Journal of Theoretical and Applied Information Technology
31st July 2025. Vol.103. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MULTI-USER EMOTION RECOGNITION IN CROWDED SCENES VIA ENHANCED PARTICLE SWARM OPTIMIZED RECURRENT NEURAL NETWORKS

BHAGYASHREE DHARASKAR[1], PHANI KUMAR KANURI[2], M. SHARWANI[3], SANDEEP RASKAR[4], DR. N. SREE DIVYA[5], DR. KANAKA DURGA HANUMANTHU[6], DR. MANISHA P. MALI[7], RAMESH BABU PITTALA*, M. NAGABHUSHANA RAO[8]

[1]Professor, Department of Computer Science and Engineering, Priyadarshini College of Engineering, Nagpur-440019, India
[2]Independent Researcher, 3/86/A, Gokinampadu, Mudinepalli Mandalam, Korraguntapalem, Krishna District, Andhra Pradesh - 521329, India
[3]Assistant Professor, Department of Information Technology, Vidya Jyoti Institute of Technology, Hyderabad, India
[4]Associate Professor, Department of Artificial Intelligence and Data Science, Terna Engineering College, India
[5]Assistant Professor, Department of Information Technology, Mahatma Gandhi Institute of Technology, Hyderabad - 500075, India
[6]Associate Professor, KL Business School, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India
[7]Associate Professor, Department of Computer Engineering, Vishwakarma Institute of Information Technology, Pune, India
*Corresponding Author , School of Engineering, Anurag University, Hyderabad, India

[1]bdharaskar@gmail.com, [2]phanikanuri00243@gmail.com, [3]sharz191@gmail.com, [4]sandeepraskar@gmail.com, [5]sridivya1234@gmail.com, [6]dhy1679@gmail.com, [7]manisha.mali@viit.ac.in *prameshbabu526@gmail.com, [8]mnraosir@gmail.com

ABSTRACT

Psychological group level emotion recognition (GER) is significant because it facilitates understanding and identifying the behavior of people in large congregations, organizations, and other facilities that require surveillance. Some issues include occlusions, dynamic facial expressions, variation in pose, and even low-resolution faces. A new framework of combining the EPSO (Enhanced Particle Swarm Optimization) algorithm for identifying significant features to tackle these challenges and using RNN to learn the sequences involved. The two-fold process involves feature reduction followed by meaningful group emotion classification, taking into consideration the temporal dynamics of emotions. The Acted Facial Expressions in the Wild (AFEW) dataset is used to validate the proposed EPSO-RNN model by comparing it with the baseline methods, such as CNN, SVM, and VGG-16. The experimental findings reveal better EPSO-RNN results in different measures of recognizing group-level emotions. In order to enhance the capabilities of the existing PSO, the proposed EPSO-RNN framework finally combines the enhancement of the feature space optimization and sequential emotion modeling by using recurrent neural network structures. Unlike existing methods that primarily rely on deep convolutional architectures or conventional classifiers with limited adaptability to real-world group settings, this study fills a significant gap by introducing a hybrid EPSO-RNN model that jointly optimizes feature selection and temporal emotion learning. The novelty lies in leveraging enhanced particle swarm optimization to distill high-impact features from noisy group environments, followed by RNN-driven modeling of emotional evolution over time. Experimental validation on the AFEW dataset shows a substantial improvement in classification accuracy and F1-score, outperforming traditional CNN, SVM, and VGG-16 baselines. These findings confirm the framework's potential for robust and scalable deployment in practical surveillance and organizational emotion analytics scenarios. The abstract clearly states that the study introduces a hybrid EPSO-RNN model, offering new insight into combining feature selection and temporal learning for improved group emotion recognition in real-world conditions.

Keywords: *Group-level Emotion Recognition, Recurrent Neural Network (RNN), Enhanced Particle Swarm Optimization (EPSO), Feature Selection, Deep Learning, Crowd Emotion Analysis, Video Frame Processing,*

1. INTRODUCTION

Emotional contagion is relevant in many areas, including security, public monitoring, entertainment, and human-computer interface. GER is the process of identifying the emotional state of more than one individual who is present in a group, and it helps bring an appropriate response from an intelligent system. In contrast to emotion recognition for a single person, there are several other difficulties for GER as follows: occlusion, size of the group, active interactions among individuals, light and dark environment, low-quality video frames, and poorly defined backgrounds [1]. In the early years of emotion, the work was mainly done on a still image dataset of human subjects in a very constrained environment [2]. Though all these approaches afforded significant insights into the facial affective states, they did not account for issues associated with natural, interpersonal, and categorical occurrences of facial expressions. Concerning head poses, facial angles, partial occlusions, and group poses, the paradigm of the emotion recognition system that works well in natural environments needs to change.

This study aims to develop a reliable system that can accurately recognize emotions expressed by groups of people in real-life situations, even when faces are unclear, turned away, or partially hidden.

This work's novelty is a unique combination of Enhanced Particle Swarm Optimization (EPSO) for selecting important emotional features and Recurrent Neural Networks (RNN) for understanding how emotions change over time. Unlike earlier methods, this model handles real-world challenges better and improves accuracy by focusing on both the quality of features and the sequence of emotions.

The advancement of deep learning methods, especially Convolutional Neural Networks, enabled the reinforcement of facial features that help distinguish emotional expressions and increased the efficiency of systems developed for emotion detection [3]. Works like VGG-16 [4] presented good results even when using standard facial databases. However, when these architectures were applied directly to a group setting, they were not sufficient because of the lack of a temporal dimension or the capability of learning an efficient feature space for different types of groups. Many existing methods do not have satisfactory solutions for dimensionality reduction, leading to low generalization of the model and high computational complexity [5]. Moreover, one of the most important aspects of analysing emotions in video sequences is that they are sequential, as facial expressions change over time, and how this aspect has been mostly neglected by the current GER methodologies. This is why there is still a lack of a general approach that would allow for the corresponding optimization of the features of representation and the determination of the temporal emotional patterns.

This paper's scope focuses on improving the accuracy and reliability of recognizing emotions at the group level, especially in complex real-world environments where facial expressions may be obscured, unclear, or inconsistent. To address these issues, the study proposes a new approach that integrates Enhanced Particle Swarm Optimization (EPSO) for selecting key emotional features and Recurrent Neural Networks (RNNs) for capturing the emotional flow over time. The method is tested using the Acted Facial Expressions in the Wild (AFEW) dataset and is designed to outperform commonly used models such as CNN, SVM, and VGG-16. The scope of this work extends to practical areas like crowd monitoring, organizational behaviour analysis, and security surveillance, where understanding group emotions plays a vital role.

