International Journal of Basic and Applied Sciences, 14 (SI-1) (2025) 615-623

SPC

International Journal of Basic and Applied Sciences

Website: www.sciencepubco.com/index.php/IJBAS
https://doi.org/10.14419/xwcnx826
Research paper

An Innovative Approach to Face Emotion Recognition Using Antlion Optimization Model

C. Shakila [1] *, T. Kamalakannan [2]

[1] *Research Scholar, Department of Computing Science and Technology, Vels Institute of Science Technology And Advanced Studies (VISTAS) – Pallavaram. Chennai,600117, TN, India*
[2] *Professor, Department of Information Technology, School of Computing Science, Vels Institute of Science Technology and Advanced Studies (VISTAS) – Pallavaram, Chennai,600117, TN, India*
**Corresponding author E-mail: sla24143@gmail.com*

Received: July 15, 2025, Accepted: July 24, 2025, Published: November 1, 2025

Abstract

Emotion analysis is an interesting area of research that contains various integrations of contextual data collected from real-life incidents. Various existing implementations of emotion detection address facial expressions that signify changing moods. The proposed system is suggested by considering the inconsistencies and rankings of various critical parameter concentrations in existing frameworks for emotion detection, such as empathic concern (intention towards others) and fortune and personal distress, which relate to the situation of comfort in response to others' emotions. Emotions need to be understood deeply to recognize the feelings of humans accurately. In the presented system, the feature extraction used for facial expression handling is implemented using a neural network with respect to facial emotions, such as happy, Fear, sad, angry, surprised, and neutral. In practical humans, they can generate various facial expressions during communication that vary in very intense manners. The proposed system-wise hybrid feature extraction and facial expression identification technique, utilizing Viola-Jones cascade object detectors and MSER feature extraction technique with speed-up robust feature extraction (SURF) technique, is represented as multimodality network features (MMNF) for classification. The identified features are further classified using an LSTM (Long Short-Term Memory) model, combined with a CNN (Convolutional Neural Network) hybrid architecture. Further, the system parameters are optimized through the Hybrid Ant-Lion optimization technique. The presented system achieved 98% accuracy compared with various state-of-the-art approaches.

Keywords: *Emotion Recognition; HALEC; Hybrid Feature Extraction; MSER; Voila-Jones Cascade Object Detectors.*

1. Introduction

Emotions are one of the common factors in humans to show interpersonal communication. Emotions are expressed in various situations and through different methods. Human behavior depends on the factor called deep emotions. Through the right tools, the real impact of the emotions is identified. Due to various changes in work life, lifestyle habits, and stressful environments, the patterns of human emotions and their expression are varied. Emotion detection is also helpful in applications such as human-computer interface animation, medicine [1], and security. Today, facial recognition is utilized in numerous real-world applications. Despite its importance, the face recognition process still presents numerous challenges, including changes in facial expressions, posture variations, occlusions, and variations in lighting. The face is a fundamentally crucial feature of the human body, used visually for identification purposes. A facial expression is a nonverbal form of communication that plays a vital role in conveying emotions and experiences. Its significance increases exponentially for people with hearing impairments, for whom facial expressions are often the only means of communication. Face area is further broken down into local and national characteristics. Due to the diverse nature of various cultures, facial expressions vary from person to person, an inborn capacity that is passed down from their ancestors, or more specifically, due to the unique facial muscle activity of individuals. It hasn't focused on classifying CNN [2]. It hasn't focused on increasing the recognition rate and classification accuracy of facial expressions. The emerging growth of computer vision analysis, face recognition systems, and various consumer products, such as driver fatigue detection, investigation of criminal cases, entertainment, and cognitive analysis, has led to the development of frameworks for facial expression detection that utilize artificial intelligence and machine learning models [3]. The face features are extracted using linear discriminant analysis (LDA) and Principal component analysis (PCA).
Automated facial expression recognition has numerous practical applications, including psychological analysis, medical diagnosis, forensic analysis (such as lie detection), evaluating learning effectiveness, and high-end sales advertising, among others. The system's capability to detect and assess human emotions offers a new perspective on the development of human-machine interfacing systems. For specific applications, smile detectors are often used to acquire the primary emotional state. Robots can derive information through facial expression detection [4]. Various existing applications in robotics consider facial expressions as primary data for reacting to humans.

- The proposed approach focuses on creating an adaptive machine learning-based emotion identification system using the Hybrid Ant Lion classification model (HALEC).
- The goal of the proposed approach is to extract the unique features from the face. From various existing state-of-the-art approaches, unique face features are extracted to get significant regions that show the maximum impact on facial expression detection.

The remainder of the paper is structured as a detailed review of existing literature and analysis in Section 2. System methodology is explored in Section 3. Various results obtained are analyzed in Section 4. With the Conclusion of the presented work, future enhancements are developed.

