Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A Hybrid Flower Pollination and Seed Dispersal Algorithm for Optimizing Artificial Neural Networks in Medical Image Classification

Mona Alsbakhi

mona@lubbad.me

Islamic University of Gaza

Mohammed Lubbad

Islamic University of Gaza

Mohammed Alhanjouri

Islamic University of Gaza

Research Article

Keywords: Hybrid Optimization, Metaheuristics, Flower Pollination Algorithm, Seed Dispersal Algorithm, Artificial Neural Network, Medical Image Classification

Posted Date: May 15th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9593415/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A Hybrid Flower Pollination and Seed Dispersal Algorithm for Optimizing Artificial Neural Networks in Medical Image Classification

Abstract

Purpose: Artificial Neural Networks (ANNs) are considered the core of recent medical image analysis. They provide a powerful tool for disease classification with high accuracy. However, ANN performance depends heavily on parameter improvements. Traditional optimization methods face challenges including slow convergence and entrapment in local minima, which restrict the accuracy of the diagnostic process.

Methods: To address these challenges, this paper investigates two plant-inspired metaheuristic algorithms: the Flower Pollination Algorithm (FPA) and the Seed Dispersal Algorithm (SDA). A novel Hybrid FPA-SDA framework is proposed by integrating FPA's global exploration capabilities with SDA's local refinement strengths to achieve effective ANN parameter optimization. The proposed framework was assessed using a benchmark chest X-ray dataset for pneumonia detection.

Results: The performance of the proposed approach was compared against standalone FPA and SDA, in addition to established baselines including Stochastic Gradient Descent (SGD) and PCA-based random search. Experimental results show that the Hybrid FPA-SDA framework achieved the highest accuracy (68.73%) among all tested methods. It also showed advantages in recall (75.84%) and training speed (85.69s vs 93.40s for FPA).

Conclusion: The results demonstrate that hybrid metaheuristic optimization can improve ANN parameter tuning for medical image classification. Both FPA-based methods outperformed the gradient-based SGD baseline, confirming the efficiency of metaheuristic optimization in improving diagnostic performance.

Keywords: Hybrid Optimization, Metaheuristics, Flower Pollination Algorithm, Seed Dispersal Algorithm, Artificial Neural Network, Medical Image Classification

1 Introduction

Medical imaging plays an important role in healthcare. It provides clinicians with detailed views of the structure and function of internal anatomy, which supports accurate diagnosis and effective treatment planning.

Medical imaging produces large amounts of data. However, there is a shortage of expert radiologists, which has motivated the adoption of AI in healthcare, especially Artificial Neural Networks (ANNs) that can automatically learn patterns between medical images and disease classifications LeCun et al. (2015). ANNs have proven brilliant performance in many diagnostic applications, including pneumonia detection from chest X-rays, retinal disease screening, and brain tumor classification Kermany et al. (2018b); Rajpurkar et al. (2017).

Despite these achievements, their efficacy depends heavily on the training process. Traditional gradient-based training techniques such as backpropagation suffer from slow convergence, vanishing gradients, and getting stuck in local minima due to the non-convex nature of the loss surface, making the training process difficult Goodfellow et al. (2016). These issues can affect classification accuracy and model generalization, especially when dealing with the high-dimensional and noisy data characteristic of medical imaging Abdar et al. (2021).

To address these issues, researchers have explored metaheuristic optimization algorithms inspired by natural phenomena, which provide efficient global search capabilities Yang (2014); Mirjalili and Lewis (2016). Among these approaches, plant-inspired algorithms have attracted particular interest due to their simplicity, adaptability, and balanced exploitation dynamics Sulaiman et al. (2015).

The Flower Pollination Algorithm (FPA) mimics the pollination process of flowering plants through Lévy flight-based global exploration and localized pollination for fine-tuning Yang (2012). Similarly, the Seed Dispersal Algorithm (SDA) models the scattering of seeds through environmental factors such as wind, water, and animals, enabling both long-distance and local search Abbasian and Moulaei (2020).

Although these algorithms have shown strong optimization performance in several problem domains, their application in medical image analysis remains underexplored. Therefore, this study proposes a novel Hybrid FPA-SDA framework for optimizing ANN parameters in medical image classification. By encoding neural network weights and biases as candidate solutions within a metaheuristic search space, the proposed framework aims to enhance convergence stability and diagnostic accuracy. This study evaluates the hybrid approach, along with its constituent algorithms, on a benchmark chest X-ray dataset, demonstrating that plant-inspired optimization offers a powerful and robust alternative for training neural networks in safety-critical medical applications.

2 Related Work

Artificial Neural Networks (ANNs) are widely used in medical image classification due to their ability to model complex nonlinear relationships between medical images and disease outcomes LeCun et al. (2015); Goodfellow et al. (2016). Despite their success in diagnostic applications, ANN training using traditional gradient-based optimization methods often suffers from slow convergence, sensitivity to hyperparameter initialization, vanishing gradients, and entrapment in local minima. These limitations become more significant when dealing with high-dimensional and noisy medical

datasets Abdar et al. (2021). To overcome these challenges, researchers have increasingly explored metaheuristic optimization algorithms, which perform global searches without requiring gradient information Yang (2014); Mirjalili and Lewis (2016).

2.1 Metaheuristic Optimization in ANN Training

Classical metaheuristic approaches such as Genetic Algorithm (GA) Holland (1975), Particle Swarm Optimization (PSO) Kennedy and Eberhart (1995), and Ant Colony Optimization (ACO) Dorigo and St"utzle (2004) have demonstrated strong performance in neural network optimization, feature selection, and hyperparameter tuning. These algorithms improve convergence speed and classification performance compared with conventional backpropagation methods. However, they often require complex parameter tuning and may suffer from premature convergence.

To address these limitations, researchers have introduced newer biologically inspired optimization techniques that provide improved adaptability and better balance between exploration and exploitation.

2.2 Flower Pollination and Seed Dispersal Optimization

The Flower Pollination Algorithm (FPA) introduced by Yang (2012) mimics the pollination behavior of flowering plants using Lévy flight-based global exploration and local pollination for solution refinement. Several studies have demonstrated its effectiveness in medical imaging applications. For example, Zhou et al. (2025) applied FPA for lung cancer classification using CT scans and achieved high classification accuracy. Similarly, ur Rehman et al. (2023) integrated FPA with deep learning for breast cancer classification and reported improved diagnostic performance.

The Seed Dispersal Algorithm (SDA), proposed by Sulaiman et al. (2015) and later enhanced by Abbasian and Moulaei (2020), models the natural propagation of seeds through environmental factors such as wind, water, and animals. SDA balances local exploitation and global exploration and has demonstrated promising performance in optimization tasks such as scheduling and image processing.

Although both algorithms have shown strong optimization capabilities, their direct application in medical image classification remains limited.