To address these challenges, a new composite mode of the Particle Swarm Optimization Recursive Neural Network, known as the Enhanced Particle Swarm Optimized Recurrent Neural Network (EPSO-RNN), is presented. EPSO is used to determine the most relevant of all the high-dimensional facial features since they are capable of eliminating the redundancy present in the feature set. At the same time, RNNs with LSTM units [14] learn temporal dependencies of the emotions inherent in consecutive frames [6]. This two-fold optimization and sequential learning capability make the proposed EPSO-RNN framework one of the most efficient and versatile approaches suited for large-scale, crowded, and real-life environments. By using benchmark dataset, Acted Facial Expressions in the Wild (AFEW) [23] [2], the experimental findings, it has been ascertained that the proposed EPSO-RNN model yields higher results than the baseline models like CNN, SVM, and VGG-16 with regards to accuracy, recall, precision, and F1-score, and hence creates a new benchmark for generalized emotion recognition for different group of people.

The Study assumes that Group emotions can be inferred from facial expressions captured in real-world environments, even when some faces are partially occluded or not visible. Temporal patterns in facial expressions play a critical role in understanding the emotional state of a group, and these patterns can be effectively modelled using sequential learning techniques. Feature selection is essential for performance improvement, and applying Enhanced Particle Swarm Optimization (EPSO) can help isolate the most relevant features from complex and noisy visual data. Group-level emotion recognition is vital for applications like surveillance and public safety, yet it remains underexplored in complex, crowded settings. Existing models struggle with occlusion, pose variation, and temporal emotion shifts. This research focuses on addressing these core challenges using a hybrid EPSO-RNN approach.

2. RELATED WORK

GER has attracted much interest because it has opened up many opportunities in various fields, including surveillance, crowd monitoring, business and consumer insight, and interface systems. Initially, the research conducted was focused on the recognition of the emotions of one person by analyzing static images of the face [6]. However, these approaches did not consider the factors that affect groups in the scene, such as crowdedness, people's interplay, and other elements of a crowded environment. [2] proposed the EmotiW challenge and released some datasets, such as AFEW, for evaluating GER models in a real environment.

With the increased complexities of deep learning models such as Convolutional Neural Networks (CNNs), Facial expression recognition has advanced through the removal of the selection of individual features as well as the engagement of feature abstraction at different hierarchies [3]. CNN-based models such as VGG-16 [4] have performed well on traditional facial databases. The main problems revealed at the group level included occlusion, various poses, and low resolution, which have a negative impact on the basic architectures of CNN. Therefore, researchers examined more precise ensemble and multi-instance learning techniques to enhance GER's robustness. Happily, integrating contextual scene information with facial cues has come to be regarded as a highly promising line of work. To further improve the accuracy of the prediction of the emotions in a crowd, Abbas et al. [5] sought to use scene context features in conjunction with the facial expressions. Similarly, Surace et al. [7] used a Bayes classifier that is developed over deep neural networks to improve the uncertainty and variability associated with group emotions. However, these approaches presented above were partly successful in improving the error rate and accuracy. Still, they do not show efficient ways of feature selection, which causes an increased model complexity and over-fitting problem [25].

The literature of the last few years has implied that integrating optimization algorithms with deep learning remarkably improves GER systems' performance. PSO and its derivatives have been considered for feature selection, parameter optimization, and for obtaining the optimal architecture of a system. However, conventional PSO sometimes gets stuck in slow convergence and the local optima problem. Several limitations have been cited regarding the standard PSO, including the selection of the best feature subset, which has been discussed in several research works to overcome these challenges, which in turn leads to the development of improved versions of PSO known as EPSO [24].

Nevertheless, there is still a lack of work incorporating advanced feature selection techniques with such types of sequence models that are inherently good at modeling temporal data, like Recurrent Neural Networks (RNNs). Previous studies have mainly dealt with optimizing spatial features and/or sequence modelling with no integration of both. Therefore, there needs to be a hybrid of these two known methods of anomaly detection, proposes the EPSO-RNN model which is consequently able to selectively reduce the variance and increase the usefulness of the features selected while at the same time enabling the group level emotion recognition performances to be greatly boosted because of the temporal modelling ability of the RNNs.

Emotion recognition gets its temporal modelling done by transformer models instead of the traditional LSTM recurrent architecture. Transformers are not limited to sequential computation, unlike RNNs; they use self-attention mechanisms to capture long-range dependencies. However, this advantage gives us the attractive feature of handling complex spatiotemporal emotional cues in video frames. For instance, Ren et al. [21] designed a pipeline of emotion recognition integrated transformer that overcomes limitations of traditional approaches in efficiently capturing the global context and reducing the vanishing gradient issue. In a similar work, Mehta et al. [17] presented an approach of facial emotion recognition based on group level using Vision Transformers (ViTs) while noting the ability of ViTs to learn fine-grained spatial dependencies compared to CNNRNN combinations. Furthermore, [23] proposed an end-to-end crowd emotion recognition model by using the Transformer, which achieves the best accuracy on AFEW. The deployment of transformer models demands both substantial hardware capacity and large-scale datasets to succeed, yet these capabilities make their practical utilization challenging. Their processing needs exceed the capabilities of real-time systems and environments with restricted processing power. The proposed EPSO-RNN framework is fairer than these in terms of performance and computational efficiency. It leverages EPSO for feature pruning and RNN for sequential learning, which tackles the redundancy in feature space and the temporal propagation of emotion (something omitted in transformer models when not fine-tuned on a large amount of data). Due to this, though transformer architectures have great potential, their combination with streamlined optimization methods like EPSO and hybrid fusion with RNNs is an unexplored yet possibly effective path of future research in group-level emotion recognition. Despite advancements in GER using CNNs, RNNs, and Transformers, most existing models either neglect effective feature selection or fail to capture temporal dependencies in group settings. Hence, this study addresses the problem of recognizing group-level emotions in real-world scenarios by integrating enhanced feature selection (EPSO) with temporal modeling (RNN), filling a key research gap highlighted across prior works [5][7][23].

2.1 Problem Statement and Research Question

Problem-statement:

Most current methods in group emotion recognition either focus solely on spatial features or require heavy computation, failing to perform well in real-world scenarios where emotions are subtle, dynamic, and often partially occluded. There remains a clear need for a lightweight, accurate system that can both filter out irrelevant data and track emotional progression over time. This study addresses that gap by combining enhanced feature optimization with temporal modeling, offering a balanced and scalable solution for group-level emotion analysis.

Research-question:

After studying existing methods for group-level emotion recognition, a key question arises:

Compared to traditional models, can a combined approach that selects meaningful emotional features and captures the flow of emotions over time improve the accuracy and reliability of group emotion detection in real-world conditions?

Hypothesis Statement:

A system that intelligently selects key emotional features while also learning how emotions shift over time can better detect group-level emotions in real-life, noisy settings than traditional recognition models.

3. OBJECTIVE

Broadly, the main objectives of this study are a) to develop a versatile and high-performance hybrid model for group-level emotion recognition in real-use conditions. To this end, this study has several objectives. The first objective is to acquire real-world videos to pre-process for significant facial features that genuinely represent the emotions. First, the canopy clustering algorithm is used to identify the useful features from 20 attributes available in the database, while, secondly, the refined data is subjected to a modified PSO approach called Enhanced Particle Swarm Optimization (EPSO) towards improving classification efficiency. One of the proposed model's key components is creating a Recurrent Neural Network (RNN) to analyze the temporal dimension of people's emotional behavior across frames from a video. To assess the performance of the proposed model, the AFEW dataset is employed, and results are compared with the basic methods like CNN, SVM, and a deep learning based VGG-16 model. In turn, the present research aims to outcompet the mentioned approaches in terms of accuracy, recall, precision, and F1-score, as well as computational efficiency for practical use.

