

Optimization of Convolutional Neural Network for Accurate Stress Detection Using Multimodal Physiological Signals

**Mr. Nilankar Bhanja, Mr. Sanjib Kumar Dhara, Dr. K Venkata Murali Mohan,
DR. Mukesh Tiwari,**

¹Research Scholar, Electronics and Communication Engineering,
Sri Satya Sai University of Technology & Medical Sciences,
Bhopal-Indore Road, Sehore (M.P), 466001, Madhya Pradesh, India
nilankarbhanja2005@gmail.com

²Research Scholar, Electronics and Communication Engineering,
Sri Satya Sai University of Technology & Medical Sciences,
Bhopal-Indore Road, Sehore (M.P), 466001, Madhya Pradesh, India
sanjibkumardhara@gmail.com

³Professor of ECE and Principal,
Teegala Krishna Reddy Engineering College,
kvmmece@gmail.com

⁴Professor of Electronics and Communication Engineering,
Sri Satya Sai University of Technology & Medical Sciences,
info@sssutms.co.in

KEYWORDS

Stress Detection,
Multimodal
Signal, EEG
Signal, Machine
Learning,
Convolutional
Neural Network,
Optimization,
Confusion Matrix,
Accuracy

ABSTRACT

Stress is an integral part of daily life that most individuals must manage regularly. However, long-term stress or high levels of stress can compromise safety and disrupt normal lifestyles. Early detection of mental stress can help prevent numerous stress-related health problems. When an individual experiences stress, noticeable changes occur in several physiological signals, including impedance, thermal, electrical, and optical signals. By analyzing these signals, stress levels can be effectively determined. Even with the use of advanced technology, existing research on stress detection has failed to produce satisfactory accuracy. To address this gap, the research proposed two ideas. First, instead of using a single physiological signal, the study used multimodal signals. Electrocardiogram (ECG) and Electroencephalogram (EEG) signals under no-stress, low-stress, and high-stress conditions were acquired from the Kaggle site. Second, the Convolutional Neural Network (CNN) hyperparameters were tuned using a bio-inspired optimization technique called the Firefly Algorithm (FA). The drawbacks of the FA were identified and further improved, leading to the development of an Improved Firefly Algorithm (IFA) to fine-tune the CNN hyperparameters. The multimodal data from Kaggle was processed to remove noise and then fed into the proposed IFA-CNN, FA-CNN, and a baseline CNN model to predict the stress levels of individuals. Additionally, the three models were also tested with ECG and EEG data separately. The outcomes of all three models, using ECG, EEG, and multimodal data, were compared using positive metrics (accuracy, recall, precision, F1-score) and negative metrics (False Negative Rate (FNR), and False Positive Rate (FPR)). The experimental results showed that the proposed IFA-CNN using multimodal data achieved the highest correct stress-level prediction, with 47 out of 48 samples correctly identified, yielding an accuracy of 97.92%. The comparison results also highlighted the advantage of using multimodal data over single-signal data. The proposed approach is highly beneficial for reliable stress-level prediction in individuals.

I. INTRODUCTION

In today's world, people's lives are more complex than ever before. A new study reveals that working professionals aged 25 to 40 are the most affected by stress in their daily lives. Recent research has demonstrated that stress can slowly but steadily damage brain cells [1]. Any mental health condition often originates from stress, which can manifest in various ways within the human body. Stress levels are rising due to factors such as financial difficulties, job losses, unemployment, and work pressures related to meeting deadlines [2]. The stability of the neurological system depends on the brain's ability to respond effectively to stress. Prolonged stress in professionals leads to irritability and reduced performance. It also results in several health issues, including sleeplessness, weakened immunity, infections, cervical impairments, headaches, and other physical ailments [3]. The subtle yet significant impact of stress on mental health often goes unnoticed. Strategic planning by healthcare organizations, the use of digital tools for stress prediction, and the growth of entrepreneurship in the medical services sector could collectively improve the standard of living and unlock greater human potential for all citizens.

The primary signs of stress include changes in core body temperature, neural activity, and ocular motility. The World Health Organization (WHO) has identified stress as a leading cause of illness globally [4]. A mental health survey conducted in India by the National Institute of Mental Health and Neuro Sciences from 2015 to 2016 revealed that the percentage of individuals with mental health issues increased from 7.5% in 2014 to 10.6% in 2016 [5]. The low- and middle-income patient-to-doctor ratio is in jeopardy. More than 150 million people in India suffer from various mental illnesses, including anxiety, depression, and personality disorders. The effects of the COVID-19 pandemic on people's lives have further exacerbated stress levels. These mental disorders require crucial mental health care; however, there is a 74% to 90% treatment gap for such services [6].

There are several approaches to recording and collecting human stress levels. Photoplethysmography (PPG), Phonocardiography (PCG), Respiration Rate (RR), Heart Rate (HR), Galvanic Skin Response (GSR), ECG, and EEG are some of the physiological signals commonly used in stress detection methods [7]. According to research, biochemical and biological processes often produce contradictory results due to hormonal instability.

Mental stress is a widely recognized concept that is gaining traction in various research fields, such as neurology, medicine, psychology, and sentiment computing. Consequently, it is vital to investigate stress through a range of techniques. Physiological signals are particularly effective at revealing stress-related patterns. However, professionals often struggle to manage and interpret large volumes of data from diverse sources, such as wearable devices, audio recordings, or facial images [8]. Stress detection typically relies on personal consultations with psychologists or counselors, which are time-consuming, costly, and not universally accessible [9]. While specialists remain essential in many cases, Artificial Intelligence (AI) offers a complementary approach that enhances stress detection by enabling real-time, scalable, cost-effective, and objective monitoring. The adoption of AI can make stress management more personalized and proactive, leading to better mental health outcomes [10].

In recent times, Deep Learning (DL) algorithms have been increasingly utilized to analyze medical signals for disease diagnosis [11, 12]. Several studies have proposed DL models for stress detection; however, challenges such as insufficient data, noise, and inefficient feature extraction due to similarities in patterns between no-stress and low-stress signals result in suboptimal accuracy. To address these limitations, multimodal signals are utilized instead of single physiological signals and an optimized CNN model is proposed to fine-tune the hyperparameters, thereby enhancing the accuracy of stress prediction. The main contributions of this research are as follows:

- Multimodal ECG and EEG signals are employed to identify the stress level of individuals.
- The CNN hyperparameters are fine-tuned using the Improved Firefly Algorithm (IFA), which aids in extracting the most significant features and improving classification accuracy.
- The proposed IFA-CNN model's performance is compared with the FA-CNN and standard CNN models without optimization using standard evaluation metrics for both positive and negative outcomes.
- The outcomes of the multimodal signals are also evaluated against single ECG and EEG signals to highlight the importance of using multimodal signals in stress prediction.

