

BRAIN. Broad Research in Artificial Intelligence and Neuroscience

e-ISSN: 2067-3957 | p-ISSN: 2068-0473

Covered in: Web of Science (ESCI); EBSCO; JERIH PLUS (hkdir.no); IndexCopernicus; Google Scholar; SHERPA/RoMEO; ArticleReach Direct; WorldCat; CrossRef; Peeref; Bridge of Knowledge (mostwiedzy.pl); abcdindex.com; Editage; Ingenta Connect Publication; OALib; scite.ai; Scholar9; Scientific and Technical Information Portal; FID Move; ADVANCED SCIENCES INDEX (European Science

Evaluation Centre, neredataltics.org); ivySCI; exaly.com; Journal Selector Tool (letpub.com); Citefactor.org; fatcat!; ZDB catalogue; Catalogue SUDOC (abes.fr); OpenAlex; Wikidata; The ISSN Portal; Socolar; KVK-Volltitel (kit.edu)

2026, Volume 17, Issue 1, pages: 26-39

Submitted: September 24th, 2025 | Accepted for publication: November 23rd, 2025

Emotion Recognition Method based on Convolutional Neural Network and Black Widow Optimisation Algorithm

Pratiksha Deshmukh

Affiliation, including city and Country
Department of Computer Engineering
Thakur College of Engineering Technology,
Mumbai, Maharashtra, India.
pratiksha.deshmukh@thakureducation.org,
<https://orcid.org/0009-0002-8172-3149>

Harshali Patil

Thakur College of Engineering Technology,
Mumbai, Maharashtra, India.
harshali.patil@thakureducation.org,
<https://orcid.org/0000-0003-1228-8737>

Abstract:

Background: The automated Emotion Recognition (ER) method has gained popularity in several applications, namely, psychology and mental health. To accomplish this goal, deep learning algorithms are employed. These algorithms extract the appropriate features from the dataset images and recognise the emotions.

Methodology: In this paper, we have developed an emotion recognition method based on the Convolutional Neural Network (CNN). Furthermore, the main focus is on hyper-tuning the learning parameters of the CNN algorithm using the metaheuristic Black Widow Optimisation (BWO) to enhance the recognition accuracy. The BWO algorithm is based on the mating process of black widow spiders and provides a better convergence rate to find the optimal solution than other metaheuristic algorithms.

Results: The proposed method was simulated on the FER2013 dataset. The dataset was split into a 70:30 ratio. This reflects that 70% of the dataset was used for training purposes, whereas 30% of the dataset was utilised for testing purposes. The proposed method shows impressive results in recognising various emotions and achieves high values for the performance metrics, such as an average accuracy value of 0.99392, a precision value of 0.98467, a recall value of 0.98117, and an F1-score value of 0.96731. Finally, we have performed the comparative analysis of the presented approach with existing studies. The result shows that we have achieved better results due to employing the hyper-tuning strategy in the proposed method.

Conclusion and Recommendation: The proposed ER method is effectively recognising the different emotions. However, the same dataset was used for training and testing purposes, which negatively impacts the robustness and generalisation. Therefore, in the future, we will combine multiple datasets and balance them to evaluate the effectiveness of the proposed method.

Keywords: BWO; classification; CNN; emotion; healthcare; hyperparameter; metaheuristic.

How to cite: Deshmukh, P. & Patil, H. (2026). Emotion recognition method based on convolutional neural network and black widow optimisation algorithm. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 17(1), 26-39. <https://doi.org/10.70594/brain/17.1/2>

1. Introduction

Human emotions are fundamental characteristics that are crucial for effective communication (Akhand et al., 2021). Humans use a variety of body language, facial expressions, and speech to convey their emotions. The most widely used and deeply researched component of emotion identification is facial expression analysis. Recently, the recognition of emotion through facial expressions has emerged as an interesting research area in the fields of psychiatry, psychology, and mental disorders (Khan, 2022). Automated ER from facial expressions is also required for healthcare systems, smart living, emotional disorder diagnosis, human-robot interaction (HRI), human-computer interaction (HCI), and HRI-based social welfare programmes. As a result, facial expression recognition (FER) has drawn the attention of researchers because it has a lot of potential uses.

The aim of the facial ER method is to map emotional states based on the different facial expressions. The two primary processes of conventional FER are emotion recognition and feature extraction (Wojciuk et al., 2024). Nowadays, comprehensive work is being conducted to develop efficient automated face recognition systems using computer vision methods, AI, image processing, ML, and DL.

In the ML-based emotion recognition method, the main stages are reading the facial expression, followed by pre-processing (face alignment, augmentation, normalisation, enhancement), feature extraction (LBP, HOG, SIFT, SURF, LTP), and classification (SVM, DT, RF, and KNN) (Khan, 2022). However, in ML algorithms, feature extraction needs to be performed manually for emotion recognition, whereas in deep learning-based methods, feature extraction is required automatically. Therefore, in this research, the deep learning area is explored to design an emotion recognition method. Furthermore, the key emphasis is on hyper-tuning the parameters of the deep learning algorithm. Finding the best settings for a deep learning model's hyperparameters throughout the learning process is very important for influencing how the model behaves.