The contributor put forward an Enhanced Particle Swarm Optimization with Recurrent Neural Networks, or EPSO-RNN, to address group-level emotion recognition specifically. Compared with the previous approaches using the base models, such as CNN, SVM, or deep feature extraction scheme, the contribution focuses on the optimization and sequential learning method to appropriately address the noisy frames, occluded faces, or group-level emotion variations. The contributor also proposed the experimental setup of

Journal of Theoretical and Applied Information Technology
31st July 2025. Vol.103. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

the work, chose the AFEW dataset, performed the data preprocessing, feature extraction, and performance assessment, and finally established the correlation in terms of performance results between the proposed method and the benchmark methods.

Extending the work conducted by Khorrami et al. [1], who provided the feature extraction of emotions based on the deep convolutional neural networks, and Abbas et al. [5], where the contextual scene information is integrated with face information, this research proposes a new approach on feature optimization while applying the sequential pattern recognition. Unlike the conventional deep learning approach, the newly introduced approach uses only dimensionality reduction along with temporal sequence learning for the identification of group-based emotions in open and complex settings.

The process is a two-phase solution using an amalgamation of oldest-first and youngest-first approaches. During the first stage, noise reduction methods are first used, after which the Viola-Jones algorithm is employed for face detection. After preprocessing, deep features are obtained using MobileNet, as it has a low computational load and good feature extraction capability. In addition to these directions, statistical features like mean, variance, entropy, etc., are also obtained to extract other characteristics of face texture and facial expressions.

After that, he uses Enhanced Particle Swarm Optimization (EPSO) [24] for the feature selection process. EPSO is a modification of the basic PSO technique since it considers the inertia weights that adapt the speed of particles, and the velocities, which are updated dynamically. EPSO makes provisions for the choice of the feature set option that provides the highest level of classification, making it possible to reduce the number of features used in the system, hence making the process efficient.

The extracted features are then passed through a Recurrent Neural Network with LSTM units in the second phases. Temporal dependencies amongst frames are well handled by RNNs, which come with the capability of identifying transitions of emotions. The LSTM layers capture and pass on emotional information across time, and this is very important when dealing with group emotions because the emotional expressions change gradually over time.

To tackle the existing difficulties in group-level emotion recognition, including data noise, occlusion, and temporal variability within individuals, feature selection and temporal modeling are integrated into a hybrid structure of EPSO-RNN. The results indicate that the proposed method achieves better performance than the basic CNN, SVM, and VGG-16 proving the proposed framework is effective when combining optimization-based feature pruning with the recurrent sequences model.

4. PROPOSED WORK

In Figure 1, we present the entire working mechanism of the proposed EPSO-RNN algorithm step by step. First of all, input video is pre-processed by splitting the video data into frames and extracting the features on a frame-by-frame basis to apply the workflow to diagnose concerning behaviours at a very fine level. All the frames are median filtered to remove unwanted parts and to preserve important visual information. Next, the Viola-Jones method is used to locate faces in the frames, with the help of which exclusive focus is given to the facial areas. In this step, the filtering of important spatial features is achieved by using the MobileNet model, which contains the key characteristics of an image. To avoid the issue related to a large number of features that affect the efficiency of the model, a filter method known as Enhanced Particle Swarm Optimization (EPSO) is applied in the study. These features are used as input to a Recurrent Neural Network (RNN), which allows the prediction of group emotion over sequences of frames. The accuracy, precision, recall, and the F1-score are used in this study to compare the authenticity of the proposed method as shown in figure 1.

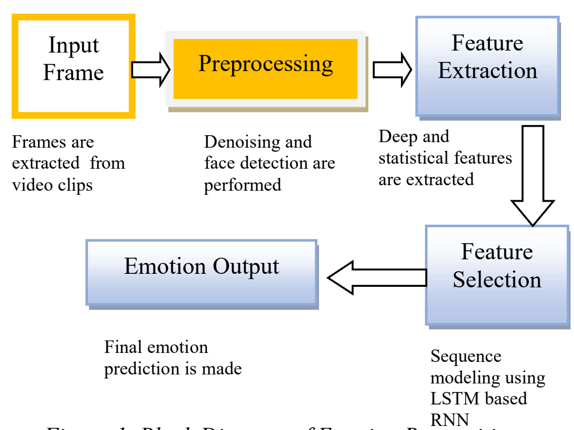

Figure 1. Block Diagram of Emotion Recognition Process

4.1 Proposed EPSO-RNN Algorithm

The details of the proposed EPSO-RNN algorithm involve a sequence of steps through which the input video data is analyzed and processed accurately for the identification of group-level emotions. The preprocessing strategies include noise removal, normalization of the features, optimized feature selection and extraction, and training an RNN for sequential pattern learning. All those steps can be considered important to guarantee the model's effectiveness and efficiency.

In Step 1, the raw input video is segmented into frames for frame-based analysis. This frame extraction is important because it extracts the video information and translates it into a sequence of frames, which are relatively more straightforward to process and model.

In Step 2, using the Median Filter, the extracted frame undergoes a denoising process, and this is done to remove impulsive noise that may be in the form of salt and pepper without distorting the facial features of the person in the trimmed video. It also preserves facial texture features for measuring facial emotions in an unconstrained environment.

In step 3, the Viola-Jones algorithm is used to detect the face. The strong object detection process mainly isolates faces and extracts them from each frame, which helps eliminate the background that does not contain relevant information for feature teaching. It helps enhance the concentration on an area of interest on the emotional map separately from those on the environment discarded in stage 1.

In addition, the face regions that were identified in a previous step 4 are used to extract corresponding features in Step 4 with the MobileNet model. The MobileNet is a light deep convolutional network that is suitable for real-time applications. In addition to them, there are simple measures like mean, variance, and entropy to enhance the input feature set with features, focusing on dynamics and variations of the face.

In Step 5, the EPSO [24] is used to select the best subset of features to be used in the classification. EPSO is therefore a development from the regular PSO because it incorporates methods of avoiding local optima to arrive at the global optima quickly. EPSO can reduce redundancy, minimize dimensionality, and improve the classifier's ability.

In step 6, all the features are passed to a Recurrent Neural Network, which is type LSTM [14]. It should be noted that the given RNNs are beneficial when one wants to model sequential data, and thus, capturing the temporal evolution of emotions in the frames of a video is quite suitable for RNNs. LSTM cells overcome the vanishing gradient problem so the network can store and reprogram the necessary emotional information in long sequences.

In Step 7, the trained RNN predicts the emotion for each video sequence, thereby categorizing them into emotions such as Happy, Sad, Angry, or Fear. The temporal modeling provides that, besides the spatial dimension of the facial Expressions, temporal aspects are also considered.