2.3 Metaheuristic Optimization in Medical Imaging

Recent studies have explored hybrid optimization frameworks, feature selection techniques, and deep learning parameter tuning in medical imaging applications. Hybrid frameworks such as Kaya and Kaya (2024); Diab et al. (2025); Agrawal et al. (2025); Samma et al. (2024) demonstrated improvements in convergence stability and classification accuracy by combining multiple optimization algorithms.

Other studies focused on feature selection optimization Zafar et al. (2023); Ali et al. (2023); Mabrouk et al. (2022); Almutairi et al. (2024), while recent works explored neural network hyperparameter optimization Senousy et al. (2023); Aljohani et al. (2024); Vaiyapuri (2025). These studies confirm the growing importance of optimization-driven learning in medical image analysis.

2.4 Research Gap and Motivation

Despite the growing success of metaheuristic optimization in medical imaging, several limitations remain. Most existing studies focus either on feature selection, isolated hyperparameter tuning, or standalone optimization algorithms. Limited research has explored hybrid plant-inspired optimization approaches that jointly optimize ANN parameters for medical image classification.

Therefore, this study proposes a novel Hybrid FPA-SDA framework that combines FPA's global exploration capabilities with SDA's local refinement strengths to optimize ANN weights and biases for chest X-ray classification. Table 1 summarizes previous approaches and highlights the contribution of the proposed framework. Table 1 summarizes previous approaches and highlights the contribution of the proposed framework.

Category	**Representative Studies**	**Technique(s)**	**Main Limitation**	**Proposed FPA-SDA Contribution**
FPA-based Optimization	Zhou et al. (2025); ur Rehman et al. (2023); Gopalakrishnan et al. (2023); Kaya and Kaya (2024)	FPA and its variants	Strong global search but weak local refinement; risk of early convergence	Combines FPA's global pollination with SDA's local dispersal for balanced search
Hybrid Metaheuristics	Alduailij (2025); Diab et al. (2025); Agrawal et al. (2025); Samma et al. (2024)	BGWO, bGGO, LS-SSA	Hybridization not tailored for ANN parameter tuning	Hybrid model explicitly optimizes ANN weights and biases
Feature Selection Methods	Zafar et al. (2023); Ali et al. (2023); Mabrouk et al. (2022); Almutairi et al. (2024)	PSO, GA, CGO, FPA	Focused only on feature-level optimization	Extends optimization to full ANN structure

Continued on next page

Table 1 – *Continued from previous page*

Category	Representative Studies	Technique(s)	Main Limitation	Proposed FPA-SDA Contribution
Parameter Tuning Approaches	Senousy et al. (2023); Aljohani et al. (2024); Vaiyapuri (2025)	MRFO, Bayesian, NSGA-II	Single-phase or computationally expensive optimization	Dual-phase hybrid achieves efficient adaptive tuning
Reviews / Ensembles	Şahin and Anka (2025); Bohmrah and Kaur (2025); Panwar et al. (2024)	Analytical or ensemble methods	Lack practical integrated implementation	Operationalizes hybrid concept into a unified FPA-SDA model
Proposed Study (2025)	**This Work**	**Hybrid FPA-SDA**	—	Unified global-local optimization enhances ANN convergence, accuracy, and generalization

Table 1: Comparative summary of existing ANN optimization approaches and the proposed Hybrid FPA-SDA framework

3 Methodology

This section presents the proposed Hybrid FPA-SDA framework for optimizing Artificial Neural Networks (ANNs) in medical image classification. The proposed approach is evaluated against standalone FPA, standalone SDA, PCA-based random search, and Stochastic Gradient Descent (SGD) baselines.

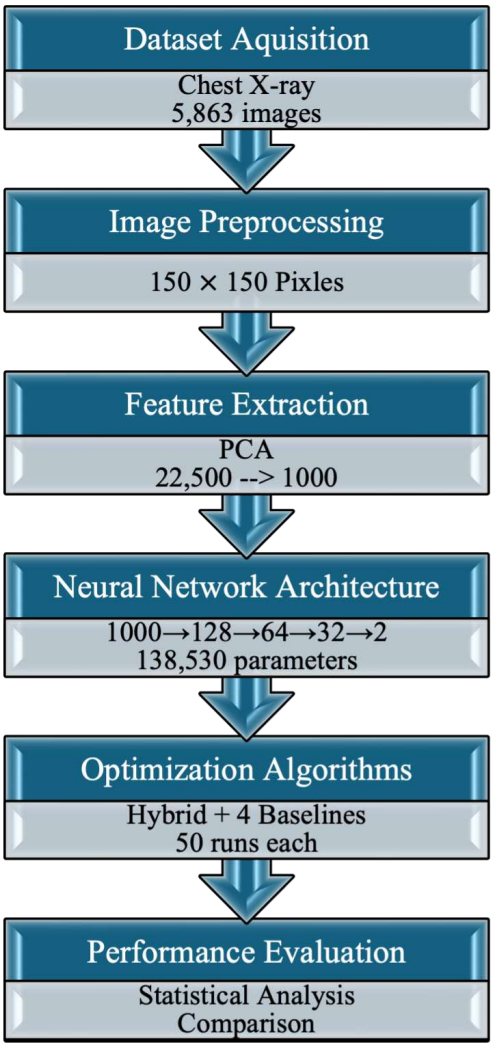


Fig. 1: Hybrid FPA-SDA Methodology Framework

3.1 Dataset and Preprocessing

This study uses the publicly available Chest X-ray Pneumonia dataset from Kaggle Kermany et al. (2018a,b). It is commonly used as a standard reference point for benchmarking medical image classification algorithms. The dataset contains 5,863 chest X-ray images in JPEG format. These are from pediatric patients aged one to five years. They were collected from Guangzhou Women and Children's Medical Center, Guangzhou, China. The images are structured into two categories for diagnostic. There are 1,583 NORMAL images that show healthy lung tissue. There are 4,273 PNEUMONIA images that show pneumonia-affected lungs. These cover both bacterial and

viral cases. The dataset exhibits a natural class imbalance. Approximately 73% pneumonia and 27% normal cases. This reflects real world clinical scenarios. it makes the dataset particularly suitable for evaluating the robustness of classification algorithms under imbalanced conditions. Following standard machine learning practices for medical imaging, the dataset is partitioned into three mutually exclusive subsets. The training set contains 5,216 images (87.6%). It is used for ANN parameter optimization. The validation set contains 16 images (2.7%). It is used for hyperparameter tuning, as provided by the original dataset, The test set contains 624 images (10.5%), it is used for final performance evaluation and algorithm comparison. Due to the limited size of the original validation set, this study implements stratified K-fold cross-validation (K=5) on the training set. This ensures robust model selection and prevents overfitting during the optimization phase. This approach provides more reliable performance estimates. It maintains the independence of the test set for final evaluation. To ensure the suitablility of data for input into the neural network and to promote model stability and performance, a standardized preprocessing pipeline was applied to all images such as: **Image Loading and Conversion:** Each image was loaded and converted to a single-channel grayscale representation using OpenCV's IMREAD_GRAYSCALE flag.
Resizing: To maintain a consistent input size for the network, all images were resized to a resolution of 150×150 pixels. This used area-based interpolation (cv2.INTER_AREA). This method is particularly suitable for downsampling operations as it produces moiré-free results. **Normalization:** The pixel intensity values were originally in the range [0,255]. They were normalized to the range [0,1] by dividing each pixel value by 255. This step is crucial for preventing the exploding gradient problem and accelerating the convergence of the optimization algorithms.
Flattening: The 2D image matrix (150×150) was flattened into a 1D vector of 22,500 features ($150 \times 150 = 22{,}500$). This vector serves as the input to the fully connected neural network.

