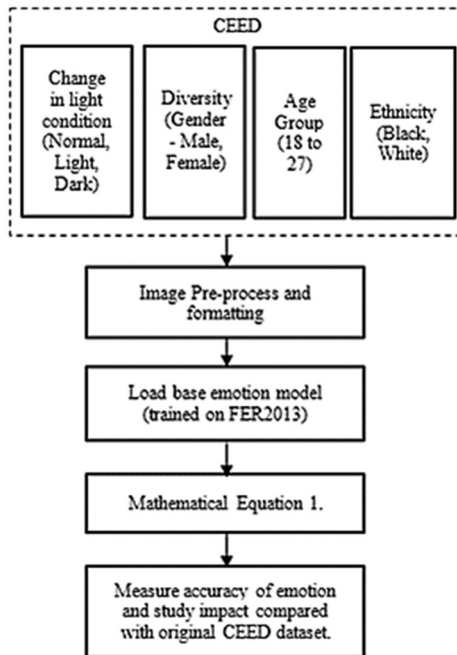


Fig. 3. Flow chart to study the impact of parameters on the CFER model.

Different variations of light conditions were created by changing the light conditions, i.e., the value for generating bright and dark images of the same faces, as shown in Figure 4. To create a different light condition dataset, images initially in

the RGB color space were converted to the HSV color space. In HSV, the Value (V) channel controls the brightness factor, making it easier to adjust brightness without affecting the colors of the image. The V channel was scaled by a factor between 1.5 and 0.5. This adjusts the brightness by scaling the V channel in HSV space with a specified factor, effectively making the image brighter or darker without altering the hue or saturation.

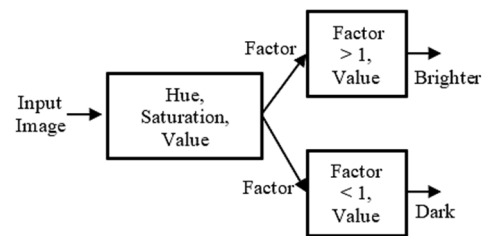


Fig. 4. Illumination.



Fig. 5. Brokenhearted emotion in Normal, Bright, and Dark conditions.

### III. RESULTS AND DISCUSSION

#### A. Impact of Light Conditions on CFER Performance

The light conditions are important to improve the robustness of FER, as they have a significant impact on the image quality and emotion identification. Normal lighting provides balanced illumination, ensuring clear visibility of facial features and serving as a baseline for analysis. Bright lighting, or overexposure, results in washed-out areas that obscure fine details, challenging the model's ability to detect subtle emotions. Dark lighting, or low light, reduces facial visibility, simulating poorly lit environments and testing the model's robustness in extracting features under such conditions. Directional lighting introduces uneven illumination and shadows, reflecting real-world scenarios, such as side lighting, where parts of the face are obscured. These lighting variations are critical for evaluating the performance and adaptability of FER systems in diverse real-world settings [3]. Directional lightning was not within the scope of this research paper. Table III shows the CES obtained under various illumination conditions for a particular complex emotion. The table has sample values of the CES scores for 10 images tested. Table IV presents the average of all images available in the category.

In this study, accuracy refers to the percentage of correctly classified complex emotions, computed by combining two dominant basic emotions predicted by the CNN and validating them against the complex label available.

Accuracy =

$$\frac{\text{No. of correctly classified emotions}}{\text{Total No. of Images}} \times 100 \quad (4)$$

TABLE III. PERFORMANCE COMPARISON OF CES FOR NORMAL, DARK, AND BRIGHT CONDITIONS

CES for original images	CES for brighter images	CES for darker images	Age	Gender	Ethnicity
89.23	84.57	85.48	26	Man	Black
96.37	95.44	91.73	26	Woman	Black
96.19	96.15	90.89	26	Woman	Asian
95.47	94.66	94.14	27	Woman	Asian
53.64	25.56	81.68	26	Man	Latino-Hispanic
96.53	96.16	89.86	27	Man	black
74.66	67.82	83.84	28	Man	Latino-Hispanic
96.20	95.77	88.77	28	Man	Black
64.98	59.39	56.79	29	Man	Indian-Latino
95.16	90.59	90.71	27	Man	Latino-Hispanic

TABLE IV. COMPARISON OF CES AND CLASSIFICATION ACCURACY ACROSS ORIGINAL, BRIGHT AND DARK IMAGES

Parameters	Original images	Bright images	Dark images
Average CES	80	73	81
Count of images whose CES > Average CES	20	18	26
Accuracy %	89	81	92

Table IV reveals that lighting conditions significantly affect both the CES and the classification accuracy. Dark images exhibit the highest CES (81) and the highest accuracy (92%), suggesting that darker conditions enhance the model's capability to detect and classify human emotions, likely due to better feature emphasis by changing the image value factor to 1.5 (Figure 4). In contrast, brighter images have the lowest CES (73) and accuracy (81%), potentially due to overexposure and loss of detail. Original images serve as a baseline with moderate CES (80) and accuracy (89%), indicating the model performs well under standard lighting but can benefit from optimization to handle varying light conditions.