The article is structured as follows: Section I discusses stress and the available stress-measuring physiological signals. Section II provides a brief overview of related work on stress prediction and its limitations. Section III explains CNN and the optimization techniques employed for hyperparameter selection. Section IV presents the data and experimental outcomes of CNN and optimized CNN models using multimodal and single physiological signals. Section V concludes the research and outlines future work.

II. RELATED WORK

Many studies have already been conducted on stress prediction using physiological signals with DL models. Some of the notable studies are given in Table I. The table helps to identify the data used, the accuracy attained by the model, along with its advantages and limitations. The literature survey helps to detect the current limitations and propose a novel model.

TABLE I. RECENT WORK ON STRESS PREDICTION USING DL MODELS

Ref	Model	Data Used	Accuracy	Advantage	Limitation
[13]	Discrete Wavelet Transform+ CNN+Bidirectional Long Short-Term Memory (BiLSTM) + Gated Recurrent Unit (GRU)	EEG	98.10%	High accuracy Effective feature extraction and classification	High complexity due to parameter tuning Time-consuming
[14]	2D CNN + LSTM	EEG	97.8%	Faster training with fewer epochs	Small datasets limit generalizability and potential for overfitting
[15]	Support Vector Machine	PPG	95.55%	Non-intrusive and readily available sensors in smartwatches; suitable for real-time stress detection.	Limited to the WESAD dataset; potential challenges in generalizing to diverse populations
[16]	LSTM + CNN	EEG	97.8%	Effective temporal and spatial feature extraction	The complex model structure may increase implementation complexity.
[17]	GRU	EEG	95%	High accuracy in multi-level stress classification	Limited to specific gaming scenarios.
[18]	Boosting Neural Network	HR, Temperature	94%	Wearable, non-invasive devices that do not require sophisticated medical equipment.	Need to improve accuracy further by optimization

[19]	CNN + BiLSTM	ECG	86.5%	Achieves real-time stress detection using just 10s of ECG data	Accuracy is really low
[20]	CNN + BiLSTM with Attention Mechanism	ECG	86.8%	The effectiveness of RNN with Attention Mechanism helps in effective feature extraction	Need to improve accuracy
[21]	CNN	PPG	92.04%	Non-invasive, wearable, and easy-to-use ear-mounted PPG sensor.	Accuracy needs to improve
[22]	Hybrid DL (LSTM, GRU, 1D-CNN)	Remote PPG	95.83%	Computational efficiency and suitability for edge devices.	Signal-to-noise ratio issues, and varying accuracy across diverse populations.

III. PROPOSED METHODOLOGY

The bio-inspired optimization algorithm is used for hyperparameter tuning of the CNN model. Hyperparameters of a CNN are crucial for improving classification accuracy and enhancing model generalizability. It is essential to select hyperparameters appropriately. Instead of using backpropagation for adjusting hyperparameter values, the IFA technique is employed. This approach produces better results compared to traditional methods. The architecture of the proposed IFA-CNN is shown in Figure 1. The CNN consists of convolution, pooling, and flattening layers for feature extraction from the multimodal signals, followed by fully connected layers and an output layer for classification. Based on the loss function in the output layer, the hyperparameters are tuned using the IFA technique. The tuning process continues until satisfactory results are achieved or the maximum number of epochs is reached. This technique has the advantage of better exploration and exploitation during the search for the optimal solution, making it a robust method for hyperparameter optimization. The detailed architecture and workings of CNN, FA, and IFA are explained in this section.