In the literature, the CNN is the most preferred DL algorithm for facial emotion recognition. Therefore, when hyperparameter optimisation techniques are used, (I) CNN performance gets better (by a few percentage points), (II) method development time goes down (when paired with distributed training for quick parallel searching), and (III) solutions are provided faster, which is very important for different applications. The hyperparameters are mainly related to network design and training techniques. They are utilised to find a set of parameters that delivers the best accuracy on a test set when used during CNN training. However, it is not easy to choose the hyperparameters that guarantee CNN's optimal performance. It is computationally costly since it often requires assessing the CNN model many times (Amou et al., 2022; Cabada et al., 2019; Lateef & Abbas, 2022; Wojciuk et al., 2024). Hyperparameter configuration is difficult, and examining the hyperparameter space is time-consuming, computationally costly, and inefficient. For hyperparameter optimisation, many methods have been developed, such as Hessian-free optimisation, Random Search, and Grid Search (Rimal, Sharma, & Alsadoon, 2024). However, previous optimisation approaches examine hyperparameter values at predefined levels, resulting in higher computational complexity. Therefore, the main research question addressed in this study is how to efficiently explore the solution space of the parameters for hyper-tuning while minimising computational complexity. To accomplish this goal, we have explored an optimisation approach, namely a metaheuristic, which outperforms the grid search and random search methods.

The metaheuristic algorithm explores the problem's solution space based on the objective function to find the optimal or near-optimal solution. In the metaheuristic algorithm, the population and iteration parameters play an important role, and their values are defined according to the computational complexity of the designed method. The population parameter is randomly initialised in the solution space according to the population size and evaluated based on the objective function to determine the best solution. After that, new populations are generated by considering the initial population values to explore new solutions in the solution space, which may outperform the current

best solution. The new population generation process is iterated according to the iteration parameter value.

The main contribution of this manuscript is to present a method for classifying different emotions using the CNN model. Furthermore, the key emphasis is given to hyper-tuning the CNN learning parameters using the BWO algorithm for enhancing the classification accuracy of the presented method. We have chosen the BWO algorithm over other metaheuristic algorithms due to its better exploration rate in searching the solution space of the parameters. Furthermore, it contains three stages, namely, procreation, cannibalism, and mutation. Out of these stages, the cannibalism stage helps in removing inappropriate solutions while exploring the solution space. Thus, it provides a better convergence rate to reach the desired solution more efficiently. Hence, the proposed method shows impressive results in classifying various emotions and achieves high performance metric values, such as accuracy, precision, recall, and F1-Score of 0.99392, 0.98467, 0.98117, and 0.96731, respectively. Finally, we have performed a comparative analysis of the presented approach with existing studies. The result shows that we achieved better outcomes due to employing the hyper-tuning strategy in the proposed method.

The remaining structure of the manuscript is as follows. Section 2 presents the related work on the advancement in emotion recognition methods based on deep learning and the limitations of previous studies. Section 3 presents the preliminaries in which an overview of the CNN and BWO algorithms is given. Section 4 presents the proposed emotion recognition classification method based on hyper-tuning the CNN model for better classification. Section 5 shows the simulation results and comparative analysis with the previous approaches. Finally, the manuscript is summarised and future directions are defined in Section 6.

2. Related Work

Khattak et al. (2021) designed a CNN-based emotion recognition method by analysing facial expressions. In addition, in their method, age and gender factors were also recognised. They evaluated their method on three datasets, namely JAFFE, UTKFace, and CK+. Furthermore, a maximum proportion of the data (90%) was used for training purposes. Therefore, the results indicate that they achieved accuracy values of 95.65%, 98.5%, and 99.14% for emotion, age, and gender recognition, respectively. S.B. et al. (2024) recognised seven emotions using the Xception-CNN and evaluated their method on the FER2013 database available on Kaggle. They achieved an accuracy value of 95.60%. Further, like the previous study, Bhagat et al. (2024) used the CNN algorithm to extract local and global facial features for emotion recognition. Furthermore, in their work, pre-trained classifiers, namely ResNet, VGGNet, EfficientNet, and a Haar face classifier, were used. They accomplished an accuracy value of 82% for the FER2013 database. Haq et al. (2024) used the MobileNet-V1 algorithm for classifying seven emotions on the FER2013 and a random database. The result indicates that they have achieved an accuracy value of 97.7%. Agung, Rifai, and Wijayanto (2024) recognised ten emotions by utilising CNN in two approaches, namely, Inception-V3 and MobileNet-V2. Furthermore, in their work, the Taguchi method was used to hyper-tune the parameters. They achieved accuracy and F1-score values of 96% and 0.95 for the video dataset considered for evaluation purposes. Karatay et al. (2023) used the Gaussian Mixture Model, CNN, and LSTM algorithm for emotion recognition by analysing the facial and body features. They evaluated their model on two datasets (FER2013 and CK+). The results indicate that they accomplished an overall accuracy of 90% across the different combinations utilised in their research. Meena et al. (2023) designed a sentiment analysis framework by analysing facial expressions using the CNN. In their work, they modified CNN layers to improve the performance of their framework. Their results show that they accomplished accuracy values of 79% and 95% for the FER2013 and CK+ datasets. F. M. Talaat (2023) designed a facial emotion recognition method for children by hyper-tuning the parameters of the CNN algorithm using GA. Besides that, in their work, they used the U-Net method in the pre-processing stage. The results indicate that they

accomplished an accuracy value of 99.99%. However, in their research, a very small dataset was used.

Previous studies have employed different architectures of the CNN algorithm for facial emotion recognition. Nevertheless, the CNN algorithm's performance is heavily reliant on its learning parameters. In order to accomplish this goal, F. M. Talaat (2023) used the GA algorithm. However, the GA algorithm has a low convergence rate in finding the best parameter value while exploring the solution space because inappropriate solutions are not removed during the search process. This issue is considered in this manuscript, and a recent BWO algorithm is employed to hyper-tune the parameters because this algorithm eliminates the inappropriate solutions while exploring the solution space of the given parameter (Hayyolalam & Kazem, 2019). In the literature, it has been successfully applied in several applications, such as segmentation (Houssein et al., 2020), feature selection (Hu et al., 2021), and distributed power systems (Sadeghi et al., 2021).