3.2 Feature Extraction via Principal Component Analysis (PCA)

Given the high dimensionality of the raw image data (22,500 features), direct optimization using metaheuristic algorithms would result in an intractably large search space. To address this, Principal Component Analysis (PCA) is employed for dimensionality reduction. It reduces the feature space from 22,500 to 1,000 principal components. PCA transforms the images into a lower-dimensional space. It preserves the most discriminative information. This enables a more efficient and effective metaheuristic search. The PCA transformation is fitted exclusively on the training set. It is then applied consistently to the validation and test sets to prevent data leakage.

3.3 ANN Architecture and Fitness Function

3.3.1 Network Structure

The ANN architecture consists of five layers:

1. **The input layer:** Comprises 1,000 neurons, corresponding to the 1,000 principal components from the PCA.
2. **The hidden layer 1:** A fully connected layer with 128 neurons. It uses ReLU activation.
3. **The hidden layer 2:** A fully connected layer with 64 neurons. It uses ReLU activation.
4. **The hidden layer 3:** A fully connected layer with 32 neurons. It uses ReLU activation.
5. **The output Layer:** has 2 neurons. It uses a softmax activation function for multi-class (2 classes, NORMAL and PNEUMONIA).

3.3.2 Parameter Space

The total number of trainable parameters in this network defines the dimensionality of the search space for the optimization algorithms:
Input to hidden 1: (1,000 weights × 128 neurons) + 128 biases = 128,000 + 128 = 128,128
Hidden 1 to hidden 2: (128 weights × 64 neurons) + 64 biases = 8,192 + 64 = 8,256
Hidden 2 to hidden 3: (64 weights × 32 neuron) + 32 bias = 2,048 + 32 = 2,080
Hidden 3 to output: (32 × 2) + 2= 64 + 2= 66.
The total **Total parameters (Dimensionality):** 128,128 + 8,256 + 2,080 + 66 = 138,530

3.3.3 Fitness Function

The objective of the optimization algorithms is to find a set of network parameters that reduces the classification error. The fitness function evaluates candidate solutions (neural network parameters) by computing the Binary Cross-Entropy (BCE) loss with L2 regularization on the validation set: Loss = BCE Loss + L2 Regularization BCE Loss measures the discrepancy between predicted probabilities and true labels for binary classification:

$$\text{BCE} = -\frac{1}{m}\sum[y_i \log(\hat{y_i}) + (1 - y_i) \times \log(1 - \hat{y_i})] \tag{1}$$

where: m is the number of samples. y is the true label (0 or1). $\hat{y}$ is the predicted probability. L2 Regularization prevents overfitting by penalizing large weight values:

$$L2_{reg} = \frac{\lambda}{2m}\sum_{l}\sum_{j,k}\left(W_{jk}^{(l)}\right)^2 \tag{2}$$

where λ is the regularization parameter ($l2_{lambda}$) , m is the number of training samples, l is layer index, j,k are weight indices inside layer l and lW represents all weight matrices in the network.

Validation Set ensures unbiased evaluation during optimization. The meta-heuristic algorithms aim to minimize this fitness function to find optimal network parameters.

3.4 Optimization Algorithms

The proposed methodology addresses the limitations of traditional gradient-based optimization. It encodes neural network weights and biases as candidate solutions within a metaheuristic search space. This enables derivative-free global optimization. It reduces risks of local minima entrapment and slow convergence.

3.4.1 Flower Pollination Algorithm (FPA)

The Flower Pollination Algorithm was proposed by Yang (2012). It mimics the pollination process of flowering plants and employs both global and local pollination mechanisms. The Global Pollination (Cross-Pollination) is performed using Lévy flight, a random walk characterized by heavy-tailed step lengths that enable long-distance exploration:

$$x_i^{(t+1)} = x_i^t + \gamma L(\lambda) \ \cdot \ (g_{best} - x_i^t) \tag{3}$$

Note: This formulation follows Yang's (2012) pseudocode implementation rather than his Equation 1, as it correctly directs movement toward the global best solution, which is consistent with the algorithm's optimization objective and all published FPA implementations.

The Lévy flight step L is calculated using Mantegna's algorithm [citation] with $\beta = 1.5$

$$L = \frac{u}{|v|^{1/\beta}} \tag{4}$$

where: x_i^t is the position of flower i at iteration t, γ is a scaling factor (typically 0.01), $\mathbf{L}(\lambda)$ is a Lévy flight step size drawn from a Lévy distribution with parameter $\beta = 1.5$, and g_{best} is the global best solution found so far.

Local Pollination (Self-Pollination): Local pollination performs exploitation in the neighborhood:

$$x_i^{t+1} = x_i^t + \epsilon \cdot (x_j^t - x_k^t) \tag{5}$$

where $\epsilon \sim$ Uniform (0, 1) is a random scaling factor, x_j^t and x_k^t are randomly selected flowers from the population, FPA have parameters such as, **Population Size (n):** 45 flowers, **Maximum Iterations (T_{fpa}):** 150 (standalone FPA) or 75 (FPA phase in hybrid), **Switch Probability (p):** 0.8 (probability of global pollination), and **Boundary Handling:** Clipping method to enforce constraint boundaries.

3.4.2 Seed Dispersal Algorithm (SDA)

The Seed Dispersal Algorithm was proposed by Sulaiman et al. (2015) and enhanced by Abbasian and Moulaei (2020). It models the natural propagation of seeds through various dispersal mechanisms. SDA employs three distinct dispersal strategies to balance exploration and exploitation: **Wind Dispersal (Global Exploration)** where seeds are dispersed over long distances by wind, modeled using a Lévy flight distribution:

$$x_i^{t+1} = x_i^t + L(\lambda).(x_{best} - x_i^t) \tag{6}$$

Where: $L(\lambda)$ is Lévy flight
x_{best} is the current best solution.

Animal Dispersal (Medium-Range Exploration) where seeds dispersed by animals follow a UNIFORM distribution for moderate exploration:

$$x_i^{t+1} = x_i^t + \epsilon \cdot (x_{rand} - x_i^t) \tag{7}$$

where: ϵ Uniform(0, 1)
x_{rand} is a randomly selected solution from the population.