2. Related Works

The rapid development of internet-enabled devices has created numerous advantages in consumer accessibility; moreover, the number of e-learners has drastically increased due to the facilities available. Similarly, people began working from home, spending more time on social media, and utilizing internet-connected systems to perform their essential tasks. Distance learning and e-learning have become an integral part of our lifestyle. Through facial expression analysis, teachers can evaluate the level of student involvement in the learning session [5]. The standard methods utilized in face expression detection are Viola-Jones-based face part detection and the HAAR cascade algorithm for object detection from the facial images. Convolutional neural network models are frequently used in image classification. The challenges in face detection and recognition systems are that most facial regions share typical pixel intensities. The unique expressions are detected through facial expressions [6], [7]. In the service industry, extracting feedback from consumers through facial expressions is considered in applications such as healthcare, tourism, hospitality, and retail. Facial expressions are sometimes captured through single image capture, but in most cases, a sequence of images is required as video input to capture the best image frame. The Viola-Jones Haar cascade model is a commonly utilized facial feature detection system, which detects the most important facial features [9], [10]. Various machine learning algorithms are employed in face feature detection, including AdaBoost training and systematic cascading classifiers.

The HALEC model addresses some of the shortcomings observed in previous studies of emotion recognition, particularly in detecting and classifying facial expressions. An example is the work by Khoeun et al. [15], where deep learning-based models demonstrated an inability to work with occlusions and low-quality images, frequently resulting in a decrease in recognition accuracy. This limitation is overcome by the HALEC model, which employs a hybrid approach to feature extraction that combines both MSER and SURF features, thereby making the model more resilient to occlusions and variations in image quality. Moreover, the model could not handle various and complicated facial expressions, as only one feature extraction method was used, which is a limitation that reduced the accuracy of emotion classification in the study by Dong et al. [14]. This is further enhanced by the HALEC model, which combines various feature extraction techniques and forms a multimodal network with more advanced and varied facial features. This combination of multiple modalities in HALEC exhibits superior performance in terms of accuracy and generalization to various facial expressions. Moreover, unlike Prabhu et al. [2], who used convolutional neural networks (CNNs) as the primary tool, HALEC employs the Ant-Lion Optimization, which optimizes both the feature selection process and the classification process, making the model significantly more efficient and accurate. Using a combination of these cutting-edge methods, HALEC not only overcomes the shortcomings of other methods but also establishes a new paradigm for emotion recognition in facial expression recognition.

Oulefkiet et al. [11] explore the image processing technique involved in a face detection system, where the Histogram equalization (CLAHE) model for face enhancement is considered. Using a Fuzzy logic interference system, the uniform intensity of face images is explored. Environmental lighting affects the detection accuracy of face images. The image preprocessing process involved dividing the image into two sub-regions: the darker area and the lighter area, where the intensity of the image is low. Normalization is the process utilized to handle the massive pixel density into sub-regions.

Bendjill et al. [12] presented a system utilizing a modified-contrast limited adaptive histogram equalization algorithm (M-CLAHE) for the detailed enhancement of face images. Subsequently, face feature extraction using the Viola-Jones model is discussed. The features are divided into training images and testing images, and then the classification process is applied to these images.

Z. Dong et.al. [13]A Novel multiple-network impacted FER network is created for applications involving high-quality face feature extraction, data augmentation processes, and cognitive analysis of various facial expressions. To establish the relationship between facial expressions and brain data analysis, a FER-based multimodal network is proposed. The major challenge in the proposed work is that the unstructured facial features need to be optimized to achieve high accuracy.

M. D'incàet et al [14] presented an effective system through a convolutional residual auto-encoder, enabling a positive and negative emotions extractor model. The study considers various existing supervised models and unsupervised feature extraction methods to perform comparisons across different domains with existing state-of-the-art approaches. The primary challenge in unsupervised systems is that features are available with noise factors that need to be cleaned.

R. Khoeun et.al [15], the authors performed a deep analysis framework through LSTM (Long-Short term memory model) with an Artificial Neural Network (ANN) using various standard benchmark face datasets such as CK+, RAF-DB, and FER2013 data storage. The proposed SLPPE-based feature analysis model and its performance, including accuracy, sensitivity, and error ratio, are compared with existing state-of-the-art implementations.

Y. Ma et al [16] presented a system for facial movement analysis through discrete wavelet transform (DWT). The Bayesian network-enabled Pearson correlation coefficient (PCC) is evaluated using information from 30 depressed patients and 30 healthy persons. The proposed system, with a PCC accuracy of 81.7%, achieves a recall value of 96.7%.

Dataset Details: The dataset used in the face emotion recognition task in this research is a combination of the JAFFE and JAFFE-2013 datasets, which are commonly applied in emotion recognition studies. FER-2013 is a collection of 35,887 labeled photos organized into seven emotional categories: Happy, Sad, Surprise, Anger, Disgust, Fear, and Neutral. The data comes from the Kaggle Facial Expression Recognition Challenge. The CK+ dataset comprises 593 image sequences of 123 subjects, encompassing the emotions of Anger, Disgust, Fear, Happiness, Sadness, Surprise, and Contempt. The two datasets are characterized by heterogeneity in terms of lighting, subject demographics, and facial expressions, which play a significant role in training powerful emotion recognition models. Preprocessing methods were employed to normalize image sizes and enhance the feature extraction process. Additionally, augmentation measures were implemented to optimize the dataset for training and testing purposes.