3. Preliminaries

This section presents the algorithms, namely, CNN and the metaheuristic BWO, that are employed to design the image-based depression classification method.

- **CNN Algorithm:**

In previous studies on image processing, authors usually use CNNs to analyse images because they are a type of deep, feed-forward artificial neural network (Cabada et al., 2019). Other image recognition methods require more complex image preparation than a CNN. This is because a CNN's convolutional component extracts an image's key properties. Three different kinds of layers make up the CNNs' multilayered architecture. The primary characteristics and patterns in an image are extracted by the first layer, known as the convolutional layer. The pooling layer reduces the number of features taken from the previous layer. The features and patterns gathered from the earlier layers are then categorised by a fully connected layer. A CNN's three layers are shown in Figure 1 (Abdullah & Abdulazeez, 2021).

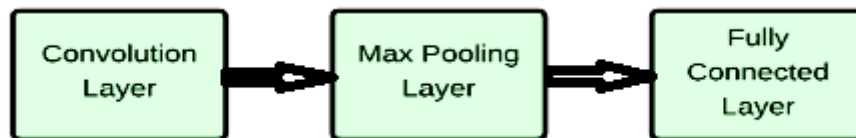


Figure 1. Three layers of CNN algorithm

The learning parameters, such as learning rate, batch size, and kernel size impact the performance of the CNN. Therefore, hyper-tuning these parameters enhances the classification accuracy using BWO.

- **BWO Algorithm:** This algorithm is designed by analysing the mating behaviour of spiders. These spiders produce a variable number of offspring, and to ensure their survival, the stronger offspring consume the weaker ones. Based on this process, three stages were designed, known as procreation, cannibalism, and mutation. The cannibalism stage is the most important, making it superior to the genetic algorithm and ensuring early convergence. At this stage, while exploring the optimal solution in the solution space, inappropriate solutions are removed (Hayyolalam & Kazem, 2019).

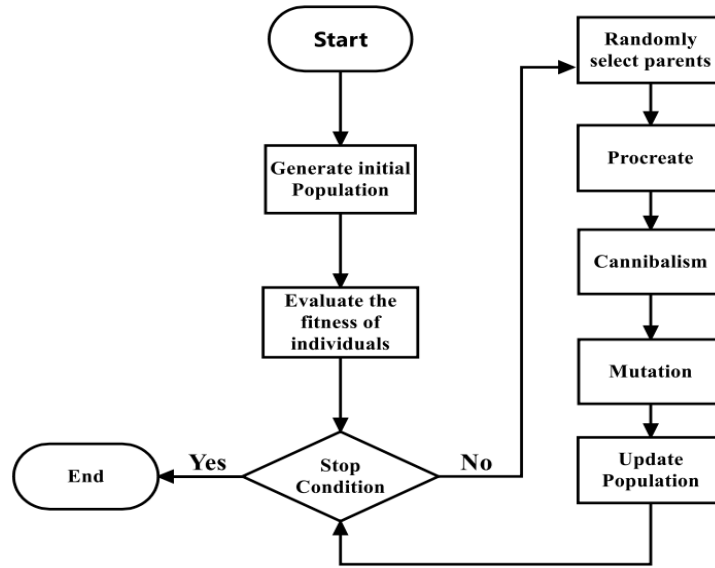


Figure 2. Flowchart of the BWO algorithm (Hayyolalam & Kazem, 2019)

- 1. Initialisation:** In this step, similar to other optimisation algorithms, an initial population is generated within the lower and upper bounds of the solution space to explore the optimal solution. The initialisation is done randomly.
- 2. Procreate:** In this step, the initial population is employed to generate new offspring using Eq. (1).

$$\begin{cases} y_1 = \alpha \times x_1 + (1 - \alpha) \times x_2 \\ y_2 = \alpha \times x_2 + (1 - \alpha) \times x_1 \end{cases} \quad (1)$$

In the above Eq. (1), y_1 and y_2 denote the generated offspring, whereas x_1 and x_2 are the initial population.

- 3. Cannibalism:** This step is utilised for removing the inferior-quality populations while exploring the solution space. To accomplish this goal, the generated offspring and initial populations are evaluated based on the objective function. Thereafter, based on the cannibalism rate, the inferior-quality populations are removed. This step aids in generating a new population that is superior to the previous one, leading to earlier convergence to the optimal solution in the next iteration.
- 4. Mutation:** The mutation step is performed to prevent the BWO algorithm from premature convergence. In this step, a random population is selected and updated by performing mutation based on the mutation rate.

4. Proposed Emotion Recognition Method

This section presents the proposed emotion recognition method by analysing facial expression using the deep learning CNN algorithm. The main novelty of the presented method is that it classifies the seven emotions. Furthermore, in this research, we enhanced the recognition accuracy by hyper-tuning the learning parameters of the CNN algorithm. Figure 3 shows the block diagram of the proposed emotion recognition method.

