Facial emotion recognition using Hybrid approach for random forest and convolutional neural network

AnjaniSuputri Devi D[1*], Suneetha Eluri[2]

Abstract

Facial expression recognition is a crucial component of emotion research and a prerequisite for human-machine interface.In general, face detection, feature extraction, and feature classification are part of a facial expression recognition system.Although traditional machine learning techniques have achieved significant success, the majority of them have difficult computational issues and cannot extract complete and abstract information.Deep learning-based techniques can achieve a greater detection accuracy for facial emotions, but they have a high hardware need and require a lot of training data and tuning parameters.In order to address the aforementioned issues, this paper suggests a method that combines features extracted by a convolutional neural network (CNN) with the C4.5 classifier to identify facial expressions. This method not only identifies the deficiencies of manually created features but also avoids the need for a deep learning model with a high hardware configuration.Random forest is also used to address several issues with C4.5classifier so it is based on the concept of ensemble learning, which is a process of combining multiple classifiers to solve a complex problem and to improve the performance of the model.By combining these two methods, the proposed system achieves high-level accuracy and these procedures enable the identification of the facial emotions: anger, disgust, surprise, sadness, fear, happiness, and neutral.Performance of the proposed system is evaluated using JAFFE and CK+ data sets.

KeyWords: Convolutional neural network(CNN),Random Forest(RF),JAFFE,CK+.

DOINumber:10.14704/nq.2022.20.8.NQ44830 **NeuroQuantology2022;20(8):8049-**

Introduction

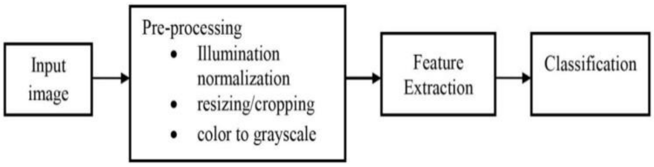
Human emotion categorization is defined as the technique of identifying human emotion through the measurement of different physiological signs as well as facial expressions, verbal expressions, gestures, and bodily movements.There is no denying the importance of human emotions in the creation of modern technology.In the modern world, human-computer interaction, automated tutoring systems, image and video retrieval, smart surroundings, and driver warning systems all place a high value on the analysis and recognition of emotion [1]. Additionally, psychologists and psychiatrists use emotion recognition to identify a variety of mental health issues.Researchers and scientists have put forth several algorithms and strategies over the last few decades to identify emotions from speech and facial data.Due to its complexity, it continues to be a difficult subject in the fields of artificial intelligence, computer vision, psychology, and physiology.Researchers and scientists generally concur that the most important factor in identifying human emotion is facial expression.However, due to the sensitivity of the external factors, such as lighting conditions and dynamic head movements, it is challenging to interpret human mood simply using facial expression features [2].In current years deep learning techniques have achieved great success as well as achieved better accuracy than traditional methods due to the inexpensive computational power.Best representation called the Convolutional Neural Network(CNN) has acquired excellent results in the field of computer vision. FER has been successfully applied to several CNN models, which have proved to be more efficient in feature learning and representations than standard techniques. The parameters of a series of filters, which capture both low-level generic features and high-level semantic information, may be determined by a well-designed CNN trained on millions of pictures. Many researchers are working to enhance the accuracy of existing deep learning models as well as their limitations.The design of the facial expression classifier, a crucial component of facial expression recognition, has a significant impact on the accuracy rate of facial expression recognition; as a result, the choice and implementation of the

Correspondingauthor: AnjaniSuputri Devi D

Address:1*Department of Computer Science and Engineering,SriVasavi Engineering College,Pedatadepalli, Andhra Pradesh, India.
2Department of Computer Science and Engineering, Jawaharlal Nehru Technological University, Kakinada (JNTUK), Kakinada, Andhra Pradesh, India.
1*E-mail:anjanihasini@gmail.com

Fig.1.The procedure of emotion recognition

classifier are crucial to the outcome.The classifier for facial expression identification should be highly computationally



efficient and capable of handling numerous data sets.Due to their straightforward operation characteristics, decision trees [3], a common approach of pattern recognition and data mining, have been extensively used in many disciplines.However, there are no sample requirements for this algorithm.These benefits apply to the classification of face expressions.The C4.5 classifier [4], which has been extensively utilised in the field of image identification, can categorise and identify face expressions. As a result, the C4.5 classifier is chosen in this paper as the classifier for expression recognition. Due to issues with overfitting and the limited generalizability of a single classifier, ensemble learning is incorporated into the decision tree algorithm to increase classification accuracy.The decision tree, which has good scalability and parallelism to high-dimensional data in classification, has a bottleneck problem that can be resolved by the random forest algorithm, which is the most typical algorithm among ensemble learning techniques.As a result, the random forest method is chosen in this research as the facial expression classifier.The rest of this paper is structured as follows.The earlier works that are relevant to our work are described in Section 2.The proposed technique is thoroughly explained in Section 3.Experiment findings and analyses are presented in Section 4.Conclusions and further work are presented in Section 5

2. Related Work

Face detection, feature extraction, and feature classification are the three processes that typically make up a face expression recognition system (Figure.1.), with feature extraction serving as the most important step.The Effectiveness of the extracted feature has a significant impact on the expression classification accuracy.

