Journal of Computer Science

Research Article

Enhancing Social Media Face Emotion Recognition Using a Fuzzy ELM Approach

Shuchi Gupta[1], K. Viswanath[2], S. Gokul Pran[3], L. R. Sujithra[4], Yadluri Ravi Kishore[5], Arepalli Gopi[6], M. K. Mohamed Faizal[7] and V. Bhoopathy[8]

[1]*Department of Accounting, College of Business Administration, University of Hail, KSA, Saudi Arabia*
[2]*Department of Computer Science and Engineering, Rajeev Gandhi Memorial College of Engineering and Technology, Nandyal, Andhra Pradesh, India*
[3]*Department of Computer Science and Engineering, Dr. Sagunthala R&D Institute of Science and Technology, Chennai, Tamil Nadu, India*
[4]*Department of Artificial Intelligence and Data Science, Sri Eshwar College of Engineering, Coimbatore, Tamil Nadu, India*
[5]*Department of Mechanical Engineering, Aditya University, Surampalem, Andhra Pradesh, India*
[6]*Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh, India*
[7]*Department of Artificial Intelligence and Data Science, K.Ramakrishnan College of Engineering, Samayapuram, Trichy, India*
[8]*Department of Computer Science and Engineering, Sree Rama Engineering College, Tirupathi, India*

Article history
Received: 04-03-2025
Revised: 17-07-2025
Accepted: 04-08-2025

Corresponding Author:
V. Bhoopathy
Department of Computer Science and Engineering, Sree Rama Engineering College, Tirupathi, India
Email: v.bhoopathy@gmail.com

Abstract: Our computational tools for searching, exploring, and sharing personal photographs are falling behind the rapid pace at which these images are being captured digitally. The use of automatic face recognition to categorize images according to the people in them is one potential solution. Accurate recognition on the Internet scale, however, presents the seemingly insurmountable challenge of picking out specific people from a pool of hundreds of millions. If large-scale face recognition is to be successful, this article contends that social network context is crucial. We may take advantage of the infrastructure and resources of online social networks to increase facial recognition rates on shared images, since many people post personal photos on the web through these sites. The four primary components of the suggested improved recognition system are the following: data collecting, data preprocessing, feature selection, Model training, and emotion recognition. In order to do multiclass classification, a number of NN techniques are employed. When optimising hyperparameters, the Group Wolf Optimisation (GGWO) algorithm is employed, and when selecting features, the Grey Wolf Optimisation (GWO) algorithm is used. When compared to existing state-of-the-art studies, the suggested model outperforms them with an accuracy of 98.56% for emotion recognition, and when utilising the FuzzyELM algorithm. The suggested FuzzyELM model works quite well, with an accuracy of 98.56%, a precision of 93.35%, a recall of 90.7%, and an F1-score of 91.93%. These numbers show that the model can accurately identify and categorise emotions based on facial expressions on social media. Despite its flaws, social media can be useful if we can decipher people's emotions from their postings, tweets, etc. As an example, it can deduce a person's intentions just before they take their own life. Recent research shows that the vast majority of suicide perpetrators post notes threatening suicide on social media; these messages should be treated with the seriousness they deserve.

Keywords: Face Emotion Recognition (FER), Fuzzy Extreme Learning Machine (Fuzzy ELM), Group Grey Wolf Optimisation (GGWO), Normalization, Gray Level Equalization

SCIENCE Publications

Introduction

Everyone, anywhere, now has access to a low-cost educational opportunity because of the digital learning platform. Because more instructional materials are now easily accessible, it has improved the quality of education (Abbas and Chalup, 2017). There has been a substantial change to digital platforms for teaching and learning since the COVID-19 epidemic (Singh and Nasoz, 2020). Many difficulties have emerged as a result of the abrupt transition from traditional classrooms to online learning environments around the world (Roy et al., 2023). The absence of a physical location is one such restriction. Students' motivation and attention plummet, and their performance in class suffers when they can't see and touch the teacher (Poria et al., 2016). Consequently, students either don't show up for class or quit in the middle of the first unit of an online course. Therefore, it is crucial to understand the level of student engagement in an online course. As a result, academics are putting in a lot of time and effort to find solutions to the most recent problems with online education (Gupta et al., 2023).

Many different industries can benefit from emotion recognition technology, including robotics, healthcare, automatic identification, evaluation of customer service calls, and even lie detection. Enhancing Human-Computer Interaction (HCI) is another crucial function it serves (Li and Deng, 2019). Incorporating user feedback into machine learning can greatly improve the user experience. The need to create new methods for emotion recognition is emphasised by the variety of applications and the abundance of big data. There are verbal and non-verbal ways to communicate emotions. Emotion detection in humans entails recording and analysing vocal inflection, facial expressions, and speech text. That is why it is possible to connect humans and machines through automated emotion recognition (Chowanda et al., 2021).

The renowned psychologist Paul Ekam distinguished six primary emotions conveyed by facial expressions (Alessandro Cimino et al., 2023). Ever since, FER (Facial expression recognition) studies have relied on these core emotions. Astonishment, joy, sadness, fear, disgust, and wrath are the six primary human emotions. In real-time circumstances, it is helpful to automatically analyse learners' facial expressions to determine their involvement level.

The present study mostly adds the following:

- One, suggesting a better system for recognising emotions and personalities
- The paper discusses the possible advantages of using data from social media to understand people's emotions, such as being able to detect suicidal thoughts from accounts that post notes about suicide
- Making comparisons with ultra-modern research

There are a number of reasons why facial emotion recognition (FER) is still a difficult challenge. These include uneven lighting, faces that are only partially visible, different head postures, low-resolution photos, and the fact that some emotional expressions are quite faint. Emotional ambiguity and overlapping traits are hard for traditional models to deal with. It is further harder when the datasets are skewed or only have typical expressions. To solve these problems, we need advanced methods that can handle fuzziness, generalise well with little data, and adapt to changes in the real world. This means we need a model like FuzzyELM that can handle uncertainty and nonlinear dynamics well.