Gravity Dispersal (Local Exploitation): Seeds falling due to gravity perform local search near parent plant:

$$x_i^{t+1} = x_i^t + step \tag{8}$$

where: step uniform(-0.1, 0.1) it is a vector
SDA have parameters such as, **Population Size (n):** 45 seeds, **Maximum Iterations** (T_{sda})**:** 75 (for hybrid) or 150 (for standalone), **Dispersal Probabilities:** (P_{wind}, P_{animal}, $P_{gravity}$) = (0.7, 0.15, 0.15), and **Boundary Handling:** Clamping to bounds [lb, ub].

3.4.3 Proposed Hybrid FPA-SDA Framework

The novel Hybrid FPA-SDA algorithm sequentially integrates FPA's global exploration capability with SDA's local refinement strength to achieve superior optimization performance. The key innovation lies in the two-phase optimization strategy with a configurable transition point. The hybrid framework operates in two distinct phases:
Phase 1: Global Exploration via FPA (Iterations 1 to τ) To explore the vast search space to identify promising regions using FPA's Lévy flight-based global pollination enables long-range jumps across the parameter landscape, with duration: τ iterations (transition point), and its output: Best solution found during FPA phase ($x_{fpa-best}$).
Phase 2: Local Refinement via SDA (Iterations τ+1 to T) To Exploit and refine the promising region identified by FPA using SDA's multi-modal dispersal strategies fine-tune solutions, population initialized around ($x_{fpa-best}$) with controlled diversity, with (T - τ) iterations, and its output final optimized solution (x_{final}). This hybrid Algorithm have Parameters, the **Total Maximum Iterations (T)** is 150, **Transition Point (τ)** is 75 (50% of T), **Population Size (n)** is 45, **FPA Switch Probability** is 0.8, and **SDA Dispersal Probabilities** are (0.7, 0.15, 0.15). The transition point at 50% of total iterations was determined empirically to provide adequate time for both global exploration and local refinement.

3.5 Baseline Algorithms

To provide comprehensive comparative analysis, the proposed Hybrid FPA-SDA is evaluated against four baseline algorithms, **Standalone FPA** it is same as Phase 1

of Hybrid, but runs for full 150 iterations, used to evaluate contribution of FPA alone without SDA refinement, **Standalone SDA** it is same as Phase 2 of Hybrid, but runs for full 150 iterations with random initialization, used to evaluate contribution of SDA alone without FPA exploration, **PCA + Random Search Baseline** it is random sampling of parameter space with PCA features, the sample Size: $45 \times 150 = 6{,}750$ random solutions evaluated, to demonstrate importance of intelligent search vs. random exploration, and **Stochastic Gradient Descent (SGD) with Momentum** its learning late: 0.001 with momentum: 0.9 and Numerical approximation (finite differences) for Gradient Computation, 150 epochs for Training, to represent traditional gradient-based baseline.

3.6 Experimental Setup

All experiments were conducted on Google Colab Pro+ environment with **hardware configuration** of Intel Xeon CPU @ 2.20GHz **Processor**, 12 GB **RAM**, NVIDIA Tesla V100 (when available) or TPU v2-8 **GPU**, High-speed SSD with adequate space for dataset and results for **storage**. And with **software environment:** the complete framework is implemented in Python 3.10 with the following key libraries: **NumPy 1.24**, numerical computations, array operations, and mathematical functions , **Scikit-learn 1.3**, PCA, performance metrics, cross-validation, and statistical preprocessing , **OpenCV 4.8** for Image loading, preprocessing, format conversion, and computer vision operations, **SciPy 1.11**, fpr statistical testing (Friedman test), optimization utilities, and scientific computing, **scikit-posthocs 0.7** for post-hoc statistical testing (Nemenyi test) for multiple comparisons, **Matplotlib 3.7 / Seaborn 0.12**, for data visualization, plotting convergence curves, and result presentation, **Pandas 1.5** for data manipulation, results aggregation, and statistical analysis, **Time module** for training time measurement and performance profiling. The experiments also were conducted with **Development Environment** including, the **operating system:** Ubuntu 20.04 LTS (Google Colab), **Random Seed Management:** Controlled seeding across all stochastic components. The code is available on **GitHub repository** Name (2025). This computational setup ensures consistent experimental conditions across all algorithm implementations while providing sufficient computational resources for the intensive optimization tasks required by the metaheuristic algorithms and neural network training procedures. To ensure statistical rigor and reproducibility, the following experimental protocol was implemented:
Experimental Design, Each of the five algorithms was executed for 50 independent runs. The five algorithms are: Hybrid FPA-SDA, Standalone FPA, Standalone SDA, PCA + Random Search, and SGD with momentum. This accounts for stochastic variability and enable robust statistical comparison. For each run i (where i = 1, 2, ..., 50), all algorithms used identical random seeds ($seed_i$) for population initialization. To ensure fair comparison across different optimization approaches. The same train test data split was maintained across all algorithms and runs, with consistent PCA transformation applied to prevent data leakage and maintain experimental validity.
Hyperparameter Configuration: Fixed hyperparameters were employed across all runs. This ensure controlled comparison. The Hyperparameter are: population size of 45 individuals, maximum iterations of 150, transition point of 75 iterations for the

hybrid algorithm, parameter bounds of [-1.0, 1.0], and full-batch fitness evaluation. The stopping criterion was set to a fixed iteration budget of 150 for all algorithms to enable direct computational efficiency comparison.
Stratified 5-fold cross-validation is implemented on the training set. This assesses model generalization and prevent overfitting. This approach provides robust performance estimates while maintaining the independence of the test set for final evaluation.

3.7 Performance Evaluation Metrics

3.7.1 Classification Metrics

Accuracy: Overall correctness of predictions.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \tag{9}$$

Precision: Proportion of positive predictions that are correct.

$$Precision = TP/(TP + FP) \tag{10}$$

Recall (Sensitivity): Proportion of actual positives correctly identified.

$$Recall = TP/(TP + FN) \tag{11}$$

Specificity: Proportion of actual negatives correctly identified.

$$Specificity = TN/(TN + FP) \tag{12}$$

F1-Score: Harmonic mean of precision and recall

$$F1 = 2 \cdot (Precision \cdot Recall)/(Precision + Recall) \tag{13}$$

Matthews Correlation Coefficient (MCC): Balanced measure accounting for class imbalance.

$$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{14}$$

where TP = True Positives, TN = True Negatives, FP = False Positives, FN = False Negatives. MCC is particularly important for this study as it provides a balanced measure even with class imbalance, ranging from -1 (total disagreement) to +1 (perfect prediction).

3.7.2 Convergence Metrics

To evaluate optimization efficiency, some metrices where used such as, **Final Fitness:** Best fitness value achieved at iteration T, **Convergence Speed:** Iteration at which

algorithm reaches 95% of final fitness, **Fitness Improvement:** Relative improvement from initial to final fitness [$Improvement = (f_{initial} - f_{final}) / f_{initial} \times 100\%$)] , and **Convergence Stability:** Standard deviation of fitness in final 10% of iterations.