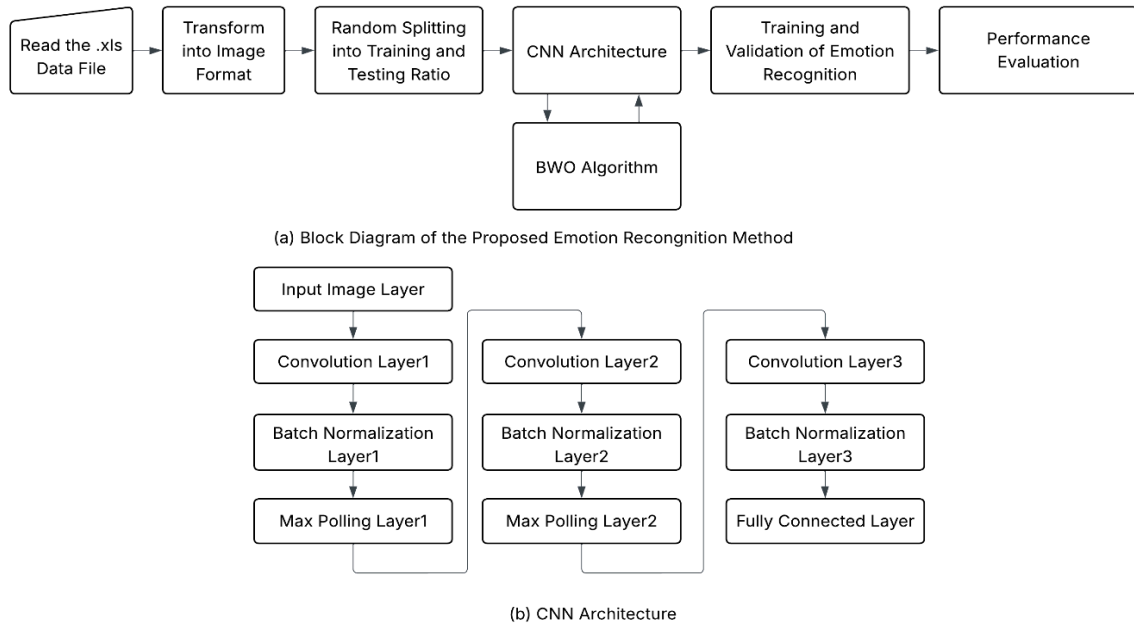


Figure 3. Block diagram of the proposed emotion recognition method

- **Data Collection and Preparation:** Initially, in this method, standard emotion datasets taken from previous studies. In previous studies, the FER2013, CK+, JAFFE, and AffectNet datasets were used by various authors. Out of these, the FER2013 dataset is the most preferred over other datasets because it contains a wide range of human emotions and the images are available in grayscale format. The dataset was available in an Excel sheet, and it contains the emotion labels and the pixel values. Therefore, the Excel sheet was read, the pixels were transformed into images, and the .png format was considered. The dataset images contain seven types of emotions (1=Angry, 2=Disgust, 3=Fear, 4=Happy, 5=Sad, 6=Surprise, 7=Neutral).
- **Random Splitting of the Data:** Next, the dataset images were randomly split into training and testing ratios. This ratio was defined to train and validate the CNN algorithm for emotion recognition. In this research, a 70:30 ratio was considered.
- **Emotion Recognition using CNN:** After that, the CNN architecture was defined for emotion recognition. The network structure of the CNN algorithm was built using multiple layers, including “input”, “convolution”, “batch normalisation”, “ReLU”, “max pooling”, and a fully connected layer. The input layer was 1-channel image data, and the resolution of the image was 32x32. After that, the image was passed through the first convolution layer with a filter size of (3,8), and batch normalisation was applied with the ReLU layer to improve training stability and speed up the convergence. Next, the data was passed through the first max pooling layer with a pool and stride size of 2X2 to perform dimensionality reduction. Similar to the first layer, the data was passed through the second convolution (3,16), batch normalisation, and a pooling layer (2x2). In the third layer, the size of the convolution filter was (3,32), and batch normalisation was applied using the ReLU layer. Finally, for classification purposes, seven fully connected layers were defined with the softmax function. After that, the learning parameters of the CNN algorithm, including learning rate and momentum rate, were hyper-tuned using the BWO algorithm. To accomplish this goal, the BWO algorithm was initialised, and a random population was generated for the parameters. The population values were considered in the CNN algorithm, and evaluation was done based on the objective function. In this work, accuracy was taken as the objective function. The BWO algorithm was iterated for a fixed number of iterations, and in each iteration, the population of inferior solutions was removed, and the optimal values were determined, and based on it, the CNN algorithm was trained and validated.

- **Performance Evaluation:** Lastly, the evaluation of the proposed method was done using the confusion matrix to determine correct and incorrect emotion recognition. Furthermore, based on this, other performance indices, including accuracy, precision, recall, and F1-score, were determined.

5. Results and Discussion

This section presents the simulation evaluation of the designed method and a comparative analysis with previous studies to validate its performance. MATLAB 2018a software was used to carry out the implementation and simulation of the proposed method. The system configuration used for MATLAB software is an i7 processor, 16 GB RAM, and a 64-bit Windows operating system. Furthermore, image processing and deep learning libraries and functions were considered for designing the ER method.

5.1 Evaluation Criteria and Indices

To evaluate the designed ER method, in the first step, the dataset images were prepared from an Excel sheet in which different emotion classes and their pixels were available (Sambare, 2013). In this step, the images were defined at a resolution of 32 x 32, and the .png format was used. In the second step, the dataset images were randomly split into training and testing ratios. In this article, we considered the 70:30 ratio. Seventy per cent of the dataset was considered for training the model, whereas testing was done on the remaining 30 per cent..

In the third step, the CNN algorithm was initialised with one input image layer of size 32 x 32, three convolution layers of size (3,8), (3,16), and (3,32), two max pooling layers, a batch normalisation layer, and a ReLU layer. Finally, in this step, the fully connected layer was defined, and the softmax layer was used to classify the different emotions.

In the fourth step, the BWO algorithm was initialised to find the learning parameter values (learning rate and momentum rate) of the CNN algorithm. Therefore, in the search for the best values, the BWO parameters were defined as a population size of (2,2) and an iteration value of 50. However, defining the BWO parameters is a challenging task because it increases the designed method's complexity in terms of time and memory space. Therefore, we chose a small population size of 2 and iterated for a fixed number of 50 iterations to search the solution space of the parameters. Furthermore, another parameter, namely the objective function, was defined. We chose accuracy as the objective function due to its simplicity and interpretability. Finally, in this section, the performance indices are defined, which are analysed in the article to evaluate the performance of the designed method against previous studies.