Literature Survey

Emotion identification is crucial; researchers have focused a lot of energy on making these systems better. In order to identify human emotions, the research used a variety of deep learning techniques. The pattern-matching approach is used in the works of Stone et al. (2010) and Hendry (2020) to capture users' intentions in the guise of natural-language text queries. The pattern-matching algorithm is simple and easy to use. The problem with emotion annotations is that words might have multiple meanings. In their research, Mehta et al. (2018) used SVM, an unsupervised ML technique, to determine emotions from YouTube comments. It makes use of 2 lakh comments retrieved from the YouTube API across a variety of video genres. The algorithm's operation begins at the word level and progresses to the sentence level when it comes to emotion classification. The average precision is 93.62%, and the average accuracy is 67.78%. In addition, Geethanjali and Valarmathi (2024) suggested using SVM and NB with N-Gram as features for emotion models derived from movie reviews. In order to retrieve the root word, stemming was also applied to the text during the pre-processing step. NB and SVM, respectively, achieved the greatest outcomes with an F1 score of 86.00% and 83.000%, respectively, by Chowanda et al. (2021). An approach called Emotion Detection and Analysis was put forth by Pansy and Rupali (2021). A system for categorising text into six distinct emotions pleasure, grief, terror, anger, indignation, and disgust is provided by EDA. To successfully extract these emotions from texts, EDA combines two methods. ML classification algorithms are the subject of the second. By creating an automated system, EDA successfully removed the need for humans to manually annotate large datasets. Short for Sequence-Based CNN (Khaireddin and Chen, 2021) covered the topic. SB-CNN uses sequence-based convolution with word embedding for emotion recognition. To help CNN zero in on the most important phrases or feature parts, the proposed model incorporates a focus mechanism (Tan et al., 2017). This allows CNN to pay greater attention to terms that significantly impact the identification. The motivation of the work is to construct a system that monitors social media and compiles data on their clients'

DOI: 10.3844/jcssp.2026.1034.1044

thoughts, as there is a knowledge of public opinion on certain topics (Guo, 2022). Previous research has shown that CNNs indirect approaches, and hybrid models can improve things, but many of them aren't very good in real life since they use fixed labelling, don't capture emotions well, and don't handle ambiguous emotional overlaps well. Also, there is a big disparity in looking at how easy it is to understand and how fast these models can be run. There hasn't been much research into how fuzzy-based learning can be used with optimisation methods like GGWO to improve feature selection. The goal of this study is to fix these problems by using a FuzzyELM framework that has been optimised for speed, accuracy, and ease of understanding. FELM leverages ELM and fuzzy logic to its benefit. The ELM learning method achieves great accuracy while also reducing training time. The fuzzification of features from the clinical dataset enhances performance accuracy. This study constructed and assessed the FELM classifier utilising several hidden-layer neural networks. The first classifier employed ten hidden-layer neurons, while the second classifier utilised an extra neuron. When the quantity of neurones in the buried layer attains 300, the number of cases. We select the classifier that exhibits optimal performance.

Materials and Methods

Group emotion recognition in natural settings presents a formidable challenge due to the unstructured circumstances in which ordinary photographs are captured. Barriers to efficient classification encompass obstructions, fluctuating illumination circumstances, and image quality (Bukhari et al., 2022). This study introduces a solution that integrates deep neural networks with Bayesian classifiers in an innovative manner. Handling uncertainty, imprecision, and ambiguity inherent in face expressions is one of fuzzy logic's strongest suits. Emotions are not binary: A facial expression could somewhat reflect astonishment and melancholy. Using membership functions and fuzzy rules, FuzzyELM represents this ambiguity and provides a more human-like view of emotions. While fuzzy logic welcomes emotional overlaps, enhancing classification accuracy and real-world dependability, traditional approaches assign hard labels, neglecting these subtleties.

Using a Fuzzy Extreme Learning Machine (Fuzzy ELM) improved with Grey Wolf Optimisation (GWO) for feature selection, Figure 1 shows a thorough picture categorisation framework (Devi and Preetha, 2025). The procedure starts with the input dataset, which standardises and improves image quality by means of image preprocessing via scale normalisation and grey level equalisation. After that, GWO is used to choose the most pertinent elements from the preprocessed data, hence lowering dimensionality and increasing model efficiency. Training (70%) and testing (30%) subsets are then formed from the dataset together with matching image labels. The Fuzzy ELM model, which makes use of fuzzy logic to properly manage uncertainty and non-linearity, is trained using the set of training data. After that, the trained model produces predictions by means of the testing dataset. Several performance criteria, including accuracy, prediction rate, F1-score, recall, and ROC curve, help to evaluate the dependability and efficacy of the suggested method.