A hybrid method of principal component analysis (PCA) and local binary pattern was proposed by Luo et al (LBP).The entire image's global grayscale features were extracted using principal component analysis, while the local features were extracted using local binary processing (LBP).The recognition of facial expressions was done using the support vector machine (SVM).Geometry normalisation, brightness normalisation, histogram equalisation, picture filtering, and facial effective area segmentation are some of the preprocessing techniques used in this research.The training set's images were all adjusted to small sizes (24x24).93.75 percent of people were recognised.These facial traits were encoded using the HOG by Chen et al. [5].After that, a linear SVM was trained to classify facial expressions.They assessed the suggested strategy on an enlarged Cohn-Kanade dataset and the JAFFE dataset.The image used in the JAFFE experiment has a size of 256x256.The size was changed to 156 x156 after the facial region from the face image was acquired.This method was tested and contrasted with the other ways using the leave-one-sample-out strategy.On this dataset, the average categorization rate was 94.3 percent.

For the purpose of recognising facial expressions, Wen et al. [6] suggested an approach that combined several convolutional neural networks with probability-based fusing.Thesoftmax classifier was applied as the final layer in all of the planned CNN models to calculate the probabilities that the testing sample would belong to each class.The probability-based fusion method was used to combine the probabilities of multiple CNNs that were created as base classifiers.This paper used the implicit method to create CNNs with rich diversity since the diversity among the basis classifiers is regarded as a critical issue in performance for any ensemble learning method. They conducted their studies using four databases: JAFFE, CK+.However, the machine learning algorithm's feature information extraction in the previous literature is not sufficient.In experiments, a lot of training process and expensive hardware are needed for the CNN model.In based on the aforementioned problems, this work proposed a method for recognising facial expressions that combines CNN feature with a machine learning classifier. This method can both shorten the training period for the model and expand the coverage of extracted feature information.

3.ProposedMethod

FER has long been a challenge in computer vision. This work combines features from the CNN model and random forest to focus on expression recognition and determines the best fusion technique based on actual experiment data. Up to seven different emotions, including neutral, sadness, happiness, disgust, fear, anger, and surprise, can be recognised by the proposed work.

3.1.Random Forest:

Supervised machine learning algorithms like random forest are frequently employed in classification and regression issues.On various samples, it constructs decision trees and uses their average for classification and majority vote for regression.Instead of relying on one decision tree, the random forest takes the prediction from each tree and bases its prediction of the final output on the majority votes of predictions. It is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset.Higher accuracy and overfitting are prevented by the larger number of trees in the forest.

3.2.Convolutional neural network:

Convolutional neural network (CNN) [7] is a supervised machine learning technique that has been extensively used in the field of computer vision. It is capable of performing the feature extraction and classification process simultaneously and can automatically discover the multiple levels of representations in data.Figure 3 depicts the basic CNN model's overall structure.In [8,9], the CNN model is introduced in depth.A combination of various pooling and convolutional layers makes up CNN.CNN features a four layer network design in proposed work.Convolutional layer and pooling layer combinations are stacked and combined to create the CNN's strength.The final layer serves as a high-level semantic representation of the input and is used for dispensation.

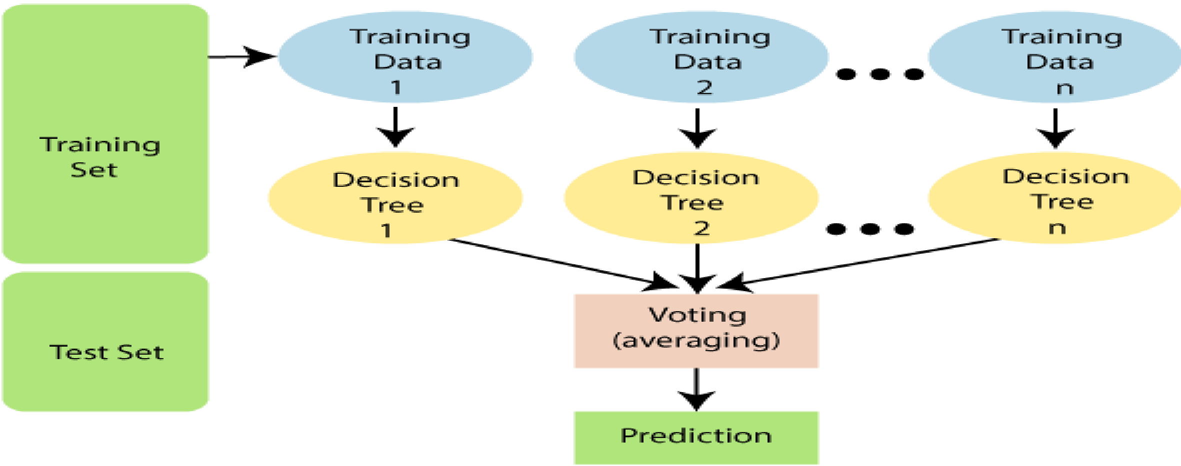


Fig.2.The random forest's general structure.

