Received: February 28, 2024. Revised: April 21, 2024. 

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

# A New Hybrid Adaptive Optimization Algorithm Based Wavelet Neural Network for Severity Level Prediction for Lung Cancer Dataset

**T. Divya[1]* J. Viji Gripsy[1]**

[1] *PSGR Krishnammal College for Women, Coimbatore, India*
* Corresponding author's Email: divyathirumurthi@gmail.com

**Abstract:** This study proposes three contributions focused on lung cancer detection and severity level identification. The absence of non-invasive technologies for predicting lung cancer necessitates faster, more efficient, and more accurate classification procedures due to the absence of non-invasive technologies for predicting lung cancer. Creating an automated and intelligent prediction system is crucial for identifying phases and predicting the possibility of a recurrence. The objective is to create an automated detection system for identifying lung cancer using an optimization-focused deep learning model. We develop an adaptive multi-swarm PSO and combine it with the firefly algorithm to determine the ideal weight values for the Wavelet Neural Network (WNN) model. We use the HAPSO-FFA-WNN method to explore problems with multiple optimal solutions. This study evaluated two lung cancer datasets, and the proposed HAPSO-FFA-WNN model achieved 97.58% accuracy for dataset 1 and 98.54% accuracy for dataset 2. Furthermore, the proposed model achieved better precision, recall, and MCC performance metrics.

**Keywords:** Risk analysis, Optimization, Lung cancer, Prediction, Wavelet neural network.

## 1. Introduction

Certain behavioural features in individuals increase the risk factors for developing lung cancer. Smoking is one of those habits. However, smoking is the primary cause of 90% of lung cancer cases. Smoking causes lung cancer, which develops over the period between smoking and the onset of the disease [1]. Cigarette smoking became more prevalent among men and women in the US and Europe throughout the world wars. The provision of complimentary cigarettes to soldiers led to the development of a nicotine addiction, which they continued to indulge in after the conflict. At that time, there was little understanding regarding the harmful consequences of smoking and the addictive nature of nicotine. Additionally, several health professionals developed the habit of smoking without being aware of its detrimental implications.

Overall survival (OS) is the duration from a random starting point to either the date of death or the most recent date the individual was still living. Researchers conduct exploratory studies to distinguish the relationship between age and CIRS-G based on fundamental characteristics and outcomes, including both PS and age [2]. Various parameters, such as age (<70 years vs. ≥70 years), severity stages 3 and 4 comorbidities, and mortality rates at severity stages 3 and 4, are analyzed in both multivariate and univariate analyses to assess treatment outcomes [3]. Quality of life (QoL) is a complex attribute including the physical state, psychological well-being, cognitive abilities, and factors related to sickness and treatment as perceived by the patient. People with lung cancer often view quality of life as a critical consequence, given their poor prognosis. Global quality of life is an important prognostic parameter for estimating lung cancer patients' survival probabilities. Patients with NSCLC cancer often struggle with anxiety and depression.

### 1.1 Problem statement

There are numerous challenges to predicting the severity of lung cancer that must be overcome in order to enhance the accuracy and reliability of predictive models. A major challenge lies in the heterogeneity of lung cancer, which comprises an extensive range of molecular profiles and histological subtypes. Predicting severity levels can be complicated by this heterogeneity, which introduces various patterns of disease progression and treatment response. The absence of exhaustive and standardized datasets presents a significant challenge as an additional difficulty. The collection of more valuable data, which comprises comprehensive molecular profiles, imaging studies, and clinical details, is critical in the construction of strong predictive models. Furthermore, the ever-changing course of lung cancer poses a challenge because the condition may develop gradually, necessitating regular monitoring and modification of prediction models. An additional issue that arises is the existence of confounding variables and comorbidities, including but not limited to age, smoking history, and lung health. These elements have the potential to impact the severity of the disease and add complexity to the predictive modeling process. In addition, the interpretation of complicated imaging studies, including PET and CT scans, necessitates specialized knowledge and may induce inconsistency in the evaluation of severity levels. Ultimately, it is imperative to carefully navigate the ethical considerations that pertain to predictive modeling in the healthcare sector. These concerns encompass patient privacy, informed consent, and the possibility of biases in algorithmic decision-making. It is imperative to confront these challenges in order to progress the prediction of severity levels in lung cancer and, consequently, enhance patient treatment and outcomes.

## 1.2 Research contribution

- A proposed hybrid adaptive particle swarm optimizer for the WNN model• The firefly technique finds the ideal weight values in a WNN model.
- Multi-swarm PSO has superior exploration performance because of scheduled regrouping.
- Due to intrinsic attraction and light intensity characteristics, the Firefly algorithm outperforms in exploitation.
- The adaptive HAPSO-FFA approach combines adaptive HAPSO and FFA methods for weight adjustment.
- Using wavelet functions as activation functions in the WNN model to produce more accurate outputs with a lower mean square error.
- We classify lung data related to lung cancer using a novel adaptive model.

Table 1. Previous Studies for Predicting Diseases using Optimization Algorithms

| Authors & Year | Dataset | Methods | Accuracy |
|---|---|---|---|
| MahaLakshmi et al. (2023) [19] | UCI Cleveland datasets for heart disease | Improved PSO | Achieved 98.41% and 97.40% accuracy. |
| Raghavendra et al. (2023) [20] | UCI ML Heart Disease dataset | DL Modified NN (DLMNN) model combined with the Pet Dog-Smell Sensing (PD-SS) algorithm | Achieved 94.21% accuracy. |
| Rajalakshmi et al. (2023) [21] | UCI heart disease dataset. | Ischemic Heart Disease Squirrel Search Optimization (IHDSSO) model with a RF classifier | Achieved 98.38% accuracy. |
| Venkatesh et al. (2022) [22] | Lung cancer in CT images | PSO and genetic algorithms | Achieved 96.97% accuracy |
| Mohamed et al. (2023) [23] | Lung cancer dataset | Ebola optimization search algorithm (EOSA) based CNN EOSA-CNN hybrid model | Achieved 0.9321 accuracy |
| Kumar et al. (2022) [24] | Lung cancer dataset | Improved Novel Genetic algorithm(GA) algorithm | Achieved 93.87% accuracy |
| Ramana et al. (2022) [25] | LUNA-16 and LIDC Lung Image datasets | Ahybrid RNN and feed-forward BPNN (Hyb-RNN-FFBPNN) optimized with the glow worm swarm algorithm (GWSA) | Achieved 94.21% accuracy |
| Swathi et al. (2023) [26] | Micro array gene dataset | Improved SALP Swarm Optimization and LSTM | Achieved 99.54% accuracy |

Received: February 28, 2024. Revised: April 21, 2024. 

- The hybrid HAPSO-FFA method addresses the early convergence problem, improving classification performance.

The remaining sections are structured as follows: previous works and the drawbacks are discussed in section 2. The proposed methodology with mathematical formulations is given in section 3. The experimental results are discussed with various parameters and the risk factors are identified in section 4. Finally, the conclusion is given in section 5.