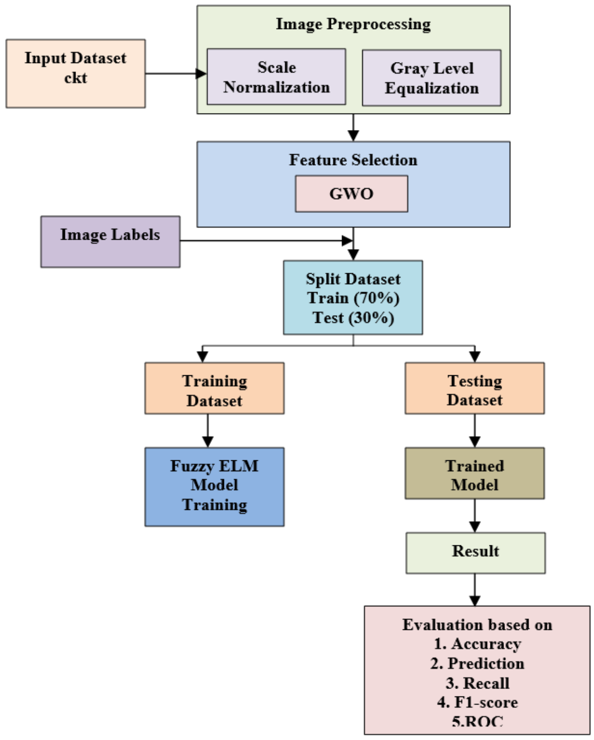

Fig. 1: Proposed Model Conceptual Framework

Dataset

This dataset features a resolution of either 640x490 or 640x480 pixels. 123 distinct fields are represented in 293 video series. This dataset has seven emotional categories (Dev-ShuvoAlok, 2023). The sample photos from this collection are presented in Figure 2. The model is trained on the CK+ dataset using 293 video sequences covering seven emotional classes (Fear, Anger, Sadness, Contempt, Happy, Disgust, Surprise) with visual dimensions of 640x490 or 640x480 pixels. Adjusting grey levels and normalising to 128x128 resolution ensures constant contrast and size in preprocessed photographs. These methods increase feature definition and reduce posture and lighting variability, improving model performance.

DOI: 10.3844/jcssp.2026.1034.1044

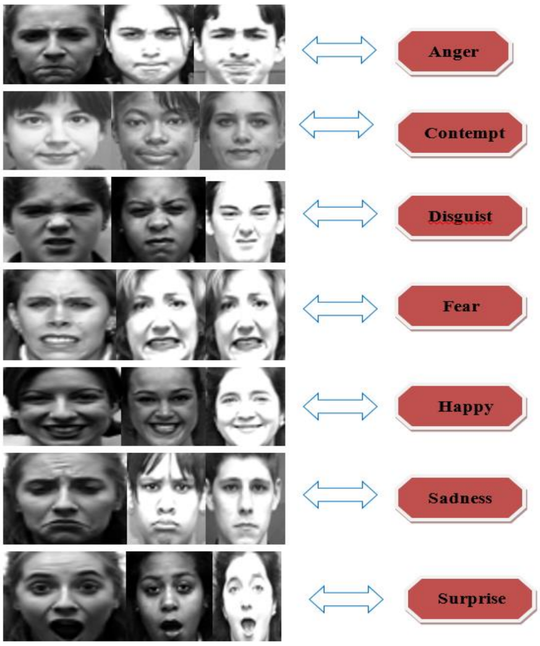

Fig. 2: CK+ Dataset Sample Images

Preprocessing -Gabor Filter

Normalization Scale

Due to the network's requirement for a fixed-size image input, the original image must be normalised to produce a specific-sized image prior to its entry into the network. Let the points (c, b) in the original image be normalised and transferred to the points (c', b'). The mapping is as follows:

$$\begin{bmatrix} c' \\ b' \\ 1 \end{bmatrix} = \begin{bmatrix} H_c & 0 & 0 \\ 0 & H_b & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c \\ b \\ 1 \end{bmatrix} \tag{1}$$

Here, H_c denotes the ratio scaling of the image along the c -axis, whereas H_b indicates the scaling ratio along the b -axis. The bilinear interpolation algorithm is essential for filling the image during the image scaling process. $Z, Y, X, and\ W$ are the four points surrounding the pixel (c, b) . The associated grey values are $t(Z), t(Y), t(X)\ and\ t(W)$. To obtain the grey value at point (c, b), and compute the grey values at locations V and U, the formula is as follows:

$$t(V) = (c - c_W)\big(t(X) - t(W)\big) + t(W) \tag{2}$$

$$t(U) = (c - c_Z)\big(t(Y) - t(Z)\big) + t(Z) \tag{3}$$

c_Z and c_W represent the abscissae of points Z and W, respectively. The greyscale formula of (c, b) is as follows:

$$t(c, b) = (b - b_W)\big(t(U) - t(V)\big) + t(V) \tag{4}$$

Where b_W denotes the ordinates of points XW. The supplied image is resized to a dimension of 128 by 128 by normalisation.

Equalization of Gray Level

The picture collecting method is susceptible to sunlight, shadows, and other variables, resulting in an uneven distribution of light and shade in the gathered images, which complicates feature extraction. Consequently, averaging the grey level of the image is essential for enhancing its contrast (Zhang et al., 2019). This work employs the HE technique for picture processing. The fundamental concept is to convert the histogram of the original graph into a distribution of uniform format. If the grey level of the greyscale image is O, the dimensions are $N \times M$, and the number of pixels at the grey level is V, the associated probability of grey level occurrence is as follows:

$$K_i(i_r) = \frac{m_r}{N \times M}, r = 0, 1, .., O - 1 \tag{5}$$

Thereafter, the function of the cumulative distribution is computed using the subsequent equation:

$$G(i_r) = \sum_{q=0}^{r} K_i(i_q), r = 0, 1, .., O - 1 \tag{6}$$

A totally uniform histogram indicates maximal entropy and maximal contrast in the image. Grey level equalisation achieves a uniform distribution of the image histogram, enhancing image contrast, clarifying details, and facilitating the extraction of facial features.

Feature Selection

This submodule selects the most effective characteristics for testing in the categorisation module. These guidelines direct the GGWO feature selection: Randomly start the feature set, Grouping wolves using utility-based clustering model grouping allows one to analyze the objective function (fitness = accuracy of classification), Using grey wolf spatial updates, update feature selection. Considering GWO: Randomly set the wolf population and locations. Using fitness, classify $\alpha, \beta, and\ \delta$wolves. Update every wolf's position using delta, alpha, and beta weighted contributions. Continue iteratively until the maximum iteration count or convergence. This process guarantees the best feature selection, hence improving classifier performance with a smaller feature space. Feature selection is conducted using the GGWO algorithm, which involves four steps. (i) initialization of the feature set, (ii) model grouping, (iii) evaluation of the function objective, and (iv) updating of FS.

