FACIAL EMOTION RECOGNITION USINGHYBRID PSO AND GA OPTIMIZED CONVOLUTIONAL NEURAL NETWORK

[1,*]KALAIARASI.P, [2]ESTHER RANI.P

[1,*]Research Scholar, Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai.

[2]Professor, Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Avadi, Chennai.

ABSTRACT

An emerging and interesting research field which focuses on the enhancement of developing an automatic emotion recognition system is Facial Expression Recognition. This technology creates a major impact on various applications such as safety, healthcare, gaming and education. Researchers in this field are working to develop methods that can extract and interpret facial expressions. Deep learning plays a vital role in the field of computer vision technology and Convolutional Neural Network (CNN) is the most significant part in the deep learning. There are few drawbacks in the training algorithm, Stochastic Gradient Descent (SGD) used in CNN. This proposed work overcomes the drawbacks of SGD by introducing an alternative algorithm based on Genetic Algorithm and Particle Swarm Optimization to improve the performance of CNN. This paper presents the novel PSO-GA algorithm and addresses the benefits of PSO-GA algorithm which challenges and outperforms other existing methods and achieves robust and state-of-the art results in various challenging datasets. This hybrid PSO-GA algorithm makes the CNN to identify the facial expressions and emotions very effectively when compared to other methodology.

Keywords: *Facial, Emotion Recognition, PSO, GA, NN*

1. INTRODUCTION

Through facial expressions, people can easily communicate their feelings and thoughts. This is one of the maximum authoritative ways for humans to express themselves. Being able to recognize other people's emotions can help people communicate better [1]. For instance, if a person is in a hospital, they might not be able to show their emotions. Artificial Intelligence (AI) has gained widespread attention due to its ability to recognize human emotions [2]. It has been used in various applications such as smart homes and hospitals [3]. AI-powered personal assistants like Siri, Cortana and other virtual assistants can communicate with humans using Natural Language Processing (NLP). However, when paired with emotional intelligence, they can improve their communication capabilities [4].

Human computer interaction (HCI) is a vital component of modern computing and facial expression recognition is an integral part of this field. It is used in various applications such as healthcare and surveillance systems [5]. Unfortunately, it is not always easy to recognize real-time facial expressions owing to a variety of circumstances such as changes in lighting and size variances [6]. Another challenge is that different people's way of expressing their emotions can lead to false predictions [7]. Another challenge is identifying facial features that can help the system distinguish between different emotions. An accurate and robust facial emotion analysis system is also necessary to perform its intended function [8].

The main challenge in developing robust and optimized discriminative facial emotions representations is ensuring that they are capable of supporting real-time facial emotion detection.

Conferring to [9], EA algorithms are commonly used for developing discriminant representations. Some of these include Genetic Algorithm, Particle Swarm Optimization, Artificial Bee Colony Optimization, Ant Colony Optimization and Firefly Optimization [10], [29], [30]. Due to the nature of these algorithms, they tend to get cluttered with their own flaws, which can result in local stagnation [11].

Face mapping is the main task of FER. It involves identifying all of the facial expressions that are related to the individual's emotional states. Aside from face mapping, other image preprocessing steps such as face detection, cropping, and resizing are also performed [12]. Feature extraction is a step in a classical FER system that involves extracting various features from an image. This step is carried out by extracting the emotions in the image. [13].

Deep Neural Networks, which are neural networks that can extract features from images, have gained increased attention due to their ability to do so. A few works have been presented using the CNN to solve FER problems [14]. However, these methods only consider the network's layers. One of the main reasons why this technique is not yet widely used is due to the complexity of the task. First, it needs a high-quality image to recognize emotion. Also, the difference between the faces in an image and those in different emotional states can be very low [15]. The number of hidden layers in a deep CNN is enormous. Even at a given amount of layering, the algorithm's accuracy does not improve [16]. Also, it makes the training more challenging. Various techniques and modifications are then introduced to the deep CNN to improve its accuracy. Some of the models that are commonly used for training are VGG-16, Inception-v3, and Resnet-152 [17].

Due to the varying training database used for different media, the performance of emotion recognition tools can be severely degraded. The classification stage is mainly focused on the features extracted from the media. Unfortunately, the large number of these features can affect the classification. Due to the complexity of the media representations of emotions, it is very important to consider the features that were extracted from the media before performing emotion recognition tests [18]. This paper presents a novel approach that uses bio-inspired heuristics.

In [19] authors goals to deliver a discriminative approach for the growth of facial representations that can be used for real-time facial recognition. In this paper, it is shown evolutionary computation algorithms can perform global search and feature selection efficiently. The PSO algorithm is commonly used to perform feature selection. It is based on the behavior of birds and can be used for fast convergence and low computational cost. However, this method can also cause it to get trapped in local optima [20].

In this research a new PSO variant with a non-replicable memory and a small-population secondary swarm was proposed. It utilizes a sub-dimensional facial feature search strategy and a velocity updating strategy to overcome the limitations of the existing algorithm. This work develops a hybrid CNN models that are based on the Genetic Algorithm and PSO. These models are designed to improve the classification of these models. This method eliminates the need for manual labor in learning discriminative features. Even though there are many hyper parameters in a CNN network, it can be hard to identify the optimal ones manually.

A hybrid algorithm is used to learn the best hyper parameter of CNN along with PSO and GA. It performs the optimization of the layer's weights. The algorithm combines the best known optimal value with the derived weights. The purpose of this algorithm is to minimize the computational cost and the length of the chromosome. The rest of the paper is divided into sections, which include the related works in section 2 and the proposed method for classifying different facial emotions recognition in section 3, results are discussed in section 4 and section 5 concludes the work.