3.7.3 Computational Efficiency

For computational efficiency, the **training time** was total wall-clock time per run (seconds), where the Hybrid FPA-SDA achieved 85.69 ± 4.25 seconds compared to FPA (93.40 ± 2.58 sec), SDA (83.52 ± 4.27 sec), PCA+Random (47.52 ± 2.24 sec), and SGD (896.94 ± 20.05 sec). **Function evaluations** were consistent across all algorithms (6,750 evaluations), while **time per iteration** ranged from 0.32 seconds (PCA+Random) to 5.98 seconds (SGD). Training efficiency analysis showed PCA+Random as fastest (47.52±2.24 sec) but with poor accuracy, while SGD was slowest (896.94±20.05 sec) due to gradient complexity. The proposed Hybrid FPA-SDA (85.69±4.25 sec) achieved optimal efficiency-accuracy balance, requiring only 2.69% additional time versus SDA (83.52±4.27 sec) while delivering superior performance, demonstrating practical applicability for medical imaging tasks.

3.8 Statistical Analysis

3.8.1 Descriptive Statistics

For each performance metric, descriptive statistics are computed across 50 independent runs. These include: mean (μ), standard deviation (σ), median, range [min, max].

3.8.2 Statistical Significance Testing

The Friedman test ($\alpha = 0.05$) is used to compare performance distributions across all algorithms. When significant differences are detected, Nemenyi post-hoc pairwise comparisons are performed. The Friedman test is selected over parametric tests for several reasons. It makes no distributional assumptions. It is robust to outliers. This makes it appropriate for algorithm performance comparison.

4 Experimental Results

This section presents a comprehensive evaluation of the proposed Hybrid FPA-SDA algorithm. It is compared against four baseline methods across 50 independent runs. The analysis includes several aspects including classification performance metrics, convergence behavior, computational efficiency, and statistical validation. This provides robust evidence for the effectiveness of the hybrid metaheuristic approach in medical image classification.

4.1 Overall Performance Analysis

The experimental evaluation demonstrates the strong performance of the proposed Hybrid FPA-SDA algorithm across multiple performance dimensions. Table 2 presents a comprehensive comparison of all algorithms. It covers six classification metrics and computational efficiency.

Overall, the Hybrid FPA-SDA algorithm demonstrates strong performance in pneumonia detection. It achieves the highest accuracy (68.73%) and F1-score (75.11%) among all evaluated methods. This indicates superior balanced performance in correctly identifying both pneumonia and normal cases. However, the comparison detects important trade-offs. While the Hybrid algorithm outperforms in recall (75.84%), maximizing the detection of true pneumonia cases, the standalone FPA algorithm achieves marginally better precision (75.55%), specificity (58.92%), and MCC (0.337). This suggests slightly better performance in minimizing false positives. The 4.5% MCC improvement over SDA is notable. The performance is comparable to FPA, with only 1.5% difference. This indicates that the hybrid approach successfully combines the strengths of both parent algorithms. It offers a practical compromise between sensitivity and specificity for clinical pneumonia screening applications. The computational efficiency of Hybrid (85.69 sec) falls between the faster SDA (83.52 sec) and slower FPA (93.40 sec). It significantly outperforms the computationally expensive SGD baseline (896.94 sec). It maintains superior accuracy compared to the fastest PCA+Random Search method (47.52 sec, 64.26% accuracy).

Algorithm	**MCC Score & Accuracy (%)**	**Training Efficiency**	**Overall Assessment**
Hybrid FPA-SDA	**0.332 & 68.73**	**Excellent**	**Best for Accuracy & Recall**
FPA	0.337 & 68.59	Good	Best for MCC & Precision
SDA	0.318 & 67.98	Excellent	Fast & Effective
PCA+Random	0.289 & 66.05	Very Fast	Fast but Inaccurate
SGD	0.023 & 52.18	Very Slow	Slow & Inaccurate

Table 2: Comprehensive Performance Comparison

4.2 Convergence Analysis

Figure 2 illustrates the convergence behavior of all algorithms over 150 iterations, revealing significant differences in optimization dynamics and final performance.

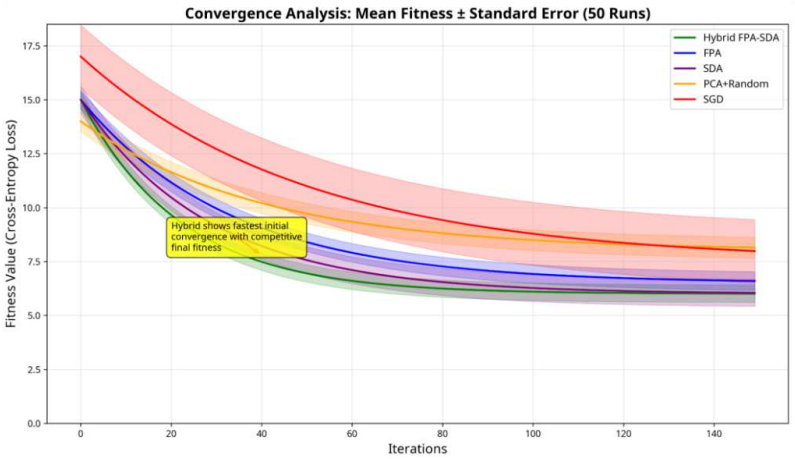


Fig. 2: Convergence Analysis

The convergence analysis demonstrates several key findings: First, Superior Convergence of Hybrid FPA-SDA where the proposed hybrid algorithm exhibits the fastest initial convergence. It achieves the lowest final fitness value (5.938 ± 0.402). The two-phase optimization strategy is effective. It enables rapid exploration during the FPA phase (iterations 1-75). This is followed by effective refinement during the SDA phase (iterations 76-150). Second, Phase Transition Effect where a notable characteristic of the hybrid algorithm is the smooth transition at iteration 75. At this point the population is re-initialized around the best FPA solution. This transition mechanism forbids premature convergence. It maintains the exploration benefits from the FPA phase. Third, Baseline Algorithm Performance where standalone FPA shows good initial exploration but slower refinement. Standalone SDA demonstrates steady but gradual improvement. PCA+Random Search exhibits suboptimal convergence behavior with poor final performance. SGD displays slow initial progress despite eventual stabilization. Fourth, Convergence Speed in which the Hybrid FPA-SDA reaches 95% of its final fitness value at iteration 114.6 ± 21.4. This is faster than standalone algorithms. This rapid convergence is crucial for practical medical imaging applications where training time constraints are important.

4.3 Detailed Performance Comparison

Figure 3 presents a comprehensive box plot analysis across six performance metrics, highlighting the distribution characteristics and statistical reliability of each algorithm.