In this paper, face emotion recognition is performed using the FER-2013 and CK+ (Cohn-Kanade) datasets. They are typically used in face recognition programs and have labeled pictures of faces of varied emotions.

- FER-2013 Dataset:

Size: 35,887 labeled images

Types of Emotions: Happy, Sad, Surprise, Anger, Disgust, Fear, Neutral.
Source: Available on Kaggle (https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data).

- CK+ Dataset:

Size: 593 series of images of 123 subjects.
Emotion Categories: Anger, Disgust, Fear, Happiness, Sadness, Surprise, and Contempt.
References: This information is publicly available in the Cohn-Kanade Facial Expression Database (https://www.pitt.edu/~emotion/).
The combination of these data was made to provide a diverse range of expressions and various lighting conditions. The dataset images were preprocessed to normalize image sizes and remove noise before feature extraction, using procedures such as SURF and MSER. The data augmentation techniques used in the preprocessing phase were employed to enhance the training and test samples.
This information on the dataset should be provided immediately after the introduction. The datasets can be introduced only after explaining the general scope of the study, which will assist in framing the research and provide a definite point of reference to the data used in the system.

3. Design Methodology

The proposed system is developed through the keen organization of processes involved in preprocessing face images from a face dataset—feature extraction from the given face images using MSER and SURF feature analysis. The system is integrated with a frequently utilized facial expression recognition (FER) model using the Viola-Jones Haar cascade-based object detectors. The features extracted from the training images are further optimized through an Ant-Lion ensemble model for feature optimization. The various processes involved in the presented system for face expression detection and analysis are illustrated through the flow diagram shown in Fig. 1.

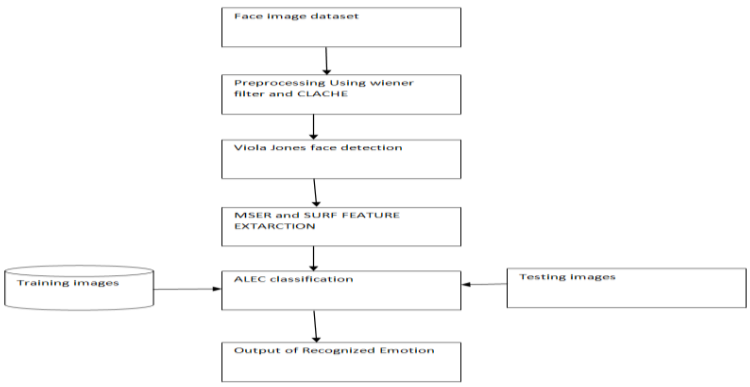

Fig. 1: System Block Diagram of the Proposed HHALEC Model.

Fig 1. Shows the system block diagram of the HHALEC model for face expression detection.

3.1. Face detection

To detect human face expressions, proper detection of face boundaries is essential. The technique is initiated through region of interest detection (ROI). The complete database images are used for training and testing, undergoing various preprocessing steps as part of this process. Every face has a similar color intensity in most areas. The facial features, such as eyes, nose, mouth, and eyebrows, are unique. To obtain a practical face recognition framework, the face features must be extracted accurately and reliably.
An approach to the multimodality feature class is developed here. A hybrid approach is designed using ensemble classification of various common traits for face expression detection. The primary goal of the system is not only to detect the face object or face expression, but a hybrid approach to emotion detection model is proposed here. The ensemble approach employed here encompasses the extraction of various features from the face, as well as the integration of different face features to achieve multimodality in face analysis. The ensemble of features is derived from the SURF technique and the MSER-based feature extraction method. The fusion of extracted features converges to form a strong pattern vector, which is helpful for classification. Neural classifiers are utilized here to establish the correlated pattern between the training feature vectors and the testing feature vectors—the classification method employed in various existing frameworks utilizes a Convolutional Neural Network (CNN) architecture. The major drawback of the CNN architecture is that it requires extensive utilization of processing space. To avoid the complexity in image classification with a CNN model, a lightweight network is developed here. The presented approach considers an ensemble of features that create a substantial impact on expression extraction. The next step after the feature extraction process is classification and optimization using the proposed Hybrid Ant-Lion optimization with a Neural network computing model. The machine learning process has a high impact on handling features, as opposed to various other methods discussed in existing works. Although the Viola-Jones-based facial feature extraction technique is commonly used in practice, the extraction of face regions, such as eyes, nose, and mouth, is often achieved using the Haar cascade model. Once the region is extracted, the unique features of the areas extracted by the Haar cascade model (HCM) are derived through MSER+SURF-based (multimodality network features) MMNF. The initial process of face detection for emotion identification is derived from positive and negative expressions, as shown in Fig. 5. After the feature extraction process, the highlighted pixels are converted into a binary format. The difference in intensities is the primary impact on the feature extraction process. Once the extracted regions are converted into binary, the white pixel regions and black pixel regions are extracted and evaluated using the existing model, which utilizes the Dlib face landmark detection identification database [17]. In the Multitask Cascade Convolutional Neural Network (MT-CNN), deep features are already extracted in the existing model [18]. Still, the drawback of occluded images, Images without faces, and low-quality images impacts the performance of the analysis model.
Various processes involved in the proposed Emotion Detection Model (EDM) are explored below.