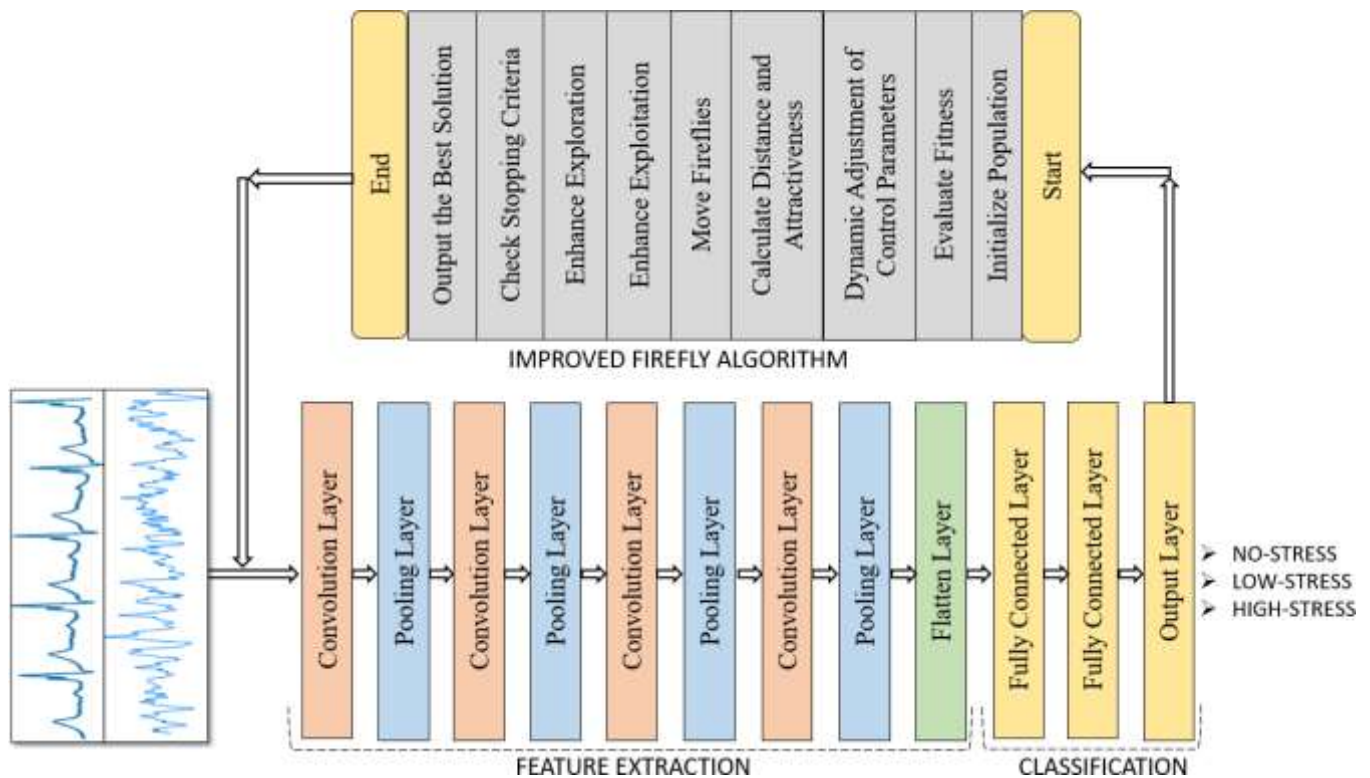


Fig. 1. Proposed Improved Firefly Algorithm for Optimization of CNN Hyperparameter Tuning.

A. CNN Architecture

The essential structure of a CNN consists of four layers: the Convolutional Layer (CL), the Pooling Layers (PL), the Activation Function (AF), and the Fully Connected Layer (FCL) [23]. In general, the signal is pre-processed before entering the network through the input layer. This is followed by a sequence of alternating CL and PL layers, and the process concludes with classification by the FCL. CNNs differ from Multi-Layer Perceptrons (MLPs) in that they include distinct CL and PL layers. When working with larger datasets, CNNs excel in both network performance and scalability [24].

Convolutional Layer

Multiple CLs can extract various input features from CNNs of a given depth [25]. The bottom CL retrieves common properties such as patterns, lines, and boundaries, while the top layer retrieves more abstract features. The CL consists of multiple convolution kernels with learnable parameters [26]. Kernels are typically denoted as matrices with learnable weights, commonly in dimensions of 7×7 , 5×5 , or 3×3 , with equal length and breadth of odd numbers. The feature maps are usually fed into the CL. Input feature maps are often denoted as $D \times B \times C$ (depth, breadth, and channels), while kernels are represented as $K \times K \times C$. This indicates that the number of kernels must match the number of input channels. Figure 3 illustrates the convolution process using input feature maps ($5 \times 5 \times 3$) and a kernel ($3 \times 3 \times 3$). The data flow in the CL can be described in Equation (1):

$$feature_surface_{out} = f(\sum_{i=3}^3 M_i * W_i + B) \quad [1]$$

Where M_i represents the feature surface of the input feature maps, W_i represents the kernel's weight matrix, M is the bias matrix, $f(\cdot)$ represents the nonlinear activation function (AF), and $feature_surface_{out}$ represents the output surface. The cross-correlation function between the feature maps and the kernel requires specific calculations within the CL. The dimension of the output feature map is given in Equation (2) and is influenced by the size of the 2-D input matrix (i), stride (s), kernel size (k), and the padding (p).

$$o = \left\lfloor \frac{i+2p-k}{s} \right\rfloor + 1 \quad [2]$$

The kernel begins in the upper-left corner of the input matrix and progresses from top to bottom and left to right. The input is a three-dimensional feature map matrix with length and width values of three. It is important to note that gradient backpropagation is used to adjust the kernel's weight parameters [27]. When the kernel examines the same set of input feature maps, the parameters remain constant. Each pixel region applies a similar kernel sliding technique, allowing the kernel to share parameters. This approach simplifies and streamlines the process, enabling it to run on large datasets while minimizing the number of training parameters and reducing the risk of overfitting.

Pooling Layer

The PL is often placed following the CL. The primary reasons for adopting the PL include the following: performing downsampling and reducing dimensionality on the input signal to minimize the connections in the CL [28], thereby lowering the computational cost of the model; ensuring that the input signal is translation, scale, and rotation invariant; and increasing the output feature map's tolerance for single-neuron distortions and inaccuracies. Average and maximum pooling are the most frequently employed techniques. These methods can significantly reduce overfitting in CNNs [29]. During the downsampling process, the general correlation between the input and output matrix dimensions in the pooling procedure is given in Equation (3):

$$o = \left\lfloor \frac{(i-k)}{s} + 1 \right\rfloor \quad [3]$$

Activation Function

The AF creates a functional relationship between input and output, which leads to a nonlinear system inside the neural network [30]. A proper nonlinear AF may significantly improve network efficiency. There are several AFs available. Saturating nonlinearities include the *sigmoid* and *tanh* equations. When the input is extremely large or extremely small, the *sigmoid* function approaches 0 or 1, while the *tanh* function reaches -1 or 1. To address the challenges highlighted by saturating nonlinearities, non-saturating nonlinearities such as ReLU were developed and given in Equation (4). The former functions require substantially longer to train with gradient descent than the latter.

$$ReLU(x) = \max(0, x) \quad [4]$$

Fully Connected (FC) Layer

The FCL is typically placed after the CL and PL, with its neurons fully connected between layers. It combines and classifies the extracted features obtained from the CL and PL. The neurons in the output layer are equal to the number of categories, which helps to categorize the signal's category. In simple terms, the FCL functions as a classifier in CNNs. Before implementing the FCL's loss calculation, the CNN output is normalized using softmax regression. When training a complex network with numerous parameters on a small database, the FCL often employs dropouts and L2 regularization. The primary objective of these techniques is to prevent the model from overfitting [31]. A typical CNN model combines ReLU with dropout to achieve reliable classification results.

Loss Function

Along with the various layers of the CNN structure described in the preceding section, the classification is performed in the output layer, which is typically the final layer of the FCL. Several loss functions influence the efficiency of the CNN design and are utilized for various tasks. For multi-class stress classification, Cross-Entropy [32] has been utilized as the CNN model's loss function. Equation (5) provides the formula:

$$Loss(y, y^*) = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^C y_{i,c} \log(p_{i,c}) \quad [5]$$

Where N denotes the samples, C represents the classes, $y_{i,c}$ represents the binary value (0 or 1) that specifies whether class c is the correct label for the i -th data, and $p_{i,c}$ represents the predicted probability of the i -th data belonging to class c .