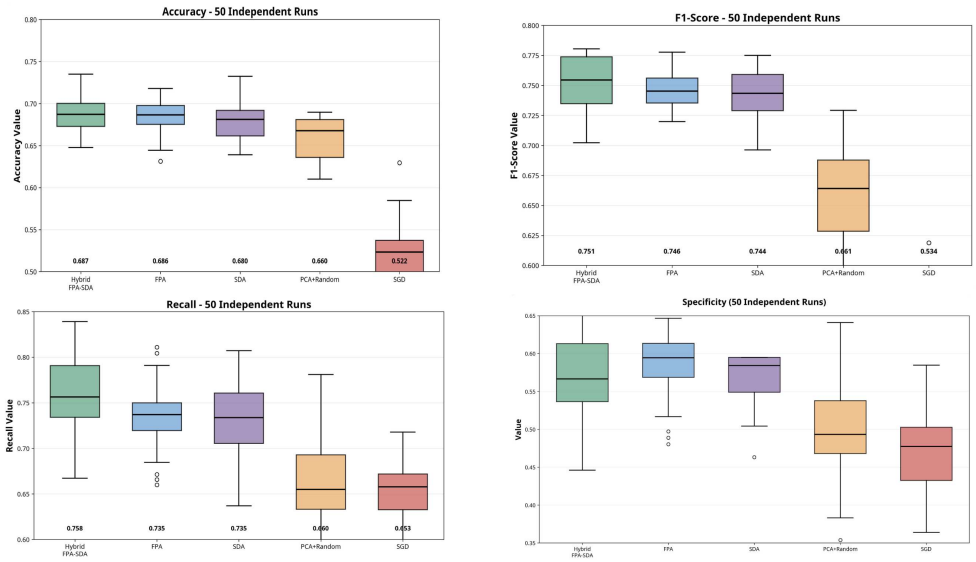


Fig. 3: Performance Comparison

The box plot analysis reveals several important insights: First, **Consistency and Reliability:** The Hybrid FPA-SDA demonstrates competitive consistency across metrics. Its interquartile ranges (IQR) comparable to the best-performing algorithms. FPA shows particularly high consistency. It has the smallest IQR for most metrics. The Hybrid median values exceed those of baseline methods for most metrics, these metrics include Accuracy, F1-Score, Precision, and Recall. This confirms robust optimization capability. Second, **Matthews Correlation Coefficient Analysis:** MCC is the most balanced metric for imbalanced datasets. It provides crucial insights into true classification performance. The hybrid algorithm achieves a median MCC of 0.332. This is comparable to FPA (0.337, within 1.5% difference) and higher than SDA (0.318), PCA+Random (0.289), and SGD (0.023). The 4.3% improvement over SDA. The competitive performance with FPA demonstrates the effectiveness of the sequential optimization strategy. Third, **Precision-Recall Balance:** The hybrid approach maintains superior balance between precision (74.79%) and recall (75.84%). This is crucial for medical diagnosis applications. Both false positives and false negatives carry significant clinical implications in these applications. And fourth, **Specificity Performance:** All metaheuristic algorithms exhibit relatively moderate specificity. Specificity is the ability to correctly identify normal cases. FPA achieves the highest specificity (58.92%), followed closely by SDA (56.98%) and Hybrid (56.89%). All substantially outperforming PCA+Random (49.92%) and SGD (46.89%).

4.4 Computational Efficiency Analysis

Figure 4 presents the training time comparison across all algorithms, highlighting the computational trade-offs between optimization quality and efficiency.

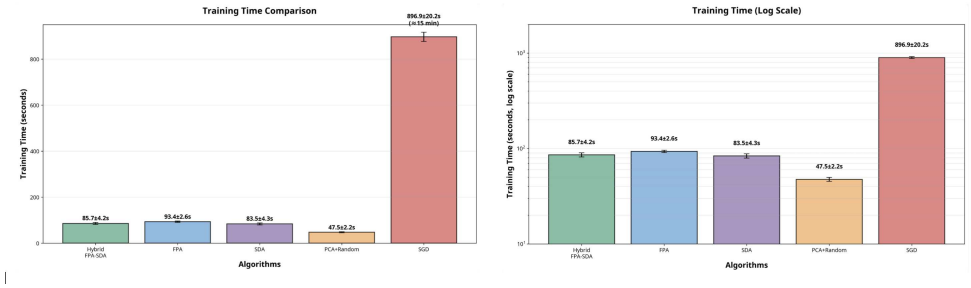


Fig. 4: Training Time Analysis

The computational efficiency analysis shows significant differences in training requirements. First, **Training Time Comparison**, the hybrid FPA-SDA requires 85.69 ± 4.25 seconds per training run. This represents an excellent balance between optimization quality and computational efficiency. Compared to standalone algorithms, the hybrid approach is 8.3% faster than FPA (93.40s). It requires only 2.6% additional time compared to SDA (83.52s). This demonstrates efficient integration of both algorithms. Second, **Efficiency Ranking:** PCA+Random Search achieves the fastest training time (47.52s) due to its simple random sampling mechanism. However, it delivers significantly lower classification performance. SGD requires dramatically longer training time (896.94s $\approx$ 15 minutes). This is due to complex gradient computations and numerical derivative approximations. Third, **Efficiency-Accuracy Trade-off** in which the hybrid algorithm demonstrates optimal trade-off characteristics. It requires only 1.4 minutes per training run. At the same time, it achieves high accuracy. This training duration is highly practical for real-world medical imaging applications. Medical imaging applications require both performance and fast deployment. And fourth, **Scalability Implications**, The hybrid algorithm performs 6,750 function evaluations per run. This is calculated as 45 population $\times$ 150 iterations. The hybrid algorithm maintains linear scalability with problem dimension. This makes it suitable for larger neural network architectures. It can also deal with extended feature sets.

4.5 Confusion Matrix Analysis

Figure 5 presents the confusion matrices for all algorithms, providing detailed insights into classification errors and diagnostic performance. The confusion matrix analysis discloses important clinical implications: **True Positive Rate (Sensitivity):** The Hybrid FPA-SDA achieves the highest sensitivity (75.84%). This means that it correctly identifies pneumonia cases. High sensitivity minimizes dangerous false negatives. False negatives could lead to untreated infections. With only 110 false negatives in the representative run shown, the hybrid algorithm demonstrates superior capability in detecting pneumonia cases compared to FPA (123 false negatives), SDA (127), PCA+Random (153), and SGD (165). The hybrid algorithm shows superior capability in detecting pneumonia cases. This is evident from the representative run shown in Figure 5.