## 2. Literature review

Lai et al. (2020) [4] developed a DNN by combining organic phenomena and clinical information from various sources. The natural science of the system was used to identify important predictive biomarkers. We leveraged the system's inherent knowledge. This technique demonstrated the highest danger magnitude ratio in comparison to both the training and testing sets. The drawback is this method did not improve categorization accuracy.

Kim et al. (2020) [5] established a DLPM. Patients diagnosed with cancer used the findings as a tool for surgical risk classification. With this approach, they were able to get repeatable analytical data. The drawback is the verification set had a relatively small number of cancer patients, resulting in a reduction in the study's statistical power.

Liao et al. (2020) [6] proposed a survival prediction model based on lymph node ratio. This model relies on many studies that suggest the positive lymph node ratio is a critical component for survival analysis. The investigation used data from the SEER database, which spans from 2010 to 2016. We acquired the training data using the X-tile software. Variable collinearity was detected by the use of Pearson's correlation coefficient, variance inflation factor, and tolerance levels. Cox regression analysis was used to determine the importance of the characteristics. Nomograms facilitate visualization. The predictive accuracy of the study is evaluated by the concordance index and decision curve analysis, which determine the study's predictive accuracy. The authors did not to classify tumor information.

Supriya et al. (2020) [7] proposed a technique in their study that could potentially identify and segment pulmonary NROI in CT lung images. The classification approach automatically detects and classifies the distinguishing features of benign and malignant nodules. This investigation was applied minimum number of samples.

As part of their investigation into lung cancer and pneumonia, Bhandary et al. (2020) [8] developed a DL model. For the purpose of assessing the described problem, there are two different deep learning methodologies for assessing the described problem, which are as follows: We used the first deep learning method, known as modified AlexNet (MAN), to classify chest X-ray images into two categories: normal and pneumonia. This work has a lack of risk factors identification.

Pradeep and Naveen (2020) [9] estimated the survival rate of patients with lung cancer by using data from EHR. Forecasting the patient's life time in months is accomplished by the use of an optimized DNN regression model. The authors identified only patients'survival rate not focused on severity level.

Ghosal et al. (2020) [10] proposed using a GAN as a data augmentation strategy in order to provide more training data for the purpose of improving CAD systems. We developed a convolutional autoencoder deep learning system to facilitate unsupervised learning of image attributes for lung nodules using unlabeled data. The drawback is therapy planning, illness detection, and the supervision of surgical therapy are all areas in which medical image processing is critical.

Ramroach et al. (2020) [11] proposed an approach that included image pre-processing methods such as median filtering for noise reduction, the MEM algorithm for image segmentation, and feature extraction to identify the area of interest of lung nodules. Lung cancer detection employs the SVM classification. They obtained minimum accuracy for the prediction.

Kar et al. (2020) [12] created a hybrid classifier model by combining GrIHS with an adaptive KNN classifier framework. The optimization approach involves two stages: generating optimum weights and estimating the ideal feature subset. This process enhances the classification accuracy significantly. The radiologists rendered their diagnosis through manual interpretation of the lung computed tomography images.

AdaBoost and genetic algorithms were used in the hybrid approach that Lu et al. (2021) [13] presented for the purpose of classifying a variety of malignancies from the UCI collection. These malignancies included breast, lung, colon, leukaemia, and brain tumours. Within the scope of this research, the concept of a decision group and a number of classifiers that are chosen at random and carried out a certain number of times in order to answer a particular problem are presented. The decision group consists of k-NN, DT, and NB, with a majority vote determining the answer. AdaBoost employs decision groups as its base classifier, with a genetic algorithm used to optimize the weights of each decision group.

The accuracy function of each choice group classifier was used to assess its fitness. Elitism is used as a classifier, along with bit-string mutation and random single-point crossover. The AdaBoost-GA methodology was compared with several ensemble methods, such as the RF and AdaBoost algorithms. It had better performance on only small samples.

Using Apache Spark designs and architecture combined with machine learning (ML) strategies, Sujitha (2021) [14] proposed a revolutionary approach to lung cancer classification. With an accuracy of 86% and an AUC of 0.88, their method fared better than others. The proposed method achieved low accuracy.

Alrahhal and Alqhtani (2021) suggested an ALCD approach using a CNN model [15]. The authors used preprocessing procedures to prepare the images for further examination. The authors used the SIFT method to select features for extraction. Their results showed that the recommended strategy achieved ratings of 96% for accuracy, 98% for sensitivity, 92% for recall, and 96% for precision. The limitations encompass the possibility of erroneous positive results, reduced sensitivity towards particular types of tumors, difficulties in accurately identifying cancerous cells, and the need for further developments to enhance precision.

Rubin et al. (2023) [16] developed and validated prediction models with the goal of assisting in the identification of individuals who are at risk of acquiring lung cancer. There is a lot of potential for register-based prediction algorithms to assist medical professionals and healthcare planners in identifying individuals who are at risk of developing lung cancer. Detailed information regarding symptom presentation, familial cancer history, and lifestyle factors such as smoking status was missing.

By fine-tuning the hyperparameters, Saghir Ahmed et al. (2023) [17] used the ensemble XGBoost and ResNet101 algorithms. Additionally, the authors were able to compare and contrast the results with traditional ML methods. The experiment results showed the recommended ensemble extreme boosting (XGBoost) model enhanced prediction performance due to its robust and expanded functionality. The results suggest that the recommended approach could significantly aid in the early detection and treatment of lung cancer, leading to a decrease in death rates and an increase in survival rates. authors did not focus on disease severity.

An innovative classifier was developed by Hussain et al. (2023) [18]. They developed an innovative classifier using a comprehensive and complete CNN. The ML-CNN is a classifier that has a broad variety of applications, including the identification and organization of biological pictures. In this work, ML-CNN is used to identify and categorize lung nodules that appear in CT scan pictures. ML CNN using PSO for the purpose of extracting features and categorizing data. Improvements in performance were accomplished by the use of the PSO approach for optimizing hyperparameters, resulting in performance improvements. The ML-CNN with PSO reaches values of 98.45 for accuracy, 98.89 for precision, 98.45 for sensitivity, 98.62 for specificity, and 98.85 for F-measure. With regard to accuracy and convergence, the hybrid methodology displays higher performance in comparison to other techniques. They struggled to diagnose lung cancer nodule because of unclear features.