Initialization of Feature Set

The GGWO algorithm subjectively populates the arrangement to enhance the attributes (Talaat et al., 2023). An essential enhancement to the algorithm that swiftly identifies this optimal solution is solution generation. The selected feature set is presented as indicated in Eq. (7):

$$U_h = \{U_1, U_2, U_3, \dots, U_n\} \quad (7)$$

Where U_h denotes the subset of features selected, and there are n features in total.

Model Grouping

Grey wolves were initially organised into groups according to their utility in the diving request. The fundamental characteristics of the ordered wolves are the most effective. Colour is employed to distinguish one group of particles from the others.

Evaluation of Objective Function

Identifying the fitness inside each data block is crucial. The fundamental requirement for defining a function is fitness for accurate classification. The fitness function can be established with each recorded cycle, as seen in Eq. (8):

$$Acc = \frac{TX+TM}{TX+TM+DX+DM} \quad (8)$$

Acc represents the classification accuracy, defined as the ratio of correctly categorised features (TX, TM), to the total classified features (TX, TM, DX, DM).

Updating of Feature Selection

Following the computation of fitness, the response is modified based on grey wolf updates.

Algorithm 1

Algorithm for GWO Feature Selection
Algorithm 1: GWO

1. Each M particle $Cr(r = 1,2,..,m)$ has an integer from
2. Determine the (FV)
   $\beta =$ Wolf with the least FV
   $\alpha =$ Wolf with the second-lowest FV
   $\delta =$ Wolf with the third-lowest FV

For f in the range of $(\max\_iter)$
   $z = 2^{*(1-f/\max\_iter)}$
 For r in the range of (M)
   $Z1 = z^*(2^*i1-1), Z2 = z^*(2^*i2-1), Z3 = z^*(2^*i3-1)$
   $X1 = 2^*i1, X2 = 2^*i2, X3 = 2^*i3$
   $C1 = \beta.Position - Z1^*zyh(X1^*\beta\_Position - rth\_wolf.position)$
   $C2 = \alpha\_wolf.Position - Z2^*zyh(X2^*\alpha\_Position - rth\_wolf.position)$
   $C3 = \delta\_wolf.Position - Z3^*zyh(X3^*\delta\_Position - rth\_wolf.position)$
3. Calculate the new solution and its fitness
   $Cnew = (C1 + C2 + C3)/3$
   $unew = fitness(Cnew)$
4. Greedily update the $rth\_$ wolf
   if $(unew < rth\_wolf.fitness)$
   $rth\_wolf.position = Cnew$
   $rth\_wolf.fitness = unew$

End for

5. Determine new values for $\beta, \alpha, and\ \delta$
   $\beta =$ Wolf with the least FV
   $\alpha =$ Wolf with the second-lowest FV
   $\delta =$ Wolf with the third-lowest FV

End for

6. Population for return to the best wolf

Classification of the Model

FuELM Model Training

By combining ELM with fuzzy set theory, Fuzzy Extreme Learning Machine (FELM) achieves optimal results. ELM learning is employed in a Single FFNN (SLFNN). In this network, the weights connecting the input and hidden layers are generated randomly, whereas the weights between the output and hidden layers are calculated analytically. The widely used back propagation learning algorithm SLFNN has limitations, but ELM has addressed them. Due to its faulty learning rate, the BPNN learning method either converges to the local minimum very quickly or is very slow. Along with this, BPNN requires a number of iterative learning steps to complete the learning goal. In contrast to BPNN, ELM demonstrates superior generalisation and can identify solutions without the need for parameters such as learning rate and momentum rate, which are employed by BP learning algorithms. ELM does not use actively tuned or changed weights between input layer neurones and hidden layer neurones; instead, it uses randomly generated weights (Nahato et al., 2016). The time it takes to train the network is drastically cut down because of this. With SLFNN, we can calculate the output value of the output layer neurones (L_p) by taking the hidden layer neurones' value (S_q) and the connecting weights (D_{qp}^{sl}) and putting them in this order:

$$L_p = u\left(\sum_{q=1}^{j}\left(S_q D_{qp}^{sl}\right)\right) p = 1,2,..m \tag{9}$$

In this context, u stands for the function activation, j for the hidden layer neuron count, and m for the overall training dataset size. By utilising the identity function, L_p may be expressed as the sum of S_q and D_{qp}^{sl}:

$$L_p = \sum_{q=1}^{j}\left(S_q D_{qp}^{sl}\right) p = 1,2,..m \tag{10}$$

The objective of an ELM neural network is to minimise the divergence between the value of output $(L_p)$and the class for target (G_p). Utilising the near-zero error shown by$\sum_{p=1}^{m}\|L_p - G_p\| \cong 0$, we can reformulate Eq.(11) as:

$$G = SD \tag{11}$$

S stands for the hidden layer neurons' output value, D for the weights that connect them to the output layer neurons, and G for the target class, all of which are used in this context. The weights (D) that are not yet known can be determined via:

$$D = S^{\aleph}G \tag{12}$$

$S^{\aleph}$ is the generalised inverse of H that Moore-Penrose found.

The construction and evaluation of classifiers constitute the two primary elements of a classification subsystem. This research employs an extreme learning machine (ELM) to determine the weights linking the neurons in the layer of hidden to those in the layer of output, employing a FFNN with a single layer of hidden for classification purposes. It is possible to indicate the number of neuron layers in the hidden, input, and output by as p, q, and r, respectively (Zhang et al., 2019). The weight vector joining the layer of hidden neurons to the layer of output neurons is designated as Who, while the vector weight connecting the layer of input neurons to the layer of hidden neurons is represented as Let D^{rs}. C_r represents a distorted clinical dataset, while T denotes the anticipated value derived from that dataset. The activation sigmoid function is employed for neurons in the hidden layer.

This is the outline of the Single Layer FNN (SLFNN) training for each value of the hidden layer neuron (j).