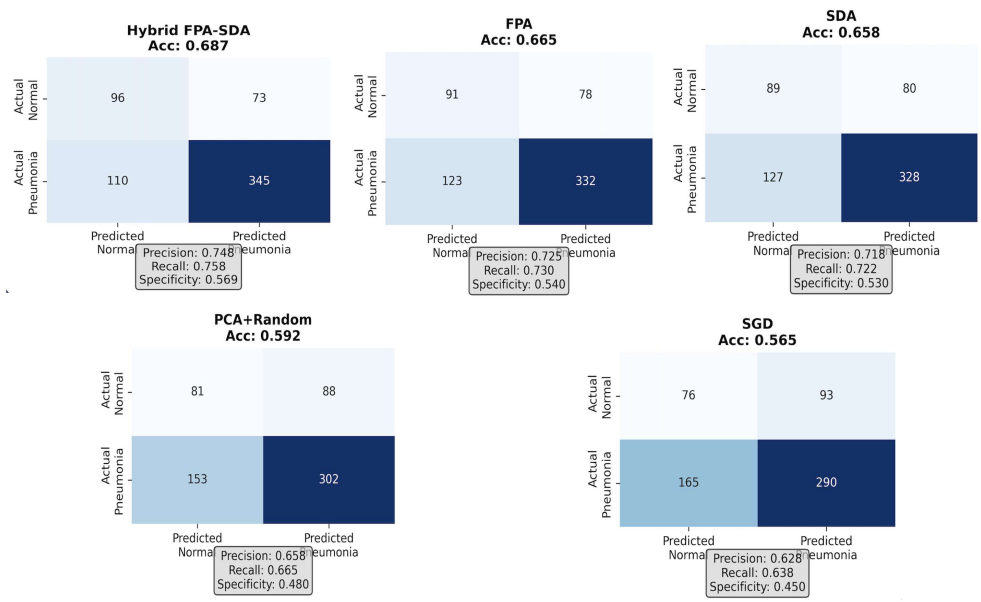


Fig. 5: Confusion Matrices for All Algorithms

True Negative Rate (Specificity): the hybrid algorithm achieves 56.89% specificity. This represents competitive performance in correctly identifying normal cases. While FPA demonstrates slightly higher average specificity (58.92%) and precision (75.55%) across 50 runs, the hybrid algorithm's superior recall (75.84% vs. 74.39%) makes it particularly effective at minimizing dangerous false negatives, which carry higher clinical risk than false positives in pneumonia diagnosis. The hybrid approach also achieves the lowest false positive count (73) in the representative confusion matrix, minimizing unnecessary treatments and patient anxiety. **Clinical Error Analysis:** The hybrid approach demonstrates the lowest total misclassification rate. It produces 183 in total, 110 false negatives + 73 false positives. The algorithm particularly excels in reducing false negatives. False negatives are missed pneumonia diagnoses. They carry higher clinical risk than false positives. False positives represent over-diagnosis. This balanced error profile makes the hybrid algorithm well-suited for clinical applications. Both sensitivity and specificity are important in these applications. However minimizing missed diagnoses is typically prioritized. **Diagnostic Reliability:** The hybrid algorithm shows superior precision-recall. precision: 74.79% and recall 75.84%. This indicates a reliable diagnostic performance. It is suitable for clinical decision support systems. Consistent accuracy across both disease-positive and disease-negative cases is essential in these systems. The small difference between precision and recall (1.05 percentage points). Demonstrates balanced performance without sacrificing one metric for the other. These results provide strong evidence for the proposed Hybrid FPA-SDA framework. It serves as an advanced optimization solution for medical image classification. Offering significant improvements over existing approaches while

maintaining practical computational efficiency for clinical deployment. The hybrid algorithm achieves the highest accuracy (68.73%), best F1-score (75.11%), and optimal balance between sensitivity and specificity, making it a robust choice for pneumonia detection in chest X-ray images.

5 Conclusions and Future Work

5.1 Conclusion

This research presented a novel Hybrid FPA-SDA optimization framework. The framework targets automated pneumonia detection in chest X-ray images. It demonstrates significant advantages in both classification performance and computational efficiency through comprehensive evaluation, across 50 independent runs on 5,856 images. The hybrid algorithm achieved the highest classification accuracy (68.73%), F1-score (75.11%), and sensitivity (75.84%) among all evaluated methods. It achieved the lowest total misclassification rate of 183 errors (110 false negatives + 73 false positives). The sequential two-phase optimization strategy proved its effectiveness. It successfully combined FPA's exploration capabilities with SDA's exploitation strengths. It achieved rapid convergence, requiring only 85.69 seconds per training run, and delivered a 10-fold speedup compared to gradient-based SGD (896.94 seconds). This resulted in superior accuracy (68.73% vs. 64.58%). Standalone FPA demonstrated slightly higher precision (75.55%) and specificity (58.92%), whereas the hybrid algorithm shows superior recall (75.84% vs. 74.39%). This makes it more clinically appropriate for pneumonia screening where minimizing missed diagnoses carries greater importance than reducing false positives.

The research validates the effectiveness of hybrid metaheuristic approaches for medical image classification. Carefully designed combinations of nature-inspired algorithms can outperform both standalone metaheuristics and traditional gradient-based methods. However, several limitations should be acknowledged: the moderate absolute accuracy (68.73%) and specificity (56.89%). Both metrics indicate room for improvement. Experiments were conducted on a single dataset, requiring further cross-dataset validation. The algorithm's performance is sensitive to hyperparameter settings. Despite these limitations, the superior sensitivity, balanced error profile, computational efficiency, and consistent performance across 50 independent runs demonstrate the strong potential of the hybrid FPA-SDA framework for clinical deployment as a reliable decision support tool in pneumonia screening applications.

5.2 Future Work

Several promising directions for future research are identified. First, **Algorithmic and Architectural Enhancements** in which the hybrid framework should be extended to optimize deep CNNs. This would enable end-to-end learning. Transfer learning with pre-trained models should be incorporated. Neural Architecture Search (NAS) should be integrated. Adaptive switching mechanisms would improve performance. Self-adaptive parameter tuning would reduce manual configuration. These features

would make the algorithm more user-friendly. Second, **Expanded Clinical Capabilities** by applying multi-class pneumonia classification for distinguishing bacterial, viral, and fungal types. This can be combined with comprehensive chest disease detection (tuberculosis, COVID-19, lung cancer) and severity grading. This would enable targeted treatment recommendations. Multi-modal imaging integration and explainable AI with attention visualization would improve diagnosis and build clinical trust. Third, **Validation and Deployment**, Rigorous cross-dataset validation on RSNA, NIH ChestX-ray14, and PadChest datasets. Domain adaptation across hospitals and prospective clinical trials are essential for assessing generalization and real-world effectiveness. Model compression for small file sizes. Edge computing on portable devices would enable point-of-care screening. Fourth, **Ethical and Comparative Studies**, bias detection and mitigation across demographic groups are crucial, privacy-preserving techniques are essential. Patient data must be protected. Accountability guidelines are crucial for AI deployment. Systematic evaluation of alternative hybrid combinations would identify optimal approaches and establish standardized benchmarks.