Table 2. Notation List

| Notation | Description |
|---|---|
| $v_i$ | velocity of particle i |
| $p_i$ | position of particle i |
| $p_{ibest}$ | the position with the 'best' fitness value by particle i |
| $g_{ibest}$ | the best fitness value obtained |
| $R_1$ and $R_2$ | the random variables in the range [0, 1] |
| $c_1$and $c_2$ | learning factors |
| I(r) | light intensity varies with respect to the distance 'r' |
| $I_0$ | the initial light intensity |
| Γ | light absorption coefficient |
| B | attractiveness of a firefly |
| R | distance between two fireflies |
| $\beta_0$ | attractiveness at r = 0 |
| $x_{i,k}$ | kth component of the longitudinal coordinate of $x_i$ the $i^{th}$ firefly |
| N | the number of dimensions |
| $\varepsilon_i$ | Gaussian distribution function |
| Γ | Factorial |
| 'α' and 'β' | random values in the range [0, 1] |
| Λ | weighting factor |
| W | the worst particles |
| $p_{std\_dev}$ | standard deviation values of the entire swarm |
| Ω | inertia coefficient |
| $ss_{initial}$ | the number of initial sub-swarms |
| $\lambda_{decrease}$ | decreasing rate control parameter |
| λ' and 't' | the dilation and translation parameters |
| $y_{out\_net}$ | the output value of *m*-th node in output layer |
| $w_t$ | weight between *i*-th hidden unit and *m*-th output node |
| $b_0$ | bias entity |

## 3. Methodology

PSO and the firefly optimisation algorithm (FFA) both have unique strengths and weaknesses when considered separately. Each approach leverages a swarm's inherent behaviours of a swarm and consists of specific update equations: PSO involves location and velocity update equations, whereas FFA involves light intensity and attractiveness. Hybrid optimisation algorithms are created to combine the strengths of several algorithms and enhance their ability to explore and exploit them efficiently. This study introduces a novel hybrid approach combining PSO and FFA to address issues related to local and global minima. The algorithm incorporates Levy flight random walk to enhance its performance.

### 3.1 PSO

PSO is a prominent swarm intelligence method that draws inspiration from the collective behaviour of bird flocks, fish schools, and fly movements. The swarm intelligence algorithm begins by initialising a population with random solutions. The search for optimum solutions is then conducted via generations (iterations) throughout the process flow. In the PSO method, each solution is referred to as a particle. Initially, a random number of particles are produced to undergo the iterative process. Every randomly produced particle may go through the multi-dimensional search space to get the best solution. The position and velocity of the particle are updated in accordance with its movement as it traverses the search space. The travelling path of the particle is considered while updating its location and velocity. This update process propels the particle towards higher fitness values, causing all surrounding particles to also move towards the solution point. During each generation of PSO, particles update their positions based on two values: their personal best solution achieved so far and the best solution achieved by any particle in the generation. PSO operates by updating location and velocity equations, as follows:

$$v_i = v_i + c_1 R_1 (P_{ibest} - p_i) + c_2 R_2 (g_{ibest} - p_i) \tag{1}$$

$$p_i = p_i + v_i \tag{2}$$

### 3.2 Firefly optimization algorithm

The firefly algorithm (FFA) was developed based on fireflies' flashing features. FFA should focus on addressing the variance in light intensity and establishing the attraction component. The goal function's brightness or light intensity determines the attraction component.

$$I(r) = I_0 e^{-\gamma r^2} \tag{3}$$

$$\beta(r) = \beta_0 e^{-\gamma r^2} \tag{4}$$

$$r_j = \sqrt{\sum_{k=1}^{n} \left(x_{i,j} - x_{j,k}\right)^2} \tag{5}$$

$$x_i(t+1) = x_i(t) + \beta_0 e^{-\gamma r^2} \left(x_j - x_i\right) + +\rho \varepsilon_i \tag{6}$$

This paper introduces a unique hybrid method, combining PSO-FFA for optimising functions in the issue area. Each approach, PSO and FFA, has its own benefits and limits. New algorithms are always being created to enhance the exploration and exploitation mechanisms in the search process. PSO is computationally efficient and exhibits strong social and cognitive behaviour, resulting in a successful exploration process. The Firefly algorithm's attractiveness is determined by the brightness coefficient, enabling effective exploitation. These algorithms are combined to create a new hybrid technique called PSO-FFA. Levy flight is incorporated to manage randomness in the algorithm. The Levy flight process manages unpredictability and enables the suggested method to reach the best solution. The levy flight is included in the hybrid PSO-FFA.

$$F_{levy\_f} = \left(\frac{\alpha \times \sqrt{\rho}}{|\beta|^{(1/\lambda)}}\right) 10^{-2} \tag{7}$$

$$with\ \rho = \left(\frac{\Gamma(\lambda+1).sin\left(\frac{\pi\lambda}{2}\right)}{\Gamma\left(\frac{\lambda+1}{2}.\lambda.2^{\left(\frac{\lambda-1}{2}\right)}\right)}\right)^{1/\lambda} \tag{8}$$

Eq. (8) uses 'Γ', with 'α' and 'β' being random values between 0 and 1. The weighting factor 'λ' guides the random walk throughout its flight, with a range of 0 to 3. Simulations were conducted using various 'λ' values, revealing that low values of 'λ' impact exploitation and high values of 'λ' impact exploration. This study contribution considers the value of 'λ' as 1.5 based on experimental experiments.

### 3.3 Hybrid PSO-FFA

Researchers use the fundamental PSO when a problem has a single solution. The study introduces an adaptive multi-swarm PSO technique and combines it with the Firefly technique to find the best weight values for the WNN model. The Hybrid Adaptive PSO (HAPSO) is used to explore scenarios with numerous optimum solutions. A novel approach called HAPSO-FFA is introduced by combining adaptive HAPSO with FFA.One of the first HAPSO optimizers One variant of PSO is known as multi-swarm optimization (MSO), which differs from PSO in that it employs many sub-swarms (SS) rather than a single swarm. In MSO, each sub-swarm concentrates on a particular region, but in diversification strategies, the placement and timing of introducing the SS are determined by a well-designed diversification strategy. A method known as multi-swarm PSO is used by this algorithm in order to deal with situations that consist of several possible optimum solutions. During the process of this research, a HAPSO was developed in order to get the most optimal parameter configuration for the wavelet learning neural model. In a typical PSO, the entities that are being searched for are referred to as particles. These particles move across the solution space in order to find solutions. The velocity and location are the two properties that are associated with each particle in the classic PSO algorithm. The following equations are provided for the purpose of updating the velocity and position as follows:

$$v_i(t+1) = \omega * v_i(t) + c_1 R_1 \big(P_{ibest} - p_i(t)\big) + c_2 R_2 (g_{ibest} - p_i(t)) \tag{9}$$

$$p_i(t+1) = p_i(t) + v_i(t) \tag{10}$$

PSO allows for stochastic searching using random variables. Learning variables c1 and c2 are required to achieve a balance between the disadvantages of exploration and exploitation. Before each update operation, the current velocity must be checked and confirmed to be within a certain limit to prevent erratic movements in the search space.

Many research studies have been conducted to enhance the variety of PSO's capabilities. The global optimum location in the conventional PSO algorithm consistently attracts all particles until they are modified. When the global best position is stuck in a local minimum, the whole swarm will struggle to escape from the local minimum. This causes the optimisation process to stall and leads to premature convergence. This work addresses this limitation by developing a unique adaptive multi-swarm PSO that partitions the whole swarm into a large number of SS. Before the search process continues, the particles of each SS are compared, and the location that is the worst is replaced with the one that is the best before the search process is continued. The worst individuals learn from the greatest, while the best does not actively engage in any actions throughout the search process. The quantity of optimal sites matches the number of SS, and this quantity adjusts according to the search procedure. Initially, a large number of SS help increase the variety of particles and enable them to explore the search space. During the optimization algorithm's execution, the number of SS decreases as the exploitation mechanism takes effect.