1. *Confusion Matrix:* $N \times N$ is the size of the square matrix that represents the confusion matrix, where N is the number of output categories. The matrix's rows reflect the number of occurrences of a predicted class, while the columns represent the number of instances of the true class. This shows how many correct and incorrect predictions a classifier produced for classification tasks, broken down by class (Sathyanarayanan, 2024; Shen et al., 2025)

		Actual Values	
		Positive	Negative
Predicted Values	Positive	True Positive (TP)	False Positive (Type I error)
	Negative	False Negative (Type II error)	True Negative

Figure 4. Confusion matrix

Predictions come in two different forms: correct and incorrect (errors).

2. *Accuracy*: Accuracy may be defined as the percentage of accurate forecasts among all predictions or the frequency of accurate predictions.

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FP} + \text{TN} + \text{FN}) \quad (1)$$

3. *Precision*: Precision is the percentage of accurately predicted favourable outcomes.

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (2)$$

This measure is helpful when there are more FPs than FNs. A high accuracy number indicates that the model produces reliable positive predictions and has a low false positive rate. On the other hand, a low precision score means that there are many false positives, meaning that the model predicts more inaccurate positives.

4. *Recall*: This assesses the accuracy of the model's positive case prediction, that is, the percentage of real positives that were properly predicted. It is computed as

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (3)$$

Recall is especially useful for minimising Type II errors, or false negatives. A high recall number means that the model does not make many false negatives, which means that it correctly identifies most favourable cases. In contrast, a low recall value means that there are more false rejections, which means that the model is missing positive cases.

5. *F-Score*: Recall and precision are combined to get the F1-score, which is significant for models when both are equally significant. Both false positives and false negatives are taken into account in this harmonic mean of precision and recall. To calculate the harmonic mean, divide the number of values in a data series by the sum of their reciprocals. The arithmetic and geometric means are always greater than or equal to the harmonic mean. This is due to the fact that the harmonic mean gives greater weight to the smaller values in the data series. The harmonic mean is maximised when recall and precision are equal. On the other hand, the harmonic mean becomes closer to the lower number when the two metrics are dissimilar. When accuracy equals recall, the F1-score reaches its peak. It balances the mistakes of Type I and Type II. When the F1-score is used, the model is not easy to understand because it is not clear whether the predictor tries to increase precision or recall. The F1-score works well for classification models with a high number of TN and identical effects associated with FP and FN. The F1-score is calculated using the following equation:

$$\text{F1-score} = (2 * \text{precision} * \text{recall}) / (\text{precision} + \text{recall}) \quad (4)$$

6. *Convergence Rate Graph*: In this article, this graph is drawn to evaluate the metaheuristic BWO algorithm. This graph is drawn between the fitness function and the iteration number, and it shows how many iterations are required to achieve the desired objective function.

5.2. Simulation Results

Figure 5 shows the confusion matrix evaluated for the proposed emotion recognition method under CNN parameter optimisation. In this paper, we recognised the different emotions. Therefore, a multi-class confusion matrix was constructed for the proposed method. The diagonal components of the matrix represent the correctly recognised emotions done by the proposed method. Based on this matrix, accuracy, precision, recall, and F1-score are determined in Table 1.

Confusion Matrix							
Output Class	0	1	2	3	4	5	6
	656 13.1%	0 0.0%	3 0.1%	4 0.1%	3 0.1%	5 0.1%	4 0.1%
	4 0.1%	78 1.6%	2 0.0%	4 0.1%	3 0.1%	1 0.0%	4 0.1%
	1 0.0%	0 0.0%	701 13.9%	3 0.1%	5 0.1%	3 0.1%	6 0.1%
	3 0.1%	0 0.0%	1 0.0%	1221 24.3%	0 0.0%	1 0.0%	4 0.1%
	4 0.1%	0 0.0%	3 0.1%	3 0.1%	858 17.1%	1 0.0%	3 0.1%
	2 0.0%	0 0.0%	2 0.0%	4 0.1%	1 0.0%	570 11.3%	5 0.1%
	1 0.0%	0 0.0%	5 0.1%	3 0.1%	3 0.1%	1 0.0%	837 16.7%
Target Class							
	0	1	2	3	4	5	6
	97.8% 2.2%	100% 0.0%	97.8% 2.2%	98.3% 1.7%	98.3% 1.7%	97.9% 2.1%	97.0% 3.0%

Figure 5. Confusion Matrix for the Proposed ER Method

Table 1 Performance Indices for the Proposed ER Method

Emotion Class	Accuracy	Recall	Precision	F1-Score
1	0.99284	0.97765	0.96898	0.97329
2	0.99642	1	0.8125	0.89655
3	0.99324	0.97768	0.97497	0.97632
4	0.99403	0.98309	0.99268	0.98786
5	0.99423	0.98282	0.98394	0.98338
6	0.99483	0.97938	0.97603	0.9777
7	0.99184	0.96756	0.98467	0.97604
Average	0.99392	0.98117	0.95625	0.96731

The result indicates that the accuracy value varies from 0.99184 to 0.99642. This reflects that the proposed method effectively recognises the different emotions for the taken dataset, as shown in Figure 6.