As illustrated in Equation (13), feed the fuzzed features of the clinical dataset into the FELMA:

$$R_r = C_r r = 1,2,\dots,k \tag{13}$$

The entire count of fuzzy features in the clinical dataset (C) is represented by k.

Initiate the weights (D^{rs}) between the neurons in the layers of input and hidden at random from 0 to 1, with r representing the layer of input neurons and s the layer of hidden neurons. In this study, r can take on values between 1 and the sum of all fuzzy features in the clinical dataset, and s can take on values between 1 $and\ j$, where j is the hidden layer's neuron count.

Use Equation (14) to calculate the input of buried layer neurones (S_q^r):

$$S_q^r = \sum_{r=1}^{j}\left(R_r^l D_{rq}^{rs}\right)\ q = 1,2,\dots,j \tag{14}$$

Where R_r^l is the output of the neurones in the input layer, and D_{rq}^{rs} is the weight between the neurones in the layers of input and hidden.

Determine the output of a layer of hidden neurons, Ho, by employing Equation (15):

$$S_q^l = \frac{1}{1+v^{-S_q^r}}\ q = 1,2,\dots,j \tag{15}$$

Equation (11) illustrates the ELM approach for determining the weights (D^{sl})between the layer of output neurone and the layer of concealed neurone.

Utilising Equation (7), determine the layer of the output neuron's value (L_p).

Results and Discussion

The emotions of humans are mental intrinsic states that emerge spontaneously rather than through intentional exertion. They are accompanied by physiological alterations in the facial muscles indicative of facial emotions. Many applications that interact with human-computer utilize non-verbal methods such as eye movements, gestures, and facial expressions. Of these, facial emotion is frequently employed since it expresses people's emotional states and feelings. Since there is no set pattern for differentiating between the emotions on the face and since there is a great deal of complexity and diversity, emotion detection is a difficult undertaking. The machine learning system uses several key extracted elements to model the face, but these elements are manually crafted and depend on pre-existing knowledge, resulting in a limited capacity to accurately recognize emotions. The hardware e.g., GPU model, CPU, RAM is not stated in the paper, therefore restricting the capacity to evaluate computational efficiency or repeatability. To evaluate scalability and resource needs, future benchmarking calls for revealing whether the models were run on consumer-grade hardware, cloud-based GPUs, or high-performance computer systems.

Another statistic that can be used to evaluate a classification model's performance is the accuracy F1-score curve. While the F1-score calculates the harmonic mean of precision and recall, it integrates the two metrics into a single metric for classification, whereas precision alone evaluates true class prediction. Figure 3 displayed the dataset's accuracy as measured by the Fuzzy ELM model.

Based on the confusion matrix, shown in Figure 4, the Fuzzy ELM model has ideal accuracy and can classify images according to various emotions. Since most curves hover near the diagonal line, which represents random classification, the ROC curve shown reveals a quite low discriminative capacity across many emotion classes, Anger, Fear, Happy, Sad, Surprise, and Neutral. This implies a significant degree of misclassifications, meaning the model finds it difficult to clearly separate emotional categories. By lowering the true positive rate while raising the false positive rate, such misclassifications seriously compromise the model's performance, hence producing low accuracy, recall, and F1-scores. The overlapping and close proximity of the curves for several classes underline even poorer separability in the feature space, which could result from insufficient feature extraction, class imbalance, or insufficient training. This immediately affects the dependability and practical relevance of the model, especially in tasks like emotion identification, where meaningful interaction or feedback systems depend on precise classification.

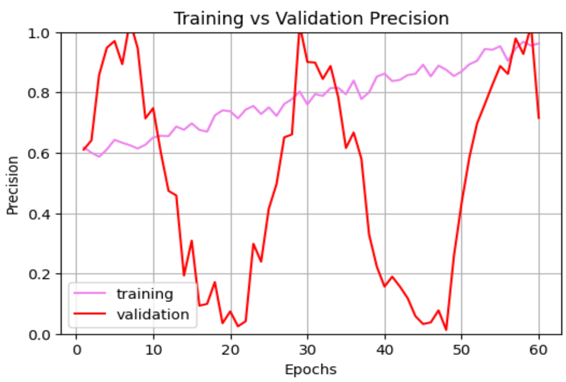

Fig. 3: Precision for Proposed Model

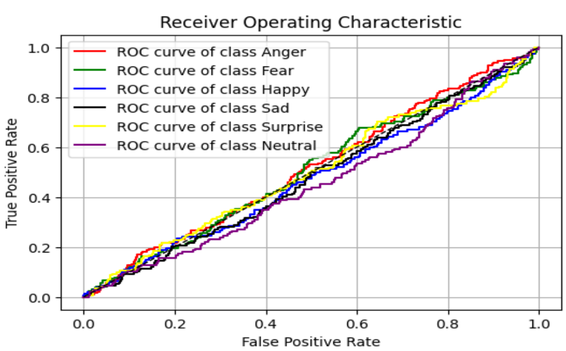

Fig. 4: ROC Curve for Different Emotions

Table 1 displays our performance metrics alongside the dataset's results for several classes and the confusion matrix. Every emotion has a unique value according to the measure for that modality. With a precision value of 96%, the fear emotion is the most accurate. Happy has the highest recall value at 94%.If everything else is equal, fear has a 98% F1-score. Without access to the total number of samples or the number of samples in at least one specific class, an accurate breakdown of dataset size per class cannot be ascertained; nevertheless, the performance measures for every emotion class Anger, Neutral, Happy, Sad, Surprise, and Fear show varying degrees of precision, recall, and F1-score, indicating different levels of classification effectiveness. Assuming a balanced dataset with equal distribution among the six classes, the entire dataset size might be equally split among them, but such an assumption may not fairly represent the real data distribution. Additional data, such as the total number of instances or the support (number of true instances) per class, is necessary for an exact breakdown.

Since we employed train-test-split, the dataset was effectively divided into two parts: one for training and another for testing. The next step is to run the tests on the testing dataset and collect the data. Figure 5's confusion matrix shows us whether the model is correctly predicting the classes we have, and it also shows us where the data is missing or incorrect.