Additionally, in every generation, there are a large number of example particles that indicate the best position when the worst position is updated. The entire swarm is not affected by the presence of a limited number of example particles that are located at certain excellent local locations does not affect the entire swarm. This enhances the diversity factor for the whole swarm. In this work, the global best solutions are prioritised above the current best solutions, unlike in the traditional PSO algorithm. The sub-swarm travels towards the best global solutions. The purpose is to calculate the standard deviation of the particle values, since they are more likely to be superior than the typical global best particle values. At this point, it has been decided to calculate the standard deviation of the position of the whole swarm's position and then attract all of the remaining particles in accordance with that estimate. All of the equations have been revised and changed properly.

$$v_{w_{-i}}(t+1) = \omega * v_{w_{-i}}(t) + c_1 R_1 \Big(P_{ibest} - p_{w_{-i}}(t)\Big) + c_2 R_2 (g_{ibest} - p_{w_{-i}}(t)) \tag{11}$$

$$p_{w_{-i}}(t+1) = p_{w_{-i}}(t) + v_{w_{-i}}(t) \tag{12}$$

The algorithm is very effective at exploring, and it makes use of an adaptive way to concentrate on exploitation. This is accomplished by establishing the number of SS using a divide and conquer strategy. There is a variation in the quantity of SS, which may range from a large amount to a small amount. Through the use of the following equation, the number of SS is dynamically determined using the following equation:

$$\delta_{coeff} = \frac{1}{2}\left(ss_{initial} - 0.5 ss_{initial}\left(1 - \left(-\frac{Fitness\_Functions}{Max\_Iteration \times \lambda_{decrease}}\right)\right)\right) \tag{13}$$

Received: February 28, 2024. Revised: April 21, 2024. 

The term '$ss_{initial}$' is used to indicate the number of initial sub-swarms, whereas the term '$\lambda_{decrease}$' is used to express the control parameter representing the rate of decline. During the first phases of optimization, the adaptive HAPSO model possesses powerful capabilities for exploration and exploitation, in addition to a rapid convergence speed. The convergence rate and accuracy in reaching optimum solutions, on the other hand, tend to decrease as the optimization process nears a close. This shortcoming is addressed by introducing a unique strategy known as HAPSO-FFA, which combines an adaptive multi-swarm particle swarm optimizer with the conventional firefly method. The goal of this approach is to give solutions that are nearly optimal. The benefits of a number of meta-heuristic algorithms are integrated via the use of this hybridization method, resulting in the production of superior responses. On standard functions, the newly developed hybrid adaptive HAPSO-FFA algorithm was put through testing. The efficiency of the swarm-based optimization method was shown by the observed results, which indicated a convergence rate that was quite near 1.

The methods in this study either do not have sufficient parameter adjustment during the optimization process or they impose a considerable computing cost. This conclusion is based on the current research perspectives and works in neural architecture models and swarm intelligence algorithms. This part combines the adaptive HAPSO-FFA with the WNN model to overcome the identified constraints.

### 3.4 WNN model

A WNN combines the principles of wavelets with neural networks. A WNN is a type of feedforward neural network that consists of a single hidden layer. In a WNN, the activations are based on orthonormal wavelet functions. The architectural architecture of a WNN resembles that of a basic neural network model, with inputs, a hidden layer, and an output layer consisting of linear combiners. The hidden and output layers are composed of neurons that use wavelets-based activation functions based on wavelets. These neurons are commonly referred to as "wavelons'."The translation and dilation coefficients of the wavelet models, together with the weights, are adjusted based on the learning technique used in the WNN models.

The wavelet function employed in this work is defined by

$$\psi_{\lambda,t}(u) = \psi\left(\frac{u-t}{\lambda}\right) \tag{14}$$

The parameters 'λ' and 't'represent the dilation and translation parameters, respectively. The wavelet function is the activation function used to calculate the output of the WNN model's output. The WNN uses a training procedure that incorporates a stochastic gradient descent learning rule to determine the network's output. Error is assessed by comparing the estimated output to the intended objective, and the learning process is adjusted to minimize the error value.

**Algorithm 1: WNN algorithm**

**Step 1:** Initialize the weights with tiny random values and provide the learning rate, momentum factor, translational factor, and dilation factor.

**Step 2:** Perform steps 3-9 as long as the termination condition is not true.

**Step 3:** For each input data samples presented to WNN model do the steps 4-10

**Step 4:** The input signal gets transmitted from the input layer to the hidden layer.

**Step 5:** Evaluate the net input of the respective layers

$$z_{in} = b_0 + \sum_{\substack{i=1 \\ j=1}}^{\substack{m \\ n}} x_i w_{ij} \tag{15}$$

**Step 6:** Utilize the activation function to determine the output.

$$\Psi_{\lambda,t} z_{in} = \Psi\left(\frac{z_{in}-t}{\lambda}\right) \tag{16}$$

**Step 7:** Examine the network model's output

$$v_i(x_i) = x_{ji} exp\left[\frac{-x_{ji}^2}{2}\right] \tag{17}$$

$$y_{out_{net}} = \sum_{t=1}^{n} w_t \phi[D_t R_t(x - t_t)] + b_0 \tag{18}$$

Where, $y_{out_net}$represents the output value of *m*-th node in outputlayer for the *n*-th incoming pattern; $w_t$ indicates the weight between *i*-th hiddenunit and *m*-th output node; $b_0$ is the bias entity.

**Step 8:** Calculate the error factor

**Step 9:** Employ stochastic gradient learning rule to perform update between the respective layers

$$w_{ij}(new) = w_{ij}(old) + \alpha\delta_j x_i \left(\delta_j = \delta_{in-j} f'(z_{inj})\right) \tag{19}$$

$$v_{jk}(new) = v_{jk}(old) + \alpha\delta_k z_j(\delta_k = target_k - y_{out_{net}})f'(y_{ink})) \quad (20)$$

Test for terminating condition. The terminating condition is the error attaining the least possible value.

### 3.5 New HAPSO-FFA based WNN model (HAPSO-FFA-WNN)

This study optimizes the WNN model's weight values using the newly developed adaptive HAPSO-FFA with levy flying. To enable faster convergence of the WNN with better solutions and metrics, the ideal weight values are chosen using adaptive HAPSO-FFA, as opposed to initializing random weight values and carrying out the training procedure. There will either be premature or delayed convergence when the network model is trained after random weight initialization. As a result, the WNN model's ideal weight values are determined by using the recently created adaptive HAPSO-FFA. Algorithm 2 presents the proposed new algorithm created using HAPSO-FFA-WNN.