B. Hyperparameter Optimization

To achieve excellent outcomes in CNN, the hyperparameter values must be selected properly. Instead of using traditional parameter tuning, bio-inspired algorithms facilitate better classification results through enhanced exploration and exploitation [33]. The hyperparameter optimization problem associated with the CNN technique for stress classification can be described in Equation (6-15):

Optimization problem:

$$x^* = \arg \max \phi(x)_x \quad [6]$$

Subject to constraints:

$$x_1 \in \{16, 32, 64, 128, 256\} \quad [7]$$

$$x_2 \in \{3, 5, 7, 9, 11\} \quad [8]$$

$$2 < x_3 < 6 \quad [9]$$

$$1 < x_4 < 5 \quad [10]$$

$$x_5 \in \{16, 32, 64, 128, 256\} \quad [11]$$

$$0.1 < x_6 < 0.5 \quad [12]$$

$$1 * 10^{-5} < x_7 < 1 * 10^{-2} \quad [13]$$

$$x_8 \in \{16, 32, 64, 128, 256\} \quad [14]$$

$$10 \leq x_9 \leq 100 \quad [15]$$

The target function, $\varphi()$, measures the CNN model's performance, which is typically the classification accuracy of the input data $x = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8]$; where x_1 represents the number of filters, x_2 represents the kernel size, x_3 represents the pooling size, x_4 represents the number of dense layers, x_5 represents the number of neurons in dense layers, x_6 represents the dropout rate, x_7 represents the learning rate, x_8 represents the batch size, and x_9 represents the epochs. This combinatorial problem, involving sets of integer and real number variables, is solved by optimizing the hyperparameters using the FA and IFA.

Firefly Algorithm

Xin-She Yang introduced the FA for optimization [34]. In this method, fireflies represent solutions. Based on Light Intensity (LI), fireflies are attracted to each other, and the solutions are directed towards the brighter ones. The FA is based on three basic principles [35]:

- A firefly is unisex and can be attracted to another firefly.
- The brightness influences the attractiveness of fireflies; as the distance between them decreases, their attractiveness increases, making them appear brighter.
- The brightness of fireflies indicates their fitness function.

Solutions are produced randomly between the upper and lower boundaries and given in Equation (16):

$$x_{i,j} = lb_j + rand. (ub_j - lb_j), rand \in 0,1 \quad [16]$$

Where, $x_{i,j}$ represents the solution, ub_j and lb_j represent the upper and lower bounds, respectively. The Euclidean distance, calculated using Equation (17), determines how close two solutions are to each other:

$$r_{i,j} = \|x_i - x_j\| = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2} \quad [17]$$

Where, $r_{i,j}$ represents the distance between the solutions of i and j , and d represents the dimension of the problem. The LI is computed using the Equation (18):

$$I(r) = \frac{I_0}{1 + \gamma r^2} \quad [18]$$

The LI is represented by I_0 , whereas the LI at distance r is represented by $I(r)$, where γ represents the light absorption coefficient. The attractiveness level of two solutions is computed using Equation (19):

$$\beta(r) = \frac{\beta_0}{1 + \gamma r^2} \quad [19]$$

The attractiveness at zero distance is represented by β_0 , while β represents the attractiveness itself. The solutions move to a new position in the search space using the Equation (20):

$$x_i^{t+1} = x_i^t + \beta_0 r^{-\gamma r_{i,j}^2} (x_j - x_i) + \alpha (rand - 0.5) \quad [20]$$

Where, x_i^{t+1} represents the new position, x_i^t denotes the current position, and the movement progresses toward the brighter solution x_j . The control parameter α is obtained from the Gaussian distribution.

Improved Firefly Algorithm

The FA's exploration and exploitation balance is primarily determined by the parameter α , which changes dynamically during each algorithm run using Equation (21) [36]:

$$\alpha(t) = \left(1 - \left(1 - (10^{-4}/9)^{1/t_{max}}\right)\right) * \alpha(t-1) \quad [21]$$

Where, t_{max} represents the maximum iterations performed in a single run. This indicates that the value of α is higher at the start of a run (providing greater exploration) and decreases throughout the search process (leading to greater exploitation power).

Although the FA has relatively steady and effective exploitation, as determined by simulation outcomes on standard unconstrained benchmarks, the intensification-diversification trade-off can be improved by integrating additional steps into the FA technique. In certain cases, the FA technique exhibits premature convergence, which occurs when the entire population converges on the current optimal solutions [37]. To address this, the exploration capability of the FA is enhanced by introducing a population randomization technique. The FA ($E^3 - FA$) incorporates two approaches to improve exploitation and exploration.

The first strategy improves the exploitation capability of the standard FA by introducing a new randomization parameter, p , into the algorithm. This parameter determines whether the solution is updated using Equation (20) or dispersed around the current optimal solution based on a normal distribution.

$$x_i^{t+1} = x_{*,j}^t + N(x_{*,j}^t, \sigma_t^2) \quad [22]$$

Where, x_i^{t+1} represents the updated solution, $x_{*,j}^t$ represents the current optimal solution, and $N(\cdot)$ represents the normal distribution with the mean being the current optimal solution and variance σ^2 at iteration t . The variance influences the step size and decreases with the iterations. The distribution of the current optimal solution is computed using Equation (23):

$$N(x_i^t; x_{*,j}^t, \sigma_t^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(x_i^t - x_{*,j}^t\right)^2 / 2\sigma^2} \quad [23]$$

By using Equation (23), the updated solutions move nearer to the optimal solution, making it easier to reach the global optimum more quickly.

As the technique approaches its maximum iteration, the step size (σ) decreases and approaches zero ($\sigma_t \rightarrow 0$). Equation (24) is used to adjust the step size at each iteration t .

$$\sigma_{t+1} = \sigma_0 - \sigma_0 t / t_{max} \quad [24]$$

Where, t denotes the current iteration, σ_0 denotes the initial step size, and t_{max} denotes the maximum iteration.

The second strategy, which improves exploration, involves sorting the solutions after each iteration based on their FF value. The weakest solutions are replaced with new random solutions generated by Equation (16). This method introduces new, random solutions to prevent the algorithm from becoming stuck in suboptimal regions, especially if it converges in the incorrect part of the search space.