2. RELATED WORKS

Sikkandar et al(2021)[21] –Identified the different kinds of expressions used by humans. This study was carried out using a method that combines various features and methods in order to improve the accuracy and time of a proposed system. The proposed method was able to recognize different kinds of expressions and recall f-measures in various datasets. Although it only performed well in real-world images, the system was able to detect expressions in each frame.

Alenazy et al (2021) [22] - In this paper, a novel model was presented that used the DBN and GSA algorithm. It is discussed that it can be derived efficiently using the appropriate feature extraction techniques. A video recording can be derived by combining 2D and 3D features extraction techniques. This model derived the facial expressions of the video and then identified the facial expressions of the video by combining the 2D-DWT and 3D features extraction techniques. This algorithm achieves a higher identification rate than other methods. This method was capable of saving the computational time and increasing the retention capacity of the data. It was tested in two widely used datasets. The model achieves 99.99% classification accuracy for the datasets, such as the JAFFE. It achieves better accuracy but it is lagging on the interpretation of action and zapping facial expressions and low resolution images

Zatarain et al(2020) [23]- This study was to improve the accuracy of a CNN by used a general algorithm (GA) for hyper parameter tuning. The results suggested that this method can improved the accuracy of the system when compared with other machine learning classifiers. Although CNN was more accurate than other deep learning algorithms, it is still not as robust as other classifiers. Because a CNN system can perform facial gestures with high accuracy using a neural network known as a convolutional network. However, one of the biggest issues with this technology is the number of hyper parameters. To test its accuracy, it was implemented with a mobile device for Android. It also recognized faces used an emotion recognition algorithm. The CNN was implemented using the Python programming language and the Java multi-sensei algorithm. Its recognition rate improved significantly when improved the classes focus and relaxed state.

Gogić et al (2020) [24] - The method was presented using a feature extraction process and a shallow neural network. The two-layer network was able to model the relationship between expressions. The method was able to improve the recognition rates of difficult expressions such as sadness and fear. It can also be used as an alternative to end-to-end CNNs. Due to its generalization capabilities, the method was utilized in situations where small data sets are needed. The method was able to boost the expressive power of facial expressions by limiting the feature space around prominent landmarks. It was also able to perform joint classification, which allowed it to improve the recognition rate of ambiguous expressions. It could also be extended to include temporal information by incorporating the use of recurrent neural networks. This method would be natural since expressions change over time.

Cheng et al (2020) [25] - The paper shown how these features can be used to identify facial expressions. Facial expressions were extracted from the facial landmarks that are associated with the movement of facial muscles. This paper presented a method that combines the features and parameter selection algorithm of a support vector machine with a genetic algorithm. The evaluation was performed on the extension of the Cohn-Kanade database and the MUG. The results indicated that the presented algorithm performed well in terms of validation accuracy. This algorithm performed well in terms of recall, precision, and F1-score. It was compared to CNN, which is widely used method for facial emotion recognition. The presented algorithm performed slightly better than CNN in terms of test accuracy for 8-class CK+ and 7-class CK+. However, it was slightly behind CNN for the MUG dataset. The presented algorithm uses less complex models and was capable of performing real-time machine vision tasks.

3. PROPOSED METHODOLOGY

In social situations, people will naturally express their feelings. Understanding their emotions helps build trust and mutual understanding. They can also communicate their feelings through various means such as facial expressions, voice, and language. Human emotion expression is a vital skill that people must master. Researchers study the various aspects of human emotion expression in order to gain a deeper understanding of the human inner emotional activities. There are six basic categories of facial expressions: anger, disgust, fear, happiness, and surprise. Due to the advancements in technology, such as the development of interactive games and robots, facial expression recognition has become a popular field of study. Different facial components are used to classify different emotions. For instance, a frown, upside-down smile, and wrinkled forehead are considered as expressions of anger. More advanced emotion detectors can now identify complex feelings by mixing the basic classifications.

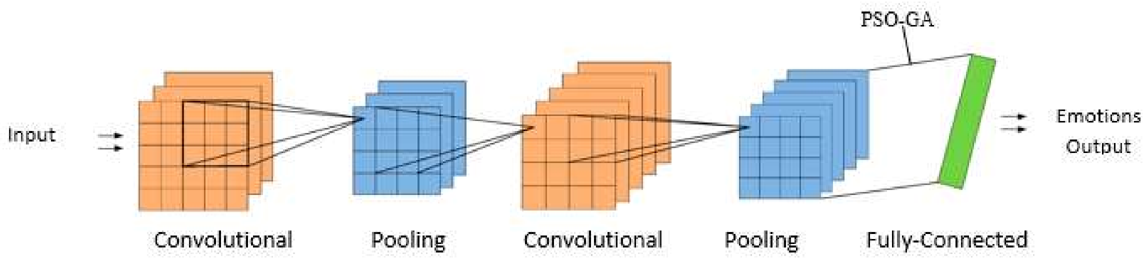


Figure 1: Block Diagram For Emotion Recognition Using Hybrid CNN-PSO-GA

In the above figure 1 developed a hybrid CNN model with a GA and PSO to improve their classification accuracy. It has been stated that CNN networks might be trained to end in a supervised manner while learning discriminative features. Even though these networks have a large variety of hyper parameters, it is typically challenging to identify the optimal ones manually. In this work, we introduce a deep learning approach for the convolutional neural network. We first presented the PSO-GA-CNN model. It was developed using the PSO algorithm to evolve the network's parameters. The second module of hybrid model uses a GA to learn the optimal hyper parameter for CNN. The results of our study show that the second order optimization method can improve convergence to the optimal value when applied to the weights that are obtained as outputs of GA. For the calculation of the optimal weights for the full convolution neural network, we used a Genetic Algorithm.