**Algorithm2: New HAPSO-FFA-WNN algorithm**

**Input Parameters:**

I = ($I_1$,$I_2$,$I_3$,…..$I_n$) - Input Image Data samples
n = Number of Input samples
W=($w_{11}$, $w_{12}$, …. $w_{nn}$)- Weightparameters
V=($v_{11}$,$v_{12}$, …..,$v_{nn}$) - Weightparameters
B=($b_{01}$,$b_{02}$,…$b_{0n}$)- Bias parameters
X=($x_1$, $x_2$, $x_3$,…, $x_n$) - individual fireflies
Y=minf(X) - Objective function

**Output Parameters:**

$Y_{best}$- global optimal solution of the objective function.
E - Mean square error, $y_{output}$ - Output of trained neural network

**Set Parameters:**

N - population size of fireflies/ particles
$I_0$ - original light intensity
γ - light absorption co-efficient, $\beta_0$ - attractiveness rate
n - no. of input neurons, m - no. of hidden neurons, k - no. of output neurons
$c_1$, $c_2$ - acceleration factors of PSO, α - learning rate parameter

**Step 1:** Start
**Step 2:** Invoke WNN model
**Step 3:** For i=1 to no. of input neurons
**Step 4:** For j=1 to no. of hidden neurons
**Step 5:** For k=1 to no. of output neurons
**Step 6:** Present the training datasets
**Step 7:** Determine the network's output.
**Step 8:** Evaluate error and perform weight update.
**Step 9:** End For
**Step 10:** End For
**Step 11:** End For

**Step 12:** ComputeError→ $MSE(\rho) = \frac{1}{N}\sum_{i=1}^{N} E(y_{out\_net}, y_{desired\_target})$ (21)

Weights_Final ← Min(MSE(ρ))
**Step 13:** Invoke Hybrid adaptive HAPSO-FFA with Levy flight
**Step 14:** Initial population←Weights_Final
**Step 15:** Generate the population of fireflies
**Step 16:** Apply FFA to obtain the optimal solution;
**Step 17:** Invoke Levy flight using Eqs. (7) and (8)
**Step 18:** Set light absorption coefficient γ
**Step 19:** While (Stopping condition not met do the following)
**Step 20:** For i=1:n of all fireflies
**Step 21:** For j=1:i of all fireflies
**Step 22:** Evaluate light intensity I(r) for generated fireflies using objective function F(x)
**Step 23:** If ($I_j$>$I_i$)
**Step 24:** Move firefly 'i' towards 'j' in all dimensions
**Step 25:** End If
**Step 26:** Attractiveness varies with distance 'r' through exp(-γr) update using Eq. (6).
**Step 27:** Compute new solutions
**Step 28:** Update light intensity using Eq.(3).
**Step 29:** End For
**Step 30:** End For
**Step 31:** Determine the current greatest firefly by ranking them.
**Step 32:** Stop the FFA iteration process, and return the result $Ybest_{FFA}$;
**Step 33:** End If
**Step 34:** End While

*//Utilize FFA to get the best solutions more quickly for the originally produced particles.*
**Step 35:** Set the initial particles as the solution attained from FFA;
**Step 36:** Randomly generate the Swarm and evaluate the objective function
**Step 37:** While (i ≥1, j≤N) Do

Received: February 28, 2024. Revised: April 21, 2024. 

**Step 38:** Evaluate the standard deviation position according to the whole swarm position
**Step 39:** Compute the size of sub-swarm ‘$\delta_{coeff}$’ employing Eq. (13).
**Step 40:** Randomly select number of particles - mod(no. of particles, ‘$\delta_{coeff}$’) for update of particles
**Step 41:** For each sub-swarm (‘$\delta_{coeff}$’) Do
**Step 42:** Compare particles based on their fitness
**Step 43:** Choose the local best particle with Eq. (11)
**Step 44:** Update Particles
**Step 45:** Evaluate Objective function
**Step 46:** End For
**Step 47:** IF (FFs≥Max_FFs)
**Step 48:** Output Global best particle
**Step 49:** ElseIf
**Step 50:** Goto the loop
**Step 51:** End If
**Step 52:** End While
**Step 53:** Return Global best particle
**Step 54:** Weights_Final←Global best particle
**Step 55:** Train WNN model
**Step 56:** Output MSE
**Step 57:** When the stop condition is met, terminate the procedure.
(The stopping condition is either reaching a specified number of iterations or attaining the minimum MAE.)

## 4. Results and discussion

The medical records examined in this research include the medical histories, habits, and demographic data of patients with varying degrees of lung cancer severity. In dataset 1, there are 16 attributes and 238 instances, while in dataset 2, there are 25 attributes and 22000 instances. The study’s participants are chosen based on the caliber and volume of information included in their medical records. Dataset 1 only includes 9 risk variables; however, medical professionals believe 16 of them are extremely likely relevant risk factors for lung cancer and dataset 2 has 17 important risk factors. We divide lung cancer severity into three categories: low, medium, and high, based on the patient’s stage of the illness. The hospital’s medical records included the risk variables.

Table 3 shows lung cancer severity level for dataset 1 and it contains 126 male patients and 112 female patients with 21 to 87 years old patients affected by lung cancer.

Table 4 shows lung cancer severity level for dataset 2 includes the medical data of 11089 women and 10911 men. The research included participants aged 14 to 73 years. Table 2 displays the total number of individuals in each of the four categories.

Table 3. Lung cancer severities for Dataset 1

| Lung Cancer Severity Level | No. of Patients |
|---|---|
| Low | 82 |
| Medium | 64 |
| High | 92 |
| **Total** | **238** |

Table 4. Lung cancer severities for Dataset 2

| Lung Cancer Severity Level | No. of Patients |
|---|---|
| Low | 7271 |
| Medium | 7295 |
| High | 7434 |
| **Total** | **22000** |

Table 5. Risk Factors Associated with Age and Gender for Dataset 1

| Risk Level | Age | Male | Female |
|---|---|---|---|
| High | 21-30 | 2 | 3 |
| | 31-40 | 3 | 4 |
| | 41-50 | 5 | 6 |
| | 51-60 | 9 | 7 |
| | 61-70 | 10 | 7 |
| | **71-80** | **14** | **15** |
| | 81-87 | 4 | 3 |
| Medium | 21-30 | 6 | 4 |
| | 31-40 | 1 | 0 |
| | 41-50 | 2 | 2 |
| | 51-60 | 6 | 4 |
| | 61-70 | 8 | 4 |
| | **71-80** | **8** | **9** |
| | 81-87 | 5 | 5 |
| Low | 21-30 | 4 | 3 |
| | 31-40 | 6 | 5 |
| | 41-50 | 7 | 5 |
| | 51-60 | 3 | 5 |
| | 61-70 | 7 | 8 |
| | **71-80** | **10** | **9** |
| | 81-87 | 4 | 6 |

Table 4 describes the risk factors which are associated with age and gender in dataset 1. The table expressed 126 male patients, and 112 female patients are affected by lung cancer. 47 male patients, and 45 female patients are in high-risk category. From that, above 71 to 80 age holders are highly affected by lung cancer. Moreover, 60 to 87 aged persons are generally affected in medium and low categories.