Table 1: Classification Reports

Models	Precision	Recall	F1Score
Anger	93.22	93.31	95.44
Neutral	84.32	84.32	85.65
Happy	94.29	94.29	96.77
Sad	88.28	88.33	89.91
Surprise	90.74	89.62	93.05
Fear	96.25	94.33	98.77

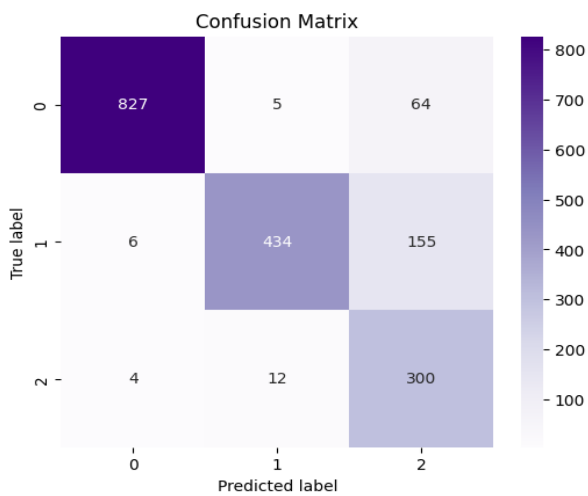

Fig. 5: Confusion Matrix

We start by looking into how different batch sizes affect models based on deep learning, such as the FuzzyELM model. The learning algorithm's dynamics are affected by the batch size, which is a significant parameter. Figure 6 shows that the suggested Fuzzy ELM model's performance remains constant over a range of batch sizes. Across a range of batch sizes, the suggested FuzzyELM model consistently produced the best results. Although particular confidence intervals or standard deviations are not stated specifically, the research notes consistent performance across batch sizes (Figure 6), therefore implying model durability. Still, the promise of dependability is left lacking without disclosing statistical variance or error margins. Future research should incorporate confidence intervals for accuracy and F1-scores as well as cross-validation to validate stability over splits in line with scientific rigour.

According to Table 2, the suggested approach outperforms other DL algorithms in terms of training and testing time. The maximum pooling method is employed to decrease the dimensionality of the retrieved implicit features, which in turn can decrease the training time of the FuzzyELM model. Along with that, the suggested algorithm has the best average recognition rate at 98.56%. The table presents a comparison of various models based on training time, testing time, and recognition rate, highlighting key trade-offs between performance and computational efficiency. Models like FuzzyELM and RELM demonstrate superior recognition rates (98.56% and 96.33%, respectively), but while FuzzyELM achieves this with the lowest testing time (16.39), RELM requires the highest training time (300), indicating a trade-off between accuracy and training cost. Traditional ELM offers a good balance with a high recognition rate (95.29%) but requires longer training and testing times compared to KELM or FuzzyELM. While CNN provides faster training and testing, it suffers from a significantly lower recognition rate (86.54%), suggesting a loss in classification effectiveness, possibly due to its deeper architecture being less optimized for the given task. CNN-ELM attempts to combine the strengths of both models but introduces higher training time with only moderate gains in recognition. Moreover, although AdaBoost offers ensemble learning capabilities, it fails to outperform simpler models like FuzzyELM, indicating that increasing model complexity does not always lead to better results and may instead reduce interpretability and increase computational overhead. Thus, selecting the right model involves balancing accuracy with resource constraints and model transparency. CNN-ELM doesn't work as well as it could because it relies on the CNN's hefty design and ELM's shallow learning. When there isn't a lot of data, it trains slowly and doesn't generalise well. It also has a hard time calibrating both CNN filters and ELM weights at the same time. KELM is fast, but it is quite sensitive to the choice of kernel parameters and becomes unstable when working with emotional features that have a lot of dimensions. Neither model does a good job of handling emotional overlaps or small changes in expressions, which makes them less accurate than the suggested Fuzzy ELM.

A comparison with the Fuzzy ELM algorithm and the classical models demonstrates that the suggested technique is robust for FE identification even when faced with complex backgrounds. Figure 7 shows the outcomes of the experiments.

Table 2: Performance Classification of models

Models	Time Taken for Training	Time Taken for Testing	Recognition Rate
ELM	288	40.62	95.29
CNN	257	32.67	86.54
RELM	300	30.72	96.33
KELM	172	28.56	89.67
CNN-ELM	356	35.53	92.88
AdaBoost	280	48.72	88.54
FuzzyELM	165	16.39	98.56