IV. RESULT AND DISCUSSION

The outcomes of the

A. Data Collection

The multimodal ECG and EEG data were collected from the Kaggle site [38] for stress detection. The Human Research Ethics Committee of Prince of Songkla University, Thailand, approved the study on July 22, 2020. Participants were selected from August 2020 to May 2021. This study comprised 40 healthy university students, 21 of whom were female (aged 18 to 25). All participants had no history of neurological or heart disease, and none were taking medications that could impact the autonomic nervous system. The Thai version of the Perceived Stress Scale was employed to assess underlying tension in the month leading up to the trial day, with each subject's stress classified into three levels. To manage the impact of circadian rhythms on physical activity, all trials were held in a quiet laboratory environment in noon (2-4 p.m.).

The ECG was recorded using a bipolar limb lead attached to an ECG amplifier (AD Instruments, New Zealand). The signals were filtered from 0.3 to 200 Hz and captured at a sampling rate of 1000 Hz. EEG electrodes were fitted to a helmet with 8 electrodes: Fp1, Fp2, P3, P4, F3, F4, T3, and T4. Monopolar recordings were performed between the reference and active electrodes, with a sampling rate of 200 Hz, and the filter was configured as low-pass with a 200 Hz cut-off frequency.

Stress was induced in this study using a mathematical challenge that is routinely utilized to cause brain activation and physiological responses (blood pressure, cortisol levels, and HR). The experimental procedure involved six steps: Step 1 was habituation, Step 2 was the eyes-open period, which is the non-stress scenario or baseline, Step 3 was the mathematical stress challenge 1, which is the low-stress scenario, Step 4 was a break, Step 5 was the mathematical stress challenge 2, which is the high-stress scenario, and Step 6 was recovery, which allowed subjects to rest for 5 minutes.

The physiological signals of EEG and ECG were collected from Kaggle and pre-processed to remove noise. Noise is introduced in both signals due to various reasons. To eliminate the noise, a band-pass filter was utilized, which helps cut off frequencies above the high-pass and below the low-pass cutoff range. This improves the results of stress detection. The EEG and ECG samples in the dataset consist of 40 no-stress samples, 40 low-stress samples, and 40 high-stress samples. However, for deep learning, the dataset was insufficient. Data augmentation was performed to duplicate the existing samples, resulting in 160 samples in each category. The augmented dataset was split into three parts: 336 samples for training, 96 for validation, and 48 for testing.

After pre-processing, the training and validation samples of both ECG and EEG signals were fed into the normal CNN and the optimized FA-CNN and IFA-CNN models for training and validation to detect the stress state of individuals. For CNN, hyperparameters were randomly fixed using the trial-and-error method. For FA-CNN and IFA-CNN, the hyperparameters were chosen using optimization techniques. After training and validation, the models were tested with 48 samples.

The confusion matrix for each model using multimodal data is shown in Table 1. The IFA-CNN correctly predicted 46 samples and misclassified 2 samples. The FA-CNN correctly predicted 44 samples and misclassified 4 samples, while the CNN correctly identified 10 samples and misclassified 8. Based on the confusion matrix, performance metrics were calculated. Table 1 also shows the accuracy of each model using multimodal data. The IFA-CNN achieved the highest accuracy of 0.9792.

In addition to accuracy, other important metrics (Precision, Recall, F1, TNR, FNR, and FPR) were calculated to validate the efficiency of the IFA-CNN model, as shown in Table II. Similarly, the metrics for FA-CNN and CNN for each stress category are tabulated in Tables III and IV. Analysis of Tables II to

IV reveals that IFA-CNN achieved the highest Precision, Recall, and F1 scores of 0.9804, 0.9792, and 0.9791, respectively, and the lowest FNR and FPR values of 0.0208 and 0.0104.

TABLE II. CONFUSION MATRIX AND ACCURACY OF CNN MODELS USING MULTIMODAL DATA

Model	Accuracy	Label	No Stress	Low Stress	High Stress
IFA-CNN	0.9792	No Stress	15	0	1
		Low Stress	0	16	0
		High Stress	0	0	16
FA-CNN	0.9167	No Stress	14	2	0
		Low Stress	0	15	0
		High Stress	1	1	15
CNN	0.8333	No Stress	13	1	0
		Low Stress	2	13	3
		High Stress	1	1	14

TABLE III. PERFORMANCE ANALYSIS OF IFA-CNN ON STRESS DETECTION USING MULTIMODAL DATA

Class	Precision	Recall	F1	FNR	FPR
No Stress	1.0000	0.9375	0.9677	0.0625	0.0000
Low Stress	1.0000	1.0000	1.0000	0.0000	0.0000
High Stress	0.9412	1.0000	0.9697	0.0000	0.0313
Macro	0.9804	0.9792	0.9791	0.0208	0.0104
Micro	0.9792	0.9792	0.9792	0.0208	0.0104

TABLE IV. PERFORMANCE ANALYSIS OF FA-CNN ON STRESS DETECTION USING MULTIMODAL DATA

Class	Precision	Recall	F1	FNR	FPR
No Stress	0.9333	0.8750	0.9032	0.1250	0.0313
Low Stress	0.8333	1.0000	0.9091	0.0000	0.0909
High Stress	1.0000	0.8824	0.9375	0.1176	0.0000
Macro	0.9222	0.9191	0.9166	0.0809	0.0407
Micro	0.9167	0.9167	0.9167	0.0833	0.0417

TABLE V. PERFORMANCE ANALYSIS OF CNN ON STRESS DETECTION USING MULTIMODAL DATA

Class	Precision	Recall	F1	FNR	FPR
No Stress	0.8125	0.9286	0.8667	0.0714	0.0882
Low Stress	0.8667	0.7222	0.7879	0.2778	0.0667
High Stress	0.8235	0.8750	0.8485	0.1250	0.0938
Macro	0.8342	0.8419	0.8343	0.1581	0.0829
Micro	0.8333	0.8333	0.8333	0.1667	0.0833

To analyze the advantages of using multimodal signals for stress detection, the same steps were repeated with ECG and EEG signals alone. This means that the CNN, FA-CNN, and IFA-CNN models were trained, validated, and tested with ECG and EEG signals separately for stress level prediction. The accuracy and confusion matrices of the models (CNN, FA-CNN, IFA-CNN) are provided in Table V. The

IFA-CNN achieved the highest number of correctly identified stress level samples (43) using ECG signals alone, followed by FA-CNN (41) and CNN (40). Based on the confusion matrix, accuracy and other metrics were calculated. The accuracy achieved by IFA-CNN, FA-CNN, and CNN was 0.8958, 0.8542, and 0.8333, respectively. The other metrics for each category of stress level prediction by IFA-CNN, FA-CNN, and CNN are presented in Tables VI and VII. When using ECG signals alone, FA-CNN achieved better precision (0.8969), recall (0.9020), and F1 (0.8969) compared to the other models. The Precision, Recall, and F1 scores for FA-CNN were 0.8556, 0.8600, and 0.8552, respectively, while for CNN, they were 0.8120, 0.8132, and 0.812.