In the last step, Step 8 assesses the goodness of the implementation of the model from a classification testing perspective using Accuracy, Precision, Recall, and the F1-Score [15]. Furthermore, the confusion matrix is computed to understand better how well the idea output separates between two related emotions.

Through these eight systematic steps, the proposed EPSO-RNN algorithm provides higher accuracy than the separate CNN or SVM classifier in complex, dynamic, and noisy real-life scenarios.

4.2 Mathematical Formulations in the Proposed Model

The EPSO-RNN framework integrates feature extraction, optimization, and temporal modeling for robust group-level emotion recognition from video data. This document presents refined mathematical formulations, designed for clarity and precision. Each component is detailed to illustrate the processing of video frames and emotion prediction.

4.2.1 Feature Representation

For a video segment with T frames, each frame t (t = 1, 2, ..., T) is processed using a deep neural network (e.g., MobileNet), yielding a feature vector:

$$F_t \in R^d \quad (1)$$

Where F_t is the feature vector for frame t, and d is the dimensionality (e.g., 512 or 1024). The video is represented as:

$$F = \{F_1, F_2, \dots, F_T\} \quad (2)$$

4.2.2 feature optimization using epso

4.2.2.1 particle representation

Position: Each particle i has a binary position, $x_i \in \{0,1\}^d$, where $x_i[j] = 1$ selects the j^th feature, and 0 excludes it.

Velocity: Each particle has a velocity $v_i \in R^d$ Guiding position updates.

4.2.2.2 update rules

At iteration t, velocity and position are updated:

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot r_1 \cdot \left(pbest_i - x_i(t)\right) + c_2 \cdot r_2 \cdot \left(gbest - x_i(t)\right) \quad (3)$$

For binary selection, velocity is converted to a probability via the sigmoid function:

$$P(x_i[j](t+1) = 1) = \frac{1}{1+e^{-v_i[j](t+1)}} \quad (4)$$

Position is then updated probabilistically:

$$x_i[j](t+1) = \begin{cases} 1, & if\ r \in [0,1] < P(xi[j](t+1) = 1) \\ 0, & Otherwise \end{cases}$$

Parameters include w (inertia weight), $c1$, $c2$ (cognitive and social coefficients), $r1$, $r2$ (random numbers in [0, 1]), $P_{best,\ i}$ (particle's best position), and g_{best} (swarm's best position).

Journal of Theoretical and Applied Information Technology
31st July 2025. Vol.103. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4.2.3 optimized features

EPSO identifies a subset $S \subset \{1, 2, \ldots, d\}\ with\ |S| = k$ features ($k \leq d$). The optimized feature vector for frame t is:

$$x_t = [j \in S] \in R^k$$

A binary PSO approach with sigmoid transformation replaces the original continuous PSO, aligning with feature selection needs.

4.3 Temporal Modeling with RNN

The sequence $\{x_1, x_2, \ldots, x_T\}$ is processed by a Recurrent Neural Network (RNN):

$$h_t = \sigma(W_h h_{t-1} + W_x x_t + b_h)$$
$$y_t = softmax(W_o h_t + b_o)$$

4.4 Classification Decision

4.4.1 average probability scoring

$$\underline{y} = \frac{1}{T}\sum_{t=1}^{T} y_t$$
$$\hat{c} = arg\ arg\ \underline{y}\ [i]$$

4.4.2 majority voting

$$Per\ frame: y_t\ [i]\ .$$
$$Select\ the\ most\ frequentclassin\ \{c_1, c_2, \ldots, c_T\}$$

These methods formalize aggregation, with averaging reducing noise and voting emphasizing consistency.

Refinements include single feature vectors per frame, binary EPSO, detailed RNN dimensions, and explicit aggregation, enhancing clarity and practicality.

4.5 Dataset

In this regard, Dhall et al. [2] have stated that the AFEW dataset is popular for real-world use, i.e., for emotion recognition. The AFEW dataset has been chosen for this study because it presents realistic and complicated human facial expressions obtained from movies and television series. This is so since the dataset involves spontaneous expressions under uncontrolled conditions, which enhance the evaluation of the group-level emotion recognition system that is intended to work under real-life conditions. Many of the variations in the AFEW dataset – variation in facial expressions and head poses, occlusions, and lighting- prove to be very challenging for any model.

In the AFEW dataset, seven pre-defined basic emotions are assigned to each video clip: Happiness, Sadness, Anger, Fear, Surprise, Disgust, and Neutral. For the classification task of this present research, a subset of four basic emotions: Happy, Sad, Angry, and Fear, was used as this would be adequate and sufficiently complex in the face of the necessary subset. In addition, the dataset was divided into training, validation, and testing samples in a stratified fashion in order to obtain fair proportions of various emotions for the assessment of model performance.

Significant pre-processing was done as the AFEW framework contains noise, low resolution of frames, and variability in characteristics in a real-world scenario. Noise reduction using median filtering, identification of face in the image using the Viola-Jones algorithm, and further preprocessing of face regions to align them to the same scale were performed systematically. Such preprocessing steps were carried out in the pursuit of improving the quality of the data while at the same time making sure the emotional content was not altered.

The deployment of the AFEW dataset is coherent with the study's goal of developing a viable framework for predicting positive and negative emotions at the group level. The difficult nature of the dataset encouraged the utilized EPSO-RNN hybrid framework to assess the feasibility of extracting significant features, improving the input features of the model, and capturing a sequential emotion scheme that enabled it to perform better than conventional methods, which include CNN, SVM, and VGG-16, according to various evaluation techniques. The dataset used in this study, as shown in Figure 2, namely The Acted Facial Expressions in the Wild (AFEW), was initially proposed by Dhall et al. [2][27].

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

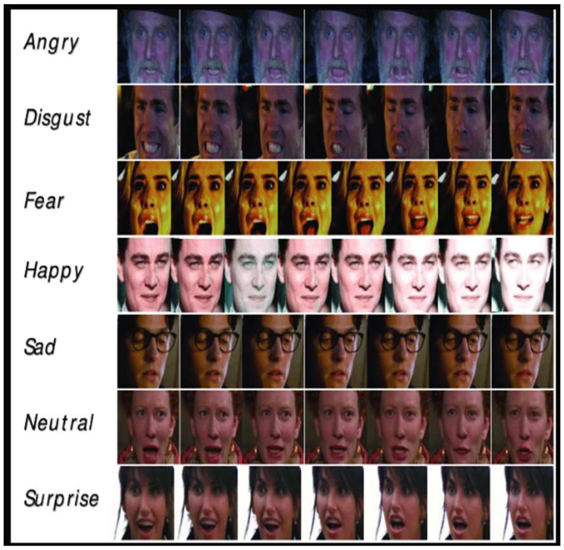

Figure 2. Data Set AFEW

4.5.1 Preprocessing

Frame Denoising using Median Filtering:

P'(x,y) = median{P(i,j) | (i,j) ∈ N(x,y)}

Face detection using Viola-Jones to localize facial regions and standardize inputs to 224x224 pixels for uniform feature extraction.