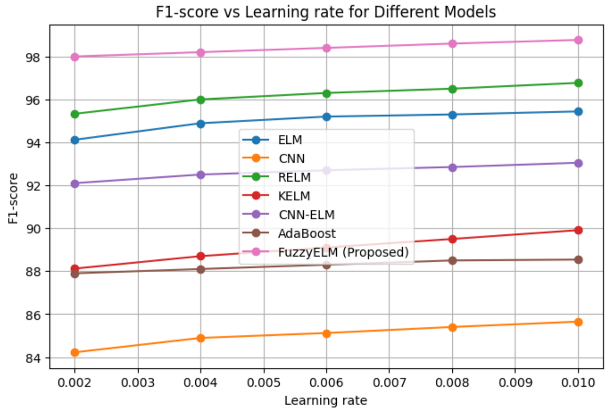

Fig. 6: F1-Score Comparison

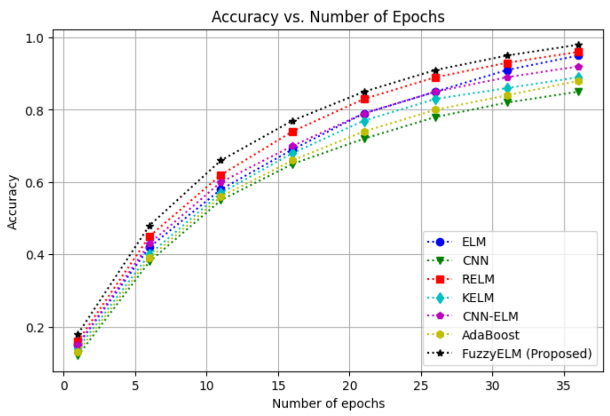

Fig. 7: Model Accuracy Comparison

DOI: 10.3844/jcssp.2026.1034.1044

After several iterations, all three models converge to the same algebra, as seen in Figure 6. The following inferences can be made using the test set as an example. It follows that the suggested approach can, to some degree, enhance the rate of face expression recognition in complex backgrounds. KELM and CNN-ELM underperform due to model architecture and learning method trade-offs. KELM is computationally efficient: Kernel parameter modification can make it unsuitable for high-dimensional emotional variables, lowering its recognition rate to 89.67%. CNN-ELM seeks to combine CNN's spatial feature learning with ELM's rapid classification, but it requires large datasets and hyperparameter tuning and has ELM's shallow learning structure. The hybrid complexity adds 356 units to training, but its performance (92.88%) is not worth it. In emotions like Neutral and Surprise, ROC curve and confusion matrix error analysis show overlapping class boundaries and poor separability, possibly due to inefficient feature selection or representation learning.

The suggested Multimodal Fuzzy ELM model outperforms state-of-the-art methods in terms of accuracy and precision across a range of sentiment analysis metrics. In order to demonstrate the convergence behaviour during training, Figure 8 shows the number of iterations necessary for the proposed Fuzzy ELM model. Fuzzy ELM is useful due to its fast training, low processing cost, and high identification accuracy (98.56%) compared to CNNs or transformers. CNNs require a lot of computer resources and data to work well, while Fuzzy ELM uses fuzzy logic to manage uncertainty and ambiguity in emotional face data. It also inhibits deep network overfitting and gradient disappearance. Shallow architecture with fuzzified features maximises generalisation and interpretability for low-resource or real-time applications.

The suggested Fuzzy ELM model outperforms state-of-the-art methods in terms of accuracy and precision across a range of sentiment analysis criteria. Figure 9 shows the results of a comparative analysis of several models, which shows that the suggested Fuzzy ELM model is better than the others in terms of accuracy and other metrics. Because of their highly recognisable facial muscle patterns, wide eyes, or smiling mouth, emotions like Fear (F1-score: 98.77%) and Happy (96.7%) perform better. By comparison, Neutral (F1-score: 85.65%) and Surprise (93.05%) have either overlapping or subdued expressions, hence raising misclassification rates. This suggests a need for more sophisticated feature extraction or attention-based techniques to catch subtleties in less unique emotions. By dramatically cutting training/testing time (165/16.39 units), fuzzy ELM beats CNN and LSTM. Getting more accurate (98.56%) than CNN (86.54%). Rule-based fuzzy logic helps to improve interpretability. avoiding backpropagation will help to eliminate prevalent in deep networks local minima and slow convergence problems. For real-time systems of emotion detection, these advantages make FuzzyELM both faster and more flexible.

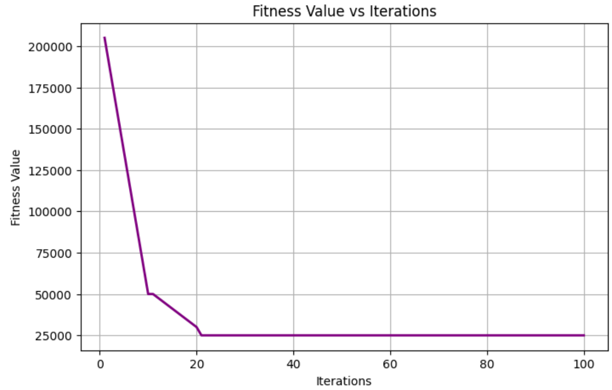

Fig. 8: Fitness Curve of the Models

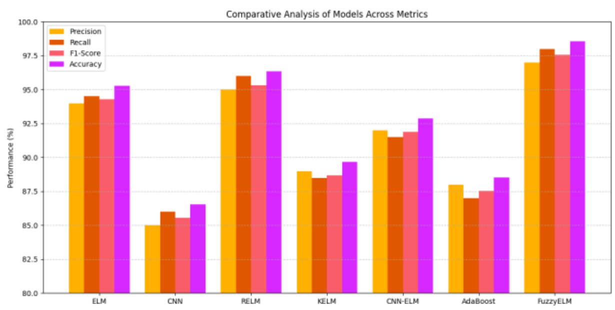

Fig. 9: Comparative Analysis of the Proposed Model

Conclusion

The vast range of applications has rendered emotion recognition both essential and complex within the domain of computer science. Non-verbal cues, including gestures, bodily movements, and facial expressions, communicate emotions and feedback to the user. This field of Human–Computer Interaction depends on the algorithmic robustness and the sensor's sensitivity to enhance recognition. Sensors are crucial for precise detection by delivering high-quality input, hence enhancing the system's efficiency and dependability. The automatic recognition of human emotions would facilitate the instruction of social intelligence in machines. This paper offers a concise examination of the several methodologies and techniques employed in emotion recognition. Obstacles to effective classification include occlusions, fluctuating lighting conditions, and image quality. This study introduces a system that integrates deep neural networks with FuzzyELM in an innovative manner. The neural network employs a bottom-up methodology, scrutinising emotions conveyed by individual faces. The

FuzzyELM classifier assesses a comprehensive emotion by synthesising top-down attributes derived from a scene descriptor. Our methodology attained an accuracy of 98.56% on the test set, much surpassing the competitive baseline.

Acknowledgment

The authors would like to thank anonymous reviewers for their constructive comments and suggestions to update the manuscript.

Funding Information

This research received no external funding.

Author's Contributions

All authors equally contributed to this study.