Table 6 describes the risk factors which are associated with age and gender in dataset 2. The table expressed 10911 male patients, and 11089 female patients are affected by lung cancer. 3663 male

Table 6. Risk Factors Associated with Age and Gender for Dataset 2

| Risk Level | Age | Male | Female |
|---|---|---|---|
| High | 14-20 | 442 | 484 |
| | 21-30 | 592 | 624 |
| | 30-40 | 605 | 623 |
| | 40-50 | 635 | 598 |
| | 50-60 | 589 | 626 |
| | **60-73** | **800** | **816** |
| Medium | 14-20 | 438 | 422 |
| | 21-30 | 604 | 622 |
| | 30-40 | 581 | 606 |
| | 40-50 | 596 | 631 |
| | 50-60 | 590 | 585 |
| | **60-73** | **801** | **819** |
| Low | 14-20 | 446 | 425 |
| | 21-30 | 602 | 664 |
| | 30-40 | 563 | 589 |
| | 40-50 | 613 | 578 |
| | 50-60 | 617 | 589 |
| | **60-73** | **797** | **788** |
| **Total** | | **10911** | **11089** |

patients and 3771 female patients are in high-risk category. From that, above 60 to 73 age holders are highly affected by lung cancer. Lung cancer is especially prevalent among people aged 60 to 73. Moreover, 60 to 73-year-olds are generally affected in the medium and low categories.

## 4.1 Performance metrics

The effectiveness of the suggested model is assessed using classification accuracy, precision, recall, and the Matthews correlation coefficient (MCC) described in Eqs. (22)-(25).

**Accuracy:** It represents the ratio of correctly recognized risk to the overall lung cancer. Eq. (22) is a mathematical expression for computing accuracy.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \tag{22}$$

**Precision:** The term refers to the proportion of correctly identified risk factors out of the total number of cases of lung cancer cases. The measurement is acquired by using Eq. (23).

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{23}$$

**Recall:** It precisely quantifies the proportion of recognized risk factors for lung cancer out of the total quantity. The sensitivity or recall is computed using Eq. (24), which follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{24}$$

**MCC:** It calculates the correlation coefficient between the observed and detected risk factors.

$$\text{MCC} = \frac{(\text{TP}\times\text{TN})-(\text{FP}\times\text{FN})}{\sqrt{(\text{TP+FP})(\text{TP+FN})(\text{TN+FP})(\text{TN+FN})}} \tag{25}$$

Table 7. Accuracy for Risk Level Prediction

| Algorithm | Dataset 1 | Dataset 2 |
|---|---|---|
| PSO | 92.24 | 93.21 |
| FFA | 92.58 | 93.15 |
| PSO-FFA | 95.56 | 95.78 |
| IDL-EGWO | 96.26 | 95.68 |
| **HAPSO-FFA-WNN** | **97.58** | **98.45** |

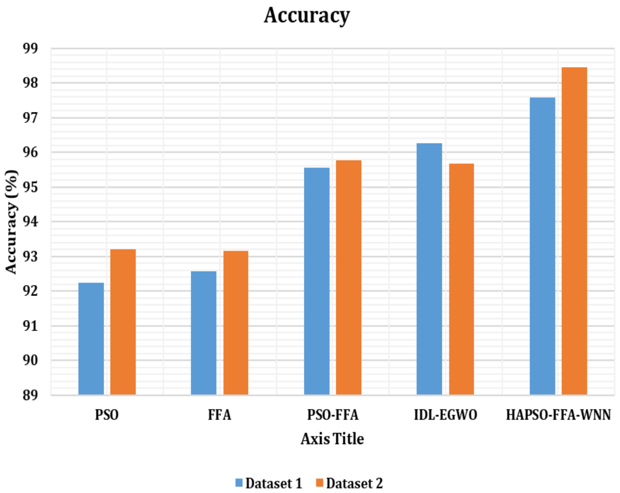


Figure. 1 Accuracy for Risk Level Prediction

Table 8. Precision for Risk Level Prediction

| Algorithm | Dataset 1 | Dataset 2 |
|---|---|---|
| PSO | 91.86 | 92.64 |
| FFA | 91.45 | 92.58 |
| PSO-FFA | 94.69 | 93.47 |
| IDL-EGWO | 94.15 | 95.28 |
| **HAPSO-FFA-WNN** | 96.39 | 97.8 |

Received: February 28, 2024. Revised: April 21, 2024. 

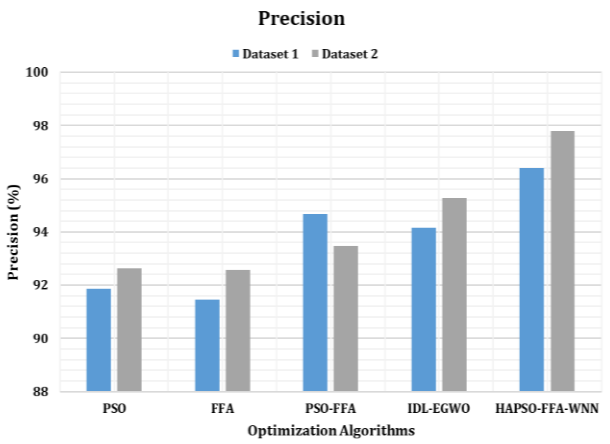


Figure. 2 Precision for Risk Level Prediction

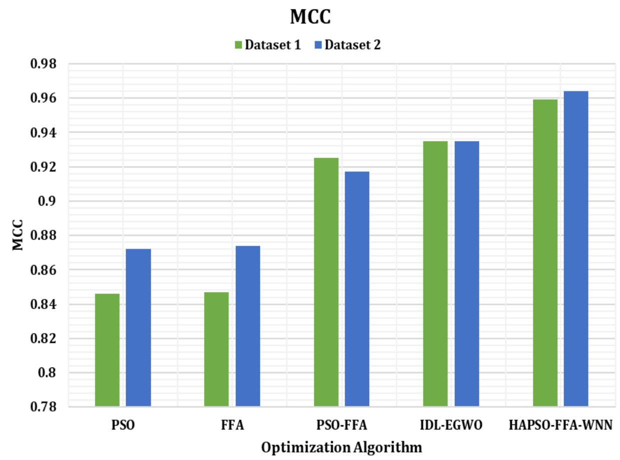


Figure. 4 MCC

Table 9. Recall For Risk Level Prediction

| Algorithm | Dataset 1 | Dataset 2 |
|---|---|---|
| PSO | 92.26 | 92.3 |
| FFA | 92.22 | 93.58 |
| PSO-FFA | 94.89 | 95.75 |
| IDL-EGWO | 95 | 96.2 |
| **HAPSO-FFA-WNN** | 97.24 | 98.4 |

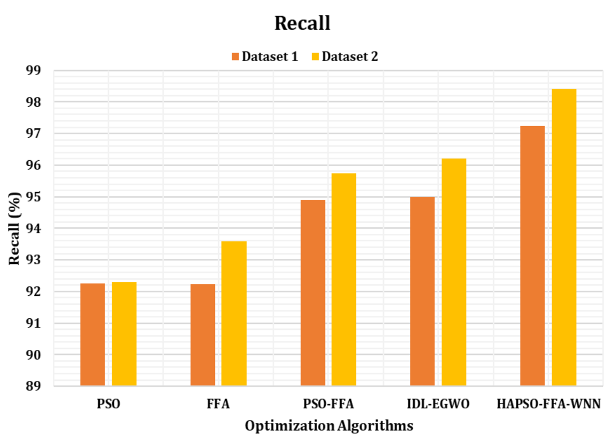


Figure. 3 Recall for Risk Level Prediction

Table 10. MCC for Risk Level Prediction

| Algorithm | Dataset 1 | Dataset 2 |
|---|---|---|
| PSO | 0.846 | 0.872 |
| FFA | 0.847 | 0.874 |
| PSO-FFA | 0.925 | 0.917 |
| IDL-EGWO | 0.935 | 0.935 |
| **HAPSO-FFA-WNN** | 0.959 | 0.964 |

