RESEARCH ARTICLE

Editorial Process: Submission:02/12/2025 Acceptance:08/14/2025 Published:08/22/2025

AI-Powered Skin Lesion Diagnosis using Whale Optimization Algorithm Enhanced ResNet 50 for Cancer Prediction

Sabura Banu Urundai Meeran*

Abstract

Objective: The primary objective of this study is to enhance the accuracy and efficiency of binary skin lesion classification by optimizing the ResNet-50 convolutional neural network using the Whale Optimization Algorithm (WOA). This involves fine-tuning key hyperparameters such as learning rate, weights, and biases to improve predictive performance. **Methods:** This study compares five CNN architectures: AlexNet, GoogleNet, VGG16, Resnet 50, and WOA-optimized Resnet 50. The dataset comprises 3,600 balanced images (224×244 resolution) of skin moles, evenly divided into 1,800 benign and 1,800 malignant cases. The models were trained on an open-access dermoscopic dataset to categorize skin lesions. WOA was applied to optimize Resnet 50's hyperparameters weight and bias learning rate. Model performance was analysed using accuracy, preci-sion, recall, F1 score, specificity, Matthews Correlation Coefficient (MCC), log loss, AUC-ROC, and infer-ence time. The confusion matrix was analyzed to assess misclassification rates. **Result:** The WOA-optimized Resnet 50 outperformed all other models, achieving 98.29% accuracy, higher than standard Resnet 50 (90.13%), GoogleNet (87.1%), AlexNet (86.53%), and VGG16 (81.18%). It also demonstrated superior recall (99.31%), specificity (97.07%), and an AUC-ROC of 99.84%, indicating excellent classification capability. The MCC score (0.9657) confirmed strong predictive reliability. Addi-tionally, the optimized model achieved the lowest log loss (0.0512), ensuring high confidence in predictions. With an inference time of 0.1488 seconds, it was significantly faster than standard Resnet 50 (1.029 seconds), making it computationally efficient. The confusion matrix confirmed its reliability, showing min-imal false positives (7) and false negatives (2). **Conclusion:** WOA-optimized Resnet 50 significantly improves accuracy, recall, specificity, and computational efficiency for binary skin lesion classification. Compared to traditional deep learning models, it offers superior predictive performance while maintaining fast inference time. These findings suggest that WOA-enhanced deep learning can enhance dermatological diagnostics, aiding early detection and clinical decision-making. Future research may explore its application for multi-class skin lesion classification and real-time medical imaging systems.

Keywords: Skin lesion Classification- Whale Optimization Algorithm (WOA)- Deep Learning in Dermatology

Asian Pac J Cancer Prev, **26 (8)**, 2919-2928

Introduction

Skin cancer represents a growing global health concern, with millions of new cases diagnosed annually. Melanoma, though less common than other skin cancers, is the deadliest due to its aggressive nature and high metastatic potential. According to global health statistics, non-melanoma skin cancers affect over 3 million people yearly, while melanoma accounts for over 132,000 new cases worldwide. Early detection through accurate classification of skin lesions can significantly improve survival rates. As manual diagnosis by dermatologists is time-consuming and prone to inter-observer variability, automated diagnostic tools powered by artificial intelligence have become increasingly important in clinical dermatology.

In this context, deep learning has revolutionized medical image analysis, particularly through convolutional neural networks (CNNs). Pioneering architectures such as AlexNet [1], GoogLeNet [2], VGG16 [3], and Inception V3 [4] have been widely adopted for skin lesion classification. AlexNet achieved 86.53% accuracy but was constrained by its shallow structure and overfitting issues. GoogLeNet introduced inception modules to improve efficiency, reaching 87.1% accuracy. VGG16, while deeper, attained only 81.18% due to high parameter counts and overfitting. Inception V3 advanced the field with label smoothing and factorized convolutions, achieving 90.13% accuracy. However, these models often suffer from suboptimal hyperparameter tuning and limited optimization strategies. Traditional optimizers like Adam [5], RMSprop [6], and SGD [7] improve convergence but

*Department of Electrical and Electronics Engineering, Saveetha Engineering College, Saveetha Nagar, Thandalam, Chennai: 602105, Tamilnadu, India. *For Correspondence: saburasec@gmail.com*

are susceptible to local minima and poor generalization in high-dimensional spaces.