TABLE VI. CONFUSION MATRIX AND ACCURACY OF CNN MODELS USING ECG DATA

Model	Accuracy	Label	No Stress	Low Stress	High Stress
IFA-CNN	0.8958	No Stress	14	0	0
		Low Stress	1	14	2
		High Stress	1	1	15
FA-CNN	0.8542	No Stress	12	1	0
		Low Stress	2	15	1
		High Stress	1	2	14
CNN	0.8333	No Stress	13	2	1
		Low Stress	2	13	1
		High Stress	1	1	14

TABLE VII. PERFORMANCE ANALYSIS OF IFA-CNN ON STRESS DETECTION USING ECG DATA

Class	Precision	Recall	F1	FNR	FPR
No Stress	0.8750	1.0000	0.9333	0.0000	0.0588
Low Stress	0.9333	0.8235	0.8750	0.1765	0.0323
High Stress	0.8824	0.8824	0.8824	0.1176	0.0645
Macro	0.8969	0.9020	0.8969	0.0980	0.0519
Micro	0.8958	0.8958	0.8958	0.1042	0.0521

TABLE VIII. PERFORMANCE ANALYSIS OF FA-CNN ON STRESS DETECTION USING ECG DATA

Class	Precision	Recall	F1	FNR	FPR
No Stress	0.8000	0.9231	0.8571	0.0769	0.0857
Low Stress	0.8333	0.8333	0.8333	0.1667	0.1000
High Stress	0.9333	0.8235	0.8750	0.1765	0.0323
Macro	0.8556	0.8600	0.8552	0.1400	0.0727
Micro	0.8542	0.8542	0.8542	0.1458	0.0729

TABLE IX. PERFORMANCE ANALYSIS OF CNN ON STRESS DETECTION USING ECG DATA

Class	Precision	Recall	F1	FNR	FPR
No Stress	0.8000	0.8000	0.8000	0.2000	0.0909
Low Stress	0.8235	0.8750	0.8485	0.1250	0.0938
High Stress	0.8125	0.7647	0.7879	0.2353	0.0968
Macro	0.8120	0.8132	0.8121	0.1868	0.0938
Micro	0.8125	0.8125	0.8125	0.1875	0.0938

The analysis of stress level prediction using EEG signals alone was also conducted. The outcome of the confusion matrix and accuracy is given in Table IX. In this scenario, the correct prediction of stress level samples by IFA-CNN, FA-CNN, and CNN was 41, 39, and 37, respectively. The misclassifications were 7 for IFA-CNN, 9 for FA-CNN, and 11 for CNN. The accuracy of IFA-CNN, FA-CNN, and CNN was 0.8542, 0.8125, and 0.7708, respectively. The remaining metrics for IFA-CNN, FA-CNN, and CNN using EEG signals are provided in Tables X to XII.

TABLE X. CONFUSION MATRIX AND ACCURACY OF CNN MODELS USING EEG DATA

Model	Accuracy	Label	No Stress	Low Stress	High Stress
IFA-CNN	0.8542	No Stress	14	1	2
		Low Stress	0	14	1
		High Stress	1	2	13
FA-CNN	0.8125	No Stress	12	1	2
		Low Stress	1	14	1
		High Stress	2	2	13
CNN	0.7708	No Stress	11	1	2
		Low Stress	1	13	2
		High Stress	2	3	13

TABLE XI. PERFORMANCE ANALYSIS OF IFA- CNN ON STRESS DETECTION USING EEG DATA

Class	Precision	Recall	F1	FNR	FPR
No Stress	0.9333	0.8235	0.8750	0.1765	0.0323
Low Stress	0.8235	0.9333	0.8750	0.0667	0.0909
High Stress	0.8125	0.8125	0.8125	0.1875	0.0938
Macro	0.8565	0.8565	0.8542	0.1435	0.0723
Micro	0.8542	0.8542	0.8542	0.1458	0.0729

TABLE XII. PERFORMANCE ANALYSIS OF FA-CNN ON STRESS DETECTION USING ECG DATA

Class	Precision	Recall	F1	FNR	FPR
No Stress	0.8125	0.8125	0.8125	0.1875	0.0938
Low Stress	0.8125	0.8125	0.8125	0.1875	0.0938
High Stress	0.8750	0.8750	0.8750	0.1250	0.0625
Macro	0.8333	0.8333	0.8333	0.1667	0.0833
Micro	0.8333	0.8333	0.8333	0.1667	0.0833

TABLE XIII. PERFORMANCE ANALYSIS OF CNN ON STRESS DETECTION USING EEG DATA

Class	Precision	Recall	F1	FNR	FPR
No Stress	0.7857	0.7857	0.7857	0.2143	0.0882
Low Stress	0.7647	0.8125	0.7879	0.1875	0.1250
High Stress	0.7647	0.7222	0.7429	0.2778	0.1333
Macro	0.7717	0.7735	0.7722	0.2265	0.1155
Micro	0.7708	0.7708	0.7708	0.2292	0.1146

Finally, to determine which model and signal type provide better stress level prediction, the accuracy of different CNN variants using single and multimodal signals was compared. The comparison chart is

presented in Figure 3. The figure clearly shows that IFA-CNN achieves better accuracy than FA-CNN and CNN models. By comparing the accuracy across all IFA-CNN models, the multimodal signal achieved the highest accuracy of 0.9792.