3.2. Preprocessing

Preprocessing of images is essential for an analysis model that involves machine learning algorithms. To ensure the quality of detection, preprocessing steps are crucial. It removes the noise present in the system at the initial stage itself. Preprocessing also acts as the initial screening process for removing low-quality images.

3.3. CLAHE

CLAHE is a significant model used in treating low-quality images impacted by environmental light differences, low contrast, etc. The role of the CLAHE technique in medical image processing is outstanding, as it offers adaptive changes in image intensity, similar to histogram equalization, but with enhanced histograms for complex images. The method retains the noise spots present in the input images. Still, a few limitations persist, including interpolation and high-impact noise spots. The histogram of the raw input image enhances the image quality, which has a significant impact on the performance of the analysis model. The suppression factor β can be explored as follows.
$\beta = MN\ G\ \{1 + \alpha\ 100\ (ASmax - 1)\}$ (1)
Where M and N are the factors representing the number of pixels in each area of the input image, g is the factor representing the grayscale levels of pixels. αact as the clipping factor. ASmax denotes the maximum value of the slope to derive enhanced quality.

3.4. Weiner filter

The Weiner filter is one of the adaptive performance filters that operates in conjunction with the basic process of Fourier iterations. The factor called variance is extracted from the images. Based on the variance value, the smoothing process takes place at the required level of iterations. The noise in the low-quality image is adaptively tuned through the high-frequency component, removing the affected area without compromising the quality of the input image. The overall Mean Squared Error (MSE) of the input images is reduced through the Wiener filter.

$$\mu = \frac{1}{NM}\sum_{n1,n2\epsilon\tau} a(n1, n2) \tag{2}$$

$$\sigma^2 = \frac{1}{NM}\sum_{n1,n2\epsilon\tau} a^2(n1, n2) - \mu^2 \tag{3}$$

Where A holds the N-by-M representative local parameters, the specialty of the Weiner filter is that the tuning process involves pixel-by-pixel processing.

$$b(n1, n2) = \mu + \frac{\sigma^2 - \vartheta^2}{\sigma^2}(a(n1, n2) - \mu) \tag{4}$$

3.5. SURF feature extraction

The standard feature extraction process utilized in image processing toolboxes, such as the SURF (Speeded-Up Robust Features) method, extracts features from the surface of the images. The descriptors are used to extract the highlighted pixels that look outstanding in the test image. A combination of various unique pixels extracted from different images during the training process is used to create a pattern of feature vectors. The SURF feature is expressed by the following formula for a unique interest point descriptor—the preprocessed input after SURF feature extraction is represented as a Hessian Matrix (H). Assuming the input image and its associated pixels as the base data, the Hessian matrix is derived using the following expression as H(X, σ), where X represents the scale, and σ is a significant factor for filtering [19].

$$H(X, \sigma) = \frac{L_{xx}(X,\sigma)L_{xy}(X,\sigma)}{L_{xx}(X,\sigma)L_{yy}(X,\sigma)} \tag{5}$$

Where $(X,)$ is the Gaussian factor in the second-order derivative with the convolution process, as $(X,)$ and $Ly(X,\sigma)$ can also be treated with the normal multiplication process.
From the SURF feature, interest descriptors are extracted using a wavelet transform. The feature descriptors are collected from the convergence of the various similar pixels in the image. The entire picture is divided into 4x4 sub-regions.

3.6. Maximally stable extremal regions (MSER)

MSER is one of the robust methods for region-based similarity extraction as well as shape extraction—the convergence of the accelerated feature to detect morphological blobs in the input image pattern. The region influenced by similar pixels is extracted through a standard connected component process. If similar pixels are extracted together, then MSER patterns are maximally stable. The impacted intensity has an average threshold that falls below or above the expected range. The intensity pixel values are incremented from 0 to the maximum value or from the maximum value to the optimal value based on the similarity ratio.

3.7. HALEC network model

The Hybrid ALEC model comprises two main phases of operation. The first phase involves image processing and enhancement, while the second phase utilizes an optimization model.
The implementation summary of the proposed novel work is explored here. The images from the database are considered the benchmark images for analysis. The input images are preprocessed, and the feature extraction process is focused. The Viola-Jones algorithm is used to extract face-detecting features. The essential facial regions, unique in nature, include the eyes, nose, mouth, eyebrows, and others. Through the Haar cascade algorithm, deep features are extracted.