To overcome these limitations, various metaheuristic optimization techniques have been integrated with CNNs to enhance performance. Genetic Algorithms (GA) [8] and Particle Swarm Optimization (PSO) [9] were among the earliest evolutionary approaches applied to CNN training, with moderate success in im-proving parameter tuning. Swarm-based algorithms like the Firefly Algorithm [10], Ant Colony Optimiza-tion (ACO) [11], and Artificial Bee Colony (ABC) [12] have also been explored in medical imaging tasks, though they often face challenges adapting to the nonlinear complexities of deep networks. In contrast, the Whale Optimization Algorithm (WOA), introduced by Mirjalili and Lewis [13], offers superior global search capabilities by mimicking the bubble-net hunting behavior of humpback whales. WOA has demon-strated competitive performance in diverse applications, including supply chain design [14], brain tumor detection [15], and global function optimization [16]. Its effectiveness in medical imaging is further vali-dated through feature selection [17], melanoma detection [18], and CNN hyperparameter tuning [19].

Recent developments have continued to improve skin cancer classification using deep learning. For in-stance, EfficientNet models trained via transfer learning achieved high accuracy in classifying dermoscopic images [20], and hybrid CNN-transformer models were proposed for robust melanoma detection [21]. Saranya et al. [22] optimized MobileNetV3 using augmentation techniques for real-time classification, while Vidhya et al. [23] enhanced DenseNet models using metaheuristic tuning and hybrid feature fusion. Meenakshi et al. [24] incorporated the Firefly Algorithm with ResNet and attention mechanisms to improve lesion detection, and Lakshmi et al. [25] conducted a comparative analysis of multiple metaheuris-tics including WOA showing its consistent advantage in convergence and performance metrics. De-spite these advancements, existing models still fall short in generalizability and optimal hyperparameter control, particularly in the context of complex, real-world skin lesion datasets.

The aim of this study was to develop a Whale Optimization Algorithm-based deep learning framework for skin cancer classification using dermoscopic images. By integrating WOA with a deep CNN architecture, specifically ResNet-50, this research seeks to overcome the limitations of traditional training optimizers, enhance feature extraction, minimize overfitting, and improve overall classification accuracy. This study fills a critical gap in existing literature by offering a robust, adaptive, and scalable approach to skin cancer diagnosis, contributing to the advancement of intelligent clinical decision-support systems.

Materials and Methods

This section details the datasets, preprocessing steps, deep learning models, optimization strategies, and evaluation metrics employed in this study. To ensure a robust and reproducible framework for skin lesion classification, the proposed approach integrates traditional convolutional neural networks with metaheu-ristic optimization. Special emphasis is placed on the Whale Optimization Algorithm (WOA) for fine-tuning hyperparameters of the ResNet 50 architecture, thereby enhancing its predictive performance. The meth-odology is organized into distinct phases data acquisition and preparation, model adaptation, optimiza-tion, training, and evaluation each contributing to the overall effectiveness and reliability of the classification system.

Materials: Dataset Description

The “Medical Imaging (CT scan, MRI, X-ray, and Microscopic Imagery) Data” dataset [26], published on July 11, 2024, by Sibtain Syed, Rehan Ahmed, and Arshad Iqbal, includes a subset dedicated to skin le-sions. The dataset is publicly available under the Creative Commons Attribution 4.0 International (CC BY 4.0) license, promoting its use in research and development of computer-aided diagnostic tools.

It comprises five distinct subsets of medical images, each targeting a specific medical condition:

1. Lung Cancer: CT scan images for lung cancer classification.
2. Bone Fracture: X ray images for detecting bone fractures.
3. Brain Tumor: MRI images for brain tumor identification.
4. Skin Lesions: Microscopic images for classifying skin lesions.
5. Renal Malignancy: CT scan images for renal malignancy detection.

Each subset includes images of both diseased and healthy tissues, facilitating binary classification tasks. This subset comprises microscopic images aimed at facilitating the classification of skin lesions into benign or malignant categories. The dataset is suitable for training deep convolutional neural network (DCNN) models, such as ResNet50, for disease classification. The contributors have utilized this dataset in their research, indicating its applicability in medical image analysis. The data set consists of 1440 Benign images and 1197 Malignant images.

This study utilizes deep learning models for skin cancer classification, leveraging publicly available datasets of dermoscopic images. Figure 1 shows the sample of the images used in this research under the category of benign and malignant. These datasets contain both benign and malignant skin lesion images, preprocessed for optimal training and validation. Image preprocessing techniques include resizing, nor-malization, augmentation (rotation, flipping, contrast enhancement), and segmentation to enhance fea-ture extraction.

Methodology

Data Preprocessing

All images are converted to a standardized 224x224 pixels to maintain consistency across deep learning models. To enhance model stability and convergence,

Figure 1. Sample of Images Used in This Research

pixel values are scaled to the range [0,1]. Addition-ally, a variety of augmentation techniques, including rotation, flipping, zooming, and brightness adjust-ment, are applied to reduce overfitting and improve the model's ability to generalize. To ensure a com-prehensive evaluation, the dataset is divided into three subsets: 80% for training, 10% for validation, and 10% for testing.