3. 1 Designing and Optimizing the CNN

A CNN is an artificial neural network that is commonly used to perform image analysis. It has the same characteristics as a feed-forward algorithm. Unlike other systems, a CNN has a layered structure that extracts the most important features of an image. The first layer, which is called the convolutional layer, is used to extract the main features of an image. The second layer, which is called the pooling layer, is used to decrease the number of features that have to be extracted. These features are then reduced in the pooling layer and passed on to a fully connected layer and finally to the classification layer.

There are is of ambiguous steps, which need to be clarified such as how can the CNN parameters be encapsulated into particles? How do they cooperate with each other? How can it justify the best particles with the ensembles of swarm? To answer these questions, how the parameters of CNN are distributed. It is obvious that the weights and biases are constituent parameters of CNN. Therefore, in this work, the weights and bias are dismantled and encapsulated into vectors as shown below:

$$W = w^1, w^2 \ldots w^p \tag{1}$$

Journal of Theoretical and Applied Information Technology
30th November 2022. Vol.100. No 22

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

$$p = b^1, b^2 \ldots b^p \quad (2)$$

$$W^n = w_1^{\ 1}, w_2^{\ 2} \ldots . w_L^{\ n} \quad (3)$$

$$b^n = b_1^{\ 1}, b_2^{\ 2} \ldots b_L^{\ n} \quad l = 1 \ldots L, n = 1 \ldots P \quad (4)$$

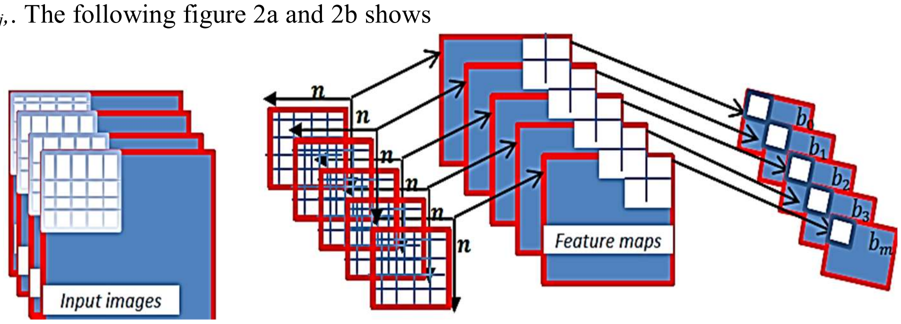
The total number of particles in a given layer is known as the layer index. The weight parameters of the layer are also known as the bias parameters of the layer.

$$W^n = w, P \quad (5)$$

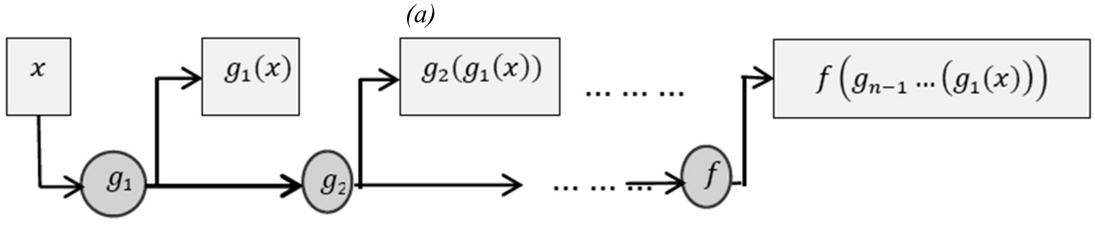
The first convolutional and max-pooling layers of CNN have set of filter dimensions n*n and 1+n2. The total weight parameters of the m filters for the lth layer are $w_1^{\ n} = mn^{\ 2}$ in eqn 3. In addition, the total bias parameters for the given layer are $b_1^{\ n} = 1 \times m$ in eqn 4. The goal of developing a PSO-GA approach was to combine the advantages of the particle swarm optimization and the genetic algorithm. This will allow the exploration and exploitation capabilities of the system to be improved. However, the two models have weaknesses and strengths. In general, the goal of the PSO-GA approach is to combine the capabilities of the genetic algorithm and the social thinking framework to improve the performance of the system. In order to achieve this, the two models should be combined with the local search capabilities of the GA. As both are population-based, the hybrid PSO-GA approach should be able to find a global solution. In order to evaluate which particle is the best among the swarm, the following concept is used. It states that each particle can give an optimal solution $W^* = \arg\min E(\theta)$.The best particle among the swarm is w*. The measured error between the model outputs.

$$E = \frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{m} (\mathrm{Re} f_{ji} - d_{ji})^2 \quad (6)$$

The number of training samples N is the output layer's number m. The output of CNN is the number of training samples which is $\mathrm{Re} f_{ji}$ is the reference d_{ij},. The following figure 2a and 2b shows the difference between the parameters used by Auto CNN - PSO and GA. The last layer $g_n\, where n = 1 \ldots L and L$ has a function f.



(a)



(b)

Fig. 2. The principle of working both algorithms (a)CNN (B) PSO

In addition, BP has two phases: the forward and backward. The activation function must respond to the input and the network parameters. PSO does not require a backward phase. This eliminates the need for the activation function to go through the backward phase.