Furthermore, the input feature extraction utilizes a multiple modality network (MMNF) that incorporates MSER feature descriptors and SURF feature descriptors. The advantages of MMNF derive from the deep features in facial images. The fused features are further divided into training patterns and testing patterns. Generally, the features extracted from the input images vary in size. These features are distributed through a probabilistic random distributor. To normalize the features, the values are randomly repeated. The process of random reputation produces a more robust similarity correlation ratio. The classification model adopted here is an LSTM-based CNN architecture, focusing on a two-level deep classification.

3.8. LSTM-CNN

The LSTM-CNN architecture holds several neurons in the memory unit to process the features extracted from the input image. The general structure of the LSTM model consists of an input layer, multiple hidden layers, and fully connected layers. Each neuron holds certain information regarding the input image feature; hence, the total number of hidden neurons holds relevant information about the input image that needs to be analyzed. The feature elements are randomly selected from the hidden layers. The deep high-level features are extracted from the LSTM-CNN model as a conjunction of image similarity compared with various images from the database. The sequence of hidden neurons represents the image feature as h = {h1, h2, h3, …, hn}. To select the relevant features from the facial recognition process, similar pixels are gathered together. The amount of correlation present between the training image features and the testing image features is then higher, as indicated by the correlation score. The correlation parameters are evaluated using the formula below.

$$s(h_n, q) = h_n^T q \tag{6}$$

S (h_n, q) is the suppression ratio attained from the function of the LSTM process.
Once the scoring threshold of the relevant space is normalized, then the hidden state is represented as below.

$$: a = \max (s(h_n, q)) = \exp (s(h_n, q)) / \sum_{j=1}^{N} \exp (s(h_n, q)) \tag{7}$$

The hidden state function, hn, represents the suppression ratio. The LSTM model adaptively adds up or suppresses the feature weights. The iterations run till 1000. The deeper the input image pixel, the process continues.

$$att(h, q) = \sum_{n=1}^{N} a_n h_n \tag{8}$$

The proposed LSTM-CNN model gets tuned by the hyperparameters. The significant values are adjusted for smoothening and normalization of the pixels. The learning pattern of the input is adjusted at every iteration to achieve better accuracy. In the proposed approach, the hyperparameters are tuned to achieve maximum stability in the analysis model. The presented system is enhanced with the help of a hybrid optimization model (HALEC). The optimization process is represented by the formula below.

$$x_{t+1}^{i} = \left(x_{t+1}^{i} - a_i\right) . \frac{d_i - x_{t+1}^{i}}{d_i^{t} - a_i} + c_i \tag{9}$$

Where a belongs to the minimum value of the input feature points.
This is the most significant value associated with the ith iteration.

4. Results and Discussions

The presented system is implemented using MATLAB software with 8 GB of hardware RAM and an NVIDIA GPU to support high-quality images. The performance of the proposed approach is evaluated using the confusion matrix formulation, which includes actual positive rate, True negative rate, False positive rate, and False negative rate, among others. The quantitative measures obtained from the proposed analysis are Accuracy and Loss rate.
The parameter Accuracy is calculated using the formula below,

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

The number of positive features present in the given data is used to determine the emotion expressed in the face image. From the complete iterative analysis, the number of correlated features is summed up to show the model's performance.

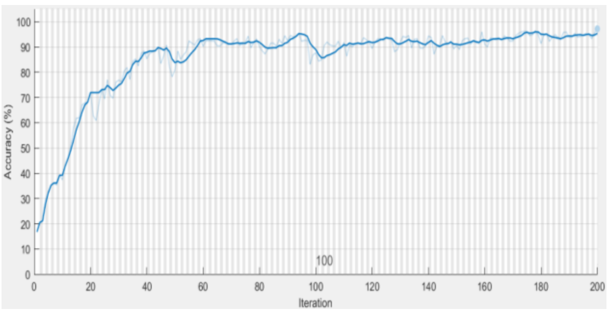

Fig. 2: Training Accuracy Score.

Fig 2. Shows the training accuracy of the proposed LSTM-CNN model with the HALEC optimization technique.
The number of training iterations conducted here is 200, and the accuracy starts rising from a value of 20 in the initial stage to a maximum of 98% by the time the iteration count reaches 200. The number of iterations depends on the complexity of the feature extraction process.

The overall images selected for testing are given to the training process. Part of the testing images is given to the testing process. The best evaluation is made once the maximum is reached.

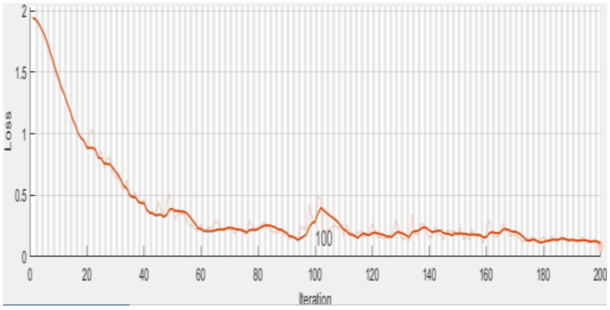

Fig. 3: Computation Loss.