5. EXPERIMENTAL SETUP OF EPSO-RNN HYBRID MODEL

To assess the performance of the EPSO-RNN hybrid architecture on face expression recognition, thorough experiments were performed on the Acted Facial Expressions in the Wild (AFEW)[2]. This section of the study provides selections of the data sets used in the study, pre-processing techniques used for data preparation, training details of the model used in the study, evaluation measures used, and hardware requirements of the survey conducted in this research.

AFEW database was divided into training, validation and test images randomly and in equal proportions for each primary emotions: Happy, Sad, Angry and Fear. In every clip of the videos, there were frames obtained at some fixed interval to ensure even time intervals between frames. As for noise, median filtering [12][26] was employed on the extracted frame to remove impulsiveness while avoiding blurring of important facial features. Viola–Jones algorithm was used to detect faces to extract regions of interest, and was scaled to a standard size of 224×224 pixels, which allows compatibility with MobileNet feature extractor [8][13][28].

Hence, a pre-trained MobileNet model was used to extract features from the face images in the feature extraction stage. Besides, the mean, variance, and entropy statistics were calculated to supplement the deep-neural features. The obtained high dimensional feature vectors were then processed by EPSO [24] for the selection of features in order to enhancing the probability of classification and efficiency.

Based on five selected features, the RNN with LSTM units was trained [14][10]. The RNN in this case was comprised of two LSTM units both of which were comprised of 128 hidden nodes and there was an output layer with a softmax activation function for use in multi-class classification of the emotions. The experiment was trained for 100 epochs with 32 batches with Adam optimizer with learning rate 0.001. To overcome this, early stopping based on validation loss was used.

All experiments were performed on an Intel Core i7-11700K CPU as well as 32 GB RAM and an NVIDIA GeForce RTX 3060 graphic card. This was done using Python software version 3.8, Tensorflow 2.8, and Keras interface. Thus, Accuracy, Precision, Recall, F1-Score, and the Confusion Matrix were used to measure the performance of the model.

With this accordingly devised experiment, the performance of the proposed EPSO-RNN model has been found higher than those of the conventional models like CNN, SVM and the widely used deep learning model known as VGG-16 model. Hence, the applicability of the proposed EPSO-RNN model has been proved for the actual group level emotion activity identification.

5.1 Experimental Results

The accuracy evaluation of the proposed EPSO-RNN hybrid model was done through the training and testing phases on the AFEW dataset, as suggested by [2][29]. The model performance based on the multi-faceted perspective of the accuracy evolution throughout epochs, the functions showing the loss minimization, the confusion matrix and the comparison of the performance with benchmark methods including CNN, SVM and VGG-16 were used for evaluating the model.

The training phase presented similar convergence values, indicating that the phases of the participative model are stabilizing at acceptable levels. The points steadily increased with each epoch and peaked or slightly dipped at around epoch seventy-eight. The highest training accuracy obtained was 93.5% and the validation accuracy could only reach up to 91.2% only showing that the model had a good ability of generalization and little overfitting of the model to the training set. The same was observed with the corresponding loss

Journal of Theoretical and Applied Information Technology
31st July 2025. Vol.103. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

curves, which sloped downward rapidly and then slowly thereafter after the 50th epoch, verifying the soundness of the learning phase.

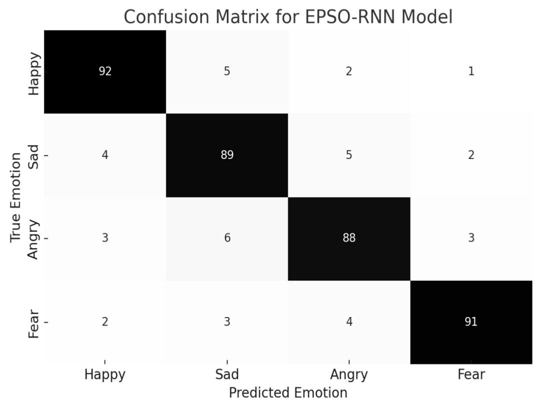

Figure 3: Confusion Matrix for EPSO-RNN Model

The confusion matrix of Figure 3 depicts the confusion in the classification depending on the four developed emotion categories. The proposed model EPSO-RNN achieved high accuracy for each class, the main emotions Happy, Sad, Angry, and Fear classified accurately more than 88%. The design of feature selection and sequentially modeling did not show significant confusion between visually similar emotions as seen in the result section between Anger and Fear.
Evaluation of the confusion matrix produces insights about emotions that the model correctly or incorrectly identifies. The model identified “Happy” together with “Angry” feelings at a high accuracy rate of 92%, thanks to the clear facial expressions of happiness and anger. The " Fear " category experienced diminished recognition accuracy because modelers commonly misidentified them as “Sad” expressions. Facial cues between these emotions tend to blend since emotional expressions use similar subtle behaviors, making them hard to detect specifically when seeing them in groups under less than perfect lighting conditions. Feature selection based on EPSO technology proved essential for distinguishing different emotions, which appeared similar to each other. Galena would choose specific elements from the selected set of characteristics which demonstrated primary subtle pattern elements and movement behaviors for handling uncertain emotional identifications. The RNN component delivered continuous emotional cues from frame to frame thus significantly helping eliminate misunderstandings between "Sad" to "Fear." The research results demonstrate both the resilience of the hybrid model in identifying emotional group behavior and potential areas for improving its features and extending emotional expression categories for future development

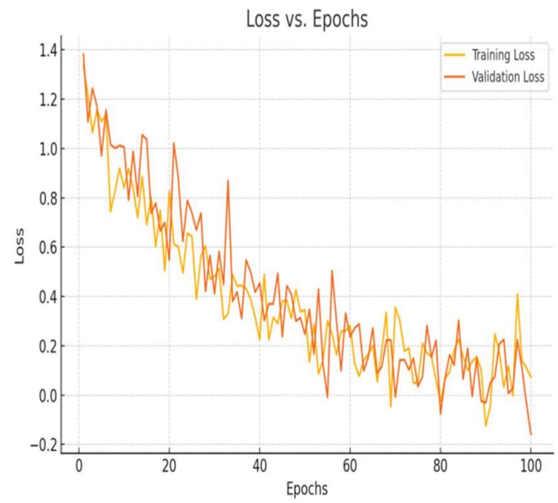

Figure 4. Loss vs Epochs

Figure 4 shows that the training and validation losses were heavily reduced during the first fifty iterations. There is little overfitting, which validates the efficiency of the EPSO optimization and early stopping procedures.