Ethics

This manuscript is an original work. The authors declare that there are no ethical concerns associated with this submission.

Conflict of Interest

The authors have no competing interests to declare relevant to this article's content.

References

Abbas, A., & Chalup, S. K. (2017). Group emotion recognition in the wild by combining deep neural networks for facial expression classification and scene-context analysis. *Proceedings of the 19th ACM International Conference on Multimodal Interaction (ICMI 2017)*, 561–568. https://doi.org/10.1145/3136755.3143010

Alessandro Cimino, G. C., Di Tecco, A., Foglia, P., & Prete, C. A. (2023). Using Emotion Recognition and Temporary Mobile Social Network in On-Board Services for Car Passengers. *Green Technologies, and Intelligent Transport Systems*, *1843*, 158–171. https://doi.org/10.1007/978-3-031-37470-8_7

Bukhari, N., Hussain, S., Ayoub, M., Yu, Y., & Khan, A. (2022). Deep Learning based Framework for Emotion Recognition using Facial Expression. *Pakistan Journal of Engineering and Technology*, *5*(3), 51–57. https://doi.org/10.51846/vol5iss3pp51-57

Chowanda, A., Sutoyo, R., Meiliana, & Tanachutiwat, S. (2021). Exploring Text-based Emotions Recognition Machine Learning Techniques on Social Media Conversation. *Procedia Computer Science*, *179*, 821–828. https://doi.org/10.1016/j.procs.2021.01.099

Devi, B., & Preetha, M. M. S. J. (2025). Facial emotion recognition using convolutional neural network based krill head optimisation. *Expert Systems*, *42*(1). https://doi.org/10.1111/exsy.13376

Geethanjali, R., & Valarmathi, A. (2024). A novel hybrid deep learning IChOA-CNN-LSTM model for modality-enriched and multilingual emotion recognition in social media. *Scientific Reports*, *14*(1), 22270. https://doi.org/10.1038/s41598-024-73452-2

Guo, J. (2022). Deep learning approach to text analysis for human emotion detection from big data. *Journal of Intelligent Systems*, *31*(1), 113–126. https://doi.org/10.1515/jisys-2022-0001

Gupta, S., Kumar, P., & Tekchandani, R. K. (2023). Facial emotion recognition based real-time learner engagement detection system in online learning context using deep learning models. *Multimedia Tools and Applications*, *82*(8), 11365–11394. https://doi.org/10.1007/s11042-022-13558-9

Hendry, N. A. (2020). Young Women's Mental Illness and (In-)visible Social Media Practices of Control and Emotional Recognition. *Social Media + Society*, *6*(4), 1–10. https://doi.org/10.1177/2056305120963832

Khaireddin, Y., & Chen, Z. (2021). Facial Emotion Recognition. *State of the Art Performance on FER2013*, 1–12. https://doi.org/10.48550/arXiv.2105.03588

Li, S., & Deng, W. (2019). Blended Emotion in-the-Wild: Multi-label Facial Expression Recognition Using Crowdsourced Annotations and Deep Locality Feature Learning. *International Journal of Computer Vision*, *127*(6–7), 884–906. https://doi.org/10.1007/s11263-018-1131-1

Mehta, D., Siddiqui, M., & Javaid, A. (2018). Facial Emotion Recognition: A Survey and Real-World User Experiences in Mixed Reality. *Sensors*, *18*(2), 416. https://doi.org/10.3390/s18020416

Nahato, K. B., Nehemiah, K. H., & Kannan, A. (2016). Hybrid approach using fuzzy sets and extreme learning machine for classifying clinical datasets. *Informatics in Medicine Unlocked*, *2*, 1–11. https://doi.org/10.1016/j.imu.2016.01.001

Pansy , N., & Rupali, V. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, *11*(1), 81. https://doi.org/10.1007/s13278-021-00776-6

Poria, S., Chaturvedi, I., Cambria, E., & Hussain, A. (2016). Convolutional MKL Based Multimodal Emotion Recognition and Sentiment Analysis. *2016 IEEE 16th International Conference on Data Mining (ICDM)*, 439–448. https://doi.org/10.1109/icdm.2016.0055

Roy, A., Shivakumara, P., Pal, U., Gornale, S. S., & Liu, C.-L. (2023). A New Lightweight Attention-Based Model for Emotion Recognition on Distorted Social Media Face Images. *Pattern Recognition*, *14407*, 243–257. https://doi.org/10.1007/978-3-031-47637-2_19

Shuvo Alo, D. (2023). CK+ Dataset. *Kaggle.Com*.

Singh, S., & Nasoz, F. (2020). Facial Expression Recognition with Convolutional Neural Networks. *Published in the IEEE Conference Proceedings, Available via IEEE Xplore*, 1–5. https://doi.org/10.1109/ccwc47524.2020.9031283

Stone, Z., Zickler, T., & Darrell, T. (2010). Toward Large-Scale Face Recognition Using Social Network Context. *Proceedings of the IEEE*, *98*(8), 1408–1415. https://doi.org/10.1109/jproc.2010.2044551

Talaat, F. M., El-Gendy, E. M., Saafan, M. M., & Gamel, S. A. (2023). Utilizing social media and machine learning for personality and emotion recognition using PERS. *Neural Computing and Applications*, *35*(33), 23927–23941. https://doi.org/10.1007/s00521-023-08962-7

Tan, L., Zhang, K., Wang, K., Zeng, X., Peng, X., & Qiao, Y. (2017). Group emotion recognition with individual facial emotion CNNs and global image based CNNs. *Proceedings of the 19th ACM International Conference on Multimodal Interaction*, 549–552. https://doi.org/10.1145/3136755.3143008

Zhang, H., Jolfaei, A., & Alazab, M. (2019). A Face Emotion Recognition Method Using Convolutional Neural Network and Image Edge Computing. *IEEE Access*, *7*, 159081–159089. https://doi.org/10.1109/access.2019.2949741