The reason why the second phase of the network is not required is that the PSO algorithm depends on the velocities and positions of the particles in the network. If the particles are P and, then the best particle m for the network satisfies the criterion of (6). The velocity of a particle is the next position in the list of CNN parameters of part $W^n(t)$ based on eqn 5.

In this paper, the concept of particle swarm optimization is presented. It is combined with the genetic algorithm approach to improve its performance. The classical PSO technique is used to describe the design variable of a given particle. The position of that particle can be updated using a formula.

$$X_{ij}^{i+1} = X_{ij}^{t} + v_{ij}^{t+1} \qquad (7)$$

Position vectors of i^{th} particles are represented by their numbers and j is the search space dimension.

$$V_{ij}^{t+1} = wV_{ij}^{t} + C_1 r_1^t (P^t{}_{best,ij} - X_{ij}^t) + C_2 r_2^t (P^t{}_{best,ij} - X_{ij}^t) \quad (8)$$

Where V_{ij}^{t+1}, r_1^t , and r_2^t are ith particle velocity vectors that are found in C1, C2, and t+ 1 iteration. They are the best positions found by the ith particle. r_1^t and r_2^t are random numbers that are between 0 and 1, and $P^t{}_{best,ij}$ is the best global position. The term inertia weight is a constant term that is introduced in the original PSO formulation. It is determined by three parameters C1, C2, and ω. The term inertia weight influences the exploration and exploitation of the PSO technique. It is shown that this term has a great impact on the performance of the technique.

The proposed solutions for the problem of finding the optimal inertia weight parameter for PSO and GA technique are presented.

$$\omega_{ij}^{t+1} = \omega_{\max,j}^{t} + (\omega_{\max,j}^{t} - w_{\min,j}^{t}) \times e^{\mu_{ij}^t + \sigma_{ij}^t} \quad (9)$$

$$\mu_{ij}^{t} = \frac{V_{ij}^{t} + P_{best,ij}^{t} + P_{gbest,j}^{t}}{3} \qquad (10)$$

$$\sigma_{ij}^{t} = \sqrt{\frac{1}{3}\left[\left(V_{ij}^{t} - \mu_{ij}^{t}\right)^2 + \left(P_{best,ij}^{t} - \mu_{ij}^{t}\right)^2 + \left(P_{gbest,j}^{t} - \mu_{ij}^{t}\right)^2\right]} \quad (11)$$

The standard deviation of the particles' velocity and the average size of the $P^t{}_{best,ij}$, $P_{gbest,j}^t$ and V_{ij}^t are the factors that determine their origin. The velocity of the particles is the minimum and maximum value of the inertia weight vector.

In this proposed formulation, different values of the inertia term are assumed for each particle, which are updated during the optimization process. The proposed formulation is based on the standard deviation and mean of the various $P^t{}_{best,ij}$ $P_g{}^t{}_{best,ij}$, and V_{ij}^t parameters.

The proposed PSO method has been enhanced through the use of a Genetic Algorithm which is commonly used to perform cross-over and mutational operations. This method is carried out in each iteration to ensure that the convergence of the results will be achieved. The use of mutational and crossover operations helps in increasing the diversity of the population and avoids trapping in the local extrema. This study aims to study the hybrid approach of these two methods. After creating a new generation of PSO characters, some facial datas are selected and then generalized by GA for each face. Since the data size of these faces is very large, this method is not used to apply the whole data. The question is how many emotions should be added to the new generation?

The goal of PSO is to provide the maximum number of generations based on eqn 9. After selecting the best individual from the population in eqn 10, the algorithm will then create a new population by implementing various interventions such as crossover and mutation operations based on eqn 11.

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

The proposed CNN along with PSO and GA algorithm is very effective and efficient for finding a near-global solution to a specific problem.The proposed technique involves analyzing the number of instances of emotion in an image and the number of labels that are used to identify it. The six labels used to classify the emotion are then used to estimate the probability of the facial expression label being assigned to the image.

4. RESULT AND DISCUSSION

The goal of this study is to develop a system that can perceive and respond to human emotions. This will allow computers to interact with humans more naturally. In the past few years, the field of image classification has advanced. This technology allows computers to classify images. Through a set of facial expressions, the researchers were able to train a machine learning system to recognize different emotions such as disgust, fear, and happiness. The system then trained itself on the facial expressions. It then performed a hybrid model to classify the facial expression. The system can now recognize facial expressions performed on a video file or a standard dataset using OpenCV.

4.1. Proposed simulation environments

In this paper, an alternative algorithm for CNN training is presented. It is called the Particle Swarm Optimization algorithm. The proposed algorithm is fully parallel and has been formulated using simple equations. It avoids generating local optimum and can be easily adaptable when used for training CNN. The PSO equation used for training weights is completely parallelized. This allows the weights of various layers to be updated without going through backward phase. GPUs can be used for training CNN efficiently. The proposed algorithm can also improve training by overcoming sluggishness and premature saturation.

4.2 Dataset

FER2013 dataset consists of 48x48 pixel images of faces. They have been registered so that the face's center is more or less centered. The faces are categorized into seven categories based on the emotion they display in their facial expressions plotted in figure 3. The training set contains 28,708 images and the test set contains 7,178 images. The compressed version of this dataset takes about 92MB space while the uncompressed version takes 140MB space. Aaron Courville and Pierre-Luc Carrier developed the facial expression recognition algorithm for the 2013 Kaggle competition. JAFFE contains 213 peak expressions collected from ten women, and the expressions are labelled as anger, disgust, fear, happiness, neutral, sadness, and surprise. Each expression is shown in approximately 30 images. Lucy and colleagues published an extended version of the CK+ dataset that contains 593 expressions that can be used to capture seven basic facial expressions. These sequences were collected from 125 subjects. Each time the sequence is captured, the subjects alter their expressions. The expressions are then analyzed using the three peak frames from the sequence. The researchers then used the collected data to create a total of 21,895 images. The 423 expressions from the 163 participants were used for testing purposes.