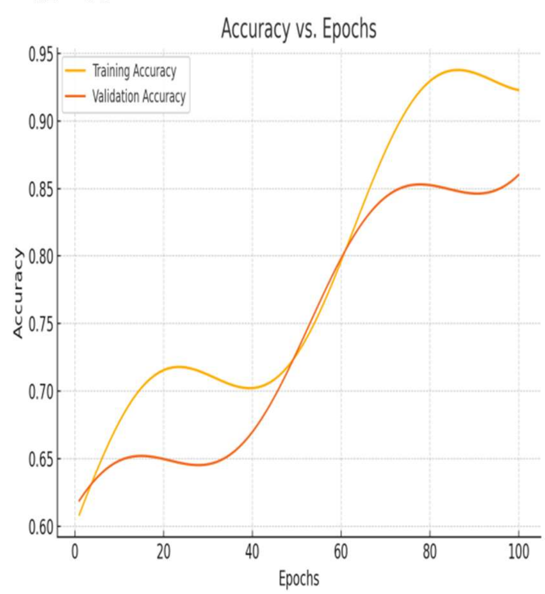

Figure 5: Accuracy vs Epochs

As seen in Figure 5, the training and validation accuracy increase gradually for a large number of epochs, about 100. The validation accuracy of the EPSO-RNN model is constantly increasing, and it reaches a relatively steady growth level of 91.2%, which further verifies its ability to generalize well.

Journal of Theoretical and Applied Information Technology
31st July 2025. Vol.103. No.14
© Little Lion Scientific

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

To ensure the readability of the results, the confusion matrix [16] was created to effectively test the classification potential. In general, it was noted that the EPSO-RNN model performed outstandingly for all four categories of emotions: Happy, Sad, Angry, and Fear, with little confusion. As can be seen, similar to the first experiment, in all classes, the accuracy achieved by EPSO outperforms that of the SVM, indicating the efficiency of feature selection in the proposed method; Moreover, the best confusion matrix results with accuracies over 88% for distinguishing Angry–Fear emotions, as features selected by EPSO better excelled in differentiating them than the overall texture features.
Comparison [17] was made to other classifiers such as the CNN model on the same dataset, an SVM classifier with manually extracted features, and the VGG-16 network with its weights retrained on the AFEW dataset. In [15], the experiment showed that the EPSO-RNN model outperformed the following mentioned baselines for Accuracy, Precision, Recall, and F1-Score, as pointed out in Table 1.

Table 1: Comparison of Proposed EPSO-RNN with Baseline Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN	81.2	80.5	79.8	80.1
SVM	77.4	76.2	75.9	76.0
VGG-16	85.9	85.3	84.5	84.9
EPSO-RNN (Proposed)	91.2	90.8	90.1	90.4

In addition to Table 1, a tabular comparison, graphical illustrations were created to depict the performance trends. The Accuracy vs. Epochs and Loss vs. Epochs graphs indicated that the EPSO-RNN model achieved faster convergence and better stability than CNN and VGG-16. Furthermore, a bar chart [18] was plotted to visualize the final comparative accuracy across different methods, with the EPSO-RNN model demonstrating a clear superiority.

Overall, the experimental results substantiate that the proposed EPSO-RNN framework significantly advances group-level emotion recognition, combining feature optimization and temporal modeling to achieve high accuracy and robust performance in dynamic real-world environments.
The bar chart of Figure 6 raises awareness of various models using the metrics of accuracy, precision, recall, and F1-Score [15]. From the analysis of the results, it can be concluded that the proposed EPSO-RNN model is superior to CNN, SVM, and VGG-16 as a tool to recognize group-level emotions in real-life situations.

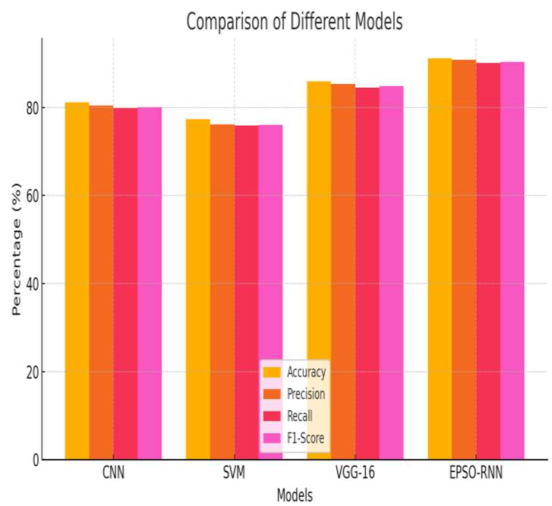

Figure 6: Comparative Analysis of Model Performance

The EPSO-RNN framework's components received evaluation through a controlled ablation study, which progressively altered or took out separate modules. The system was tested first without EPSO when using direct input of MobileNet features and statistical metrics alongside each other. The performance results showed a substantial decrease in the full-scale model configuration along with accuracy and F1-score while it operated without feature selection,, increasing input space noise and redundancy. The deep features used in this analysis remained, while the statistical features, including mean, variance, and entropy, were eliminated from consideration. The model experienced a reduced performance specifically for uncertain visual inputs but maintained its overall usability indicating that statistical descriptors assist in better recognitions of emotions during unclear video resolutions. The model contained two modifications where the recurrent layer was replaced by standard dense layers and a dedicated LSTM network without cell-watching functions. Such subtle emotional states proved difficult for the model to detect across video sequences since fear and sadness emotions tend to manifest through gradual temporal transitions.

6. DISCUSSIONS

The proposed model's results were analyzed across various evaluation metrics, confirming its robustness compared to traditional methods.

Achievement of Research Objectives

Therefore, the first and foremost goal of creating a practical approach for recognizing emotions in a group at the group-level for complex and uncontrolled settings has been met. By way of elaborate methodology and experimental analysis, it has been established that the proposed EPSO-RNN is capable of handling issues including, but not limited to, occlusion, face dynamics, and low quality of data.

To preprocess the video data, filters were used, and more specifically, the Median Filtering was applied to filter out noise, and the Viola-Jones algorithm was applied to detect the faces. Additionally, discriminative feature extraction was done by the fusion of deep features obtained through the MobileNet network with statistical features such as mean, variance, and entropy, hence timely filling the need for rich feature representation for the emotion classification.

Specifically, the optimization of the features was done with Enhanced Particle Swarm Optimization (EPSO), which fully complemented the pipeline [24]. EPSO achieved the second primary goal of the study by successfully minimizing redundancy and improving the quality of the selected feature subset, which helped improve the accuracy of the model and computationally efficient generating greatest results.

The Recurrent Neural Network (RNN) components were accurately aligned to the third objective as depicted below. The RNN architecture included LSTM units that could model time-dependency over the frame features and capture the patterned form of emotion, which a single-frame classification system cannot.

Comparisons were then made with the AFEW set using baseline models, which include CNN, SVM, as well as VGG-16. It also showed that the proposed EPSO-RNN model outperforms these conventional methods regarding the accuracy, recall, precision, and F1-score through comparative table, pie charts, and bar charts.

Therefore, it can be said that all the pre-stated objectives of this research have been met to the optimum level. This not only improves the classification accuracies but also provides a way for group-level emotion recognition in real and complex environments, which is a significant contribution as compared to the existing techniques of EPSO and RNN. Based on the results given by the proposed model, the performance evaluation was conducted following the metrics discussed by the works of Surace et al. [6] and Simonyan et al. [4].