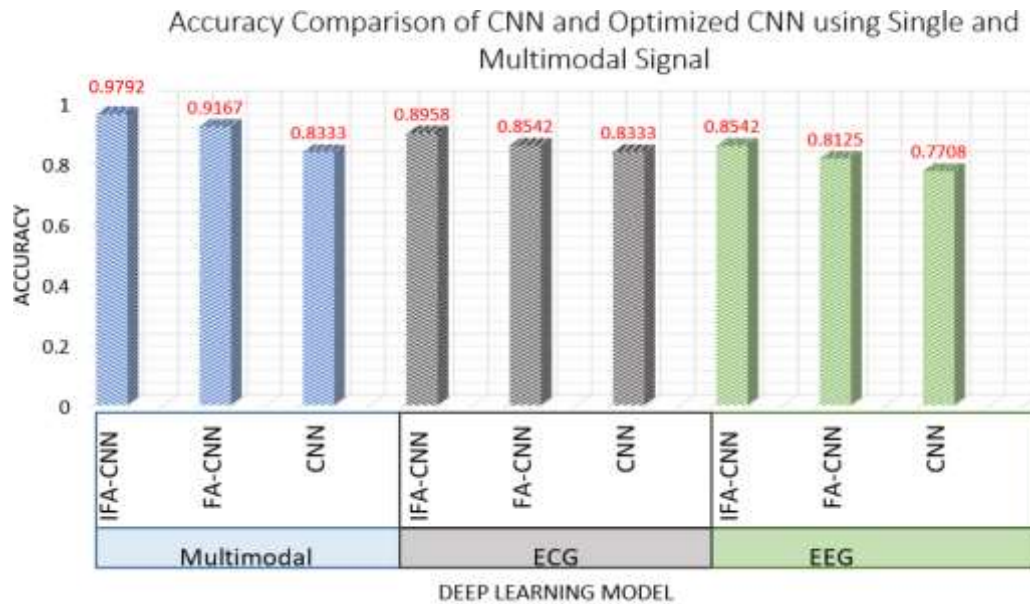


Fig. 2. Accuracy comparison of different CNN variants with single and multimodal physiological signals for stress prediction.

V. CONCLUSION

The research focuses on designing a highly accurate stress-level prediction framework. To achieve this, multimodal ECG and EEG data from Kaggle were acquired and processed. The processing techniques included noise removal and augmentation. The processed signals were then fed into three models: a CNN model without optimization, an optimized CNN using the Firefly Algorithm (FA-CNN), and an Improved Firefly Algorithm (IFA-CNN) for feature extraction and stress-level classification. Three experiments were conducted. In the first experiment, multimodal data were given to the DL models for stress prediction. In the second and third experiments, ECG and EEG signals were used separately. Standard positive and negative metrics were employed to analyze the performance of the models. For multimodal data, the IFA-CNN achieved the highest accuracy of 0.9792, while the FA-CNN and CNN yielded accuracies of 0.9167 and 0.8333, respectively. Similarly, for ECG and EEG data, the IFA-CNN outperformed the other models, achieving accuracies of 0.8958 and 0.8542, respectively. The results infer two key points: first, multimodal signals improve stress prediction accuracy, and second, the IFA-CNN performs better than the FA-CNN and the CNN without optimization techniques. The proposed model's outcomes for stress-level prediction demonstrate a promising future for medical applications.

The limitations of the research include the exclusive use of Kaggle data, which raises concerns about the model's generalizability. To address this, future work will involve testing the model on other standard datasets or real-time data. Additionally, the current analysis focuses only on accuracy, which is insufficient for real-time deployment. To determine the model's suitability for real-time applications, future research will evaluate other parameters, such as complexity and processing time. These aspects will be incorporated into future studies.

REFERENCES

- [1] De Kloet, E. Ron, Marian Joëls, and Florian Holsboer. "Stress and the brain: from adaptation to disease." *Nature reviews neuroscience* 6, no. 6 (2005): 463-475.
- [2] Michie, Susan. "Causes and management of stress at work." *Occupational and environmental medicine* 59, no. 1 (2002): 67-72.
- [3] Ovsianikova, Yanina, Diana Pokhilko, Valentyn Kerdyvar, Mykola Krasnokutsky, and Olexii Kosolapov. "Peculiarities of the impact of stress on physical and psychological health." *Multidisciplinary Science Journal* 6 (2024).
- [4] World Health Organization. "The World Health Report 2001: Mental health: new understanding, new hope." (2001).
- [5] Mikhno, Inesa, Viktor Koval, and Anton Ternavskiy. "Strategic management of healthcare institution development of the national medical services market." *Access: Access to Science, Business, Innovation in Digital Economy* 1, no. 2 (2020): 157-170.
- [6] Kohn, Robert, Shekhar Saxena, Itzhak Levav, and Benedetto Saraceno. "The treatment gap in mental health care." *Bulletin of the World health Organization* 82, no. 11 (2004): 858-866.
- [7] Soroush, Morteza Zangeneh, and Yong Zeng. "Physiological Approaches to Students' Stress Detection." *Proceedings of the Canadian Engineering Education Association (CEEAA)* (2023).
- [8] Benchekroun, Mouna. "Continuous stress detection from physiological signals." PhD diss., Université de Technologie de Compiègne, 2024.
- [9] Goyal, Aishwarya, Shailendra Singh, Dharam Vir, and Dwarka Pershad. "Automation of stress recognition using subjective or objective measures." *Psychological Studies* 61 (2016): 348-364.
- [10] Lin, Qiang, Tongtong Li, P. Mohamed Shakeel, and R. Dinesh Jackson Samuel. "Advanced artificial intelligence in heart rate and blood pressure monitoring for stress management." *Journal of Ambient Intelligence and Humanized Computing* 12 (2021): 3329-3340.
- [11] Bhanja, Nilankar, Sanjib Kumar Dhara, and Prabodh Khampariya. "A design of machine learning-based adaptive signal processing strategy for ECG signal analysis." *Multimedia Tools and Applications* (2024): 1-17.
- [12] Bhanja, Nilankar, Sanjib Kumar Dhara, and Prabodh Khampariya. "AN efficient deep learning with an optimization framework to analyse the eeg signals." *Multimedia Tools and Applications* (2024): 1-22.
- [13] Roy, Bishwajit, Lokesh Malviya, Radhikesh Kumar, Sandip Mal, Amrendra Kumar, Tanmay Bhowmik, and Jong Wan Hu. "Hybrid deep learning approach for stress detection using decomposed eeg signals." *Diagnostics* 13, no. 11 (2023): 1936.
- [14] Mane, Swaymprabha Alias Megha, and Arundhati Shinde. "StressNet: Hybrid model of LSTM and CNN for stress detection from electroencephalogram signal (EEG)." *Results in Control and Optimization* 11 (2023): 100231.
- [15] Jahanjoo, Anice, Nima TaheriNejad, and Amin Aminifar. "High-Accuracy Stress Detection Using Wrist-Worn PPG Sensors." In *2024 IEEE International Symposium on Circuits and Systems (ISCAS)*. IEEE, 2024.
- [16] Mane, Swaymprabha Alias Megha, and Arundhati Shinde. "StressNet: Hybrid model of LSTM and CNN for stress detection from electroencephalogram signal (EEG)." *Results in Control and Optimization* 11 (2023): 100231.
- [17] Gonzalez-Vazquez, Joaquin J., Lluís Bernat, Jose L. Ramon, Vicente Morell, and Andres Ubeda. "A Deep Learning Approach to Estimate Multi-Level Mental Stress from EEG using Serious Games." *IEEE Journal of Biomedical and Health Informatics* (2024).
- [18] Kumar, Sanjay, Anshuman Raj Chauhan, Akshi Kumar, and Guang Yang. "Resp-BoostNet: Mental Stress Detection from Biomarkers Measurable by Smartwatches using Boosting Neural Network Technique." *IEEE Access* (2024).
- [19] Zhang, Pengfei, Fenghua Li, Rongjian Zhao, Ruishi Zhou, Lidong Du, Zhan Zhao, Xianxiang Chen, and Zhen Fang. "Real-time psychological stress detection according to ECG using deep learning." *Applied Sciences* 11, no. 9 (2021): 3838.
- [20] Zhang, Pengfei, Fenghua Li, Lidong Du, Rongjian Zhao, Xianxiang Chen, Ting Yang, and Zhen Fang. "Psychological stress detection according to ECG using a deep learning model with attention mechanism." *Applied Sciences* 11, no. 6 (2021): 2848.