Statistical analyses were conducted to further validate the observed differences in CES. A one-way ANOVA was applied to examine the effect of lighting conditions (original, bright, dark) on CES values, while independent t-tests were used to compare differences across gender and ethnicity groups. Lighting had a statistically significant effect on CES ( $p < 0.05$ ), with dark images yielding higher scores than bright ones. Similarly, gender-based comparisons indicated that female samples maintained significantly higher CES values under bright conditions compared to male ones ( $p < 0.05$ ). Ethnicity-based analysis also showed significant differences ( $p < 0.05$ ), particularly with Latino/Hispanic faces that exhibited higher CES variability across lighting conditions. These results confirm that the observed performance variations are not due to random chance but are statistically significant.

#### B. Impact of Gender, Ethnicity, and Age on CFER

##### 1) Gender

Female images tend to maintain slightly higher CES scores in light conditions, possibly due to image features (e.g., higher contrast in facial structures) than male images. CES in dark conditions tends to drop more sharply, especially in younger individuals.

##### 2) Ethnicity

Ethnic variations significantly influence the performance of CFER due to differences in facial features, skin tones, and expression patterns across diverse populations. This study examined the influence of ethnicity on CFER accuracy, emphasizing the importance of diverse and balanced datasets for training. The findings show that CES remains high for the original images but decreases significantly in dark conditions, as shown in Figure 5.

##### 3) Age

Age is a critical factor influencing the accuracy of CFER systems, as facial features and expressions change significantly across different age groups. Younger individuals, such as children, may exhibit more exaggerated expressions, while older adults often have subtler expressions due to facial wrinkles and reduced muscle elasticity. These variations can pose challenges for cFER models, which may struggle to generalize across diverse age groups if training datasets are not age-inclusive. This study explored how age-related differences impact cFER performance, highlighting the need for age-diverse datasets to improve model robustness and ensure accurate emotion recognition across all ages.

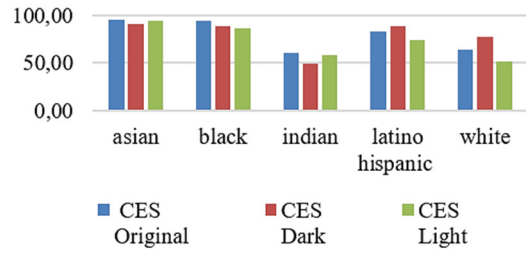


Fig. 6. Average CES by ethnicity for brokenhearted in different lighting conditions.

C. Impact of Image Type

This study considered JPEG, PNG, BMP, and TIFF file formats. Key findings reveal that CES remained consistent across all image formats, indicating that the brokenhearted emotion recognition is unaffected by the format type. Table V shows the CPU utilization and memory usage at the start and the end of the process. The results highlight the effect of appropriate preprocessing and hyperparameter tuning (Table VII) on enhancing model performance.

TABLE V. CPU AND MEMORY USAGES FOR DIFFERENT LIGHT CONDITIONS

File type	Processing stage	CPU usage (%)	Memory usage (%)
JPEG	Start	54.2	73.3
	End	49.4	73.2
BMP	Start	14.4	75.9
	End	24.3	77.7
PNG	Start	79.7	77.3
	End	33.9	79.4
TIFF	Start	38.7	76.1
	End	90.2	77.7

TABLE VI. HYPER PARAMETERS USED FOR CNN

Hyperparameter	CNN
Learning rate	0.001 (Adam)
Batch size	128
Epochs	48
Class mode	Categorical
Input shape (Gray scale)	(48, 48, 1)
Optimizer	Adam

This research highlights the significant impact of light conditions, gender, ethnicity, age, and image type on the performance and accuracy of CFER models. The findings underscore the critical role of a diverse and well-constructed dataset in building effective and reliable CFER systems. The findings underscore the crucial role of hyperparameter tuning in CNN-based FER. This study is among the first to use the CEED dataset for CFER under diverse conditions. Although it provided valuable insights by introducing variations in these parameters, it faced limitations due to the restricted availability of diverse datasets, necessitating the use of synthetic variations that may not completely encapsulate real-world complexities. Future datasets need diversity in light conditions, demographics, and image types. In addition, further advances in FER algorithms are needed to enhance accuracy and reduce latency, particularly under challenging conditions, facilitating the development of more sophisticated and effective CFER systems.