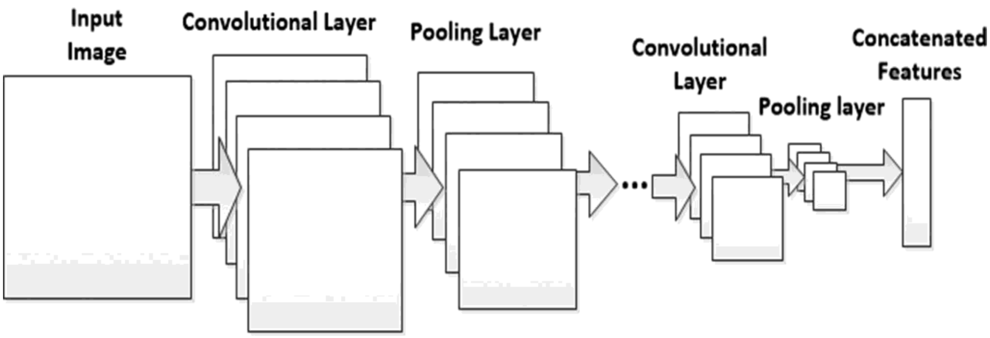


Fig.3. the convolutional neural network's general structure.Convolutional neural networks primarily consist of full connection, pooling, and convolution layers.Feature extraction is handled by the convolution layer, feature selection is handled by the pooling layer, and classification is handled by the full connection layer.

3.3.Proposed Block Diagram:

This research combines features from the CNN model and random forest to focus on expression recognition and determines the best fusion strategy based on actual experiment data.After passing through each layer of the CNN model, the original data will have undergone some information loss.Because the nature of the raw data has changed by the time it reaches the entire connection layer, random forests cannot be directly applied to the CNN structure. So the most reasonable approach, which is to incorporate the random forest into the final pool layer. The proposed model's framework shown in figure 4.

3.3.1.C4.5 DecisionTree:

Due to its great computational efficiency, capacity to handle a large number of data sets, simplicity, and ease of use, C4.5 classifier has been one of the decision tree algorithms that has seen the most use in the field of image recognition.C4.5 is therefore used in this paper as the classifier for facial expression recognition. The acquisition of information gain rate, which can be observed in Equation 1, is the most fundamental computation of the C4.5 classifier throughout the processing of the classification of continuous data (1).The decision tree's classification effectiveness is influenced by the information gain rate, which also influences the nodes that are selected throughout the decision tree construction process.

$$\begin{aligned}\mathrm{Gain}(D, a_V) &= \max_{t\in T_a}\mathrm{Gain}(D, a_V, t)\\ &= \max_{t\in T_a}\frac{\mathrm{Ent}(D)}{IV(a_v)} - \frac{\sum_{\lambda\in\{-1,1\}}\frac{|D_t^\lambda|}{|D|}\mathrm{Ent}(D_t^\lambda)}{IV(a_v)}\end{aligned} \tag{1}$$

Where

$$IV(a_v) = -\sum_{v=1}^{V} \frac{|D^V|}{|D|} \log_2 \frac{|D^V|}{|D|}$$

$$\mathrm{Ent}(D) = -\sum_{k=1}^{|y|} p_k \log_2 p_k = -\frac{m_1}{\sum_{i=1}^{7} m_i} \log_2 \frac{m_1}{\sum_{i=1}^{7} m_i} - \frac{m_2}{\sum_{i=1}^{7} m_i} \log_2 \frac{m_2}{\sum_{i=1}^{7} m_i} - \cdots - \frac{m_7}{\sum_{i=1}^{7} m_i} \log_2 \frac{m_7}{\sum_{i=1}^{7} m_i}$$

$$\mathrm{Ent}(D_t^{\lambda}) = -\frac{l_{t_1}^{\lambda}}{\sum_{i=1}^{7} l_{t_i}^{\lambda}} \log_2 \frac{l_{t_1}^{\lambda}}{\sum_{i=1}^{7} l_{t_i}^{\lambda}} - \frac{l_{t_2}^{\lambda}}{\sum_{i=1}^{7} l_{t_i}^{\lambda}} \log_2 \frac{l_{t_2}^{\lambda}}{\sum_{i=1}^{7} l_{t_i}^{\lambda}} - \cdots - \frac{l_{t_7}}{\sum_{i=1}^{7} l_{t_i}^{\lambda}} \log_2 \frac{l_{t_7}^{\lambda}}{\sum_{i=1}^{7} l_{t_i}^{\lambda}}$$

Let's begin with some notation.$[a_1, a_2, a_3, \ldots, a_v, \ldots, a_V]$stands for V attribute sets. $|y|$indicates the total number of categories. $IV(a_v)$stands for the fixed value of the attribute $a_v \cdot |D| = \sum_{i=1}^{|y|} m_i$ stands for the sample number in the whole dataset. $|D^V|$stands for the sample number in attribute value a_v. Suppose there are n different values for attribute a_v, then sort these values from smallest to largest (i.e., $\{a_1^v, a_2^v, \ldots, a_n^v\}$). D can be divided into two different subsets ($D_t^-$and D_t^+)based on partition point t. Where, $T_a = \left[\frac{a_i^v + a_{i+1}^v}{2} \mid 1 \le i \le n-1\right]$, which means that the middle point $\frac{a^i + a^{i+1}}{2}$ of the interval $[a_i^v, a_{i+1}^v)$is used as the division. These points can therefore be considered separate attribute values. Suppose $\cup_{i=1(\lambda=-1)}^{7} l_{t_i}^{\lambda}$ and $\cup_{i=1(\lambda=1)}^{7} l_{t_i}^{\lambda}$ stands for the number of seven expression labels in these two datasets $D_t^-$and D_t^+, respectively.