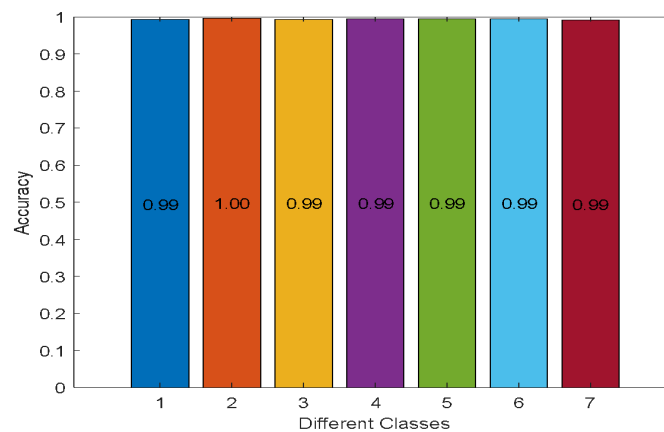


Figure 6. Accuracy Analysis for Different Classes of Emotions

Furthermore, the recall value is from 0.96756 to 1 and the precision value is from 0.81250 to 0.99268 for different classes of emotion. The higher value of recall and precision reflects that the correct estimation value is higher than the negative estimation of the recognition, as shown in Figure 7-8.

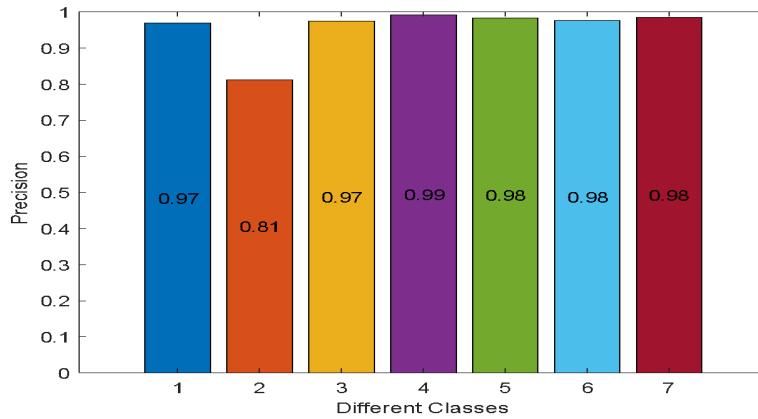


Figure 7 Precision Analysis for Different Classes of Emotions

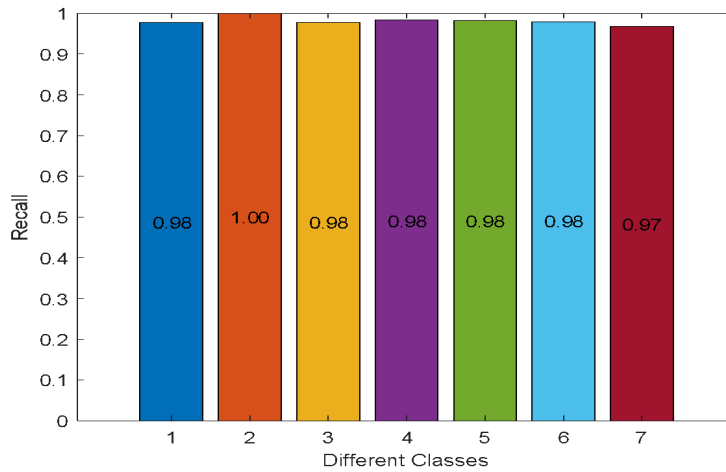


Figure 8. Recall Analysis for Different Classes of Emotions

Finally, the F1-score value varies from 0.89655 to 0.98786 for the proposed ER method. This reflects that there is a balance between recall and precision values. Furthermore, higher F1-score values indicate that the proposed method effectively minimises both false-positive and false-negative emotion classifications.

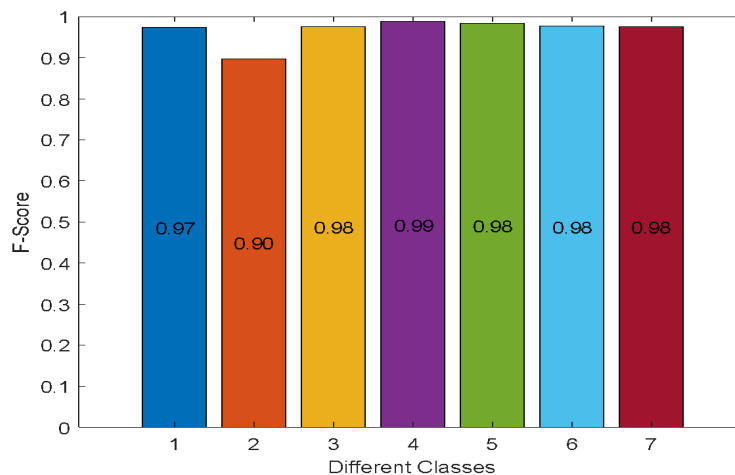


Figure 9. F-Score Analysis for Different Classes of Emotions

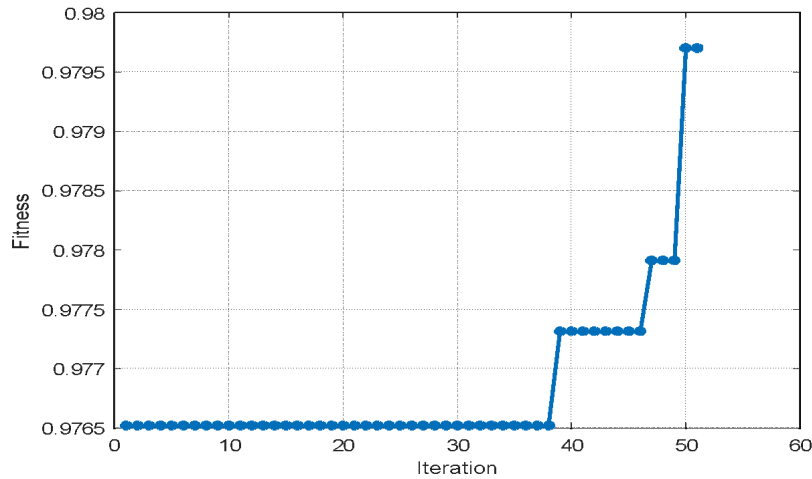


Figure 10. Evaluation of the BWO Algorithm based on the Convergence Rate (Fitness Function Vs. Iteration)

Figure 10 presents the evaluation of the BWO algorithm based on the convergence-rate graph. This graph is plotted between the fitness function and the iteration number. In this work, the primary goal is to maximise accuracy, which serves as the fitness function. The results indicate that the BWO algorithm achieves its maximum accuracy at the 50th iteration, owing to the removal of inappropriate solutions while exploring the solution space of the CNN learning parameters.