Future work should involve the expansion of dataset diversity, real-time model deployment, hybrid CNN-Transformer approaches [17], and multimodal fusion (EEG and facial data). Furthermore, future approaches should address current limitations, including limited dataset size, absence of real-world occlusions, limited cultural diversity, and variations in synthetic lighting.

IV. CONCLUSIONS

This study is one of the first to use the CEED dataset for the CFER model testing purposes, examining various conditions, such as variation in illumination intensity, and diversity factors, such as gender, age, ethnicity, and image types. CFER accuracy was better in the dark images, with a Value (V) factor of 0.5 in the HSV color space; however, CFER is most robust under normal lighting conditions. The study emphasized the critical role of balanced lighting and consistent pixel variability in achieving reliable and accurate emotion recognition. Younger individuals maintain higher CES scores, while older individuals show more variability in CES scores in brighter and darker conditions, possibly due to subtle facial feature changes. Changes in image type do not significantly influence the FER; however, slight changes in processing time were found due to different image types. The impact of demography on accuracy was studied, and the results showed that for black faces, CES remained high for the original images but dropped significantly under dark conditions.

ACKNOWLEDGMENT

The authors express their sincere gratitude to their colleagues for their valuable discussions, constructive feedback, and continuous encouragement throughout the course of this work. The authors also acknowledge the institutional support and facilities provided during this study.

REFERENCES

- [1] R. Plutchik, "The Nature of Emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," *American Scientist*, vol. 89, no. 4, 2001, Art. no. 344, <https://doi.org/10.1511/2001.4.344>.
- [2] A. M. Abdulazeez and Z. S. Ageed, "Face Emotion Recognition Based on Machine Learning: A Review," *International Journal of Informatics, Information System and Computer Engineering (INJIISCOM)*, vol. 5, no. 1, pp. 53–87, Mar. 2024, <https://doi.org/10.34010/injiiscom.v5i1.12145>.
- [3] E. S. Agung, A. P. Rifai, and T. Wijayanto, "Image-based facial emotion recognition using convolutional neural network on emognition dataset," *Scientific Reports*, vol. 14, no. 1, June 2024, Art. no. 14429, <https://doi.org/10.1038/s41598-024-65276-x>.
- [4] M. Talele, R. Jain, and P. Kulkarni, "Review of Face Emotion Recognition Using Feature Extraction Techniques," in *2023 Intelligent Computing and Control for Engineering and Business Systems (ICCEBS)*, Chennai, India, Sept. 2023, pp. 1–6, <https://doi.org/10.1109/ICCEBS58601.2023.10448632>.
- [5] H. R. Conte and R. Plutchik, "A circumplex model for interpersonal personality traits.," *Journal of Personality and Social Psychology*, vol. 40, no. 4, pp. 701–711, Apr. 1981, <https://doi.org/10.1037/0022-3514.40.4.701>.
- [6] M. Talele and R. Jain, "A Comparative Analysis of CNNs and ResNet50 for Facial Emotion Recognition," *Engineering, Technology & Applied Science Research*, vol. 15, no. 2, pp. 20693–20701, Apr. 2025, <https://doi.org/10.48084/etasr.9849>.
- [7] M. Talele and R. Jain, "Complex Face Emotion Recognition Using Convolutional Neural Networks," *Journal Européen des Systèmes*