Next, discover that the logarithm operation appears virtually constantly during the process of operating on each equation and frequently when computing the information gain rate [10].The system's computational speed will be affected by a large number of logarithm operations.By introducing Taylor series expansion, which shortens the operation time and enhances the system's real-time reaction, this article enhances the information gain formula.The new information gain rate equation looks like this.t as discrete attribute value these points.

$$\begin{aligned}\mathrm{Ent}(D) = & -\frac{m_1}{\sum_{i=1}^{7} m_i} \log_2 \left(1 - \frac{\sum_{i=2}^{7} m_i}{\sum_{i=1}^{7} m_i}\right) - -\frac{m_1}{\sum_{i=1}^{7} m_i} \log_2 \left(1 - \frac{m_1 + \sum_{i=3}^{7} m_i}{\sum_{i=1}^{7} m_i}\right) \\ & \ldots - \frac{m_7}{\sum_{i=1}^{7} m_i m_i} \log_2 \left(1 - \frac{\sum_{i=1}^{7})}{\sum_{i=1} m_i}\right. \\ = & \frac{m_1 \times \sum_{i=2}^{7} m_i}{m_2 \times (m_1 + \sum_{i=2}^{7} m_i)} + \cdots + \left.\frac{m_7 \times \sum_{i=1}^{6} m_i}{(\sum_{i=1}^{7} \ln 2} + m_i\right)^2 \ln 2\end{aligned}$$

(2),(3)

Applying Equation (3) to Equation (2) will yield the new equation for information gain rate (2).Equation (3) replaces the difficult log calculation in Equation (1) with four straightforward procedures, considerably enhancing the system's operational effectiveness and real-time performance.

3.3.2. Generation of the New Random Forest

The ensemble learning method is used to increase classification accuracy because a single classifier is prone to overfitting and has poor generalisation capabilities. Multiple decision trees are used in the 2001 proposal of random forest (RF) [11].The base learner in the ensemble learning concept should be "good but different," that is, each individual learner should have a relatively high recognition rate that is different from the others.However, when choosing a single decision tree, the decision tree's number is predetermined and randomly generated decision trees are used, and integrated voting is used to determine the outcome.The process of creating numerous single decision trees using the conventional method may result in decision trees that are not significantly different from one another or low recognition rates for the generated individual decision trees, both of which will have an impact on the outcome.This paper suggests a probability selection method that is both unique and good, while also satisfying the demands of variety.Algorithm 1 displays the specific proposed work.

Algorithm:1

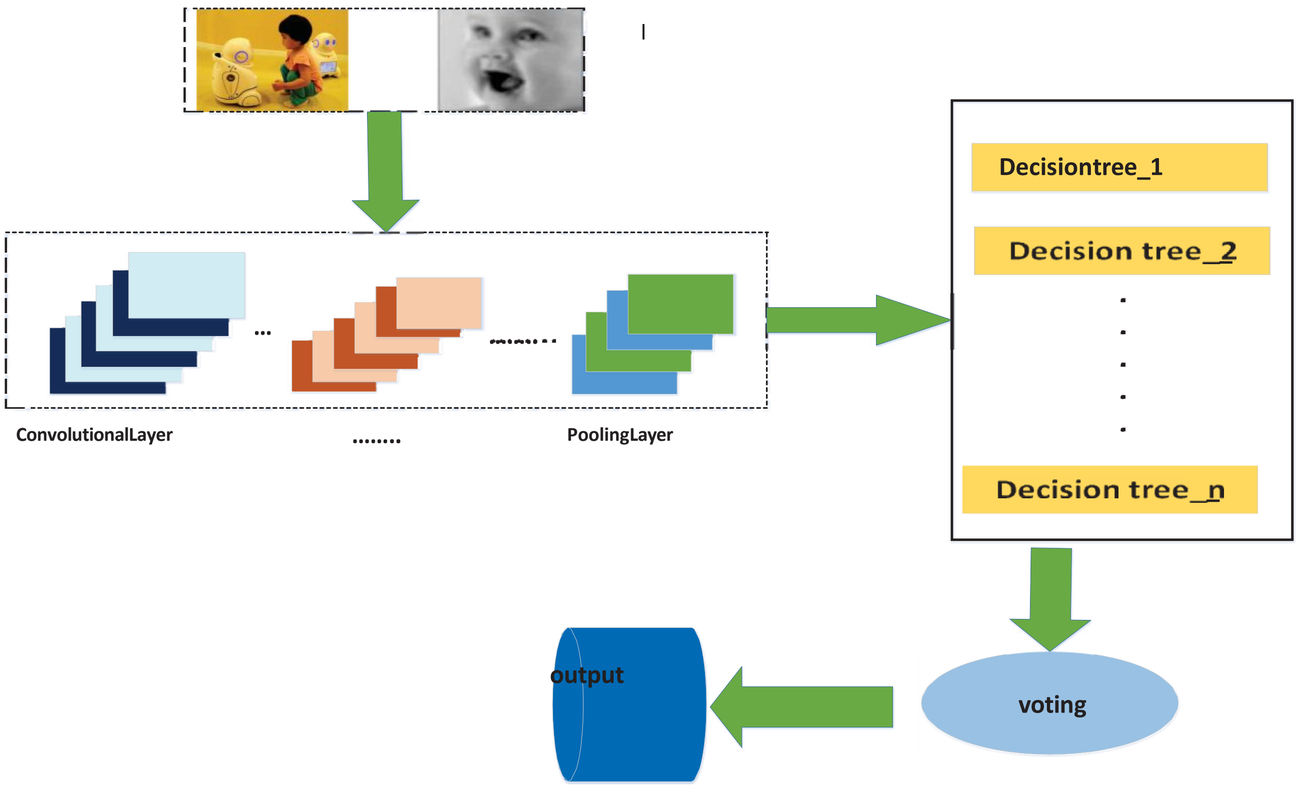
Step 1.Construct the root node
Step 2.Create the leaf nodes
Step 3. Information gain rate is calculated for each attribute
Step 4.Set the threshold value δ
Step 5.If random<δ
Step 6.Select the optimal decision trees
Step 7.number=number+1
Step 8.else
Step 9. Randomly select the decision tree
Step 10.number=number+1.
Step 11.All the decision trees are selected to form a random forest.