Journal of Theoretical and Applied Information Technology
30th November 2022. Vol.100. No 22

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

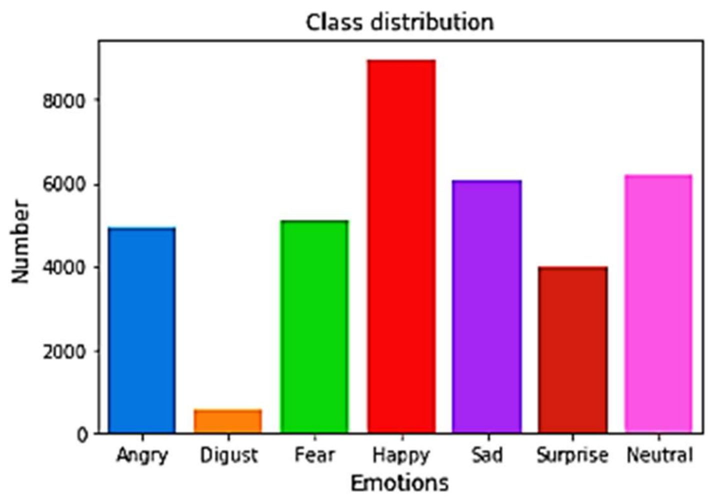


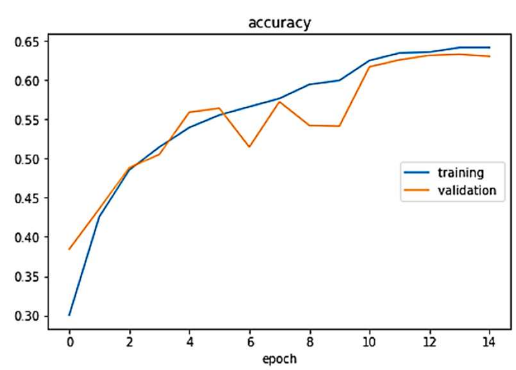
Figure 3: different emotion distribution based on dataset

4.3. The Hybrid CNN Model for Facial Emotion Recognition

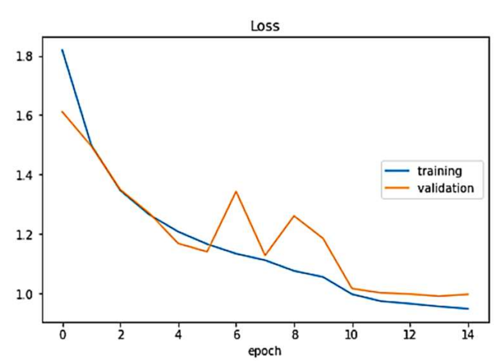
The CNN model is trained using Keras. OpenCV is used for face detection. The model is trained and saved. It is then deployed through a web interface using Flask. After the model has been trained, it is loaded into a main script and saved weights. It is then applied to a local video file and different datasets.

4.3.1. Creating training and validation batches

Keras is a tool that automatically feeds data from the test and training folder to create mini-batches for validation and training. The image size is 48x48 and the batch size is 64. The batch size is used to control the number of training samples that can be used to test and train the model in figure 4. The image data generator function of Keras generates data generators for both validation and training.



(a)



(b)

Figure 4: (a) Accuracy (b) Loss per epoch plot

4.3.2. Creating the CNN model

This project uses a sequential CNN model. The input passes through four convolution blocks.

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

The number of filters gradually increases. Each block has its own batch normalization, batch regularization, RELU, Max Pooling, and dropout regularization. At each step, the volume of the data decreases by a factor of 2. After the fourth block, the output is flattened and passed on to the two layers. A dense layer is then used to predict the output label, which will correspond to one of the seven emotions. The learning rate of the PSO with the GA optimizer is 0.0005, which speeds up the training time by about 9 minutes. Then output with all the parameters that the model will have to learn 7 emotions confusion matrix is plotted in figure 5.

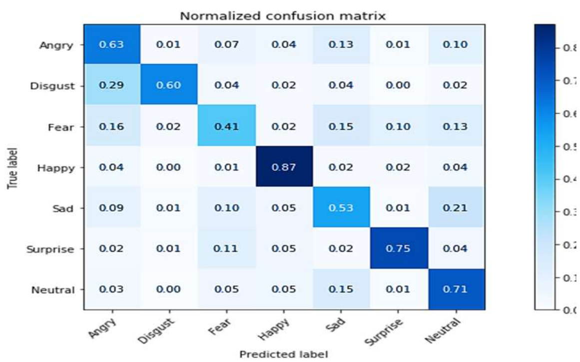


Figure 5: Confusion matrix for 7 emotional classes

4.3.3. Training and evaluating the model

The number of epochs is then used to control the number of complete passes that the training dataset generates. The number of steps per epoch is calculated by dividing the number of images in a training generator by the batch size. The validation set is followed by three callbacks. These objects can perform actions at various phases of training. The first callback is called before or after a single batch. It can reduce the learning rate if the validation loss does not improve after two epochs. The model checkpoint callback saves the weights of the training set in HDF5 format. Plot Losses is used to monitor the training loss and the accuracy per epoch. Each epoch takes around 9 to 10 minutes to complete. Due to the complexity of the process, the total time taken is about 2.5 hours.