- Automatisés*, vol. 58, no. 1, pp. 171–178, Jan. 2025, <https://doi.org/10.18280/jesa.580119>.
- [8] M. Mascaró-Oliver, E. Amengual-Alcover, M. F. Roig-Maimó, and R. Mas-Sansó, "UIBFEDPlus-Light: Virtual facial expression dataset with lighting," *PLOS ONE*, vol. 18, no. 9, Sept. 2023, Art. no. e0287006, <https://doi.org/10.1371/journal.pone.0287006>.
- [9] C. C. F. Or, D. Y. Lim, S. Chen, and A. L. F. Lee, "Face Recognition Under Adverse Viewing Conditions: Implications for Eyewitness Testimony," *Policy Insights from the Behavioral and Brain Sciences*, vol. 10, no. 2, pp. 264–271, Oct. 2023, <https://doi.org/10.1177/23727322231194458>.
- [10] S. Saxena, S. Tripathi, and T. S. B. Sudarshan, "Deep Dive into Faces: Pose & Illumination Invariant Multi-Face Emotion Recognition System," in *2019 IEEE/RISJ International Conference on Intelligent Robots and Systems (IROS)*, Macau, China, Nov. 2019, pp. 1088–1093, <https://doi.org/10.1109/IROS40897.2019.8967874>.
- [11] X. Xie *et al.*, "Affective Impressions Recognition under Different Colored Lights Based on Physiological Signals and Subjective Evaluation Method," *Sensors*, vol. 23, no. 11, June 2023, Art. no. 5322, <https://doi.org/10.3390/s23115322>.
- [12] A. Bonassi *et al.*, "The Recognition of Cross-Cultural Emotional Faces Is Affected by Intensity and Ethnicity in a Japanese Sample," *Behavioral Sciences*, vol. 11, no. 5, Apr. 2021, Art. no. 59, <https://doi.org/10.3390/bs11050059>.
- [13] D. Schneevogt and P. Paggio, "The Effect of Gender and Age Differences on the Recognition of Emotions from Facial Expressions," in *Proceedings of the Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media (PEOPLES)*, Osaka, Japan, Sept. 2016, pp. 11–19.
- [14] X. Wang, J. Jia, P. Hu, S. Wu, J. Tang, and L. Cai, "Understanding the emotional impact of images," in *Proceedings of the 20th ACM International Conference on Multimedia*, Nara, Japan, Oct. 2012, pp. 1369–1370, <https://doi.org/10.1145/2393347.2396489>.
- [15] "Face expression recognition dataset." Kaggle, Available: <https://www.kaggle.com/datasets/jonathanoheix/face-expression-recognition-dataset>.
- [16] M. S. Benda and K. S. Scherf, "The Complex Emotion Expression Database: A validated stimulus set of trained actors," *PLOS ONE*, vol. 15, no. 2, Feb. 2020, Art. no. e0228248, <https://doi.org/10.1371/journal.pone.0228248>.
- [17] M. Mujiyanto, A. Setyanto, K. Kusriani, and E. Utami, "Swin Transformer with Enhanced Dropout and Layer-wise Unfreezing for Facial Expression Recognition in Mental Health Detection," *Engineering, Technology & Applied Science Research*, vol. 14, no. 6, pp. 19016–19023, Dec. 2024, <https://doi.org/10.48084/etasr.9139>.

## AUTHORS PROFILE



Dr. Rajashree Jain holds a Doctor of Philosophy in the Faculty of Computer Studies, Symbiosis International University, India, in 2014. She also received her M.Phil. and M.Sc. from Gulbarga University, Gulbarga, India, in 1990 and 1989, respectively. She is a Professor at Symbiosis Institute of Computer Studies and Research, Symbiosis International University, Pune, India. She has more than 25 years of experience in teaching, research, and academic administration.

Her research areas are microwave techniques, antennas, AI and Sensor networks, IoT, data analysis, entrepreneurship, and education. She also has more than 30 research papers to her credit, published in reputed journals, and has presented research papers at various national and international conferences. She has served as chair/member of the organizing committees of various technical conferences and leadership summits.



Dr. Milind Talele is a Doctor of Philosophy in Computer Studies (Artificial Intelligence) with 22 years of IT industry experience in diverse domains and technologies. He has published several research papers on face emotion recognition and time series forecasting. He is proficient in programming, project management, and agile implementation across various industries. He holds a B.Sc. degree in Electronics Engineering and an MBA in IT from Pune University. He also holds multiple industry-recognized certifications, including PMP, PRINCE2, Cloud Architect, TOGAF 9, Java Certified, SAFe Agilist, Certified Scrum Master, and Kanban Coach. His diverse credentials showcase his expertise in project management, agile methodologies, enterprise architecture, and cloud solutions.



Dr. Taher Muhammad Ali is a seasoned academic and IT professional with over two decades of experience in both higher education and IT management. Currently, he serves as a Senior IT Manager and Assistant Professor at Gulf University for Science and Technology in Kuwait. He has held leadership roles such as Head of the Computer Science Department and Associate Dean of the College of Arts & Sciences. His expertise spans IT management, curriculum development, educational technology, and interdisciplinary collaboration. Passionate about bridging the gap between academia and technology, his research focuses on how emerging technologies can enhance teaching, learning, and institutional management. Through his extensive experience in leadership, project management, and academic research, he continues to drive technological advancements in education while preparing students for the evolving digital landscape.