4.Result anddiscussion

Using the suggested FER in MATLAB, the outcomes are assessed.Japanese Female Facial Expressions (JAFFE) and Cohn-Kanade+ (CK+) database images are used as input [12].In the JAFFE, there are 213 [256x256] pixel pictures depicting 7 different facial expressions.JAFFE database contains "7" primary expressions from 10 female Japanese models, along with 30 images of anger, 29 images of disgust, 32 images of fear, 31 images of happiness, 30 images of neutrality, 31 images of sadness, and 30 images of surprise.In contrast, the CK+ contains 486 ,640 x 490 pixel photos in 97 poses.Images were taken in time frames where the initial ones were neutral, then they were transformed into the desired emotion in the last frame. CK+ typically has Euro-Americans along with Afro-Americans (the peak frame).In these "2" datasets, 80% of the images are used for training and 20% are used for testing.The next sections elaborate on the comparison based on these two datasets.

$$\begin{aligned}
\mathrm{Gain}(D, a_V) &= \max_{t \in T_a} \mathrm{Gain}(D, a_V, t) \\
&= \max_{t \in T_a} \frac{\frac{1}{\left(\sum_{i=1}^{7} m_i\right)^2 \ln 2}\left[m_1 \times \sum_{i=2}^{7} m_i + m_2 \times \left(m_1 + \sum_{i=2}^{7} m_i\right) + \cdots + m_7 \times \sum_{i=1}^{6} m_i\right]}{\frac{|D^V| \times \sum_{v=1}^{V-1} |D^v|}{|D|^2 \times \ln 2}} \\
&- \frac{\sum_{\lambda \in \{-1,1\}} \frac{|D_t^\lambda|}{|D|} \frac{1}{\left(\sum_{i=1}^{7} l_{t_i}^\lambda\right)^2 \ln 2}\left[l_{t_1}^\lambda \times \sum_{i=2}^{7} l_{t_i}^\lambda + l_{t_2}^\lambda \times \left(l_{t_1}^\lambda + \sum_{i=3}^{7} l_{t_i}^\lambda\right) + \cdots + l_{t_7}^\lambda \times \left(\sum_{i=1}^{6} l_{t_i}^\lambda\right)\right]}{\frac{|D^V| \times \sum_{v=1}^{V-1} |D^v|}{|D|^2 \times \ln 2}} \\
&= \max_{t \in T_a} \frac{|D|^2}{\left(\sum_{i=1}^{7} m_i\right)^2 \times |D^V| \times \sum_{v=1}^{V-1} |D^v|}\left[m_1 \times \sum_{i=2}^{7} m_i + m_2 \times \left(m_1 + \sum_{i=2}^{7} m_i\right) + \cdots + m_7 \times \sum_{i=1}^{6} m_i\right] \\
&- \sum_{\lambda \in \{-1,1\}} \frac{|D_t^\lambda|}{|D|} \frac{|D|^2}{\left(\sum_{i=1}^{7} l_{t_i}^\lambda\right)^2 \times |D^V| \times \sum_{v=1}^{V-1} |D^v|}\left[l_{t_1}^\lambda \times \sum_{i=2}^{7} l_{t_i}^\lambda + l_{t_2}^\lambda \times \left(l_{t_1}^\lambda + \sum_{i=3}^{7} l_{t_i}^\lambda\right) + \cdots + l_{t_7}^\lambda \times \left(\sum_{i=1}^{6} l_{t_i}^\lambda\right)\right]
\end{aligned}$$

Fig.4. The proposed model's framework.Convolutional neural network (CNN) features are acquired in this model's previous work, and its later work connects CNN features to an upgraded random forest to classify face expressions.



eISSN 1303-5150

4.1.JAFFE database performance analysis

Figure 5 illustrates sample face images from the suggested JAFFE. It displays example 7 expression photographs.The input image from the JAFFE dataset is shown in Figure 5a.The pre-processed image using the Gamma-HE approach is shown in Figure 5b.The face points that were retrieved using PHOG-centric SMD are shown in Figure 5c.The segmented parts, including the left eye, face, nose, right eye, and mouth, are shown in Figure 5d.

4.1.1.Classification performance for overall emotions

JAFFE compares the proposed classifier to the current CNN[13], NN-Levenberg Marquardt (NNLM) [14], NN-Gradient Descent (NNGD) [15], NN-Evolutionary Algorithm (NN-EA) [16], NN-Firefly (NNFF) [17], and NN-Particle Swarm Optimization (NNPSO) [18], NN-Gray Wolf Optimization (NNGWO) [19] are a few examples of the algorithms used and methods in terms of specificity,accuracy,sensitivity. The proposed classifier and the existing classifiers are explained in relation to sensitivity, specificity, and accuracy in the comparison analysis table in table 1.