Table 1: accuracy ck+ dataset [26]

Algorithm	Accuracy (%)
GA	82.95
PSO	83.25
CSO	84.35
CSO+PSO+GA	90.07
LBP + SVM	91.40
CNN	95.75
DCNN + SVM	96.02
Proposed	96.78

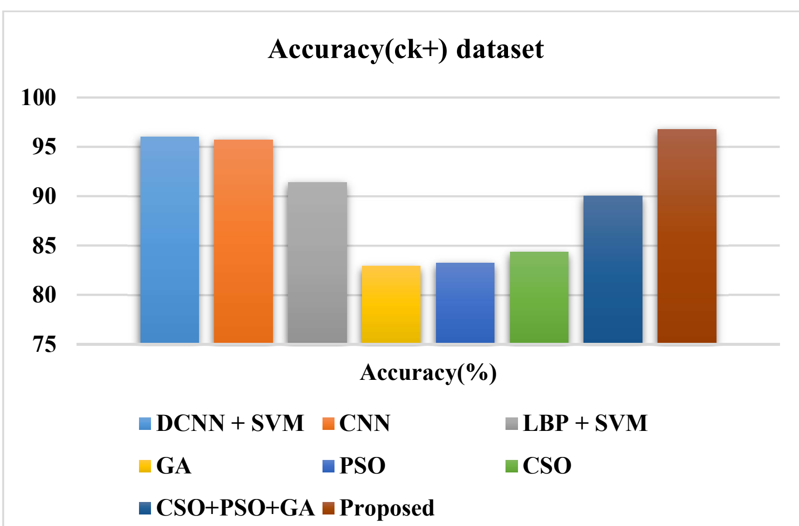


Figure 6: Accuracy ck+ dataset

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

Figure 6 shows comparison shows that our proposed method has a clear margin of improvement over other bio inspired algorithms [26]. We tried each algorithm 10 times to find the best possible result. Compared to other algorithms, our proposed method gives better accuracy values are illustrated in table 1.

Table 2: Accuracy JAFFE Dataset

Algorithm	Accuracy (%)
V. Chernykh et al	73
Y. Fan et al	79.16
P. Khorrami et al	82.43
T. Zhang et al	94.89
N,jain et al	94.90
proposed	96.89

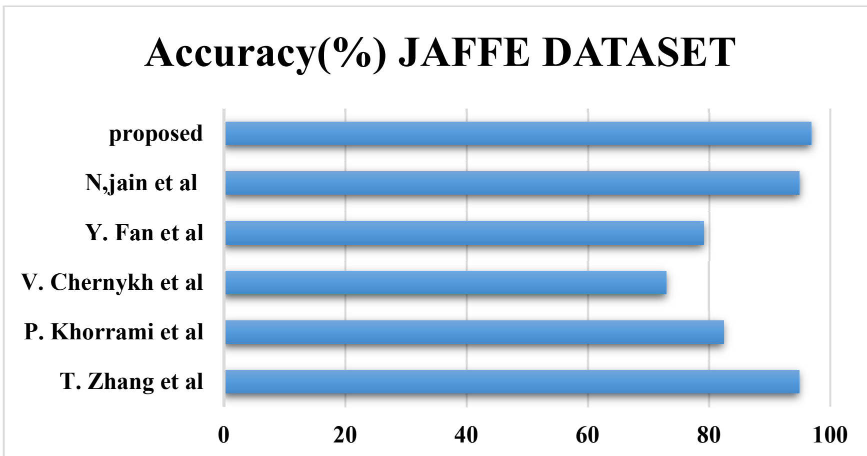


Figure 7: JAFFE Accuracy Dataset

Table 2 shows results of this study demonstrate the performance of the proposed Hybrid CNN-PSO-GA model compared to other state-of-the-art methods. The proposed model performed better than other state-of-the-art methods in terms of its performance on the JAFFE datasets[27] as showed in figure 7. Compared to all other algorithm our proposed methodology gives better accuracy

Table 3: Accuracy FER2013 Dataset

Algorithm	Accuracy (%)
Gabor filter	84.80
dual-branch CNN	85.71
LBP	87.20
HOG	89.70
C-CNN	91.64
Light-CNN	92.86
CNN	93.20
Pre-trained CNN	95.29
Proposed	97.03