Table 7 shows the accuracy comparison results of the proposed PSO [27], FFA [28], PSO-FFA[29], IDL-EGWO (previous work), and HAPSO-FFA-WNN algorithms. The accuracy values of PSO, FFA, PSO-FFA, and IDL-EGWO proved to be significantly lower compared to the accuracy values of the suggested HAPSO-FFA-WNN. The proposed HAPSO-FFA-WNN model achieves 5.34% high accuracy than PSO, 5% higher than FFA, 2.02% higher than PSO-FFA, and 1.32% higher than IDL-EGWO. Fig. 1 illustrates accuracy for risk level prediction.

Table 8 shows the precision comparison results of the proposed PSO, FFA, PSO-FFA, IDL-EGWO, and HAPSO-FFA-WNN algorithms. The precision values of PSO, FFA, PSO-FFA, and IDL-EGWO were significantly lower than those of the suggested HAPSO-FFA-WNN. The proposed HAPSO-FFA-WNN model achieves 5.34% higher precision value than PSO, 5% higher than FFA, 2.02% higher than PSO-FFA, and 1.32% higher than IDL-EGWO. Fig. 2 shows the precision value for predicting risk levels.

Table 9 shows the recall comparison results of the proposed PSO, FFA, PSO-FFA, IDL-EGWO, and HAPSO-FFA-WNN algorithms. The recall values of PSO, FFA, PSO-FFA, and IDL-EGWO were significantly lower than those of the suggested HAPSO-FFA-WNN. The proposed HAPSO-FFA-WNN model achieves 5.34% higher recall value than PSO, 5% higher than FFA, 2.02% higher than PSO-FFA, and 1.32% higher than IDL-EGWO. Fig. 2 illustrates the recall value for risk level prediction.

Table 10 shows the MCC comparison results of the proposed PSO, FFA, PSO-FFA, IDL-EGWO, and HAPSO-FFA-WNN algorithms. The MCC values of PSO, FFA, PSO-FFA, and IDL-EGWO proved to be significantly lower than those of the suggested HAPSO-FFA-WNN. The proposed HAPSO-FFA-

WNN model achieves 5.34% higher MCC value than PSO, 5% higher than FFA, 2.02% higher than PSO-FFA, and 1.32% higher than IDL-EGWO. Fig. 4 illustrates the MCC value for risk level prediction.

## 5. Conclusion

This research effort involves the construction of a unique adaptive multi-swarm PSO, which is then hybridized with the firefly technique. The goal of this research is to carry out weight optimization on the WNN model. During this research process, we developed HAPSO to enhance the exploration mechanism by incorporating a large number of SS, with the aim of identifying solutions within the search space. The hybrid version of the multi-swarm PSO, which incorporates the Firefly algorithm, has enhanced the weight values of the WNN model under consideration. We achieved this by precisely balancing the exploration and exploitation processes. The WNN model's network output has improved due to the application of wavelet functions and their corresponding translation and dilation coefficients, which has greatly reduced the mean square error. When applied to the lung cancer dataset and clinical datasets from hospitals, the hybrid version of the HAPSO-FFA-based WNN model achieved 97.58% accuracy, 96.39% precision, 97.24% recall, 0.959 MCC for dataset 1, and 98.45% classification accuracy, 97.8% precision, 98.4% recall, and 0.964 MCCfor dataset 2. We constructed a hybrid WNN model that converged, overcoming instances of both global and local minima. It has also shown improved risk-level classification in lung cancer datasets.

## Conflicts of Interest

The authors declare no conflict of interest.

## Author Contributions

The background work, conceptualization, methodology, dataset collection, implementation, result analysis and comparison, preparation and editing of the draft, and visualization have been done by the first author.

The supervision, review of work, and project administration has been done by the second author.

## References

[1] K. Priyadharshini, M. Alagarsamy, K. Sangeetha and D. K. Thangaraju, "Hybrid RNN-FFBPNN Optimized with Glowworm Swarm Algorithm for Lung Cancer Prediction", *IETE Journal of Research*, 2023, doi: 10.1080/03772063.2023.2233465.

[2] S. Solomon, and W. Mulugeta, "Diagnosis and Risk Factors and Advanced Cancers in Ethiopla", *Journal of Cancer Prevention*, Vol. 24, No. 3, September 2019, doi: 10.15430/JCP.2019.2.3.163.

[3] D. Endalie, and W. T. Abebe, "Analysis of Lung Cancer risk factors from medical records in Ethiopia using machine learning", *PLOS DigitalHealth*, Vol. 2, 2023, doi: 10.1371/journal.pdig.0000308.

[4] Y. H. Lai, W. N. Chen, T. C. Hsu, C. Lin, Y. Tsao, and S. Wu, "Small cell lung cancer by integrating microarray and clinical data with deep learning", *Scientific Reports*, 2020, doi: 10.1038/s41598-020-61588-w.

[5] H. Kim, J. M. Goo, K. H. Lee, Y. T. Kim, and C. M. Park, "Preoperative CT-Based Deep Learning Model for Predicting Disease-Free Survival in Patients with Lung Adenocarcinomas", *Radiology*, Vol. 296, No. 1, pp. 235, 2020, doi: 10.1148/radiol.2020192764.

[6] Y. Liao, G. Yin, and X. Fan, "The positive Lymph Node Ratio Predicts Survival in $T_{1-4}N_{1-3}M_0$ Non- Small Cell Lung Cancer: A Nomogram Using the SEER Database", *Frontiers in oncology*, Vol. 10, pp. 1356, 2020, doi: 10.3389/fonc.2020.01356.

[7] S. Supriya, and M. Subaji, "Deep learning the features maps for automated tumor grading of lung nodule structures using convolutional neural networks", *Intelligent Decision Technologies*, Vol. 14, No. 1, pp. 101-118,2020, doi: 10.3233/IDT-190083.

[8] A. Bhandary, G. Ananth Prabhu, V. Rajinikanth, K. P. Thanaraj, S. Chandra Satapathy, D. E. Robbins, C. Shasky, Y. D. Zhang, J. M. R.S. Tavares, and N. Sri Madhava Raja, "Deep-learning fframework to detect lung abnormality- A study with chest X-Ray and lung CT- scan images", *Pattern Recognition Letters*, Vol. 129, pp. 271-278, 2020, doi: 10.1016/j.patrec.2019.11.013.