The proposed has an accuracy of 0.9826, whereas the current NNGWO has a lesser accuracy of 0.9047 when compared to the proposed classifier.Similar to how the suggested FER utilising the classifier is more accurate than the classifiers already in use,The existing NNLM, CNN, NNEA, NNGD, NNFF, NNGWO, and NNPSO have specificities of 0.9206, 0.9603, 0.9563, 0.8888, 0.9563, 0.9444, and 0.9404, respectively, while the suggested classifier reaches 0.9845-specificity based on the specificity metric.Due to this fact, the proposed has the highest level of specificity when taking into account the current methods.Regarding the sensitivity metric, the suggested one performed better because it has a sensitivity of 0.9048, when all the current classifiers had sensitivity values below 0.80.As a result, it is discovered that the proposed offers superior performance in comparison to the current classifiers.

4.1.2 The performance of classification for various emotions

As demonstrated in the following table , the performances of the proposed and existing classifiers are compared based on their accuracy for a variety of expressions, including neutral, sad, surprised, furious, pleased, fearful, and disgusting.Table 2 compares theaccuracy-based performance of the proposed classifier and the current classifiers for various emotions. The suggested classifier obtains the maximum accuracy of 95.33% for the happy expression, while the existing NNLM, CNN, NNGD, NNEA, NNFF, NNGWO, and NNPSO reach accuracy of 91.67%, 92.98%, 79.77%, 89.59%, 84.53%, 91.68%, and 83.65%, respectively.The suggested classifier achieves 99.99% accuracy for the expression of disgust, whereas the existing classifiers provide lower accuracy.The findings suggest that the suggested classifier performs at a high level while taking into account the current techniques.

4.2.CK+ database performance analysis

Figure 6 illustrates the sample facial images of the CK+.

4.2.1Emotional classification performance overall

In terms of specificity, accuracy, sensitivity, metrics for the CK+, the proposed classifier is compared to the current NNLM, CNN, NNGD, NNFF, NNEA, NNGWO, and NNPSO. Table 3 contrasted the proposed classifier performances with those of the current classifiers in terms of their specificity, sensitivity, and accuracy for the CK+.The proposed and other classifiers perform better than the existing NNGD based on the specificity metric. The proposed has the highest overall specificity (0.9943) in the

comparison.The suggested classifier obtains 0.9985 accuracy according to the accuracy metric, however the existing NNLM, CNN, NNGD, NNEA, NNFF, NNGWO, and NNPSO only reach 0.8416, 0.87775, 0.7828, 0.8987, 0.8987, 0.8938, and 0.8742 accuracy, respectively.
Similar to this, the proposed classifier performs better when looking at the sensitivity measure.As a result, it can be concluded from the discussion that the proposed classifier gives higher performance than the current classifiers.

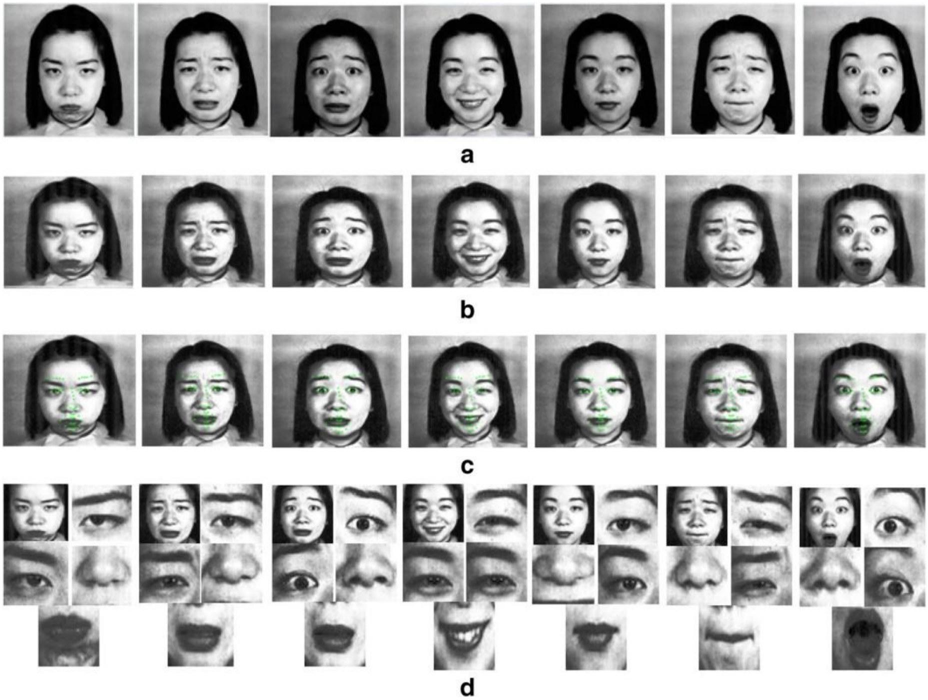


Fig. 5 Show the sample seven expression photos obtained from the JAFFE input images, preprocessed images, facial point images, and segmentation images.