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

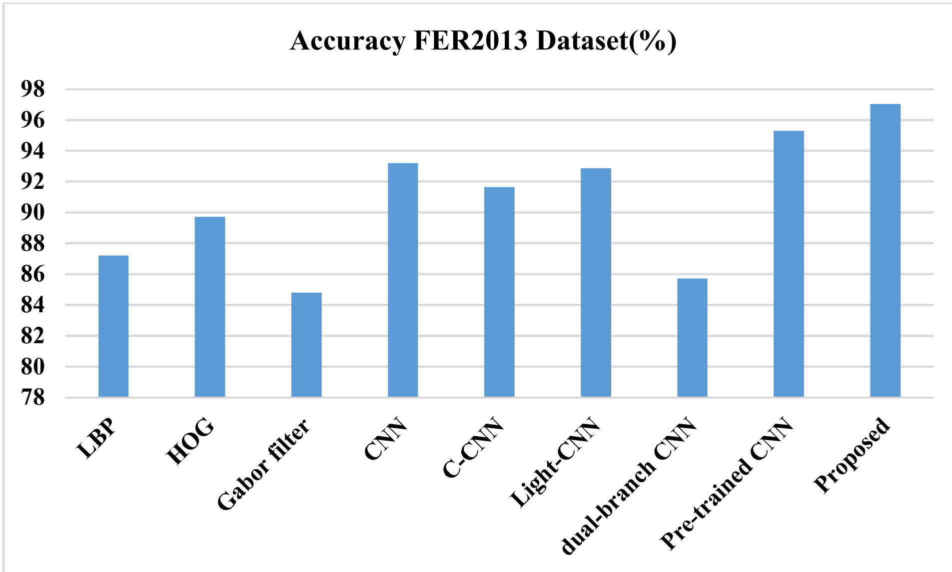


Figure 8: FER2013 Dataset Accuracy graph

The results of FER2013 Emotional database comparison are presented in Table 3, which shows the accuracy of different methods for capturing low-level transition patterns on facial expressions in figure 8. The methods that are presented in this table include CNN, LBP, and LGBP [28]. The results of the comparison show that the proposed method is capable of achieving higher accuracy.

(a) (b)

Figure 9: The input (a) and output (b) images of the FER model

Journal of Theoretical and Applied Information Technology
30th November 2022. Vol.100. No 22

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

In order to test the proposed framework's performance in real-world applications, we captured a photo of 27 participants from a meeting that was held on Tencent Meeting. The photo was taken before the meeting ended. The moderator was making a pleasant speech. The happiness of the students will increase significantly within a few minutes after the lecture ends. In this case, the faces in the photo are most likely to be neutral or happy. The output and input images of the CNN model show the results of the algorithm in figure 9. It was able to identify all the faces in the photo and mark their responses with the rectangle outlines. Ten of the 27 faces were labeled as happy, while 15 were labeled as neutral. 2 of the faces were labeled as sad held in FER2013 Dataset. The reason for this is that the 3rd and 2nd images were not detected properly by the CNN model. In addition, the 2 face images are so small that it is not possible to recognize them properly.

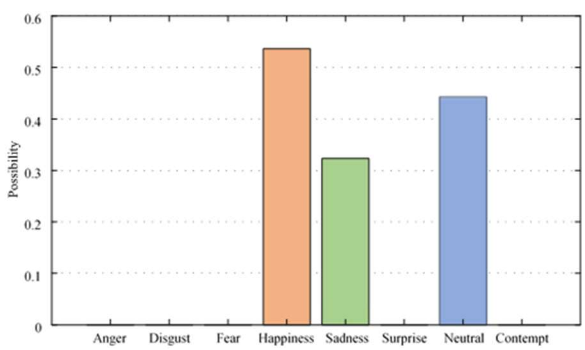


Figure 10: The probability distribution of emotions

Figure 10 shows the probability histogram is a statistical tool that allows us to observe the overall emotional state of class. While the probability of happiness is higher than that of neutral in Figure 10, the faces labeled as happy are less likely to be regarded as neutral. The difference between neutral and happy faces can be explained by the various features that appear on a face at the same time. For instance, if there are multiple expressions on a face, the most likely one will be labeled as neutral. The overall expression of an image depends on the number of features that appear on a face at the same time. The probability of happiness is higher in certain faces than in others. For instance, in some cases, the probability of happiness is higher than neutral. The results of this experiment can support the performance of a machine learning model when it is applied to real environments.

4.4. DISCUSSION

Despite the various advantages of this framework, there are still areas where it can be improved. For instance, deep learning models and computer simulations will continue to be developed to improve the performance of the frameworks. The image preprocessing tools used in this framework are mainly designed to detect and align the face, but they can also face problems when dealing with complex environments. In the future, these shortcomings could be solved. Although the CNN model currently performs well, it will eventually be replaced by models with better learning capabilities and higher accuracy. To maintain its competitiveness, the framework should be regularly updated and improved. In addition, with the increasing number of students participating in online courses, it is also not possible to ensure that everyone's expressions are fully representative of their emotions. To improve the efficiency of the framework, various measures can be taken to identify invalid information.

5. CONCLUSION

This study aims to develop a framework that can analyze the facial expressions of students using a computer simulation model. It is based on the CNN architecture. The framework can be presented in a histogram, which can then be used by teachers to adjust their teaching strategies. An algorithm for emotion recognition uses optical flow technology to recognize the feelings of people with varying backgrounds and skin tones. It can also help people with special needs, such as those with autism. It can also help organizations improve their business outcomes by analyzing the emotional responses of their audience. The system can recognize six different emotions in real time using facial landmarks and raw data. It can also perform various analysis and visualization techniques. The results of the study revealed that the deeper model, which is known as Auto-PSO CNN, performed better on emotion classification and facial feature learning.

REFERENCES

[1] Shan, C., Gong, S. and McOwan, P.W., 2009. Facial expression recognition based on local binary patterns: A comprehensive study. Image and vision Computing, 27(6), pp.803-816.

[2] Barrett, L.F., Adolphs, R., Marsella, S., Martinez, A.M. and Pollak, S.D., 2019. Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. Psychological science in the public interest, 20(1), pp.1-68.

[3] Zhang, T., Liu, M., Yuan, T. and Al-Nabhan, N., 2020. Emotion-Aware and Intelligent Internet of Medical Things Toward Emotion Recognition During COVID-19 Pandemic. IEEE Internet of Things Journal, 8(21), pp.16002-16013.

[4] Moiseeva, A., 2020. Statistical natural language processing methods for intelligent process automation (Doctoral dissertation, lmu).

[5] Martinez, B. and Valstar, M.F., 2016. Advances, challenges, and opportunities in automatic facial expression recognition. In Advances in face detection and facial image analysis (pp. 63-100). Springer, Cham.