[9] K. R. Pradeep, and N. C. Naveen, "Optimized - Deep Neural Network (Op-DNN) Regression Technique for Lung Cancer Survivability Prediction", *International Journal of Advanced Science and Technology*, Vol. 29, No. 3, pp. 4754-4772, 2020.

[10] S. S. Ghosal, I. Sarkar, and I. E. Hallaoui, "Lung Nodule Classification Using Convolutional Autoencoder and Clustering Augmented Learning Method (CALM)", In: *Proc. of 13th ACM International WSDM Conf. Houston, TX, USA*, Vol. 2551, pp. 19-26, 2020.

[11] S. Ramroach, M. John, and A. Joshi, "Lung cancer type classification using differentiator genes", *Gene Reports*, Vol. 19, pp.100647, 2020, doi: 10.1016/j.genrep.2020.100647.

[12] S. Kar, K. Das Sharma, and M. Maitra, "Adaptive weighted aggregation in Group Improvised Harmony Search for lung nodule classification", *Journal of Experimental & Theoretical Artificial Intelligence*, Vol. 32, No. 2, pp. 219-242, 2020, doi: 10.1080/0952813X.2019.1647561.

[13] H. Lu, H. Gao, M. Ye, and X. Wang, "A hybrid ensemble algorithm combining AdaBoost and genetic algorithm for cancer classification with gene expression data", In: *Proc. of International Conf. IEEE/ACM Transactions on Computational Biology and Bioinformatics*, Vol. 18, No. 3, pp. 863-870,2021, doi: 10.1109/TCBB.2019.2952102.

[14] R. Sujitha, and V. Seenivasagam, "Classification of lung cancer stages with machine learning over big data healthcare framework", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 13, pp. 5639-5649, 2021, doi: 10.1007/s12652-020-02071-2.

[15] M. S. Alrahhal, and E. Alqhtani, "Deep Learning-Based System for Detection of Lung Cancer Using Fusion of Features", *international Journal of Computer Science and Mobile Computing*, Vol. 10, pp. 57-67, 2021, doi: 10.47760/ijcsmc.2021.v10i02.009.

[16] K. H. Rubin, P. F. Haastrup, A. Nicolaisen, S. Moller, S. Wehberg, S. Rasmussen, K. Balasubramaniam, J. Sondergaard, and D. E. Jarbol, "Developing and Validating a Lung Cancer Risk Prediciton Model: A Nationwide Population-Based Study", *Cancers*, Vol. 15, pp. 457,2023, https://doi.org/10.3390/cancers15020487.

[17] S. Ahmed, B. Raza, L. Hussain, A. Aldweesh, A. Omar, Ml Shahbaz Khan, E. T. Eldin, and M. A. Nadim, "The Deep Learning ResNet101 and Ensemble XGBoost Algorithm with Hyperparameters Optimization Accurately Predict the Lung Cancer", *Applied Artificial Intelligence*, Vol. 37, No. 1, pp. 1696, 2023, doi: 10.1080/08839514.2023.2166222.

[18] Y. Hussian Ali, V. Sabu Chooralil, K. Balasubramanian, R. R. Manyam, S. K. Raju, A. T. Sadiq, and A. K. Farhan, "Optimization System Based on Convolutional Nerual Network and Internet of Medical Things for Early Diagnosis of Lung Cancer", *Bioengineering*, Vol. 10, pp. 320, 2023, doi: 10.3390/bioengineering10030320.

[19] N. V. Mahalakshmi, and R. K. Rout, "Effective heart disease prediction using improved particle swarm optimization algorithm and ensemble classification technique", *Soft Computing A Fusion of Foundations, Methodologies and Applications*, Vol. 27, 2023, doi: 10.1007/s00500-023-08388-2.

[20] S. Raghavendra, V. Parvati, R. Manjula, A. K. Nanda, R. Singh, D. Lakshmi, and S. Velmurugan, "DLMNN Based Heart Disease Prediction with PD-SS Optimization Algorithm", *Intelligent Automation & Soft Computing*, Vol. 35, No. 2, 2023, doi: 10.32604/iasc.2023.027977.

[21] S. Rajalakshmi, R. Ragulraja, K. Kalvi, S. Dhinesh Kumar, and B. Murugan, "Heart Disease Prediction Using Improved Squirrel Search Algorithm", *Journal for Research in Applied Science and Engineering Technology*, Vol. 11, 2023, doi: 10.22214/ijraset.2023.51836.

[22] C. Venkatesh, K. Ramana, S. Y. Lakkisetty, S. S. Band, S. Agarwal, and A. Mosavi, "A Neural Network and Optimization Based Lung Cancer Detection System in CT Images", *Frontiers in public health*, Vol. 10, 2022, doi: 10.3389/fpubh.2022.769692.

[23] T. I. A. Mohamed, O. N. Oyelade, and A. E. Ezugwu, "Automatic detection and classification of lung cancer CT scans based on deep learning and ebola optimization search algorithm", *PLOS ONE*, Vol. 18, 2023, doi: 10.1371/journal.pone.0285796.

[24] K. K. Kumar, and K. V. Kanimozhi, "Detection of Lung Disease using Novel Genetic Algorithm in comparison with Particle Swarm Optimization to improve Accuracy", In: *Proc. of 2022 International Conf. on Business Analytics for Technology and Security (ICBATS)*, Dubai, United Arab Emirates, pp. 1-4, 2022, doi: 10.1109/ICBATS54253.2022.9759044.

[25] K. Ramana, M. R. Kumar, K. Sreenivasulu, T. R. Gadekallu, S. Bhatia, P. Agarwal, and S. M. Idrees, "Early Prediction of Lung Cancer Using Deep Saliency Capsule and Pre-Trained Deep Learning Framework", *Frontiers in Oncology*, Vol. 12, 2022, doi: 10.3389/fonc.2022.886739.

[26] K. Swathi, and S. Kodukula, "Lung Cancer Prediction Using Improved SALP Swarm Optimization and LSTM Human Gene Classification", *International Journal of Intelligent Systems and Applications in Engineering*, Vol. 12, pp. 142151, 2023.

[27] J. Seth, Parma Nand, and S. Prabhdeep, Rajbir Kaur, "Particle Swarm Optimization Assisted Support Vector Machine Based Diagnostic

System for Lung Cancer Prediction at The Early Stage", *PalArch's Journal of Archaeology of Egypt / Egyptology*, Vol. 17, pp. 6202-6212, 2020, doi: archives.palarch.nl/index.php/jae/article/view/5179.

[28] A. Pati, A. Panigrahi, B. Sahu, G. Sahoo, M. Dash, M. Parhi, and B. K. Pattanayak, "FOHC: Firefly Optimizer Enabled Hybrid approach for Cancer Classification", *International Journal on Recent and Innovation Trends in Computing and Communication*, Vol. 11(7s), pp. 118-125, doi: 10.17762/ijritcc.v11i7s.6983.

[29] H. Xiao, and T. Hatanaka, "Model Selecting PSO-FA Hybrid for Complex Function Optimization", *International Journal of Swarm Intelligence Research (IJSIR)*, IGI Global, Vol. 12, No. 3, pp. 215232, 2021, doi: 10.4018/IJSIR.2021070110.