Fig 3. Shows the computation loss of the proposed model after the HALEC optimization network. The primary goal of utilizing the optimization process is to tune the network until it reaches the best results in terms of accuracy. The most intense feature extraction through the MMNF model provides better accuracy and reduced loss value.

Table 1: Evaluation Metrics of the LSTM-CNN-HALEC Model

Metrics	LSTM_CNN_HALEC
Accuracy	97.91
Precision	85.38
Recall	94.84
F1-score	89.865

Table 1 shows the evaluation metrics of the proposed model, indicating an accuracy of approximately 97.97%, with a precision of 85.38%, a Recall value of 94.84%, and an F1 Score of 89.865%. These metrics indicate that the proposed model outperforms in terms of accuracy.

Table 2: Comparison of Novel Model Built-in Parameters

Metrics	ALEC	CNN	IBCNN	PPDN
Augmentation	No	No	No	No
Batch size	32	32	32	32
Epoch	200	100	100	100
Activation	ReLu	ReLu	ReLu	ReLu
Accuracy	97.91	93.20	95.10	97.30

Table 2 shows the comparison of various model architectures utilized in state-of-the-art approaches, such as the conventional CNN model, IBCNN [22], PPDN [22], Ant-lion optimization, and CNN [21], etc.

Fig. 4: Face Detection Output.

Fig 4. Shows the face detection output generated by the Viola-Jones, Haar cascade model.

Fig. 5: Region Extraction.

Figure 5 shows the region of interest extraction before applying the MMNF technique.

Fig. 6: MSER-MMNF feature.

Fig 6. Shows the MSER in the MMNF technique after the region extraction process.

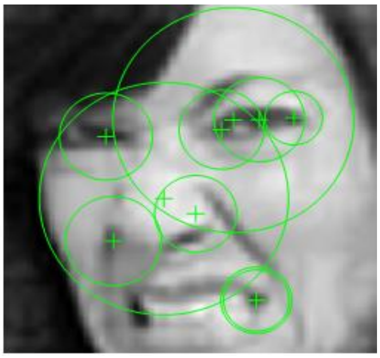

Fig. 7: SURF-MMNF Feature.

Fig 7. Shows the SURF-MMNF feature extracted from the regional extraction done from the input image.

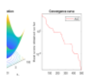

Fig. 8: HALEC-Based Optimized Results.

Fig 8. Shows the HALEC model optimization curve generated from the proposed method. After 500 iterations, the best score values are extracted from the analysis. The system's performance is improved with respect to the optimized values. The data augmentation performed here is one of the reasons for the optimal results obtained. The descriptors help read the deep parameters. Comparing it with state-of-the-art approaches in [21] and [22], the proposed model achieves better accuracy.

Limitations: Although the suggested HALEC model exhibits excellent performance in emotion recognition, several issues still exist that may compromise its application in certain situations. The main weakness lies in the computational complexity of the feature extraction methods involved, including SURF and MSER. These algorithms are powerful in their ability to capture facial features, and thus require ample computational resources, particularly when dealing with high-resolution images or massive datasets. It can cause increased memory usage and longer training times, which might restrict the model's usability in real-time or on devices with limited processing power.

Furthermore, even with the superior feature extraction and optimization algorithms used, the HALEC model can still be challenged in cases of occlusions or partial facial expression images. Accessories (e.g., glasses, masks) or unfavorable lighting conditions can lead to occlusions of essential facial features, resulting in the loss of vital information necessary to identify emotions accurately. Although HALEC implements techniques to address changes in image quality, the model may not perform well in cases of severe occlusion or highly adaptive environments, and further work is needed to enhance robustness and adaptation measures. The above shortcomings suggest that future research can focus on making the model more computationally efficient and effectively handling the occurrence of occlusions, particularly in real-world conditions where control is not easily achieved.

5. Conclusion

Emotion investigation is an intriguing area of exploration that contains different combinations of relevant information gathered from genuine episodes. Different existing implementations of emotion recognition manage to consider the evolving temperament. The proposed framework is recommended in consideration of the irregularity and position of different significant boundary points in existing systems for emotion recognition, such as empathic concern, specifically expectations towards others and fortune, as well as individual troubles, which are influenced by others' emotions. Emotions should be viewed profoundly to perceive people's inclinations accurately. In the introduced framework, the component extraction used for look taking care of is executed utilizing a neural network concerning facial emotions, such as cheerful, Trepidation, miserable, irate, shocked, and impartial. In real people, different looks can be produced during communication that vary with extremely distinct behaviors. The proposed framework-wise, half-and-half component extraction and look-recognizable proof strategy, utilizing Viola-Jones object detectors and MSER feature extraction method with accelerated strong feature extraction (SURF) procedure, addressed as multimodality network features (MMNF), is used for characterization. The recognized highlights are additionally arranged utilizing an LSTM long-term memory model with a CNN convolutional neural network crossover design. Further, the framework boundaries are streamlined through the Mixture Insect Lion improvement method. The introduced framework achieved 98% accuracy compared to various conditions of craftsmanship approaches.