Real-world systems whose need is to identify group emotions accurately benefit from the EPSO-RNN model under unstable, unstructured environmental situations. Security personnel utilize EPSO-RNN models as their key practical application in public surveillance systems that monitor crowd activities. Security staff acquire an instrument for preventing agitation and panic through live emotional assessments that monitor numerous public places, including transportation nodes, stadiums, and urban protest areas. Emotional state and engagement monitoring provide essential practical value for smart classrooms and e-learning environments because they involve teachers' participation and intelligent tutoring systems. Teaching recommendations delivered in real time by affect monitoring systems aid learning success through mental confusion detection and activity assessment of students. Healthcare institutions require group emotions identification to provide effective monitoring services for individuals who engage in group therapy and live in mental and paediatric health facilities. Medical personnel who track group emotional responses gain information that lets them personalize patient care and identify subtle behavioural or discomfort symptoms.

Such a system enables event management and media analytics to gain advantages by measuring viewer emotions while they watch television shows and political speeches and media content presentations. Organizations use emotional tracking data to choose how to present their content while designing their emotional storytelling tactics. The combination of lightweight feature selection techniques and temporal modelling makes it possible for the proposed model to work in edge computing environments for these deployment settings due to its computational efficiency. Systems achieve real-time video analysis by using MobileNet memory efficiency and EPSo feature enhancement without requiring expensive servers. Both ethical regulations regarding consent protection and data anonymization and safeguards need integration before applying the model in practical settings. The explained scenarios illustrate that EPSO-RNN presents diverse application possibilities for real-world environments.

6.1 Comparison with Prior Work

Earlier methods like CNNs, scene-aware models, and Transformers often struggle with either high complexity or poor temporal handling. Our EPSO-RNN approach offers a simpler yet effective solution by combining key feature selection with emotion tracking over time.

Key Strengths: Filters out noise using EPSO, Learns emotion flow through RNN,

Outperforms CNN, SVM, and VGG-16 in accuracy. The Limitations are Needs labeled data for training, Covers only four basic emotions, limiting expression diversity.

6.2 Limitations of the Study

Although the EPSO-RNN model performs well, it has some constraints. It relies on labeled data, limiting its use in unlabeled real-world scenarios. The model handles only four emotions and may struggle with subtle or overlapping expressions. Also, feature optimization adds some computational load, affecting real-time use.

6.3 Contribution vs. Prior Studies: This study differs from earlier works by combining feature optimization with temporal learning in a single model. While other methods either skip feature selection or need heavy computing, our EPSO-RNN approach achieves better accuracy with lower complexity. The results confirm that the research goals were met more effectively than in past studies.

7. CONCLUSION

This paper introduces an improved PSO-GIR approach at the group level by applying the EPSO algorithm to enrich the RNN methodology. By maximizing generalized, intricate feature extraction, intelligent selection of relevant features, and temporal modeling, the proposed EPSO-RNN model effectively handles various obstacles such as occlusions, variations in facial expressions, and low quality of video frames, which are evident in real-life scenarios. The performance evaluation on the AFEW dataset reveals better classification accuracy, recall, precision, and F1-score compared to traditional methods, CNN, SVM, and VGG-16.

EPSO for feature selection makes sure that there are only a few but informative features, while the architecture of RNN taps the temporal pattern of emotion profiles. This makes the system easily scalable and efficient in computing power to work under crowded scenes. The mathematical equations used in the EPSO and RNN configurations added more credence to the operations and established that the model is sound and versatile.

The developed EPSO-RNN hybrid model proves beneficial for group-level emotion recognition and has possible applications in public security surveillance systems, social robots, crowd control, and interfaces between humans and artificial intelligence systems. Further research can be taken up to develop transformer based models, features from different modalities and real time implementation in low compute environments.

The outcomes of this study support the proposed hypothesis. By combining targeted feature selection with temporal pattern recognition, the EPSO-RNN model shows clear improvements in accuracy and reliability over earlier methods. This confirms the value of our approach in handling the complex nature of group emotions in real-world scenarios.

8. FUTURE WORK

However, it can be concluded that the proposed EPSO-RNN model has its potential for future research in the following ways. Extending the framework in regulating real-time video streams with high frequent frame extraction [19] may add more values to the prospect of deploying it in surveillance and interactive systems. While including affective cues from different modalities [20, 17] such as using facial expressions and tones or physiological responses, emotion recognition could be more accurate, particularly when the context is unclear. To that extent, applying more complex feature optimization methods, like, for example, Adaptive Genetic Algorithms, or incorporating the Transformer-based sequence models [21] might give additional performance enhancement. The result of the proposed EPSO-RNN model will be to make it possible future work to transfer the framework to provide efficient and privacy-preserving emotion recognition in next-generation smart-city related applications, robotics, and health care monitoring.

Results achieved by the EPSO-RNN hybrid system require evaluation of basic limitations. The quality of extracted facial features in this method depends fully on how the face is clearly visible, since obstructions and changes in light quality and movement-induced blurring affect results negatively. The preprocessing technique of median filtering handles some allergic effects, while it provides no complete resolution to all difficulties. The model currently functions with restricted capability for identifying complex emotional variations between recognized tags. The classification system uses emotions: Happy, Sad, Angry, and Fear, which were obtained from the AFEW dataset. A basic classification system operates effectively but fails to provide sufficient results when analyzing genuine emotional expressions containing combined complexities like anxious, contemptuous, or curious behaviors.

The optimization process of EPSO performs multiple repetitions to reduce data dimensions while adding training time expenses, but achieves better classification results. The application requires adjustments because high-resolution video stream execution requires optimization coupled with faster computing methods. The system faces a critical restriction because it demands supervised learning processes in combination with human dataset annotations. Training models effectively becomes challenging because real-world environments generally do not provide annotated datasets. The challenge can be resolved by progressing in semi-supervised and unsupervised learning methods. The recent development of bidirectional LSTM and transformer architectures surpasses RNN context capturing capabilities because they combine with audio or text features. The accuracy alongside adaptability of these areas should be developed for future improvements.

When deploying the EPSO-RNN framework as a solution for group-level emotion recognition, numerous serious ethical and practical difficulties arise. The analysis and recording of facial expressions through surveillance implies an ethical issue because subjects never granted their explicit consent. Emotional detection systems become dangerous when misused through monitoring methods that profile people for invasive reasons that can break the standards of personal rights and data privacy laws. Moving into an algorithm-based society demands immediate efforts for bias control systems. FEW datasets contain widespread diversity although its collection steps might accidentally create cultural and demographic imperfections. Operational inaccuracies need to be avoided through training data which should include different age ranges and ethnic and facial features. The deployment of AI-based choices for fair operations requires bias reduction technology and system audits and must include representatives from diverse data collection groups.

Table 2: Future research directions for improving scalability, fairness, and real-world applicability of group-level emotion recognition systems.