[6] Deshmukh, S., Patwardhan, M. and Mahajan, A., 2016. Survey on real-time facial expression recognition techniques. Iet Biometrics, 5(3), pp.155-163.

[7] Marsh, A.A., Yu, H.H., Pine, D.S. and Blair, R.J.R., 2010. Oxytocin improves specific recognition of positive facial expressions. Psychopharmacology, 209(3), pp.225-232.

[8] Martinez, A. and Du, S., 2012. A model of the perception of facial expressions of emotion by humans: research overview and perspectives. Journal of Machine Learning Research, 13(5).

[9] Xue, B., Zhang, M. and Browne, W.N., 2012. Particle swarm optimization for feature selection in classification: A multi-objective approach. IEEE transactions on cybernetics, 43(6), pp.1656-1671.

[10] Mlakar, U., Fister, I., Brest, J. and Potočnik, B., 2017. Multi-objective differential evolution for feature selection in facial expression recognition systems. Expert Systems with Applications, 89, pp.129-137.

[11] Moradi, P. and Gholampour, M., 2016. A hybrid particle swarm optimization for feature subset selection by integrating a novel local search strategy. Applied Soft Computing, 43, pp.117-130.

[12] I.Michael Revina, W.R. Sam Emmanuel,A Survey on Human Face Expression Recognition Techniques,Journal of King Saud University - Computer and Information Sciences,Volume 33, Issue 6,2021,Pages 619-628,ISSN 1319-1578,

[13] Li, S. and Deng, W., 2020. Deep facial expression recognition: A survey. IEEE transactions on affective computing.

[14] Johnson, J.M., 2019. An Evaluation of Deep Learning with Class Imbalanced Big Data (Doctoral dissertation, Florida Atlantic University).

[15] Xie, S., Hu, H. and Wu, Y., 2019. Deep multi-path convolutional neural network joint with salient region attention for facial expression recognition. Pattern recognition, 92, pp.177-191.

[16] Wen, G., Hou, Z., Li, H., Li, D., Jiang, L. and Xun, E., 2017. Ensemble of deep neural networks with probability-based fusion for facial expression recognition. Cognitive Computation, 9(5), pp.597-610.

[17] Rahman, Z., Hossain, M.S., Islam, M.R., Hasan, M.M. and Hridhee, R.A., 2021. An approach for multiclass skin lesion classification based on ensemble learning. Informatics in Medicine Unlocked, 25, p.100659.

[18] Adjabi, I., Ouahabi, A., Benzaoui, A. and Taleb-Ahmed, A., 2020. Past, present, and future of face recognition: A review. Electronics, 9(8), p.1188.

[19] Li, S. and Deng, W., 2020. Deep facial expression recognition: A survey. IEEE transactions on affective computing.

[20] Pena, M.L.C., Carballal, A., Rodríguez-Fernández, N., Santos, I. and Romero, J., 2021. Artificial intelligence applied to conceptual design. a review of its use in architecture. Automation in Construction, 124, p.103550.

[21] Sikkandar, H. and Thiyagarajan, R., 2021. Deep learning based facial expression recognition using improved Cat Swarm Optimization. Journal of Ambient Intelligence and Humanized Computing, 12(2), pp.3037-3053.

[22] Alenazy, W.M. and Alqahtani, A.S., 2021. Gravitational search algorithm based optimized deep learning model with diverse set of features for facial expression recognition. Journal of Ambient Intelligence and Humanized Computing, 12(2), pp.1631-1646.

[23] Zatarain Cabada, R., Rodriguez Rangel, H., Barron Estrada, M.L. and Cardenas Lopez, H.M., 2020. Hyperparameter optimization in CNN for learning-centered emotion recognition for intelligent tutoring systems. Soft Computing, 24(10), pp.7593-7602.

[24] Gogić, I., Manhart, M., Pandžić, I.S. and Ahlberg, J., 2020. Fast facial expression recognition using local binary features and shallow neural networks. The visual computer, 36(1), pp.97-112.

[25] Liu, X., Cheng, X. and Lee, K., 2020. Ga-svm-based facial emotion recognition using facial geometric features. IEEE Sensors Journal, 21(10), pp.11532-11542.

[26] Li, K., Jin, Y., Akram, M.W. et al. Facial expression recognition with convolutional neural networks via a new face cropping and rotation strategy. Vis Comput 36, 391–404 (2020).

[27] Neha Jain, Shishir Kumar, Amit Kumar, Pourya Shamsolmoali, Masoumeh Zareapoor,Hybrid deep neural networks for face emotion recognition,Pattern Recognition Letters,Volume 115, 2018, Pages 101-106,ISSN 0167-8655,

[28] Jie Shao, Yongsheng Qian,Three convolutional neural network models for facial expression recognition in the wild, Neurocomputing, Volume 355,2019,Pages 82-92,ISSN 0925-2312.

[29] Wang, FS., Chen, LH. (2013). Heuristic Optimization. In: Dubitzky, W., Wolkenhauer, O., Cho, KH., Yokota, H. (eds) Encyclopedia of Systems Biology. Springer, New York, NY. https://doi.org/10.1007/978-1-4419-9863-7_411

[30] M. Nazari-Heris, B. Mohammadi-Ivatloo, G.B. Gharehpetian, "A comprehensive review of heuristic optimization algorithms for optimal combined heat and power dispatch from economic and environmental perspectives", Renewable and Sustainable Energy Reviews, Volume 81, Part 2, 2018, Pages 2128-2143