5.3. Comparative Analysis

This section presents a comparative analysis with previous studies using the same dataset and configuration. In the proposed emotion recognition method, the key emphasis is placed on hyper-tuning the parameters of the CNN algorithm using the metaheuristic BWO algorithm. Table 2 shows the comparative analysis with the previous studies based on the accuracy parameter. The table shows that the NB algorithm achieves the lowest accuracy value of 72.50% and the highest value of 99.39% by the proposed method. On the other hand, the other algorithms achieve 80.00% with the RNN algorithm, 82.40% by RNN-LSTM, and 82.80% by the SVM algorithm.

Table 2. Comparative Analysis with the Previous Studies

Methods	NB	RNN	RNN-LSTM	SVM	Proposed Method
Accuracy	72.50%	80.00%	82.40%	82.80%	99.39%

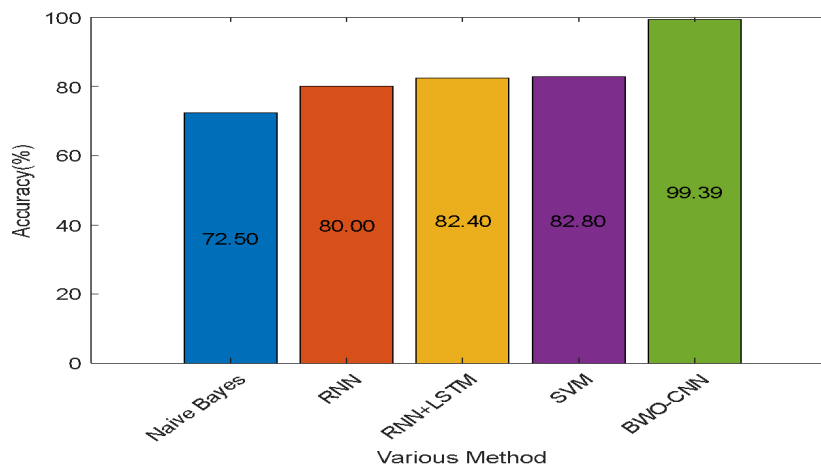


Figure 11. Comparative Analysis with Previous Studies based on Accuracy Parameter

Figure 11 indicates that the presented method (BWO-CNN) achieves the highest accuracy of 99.39% compared with the previous approaches, owing to the hyper-tuning of the learning parameters of the CNN algorithm for different emotion recognition.

6. Conclusion and Future Scope

In this article, we designed an ER method to classify multiple emotions using the CNN algorithm. Furthermore, the primary focus lies in hyper-tuning the learning parameters of the CNN algorithm. To accomplish this objective, the metaheuristic BWO algorithm is employed, and accuracy is taken as the objective function. The benefit of the BWO algorithm is that due to the cannibalism step, the inappropriate solutions are eliminated while searching for the optimal parameter values, and its performance is validated using the convergence-rate graph. Next, the simulation results indicate that the presented method effectively classifies the emotions and achieves an average accuracy of 0.99392, an average recall of 0.98117, an average precision of 0.95625, and an F1-score of 0.96731. The limitation of the designed ER method is that it is evaluated for a specified dataset (FER2013), and the same dataset is used for training and testing purposes. Therefore, in the future, to evaluate the robustness and generalisation, the proposed ER method will be evaluated by combining the different datasets. The other limitation of the designed method is that the dataset is unbalanced. Therefore, in the future, we will balance it using SMOTE and data augmentation techniques. In addition, we will optimise the layers of the CNN algorithm to design a lightweight CNN algorithm for the presented method.

References

- Abdullah, S. M. S., & Abdulazeez, A. M. (2021). Facial expression recognition based on deep learning convolution neural network: A review. *Journal of Soft Computing and Data Mining*, 2(1). <https://doi.org/10.30880/jscdm.2021.02.01.006>
- Akhand, M. A. H., Roy, S., Siddique, N., Kamal, M. A. S., & Shimamura, T. (2021). Facial emotion recognition using transfer learning in the deep CNN. *Electronics*, 10(9), 1036. <https://doi.org/10.3390/electronics10091036>
- Amou, M. A., Xia, K., Kamhi, S., & Mouhafid, M. (2022). A novel MRI diagnosis method for brain tumor classification based on CNN and Bayesian optimization. *Healthcare*, 10(3), 494. <https://doi.org/10.3390/healthcare10030494>
- Agung, E. S., Rifai, A. P., & Wijayanto, T. (2024). Image-based facial emotion recognition using convolutional neural network on emognition dataset. *Scientific Reports*, 14(1), Article 65276. <https://doi.org/10.1038/s41598-024-65276-x>
- Bhagat, D., Vakil, A., Gupta, R. K., & Kumar, A. (2024). Facial emotion recognition (FER) using convolutional neural network (CNN). *Procedia Computer Science*, 235, 2079–2089. <https://doi.org/10.1016/j.procs.2024.04.197>
- Cabada, R. Z., Rangel, H. R., Estrada, M. L. B., & Lopez, H. M. C. (2019). Hyperparameter optimization in CNN for learning-centered emotion recognition for intelligent tutoring systems. *Soft Computing*, 24(10), 7593–7602. <https://doi.org/10.1007/s00500-019-04387-4>
- Haq, H. B. U., Akram, W., Irshad, M. N., Kosar, A., & Abid, M. (2024). Enhanced real-time facial expression recognition using deep learning. *Acadlore Transactions on AI and Machine Learning*, 3(1), 24–35. <https://doi.org/10.56578/ataiml030103>
- Hayyolalam, V., & Kazem, A. A. P. (2019). Black widow optimization algorithm: A novel meta-heuristic approach for solving engineering optimization problems. *Engineering Applications of Artificial Intelligence*, 87, Article 103249. <https://doi.org/10.1016/j.engappai.2019.103249>
- Houssein, E. H., Helmy, B. E.-D., Oliva, D., Elngar, A. A., & Shaban, H. (2020). A novel black widow optimization algorithm for multilevel thresholding image segmentation. *Expert Systems With Applications*, 167, Article 114159. <https://doi.org/10.1016/j.eswa.2020.114159>