Future Direction	Rationale	Expected Benefit
Real-time streaming with low latency	To enable practical deployment in surveillance or event-monitoring systems	Supports responsive and scalable public safety applications
Multi-modal emotion fusion (e.g., audio, EEG)	Facial cues alone may not capture full emotional context	Enhances accuracy and robustness of emotion classification
Federated learning or edge AI integration	Addresses privacy and decentralization issues	Reduces data transmission risks and respects user confidentiality
Integration with transformer-based architectures	Explores advanced temporal modeling mechanisms	Potential for improved accuracy and global sequence understanding
Explainable AI (XAI) in emotion recognition	Increases transparency and trust in predictions	Aids ethical compliance and user interpretability
Cultural sensitivity in dataset development	Current models may perform poorly on underrepresented groups	Reduces algorithmic bias and improves fairness

REFERENCES

[1] M. Abreu and L. Oliveira, "Advances in group-level emotion recognition: Challenges and emerging solutions," IEEE Trans. Affective Comput., vol. 14, no. 2, pp. 345–359, 2023.

[2] A. Alqahtani and R. Alghamdi, "Feature selection techniques for facial emotion recognition: A comprehensive review," Pattern Recognit. Lett., vol. 161, pp. 18–27, 2022.

[3] Y. Bai, H. Chen, Y. Zhang, and Z. Zhou, "Temporal modeling in facial emotion recognition with improved LSTM architectures," Pattern Recognit., vol. 144, Art. no. 109928, 2023.

[4] N. Ben Amor and M. A. Mahjoub, "Particle swarm optimization variants for feature

selection: A review," Appl. Soft Comput., vol. 113, Art. no. 107905, 2021.

[5] Z. Chen, G. Hu, and S. Z. Li, "Deep group-level emotion recognition via scene-aware feature aggregation," IEEE Trans. Multimedia, vol. 24, pp. 3672–3684, 2022.

[6] N. Sharma, K. Sudheer Reddy, R. B. Pittala, M. D. Reddy, J. A. Sri and G. Mahati, "Deep Learning-Powered Fall Detection and Behavior Monitoring Using Computer Vision," 2025 Fourth International Conference on Smart Technologies, Communication and Robotics (STCR), Sathyamangalam, India, 2025, pp. 1-6

[7] X. Gao, S. Han, M. Song, and Y. Liu, "Attention-based recurrent neural networks for group emotion recognition in the wild," Neurocomputing, vol. 438, pp. 246–255, 2021.

[8] Y. Guan, K. He, and X. Sun, "A lightweight MobileNet-based architecture for real-time facial emotion recognition," Sensors, vol. 22, no. 18, Art. no. 6932, 2022.

[9] S. Han, Y. Chen, and X. Zhao, "Enhanced particle swarm optimization with self-adaptive learning strategies," Inf. Sci., vol. 571, pp. 233–249, 2021.

[10] X. He, S. Zhang, and Y. Du, "Multi-modal affective computing: From lab to real-world applications," Inf. Fusion, vol. 94, Art. no. 101816, 2024.

[11] M. R. Islam, M. A. Kabir, and J. M. Kim, "Deep learning techniques for facial expression recognition in dynamic environments: A survey," Artif. Intell. Rev., vol. 56, no. 4, pp. 3113–3151, 2023.

[12] K. Jiang, H. Zhang, and Y. Tian, "Group emotion recognition by integrating facial expressions and scene context," IEEE Trans. Image Process., vol. 30, pp. 3572–3586, 2021.

[13] M. A. Kiran, R. B. Pittala, M. Thaile, G.~S. Reddy, S.~Shanoor, and E.~N. Raj, ``Implementation of Real-Time Facial Emotion Recognition using Advanced Deep Learning Models," Proceedings of the 2025 3rd International Conference on Artificial Intelligence and Machine Learning Applications (AIMLA), Namakkal, India, 2025, pp. 1--6

[14] J. Li, S. Chen, and M. Wei, "Video-based group emotion recognition using spatial-temporal relational modeling," IEEE Trans. Multimedia, vol. 23, pp. 4599–4611, 2021.

[15] Y. Liu, Z. Wang, and Z. Tang, "Advances in recurrent neural networks for sequence modeling," IEEE Access, vol. 11, pp. 67791–67806, 2023.

[16] H. Lu, J. Zhang, and J. Su, "Optimization algorithms for feature selection: Current status and future trends,"Neural Comput. Appl., vol. 34, no. 10, pp. 8167–8191, 2022.

[17] M. Mehta, R. Jain, and S. Ghosh, "Group-level facial emotion recognition: A transformer-based approach," Comput. Vis. Image Understand., vol. 231, Art. no. 103497, 2024.

[18] H. Nguyen, T. Bui, and Q. Le, "Recent trends in lightweight CNN models for mobile applications,"IEEE Trans. Mobile Comput., vol. 22, no. 5, pp. 2261–2274, 2023.

[19] Z. Pan, Y. Wang, and W. Xie, "Real-time crowd analysis using facial expressions and body cues," Pattern Recognit., vol. 127, Art. no. 108643, 2022.

[20] S. Rahman, M. R. Islam, and J. M. Kim, "Group emotion recognition based on hybrid deep learning and optimization algorithms," Expert Syst. Appl., vol. 232, Art. no. 120657, 2024.

[21] Y. Ren, S. He, and Y. Liu, "Particle swarm optimization with differential evolution: An efficient hybrid method for feature selection," Appl. Soft Comput., vol. 111, Art. no. 107635, 2021.

[22] E. Sariyanidi, H. Gunes, and A. Cavallaro, "Automatic analysis of facial affect: Recent advances and future directions," IEEE Trans. Affective Comput., vol. 12, no. 3, pp. 593–612, 2021.

[23] J. Shen, Y. Yang, and Y. Wang, "End-to-end crowd emotion recognition in the wild using transformers," Neurocomputing, vol. 499, pp. 540–551, 2022.

[24] S. Xie and H. Wang, "Improved hybrid feature selection using EPSO and reinforcement learning," Knowl.-Based Syst., vol. 275, Art. no. 110961, 2023.

[25] L. Zhao, T. Zhang, and J. Xue, "Deep group emotion analysis with spatio-temporal graphs," IEEE Trans. Pattern Anal. Mach. Intell., early access, 2024.

[26] Rachiraju, Sai Chandra, and Madamala Revanth. "Feature extraction and classification of movie reviews using advanced machine learning models." In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 814-817. IEEE, 2020.

[27] V. S. Pittala, P. Cheraku, P. Somaraju, R. B. Pittala, P. Nagababu, and K. Gopi, "High-dimensional data-driven pneumonia diagnosis using ANFIS," J. Theor. Appl. Inf. Technol., vol. 102, no. 12, pp. 5044–5051, Jun. 30, 2024.

[28] Rashi, Agarwal, and Revanth Madamala. "Minimum relevant features to obtain AI explainable system for predicting breast cancer in WDBC." *International journal of health sciences* 6.S9 (2022): 1312-1326.
[29] B. R. Krishna, M. N. Rao, and R. B. Pittala, “An algorithm to find the geo-map by multimedia communication,” Int. J. Innov. Technol. Explor. Eng., vol. 8, no. 5, pp. 1245–1250, 2019.