It is possible to expand on the strengths of the HALEC model by considering several interesting areas of research that can be leveraged to enhance its capabilities in the future. The creation of real-time video analysis of emotion recognition is a promising field. Although the existing model is based on still images of faces, further development of the system to handle video streams in real-time would make it significantly more applicable in dynamic settings, including live interactions and video conferencing. This would entail optimizing the model to handle continuous video frames, reducing the computational load, and ensuring real-time execution without compromising accuracy. Another research direction of vital importance is investigating the cross-cultural applications of the model. Facial displays may vary significantly across different cultures, and ensuring that the model can be generalized to other ethnic and cultural groups is crucial for its universal usability. Future research would be required to gather and examine data on different populations to capture the cultural differences in facial expression interpretation in the model. This would aid in the creation of emotion recognition systems that are not only culturally inclusive but also more applicable to a wide range of demographic groups. These guidelines are designed to address the shortcomings of the existing system, and new directions become available through real-life implementation in various fields, including healthcare, human-computer interaction, and global communication platforms.

References

[1] Chu, Hui-Chuan, William Wei-Jen Tsai, Min-Ju Liao, and YuhMin Chen. ”Facial emotion recognition with transition detection for students with high-functioning autism in adaptive e-learning.” Soft Computing: 1-27, 2017.

[2] K. Prabhu, S. SathishKumar, M. Sivachitra, S. Dineshkumar, and P. Sathiyabama, “Facial expression recognition using enhanced convolution neural network with attention mechanism,” Computer Systems Science and Engineering, vol. 41, no. 1, pp. 415–426, 2022. https://doi.org/10.32604/csse.2022.019749.

[3] S. Li and W. Deng, “Deep facial expression recognition: a survey,” IEEE Transactions on Affective Computing, 2020.

[4] P. Foggia, A. Greco, A. Saggese, and M. Vento, "Multitask learning on the edge for effective gender, age, ethnicity, and emotion recognition," Engineering Applications of Artificial Intelligence, vol. 118, Article ID 105651, 2023. https://doi.org/10.1016/j.engappai.2022.105651.

[5] P. Sharma, "Student engagement detection using emotion analysis, eye tracking, and head movement with machine learning," in Proceedings of the Technology and Innovation in Learning, Teaching and Education: Third International Conference, TECH-EDU 2022, Lisbon, Portugal, September 2022. View at: Google Scholar. https://doi.org/10.1007/978-3-031-22918-3_5.

[6] M. Murugappan and A. Mutawa, “Facial geometric feature extraction based emotional expression classification using machine learning algorithms,” PLoS One, vol. 16, no. 2, Article ID e0247131, 2021.View at: Publisher Site | Google Scholar. https://doi.org/10.1371/journal.pone.0247131.

[7] E. R. Kimonis, B. Le, G. E. Fleming et al., “Facial reactions to emotional films in young children with conduct problems and varying levels of callous unemotional traits,” Journal of Child Psychology and Psychiatry, vol. 64, no. 3, pp. 357–366, 2023.View at: Publisher Site | Google Scholar. https://doi.org/10.1111/jcpp.13701.

[8] M. Sajjad, S. Zahir, A. Ullah, Z. Akhtar, and K. Muhammad, "Human behavior understanding in big multimedia data using CNN-based facial expression recognition," Mobile Networks and Applications, vol. 25, no. 4, pp. 1611–1621, 2019. https://doi.org/10.1007/s11036-019-01366-9.

[9] H. Joseph and B. K. Rajan, "Real-time drowsiness detection using Viola-Jones & KLT," in Proceedings of the 2020 International Conference on Smart Electronics and Communication (ICOSEC), IEEE, Trichy, India, September 2020. https://doi.org/10.1109/ICOSEC49089.2020.9215255.

[10] T. H. Obaida, "Real-time face detection in digital video based on Viola-Jones supported by convolutional neural networks," International Journal of Electrical and Computer Engineering, vol. 12, no. 3, pp. 2088–8708, 2022. https://doi.org/10.11591/ijece.v12i3.pp3083-3091.

[11] Oulefki, Adel, Mustapha Aouache, and Messaoud Bengherabi. "Low-Light Face Image Enhancement Based on Dynamic Face Part Selection." Iberian Conference on Pattern Recognition and Image Analysis, Springer, Cham, 2019. https://doi.org/10.1007/978-3-030-31321-0_8.

[12] Bendjillali, Ridha Ilyas, et al. "Illumination-robust face recognition based on deep convolutional neural networks architectures." Indonesian Journal of Electrical Engineering and Computer Science 18.2 (2020): 1015-1027. https://doi.org/10.11591/ijeecs.v18.i2.pp1015-1027.