- Hu, G., Du, B., Wang, X., & Wei, G. (2021). An enhanced black widow optimization algorithm for feature selection. *Knowledge-Based Systems*, 235, Article 107638. <https://doi.org/10.1016/j.knosys.2021.107638>
- Karatay, B., Beştepe, D., Sailunaz, K., Özyer, T., & Alhajj, R. (2023). CNN-transformer based emotion classification from facial expressions and body gestures. *Multimedia Tools and Applications*, 83(8), 23129–23171. <https://doi.org/10.1007/s11042-023-16342-5>
- Khan, A. R. (2022). Facial emotion recognition using conventional machine learning and deep learning methods: Current achievements, analysis and remaining challenges. *Information*, 13(6), 268. <https://doi.org/10.3390/info13060268>
- Khattak, A., Asghar, M. Z., Ali, M., & Batool, U. (2021). An efficient deep learning technique for facial emotion recognition. *Multimedia Tools and Applications*, 81(2), 1649–1683. <https://doi.org/10.1007/s11042-021-11298-w>
- Lateef, R., & Abbas, A. (2022). Tuning the hyperparameters of the 1D CNN model to improve the performance of human activity recognition. *Engineering and Technology Journal*, 40(4), 547–554. <https://doi.org/10.30684/etj.v40i4.2054>
- Meena, G., Mohbey, K. K., Indian, A., Khan, M. Z., & Kumar, S. (2023). Identifying emotions from facial expressions using a deep convolutional neural network-based approach. *Multimedia Tools and Applications*, 83(6), 15711–15732. <https://doi.org/10.1007/s11042-023-16174-3>
- Rimal, Y., Sharma, N., & Alsadoon, A. (2024). The accuracy of machine learning models relies on hyperparameter tuning: Student result classification using random forest, randomized search, grid search, Bayesian, genetic, and Optuna algorithms. *Multimedia Tools and Applications*, 83(30), 74349–74364. <https://doi.org/10.1007/s11042-024-18426-2>
- Sadeghi, B., Shafaghatian, N., Alayi, R., Assad, M. E. H., Zishan, F., & Hosseinzadeh, H. (2021). Optimization of synchronized frequency and voltage control for a distributed generation system using the black widow optimization algorithm. *Clean Energy*, 6(1), 105–118. <https://doi.org/10.1093/ce/zkab062>
- Sambare, M. (2013). *Emotion recognition dataset (FER2013)* [Data set]. Kaggle. <https://www.kaggle.com/datasets/msambare/fer2013>
- Sathyanarayanan, S. (2024). Confusion matrix-based performance evaluation metrics. *African Journal of Biomedical Research*, 27(4S), 4023–4031. <https://doi.org/10.53555/ajbr.v27i4s.4345>
- Shen, L., Shen, L., Zhang, H., Zhu, C., Li, R., Qian, K., Tian, F., Hu, B., Schuller, B. W., & Yamamoto, Y. (2025). Enhancing emotion regulation in mental disorder treatment: An AIGC-based closed-loop music intervention system. *IEEE Transactions on Affective Computing*. Advance online publication. <https://doi.org/10.1109/TAFFC.2025.3557873>
- Sowmya, B., Alex, S. A., Kanavalli, A., Supreeth, S., Shruthi, G., & Rohith, S. (2024). Machine learning model for emotion detection and recognition using an enhanced convolutional neural network. *Journal of Integrated Science and Technology*, 12(4), 786–786. <https://doi.org/10.62110/sciencein.jist.2024.v12.786>
- Talaat, F. M. (2023). Real-time facial emotion recognition system among children with autism based on deep learning and IoT. *Neural Computing and Applications*, 35(17), 12717–12728. <https://doi.org/10.1007/s00521-023-08372-9>
- Wojciuk, M., Swiderska-Chadaj, Z., Siwek, K., & Gertych, A. (2024). Improving classification accuracy of fine-tuned CNN models: Impact of hyperparameter optimization. *Heliyon*, 10(5), e26586. <https://doi.org/10.1016/j.heliyon.2024.e26586>

Appendix

Evaluation of the proposed method without optimisation of the parameters of the CNN algorithm

Table A1 shows the accuracy value of the proposed method without optimised CNN parameters. The result indicates that the proposed method accomplishes the lowest accuracy value of 0.33148 for emotion type 4 and the highest accuracy value of 0.98448 for emotion type 2.

Table A1. Accuracy of the Proposed ER Method

N-Classes	Accuracy
1	0.8661
2	0.98448
3	0.85734
4	0.33148
5	0.74075
6	0.8842
7	0.82829