- [21]Barki, Hika, and Wan-Young Chung. "Mental stress detection using a wearable in-ear plethysmography." *Biosensors* 13, no. 3 (2023): 397.
- [22]Fontes, Laura, Pedro Machado, Doratha Vinkemeier, Salisu Yahaya, Jordan J. Bird, and Isibor Kennedy Ihianle. "Enhancing Stress Detection: A Comprehensive Approach through rPPG Analysis and Deep Learning Techniques." *Sensors* 24, no. 4 (2024): 1096.
- [23]Li, Zewen, Fan Liu, Wenjie Yang, Shouheng Peng, and Jun Zhou. "A survey of convolutional neural networks: analysis, applications, and prospects." *IEEE transactions on neural networks and learning systems* 33, no. 12 (2021): 6999-7019.
- [24]Zhao, Yucheng, Guangting Wang, Chuanxin Tang, Chong Luo, Wenjun Zeng, and Zheng-Jun Zha. "A battle of network structures: An empirical study of cnn, transformer, and mlp." *arXiv preprint arXiv:2108.13002* (2021).
- [25]Jogin, Manjunath, M. S. Madhulika, G. D. Divya, R. K. Meghana, and S. Apoorva. "Feature extraction using convolution neural networks (CNN) and deep learning." In *2018 3rd IEEE international conference on recent trends in electronics, information & communication technology (RTEICT)*, pp. 2319-2323. IEEE, 2018.
- [26]Mairal, Julien, Piotr Koniusz, Zaid Harchaoui, and Cordelia Schmid. "Convolutional kernel networks." *Advances in neural information processing systems* 27 (2014).
- [27]Zhou, XueFei. "Understanding the convolutional neural networks with gradient descent and backpropagation." In *Journal of Physics: Conference Series*, vol. 1004, p. 012028. IOP Publishing, 2018.
- [28]Nirthika, Rajendran, Siyamalan Manivannan, Amirthalingam Ramanan, and Ruixuan Wang. "Pooling in convolutional neural networks for medical image analysis: a survey and an empirical study." *Neural Computing and Applications* 34, no. 7 (2022): 5321-5347.
- [29]Santos, Claudio Filipi Gonçalves Dos, and João Paulo Papa. "Avoiding overfitting: A survey on regularization methods for convolutional neural networks." *ACM Computing Surveys (CSUR)* 54, no. 10s (2022): 1-25.
- [30]Lederer, Johannes. "Activation functions in artificial neural networks: A systematic overview." *arXiv preprint arXiv:2101.09957* (2021).
- [31]Xie, Xiaoyun, Ming Xie, Ata Jahangir Moshayedi, and Mohammad Hadi Noori Skandari. "A hybrid improved neural networks algorithm based on L2 and dropout regularization." *Mathematical Problems in Engineering* 2022, no. 1 (2022): 8220453.
- [32]Mao, Anqi, Mehryar Mohri, and Yutao Zhong. "Cross-entropy loss functions: Theoretical analysis and applications." In *International conference on Machine learning*, pp. 23803-23828. PMLR, 2023.
- [33]Tuba, Eva, Nebojša Bačanin, Ivana Strumberger, and Milan Tuba. "Convolutional neural networks hyperparameters tuning." In *Artificial intelligence: theory and applications*, pp. 65-84. Cham: Springer International Publishing, 2021.
- [34]Yang, Xin-She. *Nature-inspired metaheuristic algorithms*. Luniver press, 2010.
- [35]Yang, Xin-She, and Xingshi He. "Firefly algorithm: recent advances and applications." *International journal of swarm intelligence* 1, no. 1 (2013): 36-50.
- [36]Strumberger, Ivana, Nebojsa Bacanin, and Milan Tuba. "Enhanced firefly algorithm for constrained numerical optimization." In *2017 IEEE congress on evolutionary computation (CEC)*, pp. 2120-2127. IEEE, 2017.
- [37]Wang, Jiquan, Mingxin Zhang, Haohao Song, Zhiwen Cheng, Tiezhu Chang, Yusheng Bi, and Kexin Sun. "Improvement and application of hybrid firefly algorithm." *IEEE Access* 7 (2019): 165458-165477.
- [38]ECG&EEG Stress Features from Kaggle website. Available at <https://www.kaggle.com/datasets/apithm/ecg-and-eeg-stress-features>. Access on January 2025.