Classifiers	Performance Matrics		
	Specificity	Sensitivity	Accuracy
CNN [13]	0.9603	0.7863	0.9586
NNLM [14]	0.9206	0.5526	0.8952
NNGD [15]	0.8888	0.3526	0.8120
NNEA [16]	0.9563	0.7521	0.9124
NNFF [17]	0.9563	0.7546	0.9124
NNPSO [18]	0.9404	0.6538	0.9121
NNGWO	0.9444	0.6628	0.9125

[19]			
Proposed C4.5	0.9845	0.9048	0.9826

Table.1 Analysis of the proposed and current classifiers in comparison

Emotions	Proposed classifier	CNN [13]	NNLM [14]	NNGD [15]	NNEA [16]	NNFF [17]	NNPSO [18]	NNGWO [19]
Neutral	96.23	96.23	93.85	89.19	93.85	91.47	95.63	92.94
Sad	98.71	93.85	89.09	84.33	97.72	98.68	92.95	92.95
Happy	95.33	92.98	91.67	79.77	89.59	84.53	86.65	91.68
Surprise	98.61	95.43	80.97	85.79	90.51	97.72	85.73	85.77
Fear	97.72	90.57	83.43	66.45	92.65	90.57	85.81	83.43
Angry	97.72	90.55	80.99	78.65	95.35	92.95	92.96	90.56
Disgust	99.99	95.33	88.21	85.85	91.55	96.23	96.24	97.85

Table2Comparedtheperformanceoftheproposedclassifierwiththeexistingclassifiersfordifferente motions based on theaccuracy

Classifiers	Performance Metrics		
	Specificity	Sensitivity	Accuracy
CNN [13]	0.9703	0.8863	0.87775
NNLM [14]	0.9317	0.6566	0.8416
NNGD [15]	0.8896	0.3539	0.7828
NNEA [16]	0.9573	0.7535	0.8987
NNFF [17]	0.9575	0.7575	0.8987
NNPSO [18]	0.9505	0.6693	0.8742
NNGWO [19]	0.9526	0.6785	0.8938
Proposed	**0.9943**	**0.9048**	**0.9985**

Table.3 Analysis of the proposed and current classifiers in comparison

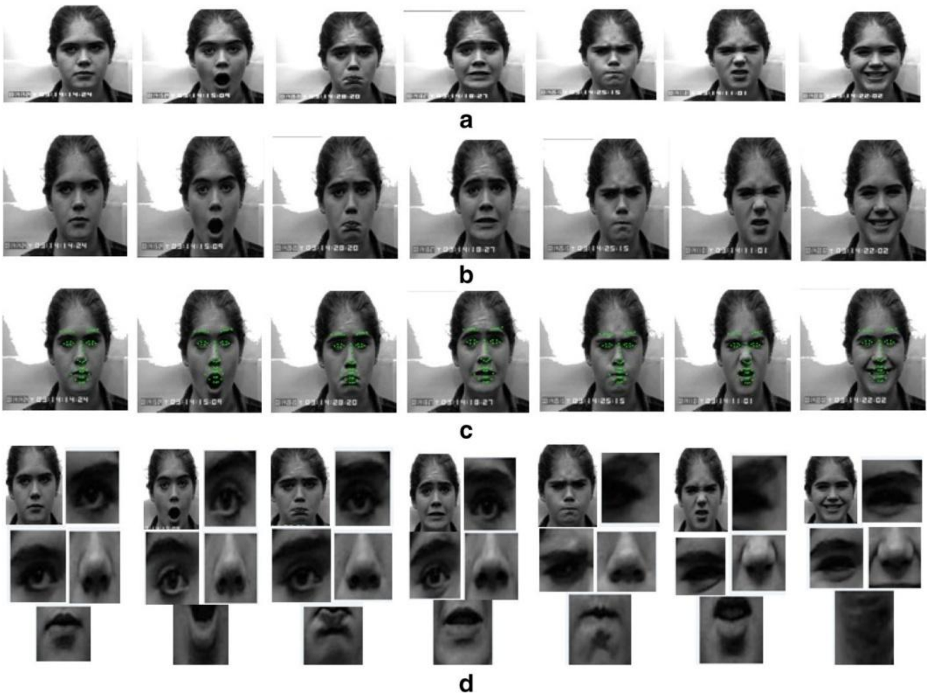


Fig. 6 Sample 7 expression images from CK+ a input images, b pre-processed images, c extracted facial point images, and d segmented images

Emotions	Proposed classifier	CNN [13]	NNLM [14]	NNGD [15]	NNEA [16]	NNFF [17]	NNPSO [18]	NNGWO [19]
Neutral	96.23	91.56	90.87	86.86	89.55	94.95	89.82	91.56
Sad	98.88	89.82	92.33	91.45	89.77	97.45	89.86	92.59
Happy	99.53	85.92	83.42	73.82	88.65	93.24	85.85	85.88
Surprise	98.44	90.95	80.51	56.88	84.59	89.82	85.78	85.88
Fear	100	87.56	77.72	70.95	92.33	92.11	88.98	86.35
Angry	100	82.95	79.45	85.56	90.38	90.85	85.56	91.55
Disgust	99.56	86.33	85.71	84.56	95.58	92.11	88.21	92.65

4.Conclusion

Human emotions are expressed through behaviours, speech, positions, facial expressions, and actions.Research is being done to examine how those emotions and channels interact.This paper proposes a hybrid strategy for random forest and convolutional neural network facial emotion identification.Based on this research, the eight most common facial emotions—anger, disgust, surprise, sorrow, fear, happiness, and neutral are identified.Performance of the suggested system is evaluated using JAFFE and CK+.The performance of the suggested system is contrasted with that of the existing systems for these two datasets.In terms of specificity, sensitivity, and accuracy, the suggested classifier is compared to the current NNLM, CNN, NNGD, NNFF, NNEA, NNGWO, and NNPSO.When compared to the current methods, the proposed classifier performed better, achieving 0.9985 accuracy for the CK+ and 0.9826 accuracy for the JAFFE.The effectiveness of the novel and established approaches is then evaluated for each emotion in the two datasets.As a result, it implies that the proposed system performs at a high level accuracy.Higher training time is the main problem with deep learning-centered FER, though.Due to the system's intricacy and other problems, this research does not conduct trials under unique circumstances like makeup and occlusion.How to recognise facial expressions in these challenging circumstances requires more study.Additionally, to ensure that the trained network performs well at generalisation, as many samples as feasible must be collected for the convolutional neural network.