[13] Z. Dong, X. Li, B. Lin, and F. Xie, "Multi-Net Fusion: Exploring a Brain-Inspired Neural Network Model for Facial Expression Recognition," in IEEE Access, vol. 11, pp. 93804-93814, 2023, https://doi.org/10.1109/ACCESS.2023.3310398.

[14] M. D’incà, C. Beyan, R. Niewiadomski, S. Barattin, and N. Sebe, "Unleashing the Transferability Power of Unsupervised Pre-Training for Emotion Recognition in Masked and Unmasked Facial Images," in IEEE Access, vol. 11, pp. 90876-90890, 2023, https://doi.org/10.1109/ACCESS.2023.3308047.

[15] R. Khoeun, W. Yookwan, P. Chophuk, A. Rodtook, and K. Chinnasarn, "Emotion Recognition of Partial Face Using Star-Like Particle Polygon Estimation," in IEEE Access, vol. 11, pp. 87558-87570, 2023, https://doi.org/10.1109/ACCESS.2023.3305514.

[16] Y. Ma et al., "What Can Facial Movements Reveal? Depression Recognition and Analysis Based on Optical Flow Using Bayesian Networks," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 31, pp. 3459-3468, 2023, https://doi.org/10.1109/TNSRE.2023.3305351.

[17] Kazemi, V.; Sullivan, J. One millisecond face alignment with an ensemble of regression trees. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; 2014; pp. 1867–1874. https://doi.org/10.1109/CVPR.2014.241.

[18] Zhang, K.; Zhang, Z.; Li, Z.; Qiao, Y. Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Process. Lett. 2016, 23, 1499–1503. https://doi.org/10.1109/LSP.2016.2603342.

[19] F. M. Javed Mehedi Shamrat, Z. Tasnim, P. Ghosh, A. Majumder, and M. Z. Hasan, "Personalization of Job Circular Announcement to Applicants Using Decision Tree Classification Algorithm," 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangluru, India, 2020, pp. 1-5 https://doi.org/10.1109/INOCON50539.2020.9298253.

[20] Raúl Cid Carro, Juan-Manuel AhuactzinLarios, Edmundo Bonilla Huerta, Roberto Morales Caporal and Federico Ramírez Cruz "Face Recognition Using SURF" ICIC 2015, Part I, LNCS 9225, pp. 316–326, 2015, https://doi.org/10.1007/978-3-319-22180-9_31.

[21] Minaee S, Minaei M, Abdolrashidi A. Deep-Emotion: Facial Expression Recognition Using Attentional Convolutional Network. Sensors 2019;21. https://doi.org/10.3390/s21093046.
[22] Rio Febriana *, Benedic Matthew Halima, Maria Christinaa, Dimas Ramdhana, Andry," Facial expression recognition using bidirectional LSTM – CNN", 1877-0509 © 2023 The Authors. Published by Elsevier B.V.
[23] Ogbonnaya, E., &Wai, Y. M. (2024). Design and Optimization of Energy Harvesting Circuits for Ultra-Low-Power Wearable Electronics. Electronics, Communications, and Computing Summit, 2(1), 123–130.
[24] Vishnupriya, T. (2025). Grid-Connected Inverter Topologies for High-Penetration Solar PV Systems: A Comparative Study. Transactions on Power Electronics and Renewable Energy Systems, 44-50.
[25] Alwetaishi, N., &Alzaed, A. (2025). Smart construction materials for sustainable and resilient infrastructure innovations. Innovative Reviews in Engineering and Science, 3(2), 60–72.
[26] Kavitha, M. (2024). Enhancing Security and Privacy in Reconfigurable Computing: Challenges and Methods. SCCTS Transactions on Reconfigurable Computing, 1(1), 16-20.
[27] Carvalho, F. M., & Perscheid, T. (2025). Fault-tolerant embedded systems: Reliable operation in harsh environments approaches. SCCTS Journal of Embedded Systems Design and Applications, 2(2), 1–8.
[28] Usikalu, M. R., Okafor, E. N. C., Alabi, D., & Ezeh, G. N. (2023). Data Distinguisher Module Implementation Using CMOS Techniques. Journal of VLSI Circuits and Systems, 5(1), 49–54. https://doi.org/10.31838/jvcs/05.01.07.
[29] Reginald, P. J. (2025). Wavelet-based denoising and classification of ECG signals using hybrid LSTM-CNN models. National Journal of Signal and Image Processing, 1(1), 9–17.
[30] Madhanraj. (2025). Design and Simulation of RF Sensors for Biomedical Implant Communication. National Journal of RF Circuits and Wireless Systems, 2(1), 44–51.
[31] Sindhu, S. (2025). Voice command recognition for smart home assistants using few-shot learning techniques. National Journal of Speech and Audio Processing, 1(1), 22–29.
[32] Velliangiri, A. (2025). Low-Power IoT Node Design for Remote Sensor Networks Utilizing Deep Sleep Protocols. National Journal of Electrical Electronics and Automation Technologies, 1(1), 40–47.