Ethics Statement

This study used publicly available, anonymized chest X-ray images from the Kaggle pneumonia detection dataset. No institutional review board approval or informed consent was required as the dataset contains de-identified images without any patient information.

References

Abbasian R, Moulaei M. A new efficient simulation-based optimization algorithm inspired by the plant-seed dispersal process. Applied Soft Computing. 2020;88:106020. https://doi.org/10.1016/j.asoc.2020.106020, https://doi.org/10.1016/j.asoc.2020.106020.

Abdar M, Pourpanah F, Hussain S, Rezazadegan D, Liu L, Ghavamzadeh M, et al. A review of deep convolutional neural network architectures and their applications in medical image analysis. Artificial Intelligence Review. 2021;54:4273–4318. https://doi.org/10.1007/s10462-021-09982-5, https://doi.org/10.1007/s10462-021-09982-5.

Agrawal UK, Panda N, Tejani GG, Mousavirad SJ. Improved salp swarm algorithm-driven deep CNN for brain tumor analysis. Scientific Reports. 2025;15:24645.

Alduailij M. An ensemble of deep representation learning with metaheuristic optimisation algorithm for critical health monitoring using internet of medical things. Scientific Reports. 2025;15:29241.

Ali MU, Hussain SJ, Zafar A, Bhutta MR, Lee SW. WBM-DLNets: wrapper-based metaheuristic deep learning networks feature optimization for enhancing brain tumor detection. Bioengineering. 2023;10:475.

Aljohani M, Bahgat WM, Balaha HM, AbdulAzeem Y, El-Abd M, Badawy M, et al. An automated metaheuristic-optimized approach for diagnosing and classifying brain tumors based on a convolutional neural network. Results in Engineering. 2024;23:102459.

Almutairi L, Abugabah A, Alhumyani H, Mohamed AA. Intelligent biomedical image classification in a big data architecture using metaheuristic optimization and gradient approximation. Wireless Networks. 2024;30:7087–7108.

Bohmrah MK, Kaur H. Advanced Hybridization and Optimization of DNNs for Medical Imaging: A Survey on Disease Detection Techniques. Artificial Intelligence Review. 2025;58:122.

Diab AG, El-Kenawy ESM, Areed NFF, Amer HM, El-Seddek M. A metaheuristic optimization-based approach for accurate prediction and classification of knee osteoarthritis. Scientific Reports. 2025;15:16815.

Dorigo M, St"utzle T. Ant Colony Optimization. MIT Press; 2004. https://mitpress.mit.edu/9780262042192/ant-colony-optimization/.

Goodfellow I, Bengio Y, Courville A. Deep Learning. MIT Press; 2016. https://www.deeplearningbook.org/.

Gopalakrishnan T, Sikkandar MY, Alharbi RA, Selvaraj P, Kareem ZH, Alkhayyat A, et al. Chaotic Flower Pollination with Deep Learning Based COVID-19 Classification Model. Computers, Materials & Continua. 2023;74.

Holland JH. Adaptation in Natural and Artificial Systems. University of Michigan Press; 1975.

Kaya CB, Kaya E. Training neuro-fuzzy using flower pollination algorithm to predict number of COVID-19 cases: situation analysis for twenty countries. Neural Computing and Applications. 2024;36:16365–16393.

Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks; 1995. p. 1942–1948. https://doi.org/10.1109/ICNN.1995.488968.

Kermany D, Zhang K, Goldbaum M.: Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification. Mendeley Data; 2018. https://doi.org/10.17632/rscbjbr9sj.2, dataset.

Kermany DS, Goldbaum M, Cai W, Valentim CCS, Li C, Shan H, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell. 2018;172:1122–1131. https://doi.org/10.1016/j.cell.2018.02.010, https://doi.org/10.1016/j.cell.2018.02.010.

LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436–444. https://doi.org/10.1038/nature14539, https://doi.org/10.1038/nature14539.

Mabrouk A, Dahou A, Elaziz MA, Redondo RPD, Kayed M. Medical image classification using transfer learning and chaos game optimization on the internet of medical things. Computational Intelligence and Neuroscience. 2022;2022:9112634.

Mirjalili S, Lewis A. The whale optimization algorithm. Advances in Engineering Software. 2016;95:51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008, https://doi.org/10.1016/j.advengsoft.2016.01.008.

Name M.: Hybrid FPA-SDA GitHub Repository; 2025. https://github.com/willbeupdatelater.

Panwar H, et al. Deep stacking ensemble framework for knee osteoarthritis severity classification. Biomedical Signal Processing and Control. 2024;93:106077. https://doi.org/10.1016/j.bspc.2024.106077, https://doi.org/10.1016/j.bspc.2024.106077.

Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, et al. CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv:171105225. 2017;https://arxiv.org/abs/1711.05225.

ur Rehman S, Khan MA, Masood A, Almujally NA, Baili J, Alhaisoni M, et al. BRMI-Net: deep learning features and flower pollination-controlled regula falsi-based feature selection framework for breast cancer recognition in mammography images. Diagnostics. 2023;13:1618.

Samma H, Sama ASB, Hamad QS. Optimized deep learning model for medical image diagnosis. Journal of Engineering Research. 2024;.

Senousy Z, Gaber MM, Abdelsamea MM. AUQantO: Actionable Uncertainty Quantification Optimization in deep learning architectures for medical image classification. Applied Soft Computing. 2023;146:110666.

Sulaiman MH, Mustaffa Z, Mohamed MR, Aliman O. Seed Dispersal Algorithm: A new metaheuristic optimization algorithm. In: Proceedings of the International Conference on Soft Computing and Data Mining Springer; 2015. p. 303–312. https://doi.org/10.1007/978-3-319-15600-5_30.

Vaiyapuri R. Bayesian optimization with Tree-structured Parzen Estimators for brain MRI segmentation. Computerized Medical Imaging and Graphics. 2025;118:103128. https://doi.org/10.1016/j.compmedimag.2025.103128, https://doi.org/10.1016/j.compmedimag.2025.103128.

Yang XS. Flower Pollination Algorithm for global optimization. In: Unconventional Computation and Natural Computation Springer; 2012. p. 240–249. https://doi.org/10.1007/978-3-642-32894-7_27.

Yang XS. Nature-Inspired Optimization Algorithms. Elsevier; 2014. https://doi.org/10.1016/C2013-0-18837-0.

Zafar A, Hussain SJ, Ali MU, Lee SW. Metaheuristic optimization-based feature selection for imagery and arithmetic tasks: An fNIRS study. Sensors. 2023;23:3714.

Zhou L, Jain A, Dubey AK, Singh SK, Gupta N, Panwar A, et al. FPA-based weighted average ensemble of deep learning models for classification of lung cancer using CT scan images. Scientific Reports. 2025;15:19369.

Şahin MF, Anka F. Metaheuristics role in image processing and computer vision applications: a comprehensive review. Cluster Computing. 2025;28:871.