References

1. Shojaeilangari, S.; Yau, W.Y.; Nandakumar, K.; Li, J.; Teoh, E.K. Robust representation and recognition of facial emotions using extreme sparse learning. IEEE Trans. Image Process. 2015, 24, 2140–2152.

2. Ko, K.E.; Sim, K.B. Development of a Facial Emotion Recognition Method based on combining AAM with DBN. In Proceedings of the 2010 International Conference on Cyberworlds (CW), Singapore, 20–22 October 2010; pp. 87–91.

3. Quinlan, J.R. Induction on decision tree. Mach. Learn. 1986, 1, 81-106. [CrossRef]

4. Setsirichok, D.; Piroonratana, T.; Wongseree, W.; Usavanarong, T.; Paulkhaolarn, N.; Kanjanakorn, C.; Sirikong, M.; Limwongse, C.; Chaiyaratana, N. Classification of complete blood count and haemoglobin typing data by a C4.5 decision tree, a naïve Bayes classifier and a multilayer perceptron for thalassaemia screening. Biomed. Signal Process. Control 2012, 7, 202-212. [CrossRef]

5. Chen, J.; Chen, Z.; Chi, Z.; Fu, H. Facial expression recognition based on facial components detection and hog features. In Proceedings of the International Workshops on Electrical and Computer Engineering Subfields, Istanbul, Turkey, 22-23 August 2014; pp. 884-888.

6. Wen, G.; Zhi, H.; Li, H.; Li, D.; Jiang, L.; Xun, E. Ensemble of Deep Neural Networks with Probability-Based Fusion for Facial Expression Recognition. Cognit. Comput. 2017, 9, 1-14. [CrossRef]

7. Lawrence, S.; Giles, C.L.; Tsoi, A.C.; Back, A.D. Face recognition: a convolutional neural-network

approach. IEEE Trans. Neural Netw. 1997, 8, 98-113. [CrossRef] [PubMed]

8. Kim, P. Convolutional Neural Network. In MATLAB Deep Learning; Apress: Berkeley, CA, USA, 2017.

9. Ming, L.; Hu, X. Recurrent

10. convolutional neural network for object recognition. In Proceedings of the Computer Vision and Pattern Recognition, Boston, MA, USA, 7-12 June 2015; pp. 3367-3375. 22. Xu, M.; Wang, J.L.; Chen, T. Improved Decision Tree Algorithm: ID3 +. In Intelligent Computing in Signal Processing and Pattern Recognition; Springer: Berlin/Heidelberg, Germany, 2006; pp. 141-149.

11. Butaye, J.; Jacquemyn, H.; Hermy,

M. Differential colonization causing non-random forest plant community structure in a fragmented agricultural landscape. Ecography 2001, 24,369-380. [CrossRef]

12. Oza, N.C. Online Ensemble Learning. In Proceedings of the Seventeenth National Conference on Artificial Intelligence and Twelfth Conference on on Innovative Applications of Artificial Intelligence, Austin, TX, USA, 30 July-3 August 2000.

13. Xie S, Hu H (2018) Facial expression recognition using hierarchical features with deep comprehensive multipatches aggregation convolutional neural networks. IEEE Trans Multimedia21(1):211–220

14. YangM,LiuY,YouZ (2017) TheEuclideanembeddinglearning based onconvolutionalneuralnetwork for stereo matching. Neurocomputing 267:195–200.https://doi.org/10.1016/j.neucom.2017.06.007

JianaJ,LinJ,Xiao-huaaZ,HaoaL(2011)Inversionofneuralnetworkrayleighwavedispersionbasedon LM algorithm. Adv Control EngInfSci15:5126–5132.https://doi.org/10.1016/j.proeng.2011.08.951

15. KobayashiM(2017)GradientdescentlearningforquaternionicHopfieldneuralnetworks.Neurocomputing 260:174–179.https://doi.org/10.1016/j.neucom.2017.04.025\

16. Boutorh A, GuessoumA (2016) Complex diseases SNP selection and classification by hybrid associationruleminingandartificialneuralnetwork—basedevolutionaryalgorithms.EngApplArtifIntell51:58–70. https://doi.org/10.1016/j.engappai.2016.01.004

17. Sánchez D, Melin P, Castillo O (2017) Optimization of

modular granular neural networks using a firefly algorithm for human recognition. EngApplArtifIntell64:172–186. https://doi.org/10.1016/j.engappai. 2017.06.007

18. LinC-H(2016)Novelapplicationofcontinuouslyvariabletransmissionsystemusingcompositerecurrent Laguerreorthogonalpolynomialsmodified PSONNcontrolsystem.ISATrans64:405–417.https://doi.org/ 10.1016/j.isatra.2016.05.013

19. Sreedharan NPN, GanesanB, Raveendran R, SaralaP, Dennis B (2018) Grey wolf optimisation-based feature selection and classification for facial emotion recognition.IET Biom7(5):490–499

eISSN 1303-5150