Pretrained Deep Learning Models for Skin Lesion Classification

AlexNet, GoogleNet, ResNet-50, and VGG16 are widely used deep learning models for binary skin lesion classification, leveraging their convolutional architectures to predict the skin lesions. In this study, these models are compared with the proposed Whale Optimization Algorithm (WOA)-tuned ResNet-50, which aims to enhance classification performance by optimizing hyperparameters and network weights. The comparative analysis evaluates key performance metrics, demonstrating the effectiveness of the WOA-optimized ResNet-50 in improving accuracy, precision, and overall model robustness for skin lesion detec-tion.

Model Tuning and Training

To enhance the reproducibility and clarity of the methodology, we explicitly structured the workflow into four key phases: (1) Data Preprocessing, where images were resized to 224×224 pixels, normalized to [0,1], and augmented using rotation, flipping, and brightness adjustments; (2) Model Adaptation, where the final fully connected layers of the pretrained networks (AlexNet, GoogleNet, VGG16, ResNet-50) were replaced with a binary classification layer; (3) Hyperparameter Optimization using Whale Optimization Algorithm (WOA), where weight and bias learning rates were dynamically adjusted through WOA's encir-cling, bubble-net attacking, and search mechanisms; and (4) Performance Evaluation, where metrics such as accuracy, recall, precision, specificity, AUC-ROC, and inference time were calculated using the test set. This structured approach ensures that each experimental step is transparent and replicable.

The last fully connected layers of each model are replaced with new layers tailored for binary classifica-tion (benign vs. malignant). The Adam optimizer is employed with a learning rate of 0.0001 to facilitate stable and efficient training. Binary Cross-Entropy loss is used to handle the two-class classification prob-lem. The experiments are conducted on a high-performance GPU to accelerate the training and inference processes.

Walrus Optimization Algorithm (WaOA) for Hyperparameter Tuning of ResNet50 in Renal Malignancy Detection

Skin Lesion prediction using deep learning models requires high accuracy and robust generalization to minimize false positives and false negatives. While ResNet-50 is a powerful deep CNN, its performance heavily depends on optimal hyperparameter selection (learning rate, batch size, weight decay, etc.). Man-ual tuning or traditional optimizers (e.g., Adam, SGD) often get stuck in local minima, leading to subopti-mal performance. The Whale Optimization Algorithm (WOA), a bio-inspired metaheuristic technique, op-timizes ResNet-50 by automatically tuning hyperparameters, leading to enhanced feature extraction, fast-er convergence, and improved classification accuracy. Modeled after the bubble-net hunting strategy of humpback whales, WOA efficiently refines the network's performance.

The algorithm consists of three main phases:

• Encircling Prey: Whales identify the optimal solution (best hyperparameter set) and ad-just their position accordingly.

• Bubble-Net Attacking Mechanism: Exploits spiral motion to balance exploration and ex-ploitation during training.

• Search for Prey: Ensures global search capability by moving towards new potential solu-tions in the hyperparameter space.

Mathematically, WOA updates the position of whales (candidate solutions) using:

$$X(t+1) = X(t) - AD$$

where:

• X is the best-known solution.

• A and D are adaptive coefficients controlling exploration and exploitation.By integrating WOA with ResNet-50, hyperparameters such as learning rate, batch size, and optimizer settings are fine-tuned dynamically, reducing misclassification errors in skin lesion detection.

Flowchart of the Whale Optimization Algorithm (WOA)-Tuned ResNet-50 Pretrained Network shown in Figure 2 for Skin Lesion Classification outlines a structured approach to optimizing deep learning based skin cancer detection. It begins with data preprocessing, where images are resized, normalized, and aug-mented to enhance model generalization.

The ResNet-50 architecture, pretrained on ImageNet, is then loaded, with modifications to its fully con-nected layer for skin lesion classification. To improve performance, WOA optimizes key hyperparameters such as learning rate, batch size, and weight decay through iterative encircling, bubble-net attacking, and search phases. Once optimal hyperparameters are determined, the fine-tuned ResNet-50 model is trained and validated, ensuring reduced misclassification rates. The model is then evaluated on unseen test data, where metrics like accuracy, precision, recall, and AUC-ROC confirm its superiority over baseline models

Technologies Used

• Deep Learning Frameworks: TensorFlow, Keras, PyTorch.

• Hardware: NVIDIA GPUs for accelerated training.

• Optimization Algorithms: Adam, SGD, and WOA.

• Evaluation Metrics: Accuracy, Precision, Recall, F1 Score, AUC-ROC, Log Loss, MCC.

Results

The accurate and efficient classification of skin lesions is crucial for early skin cancer detection and effec-tive treatment. Traditional deep learning models such as AlexNet, GoogleNet, VGG16, and ResNet-50

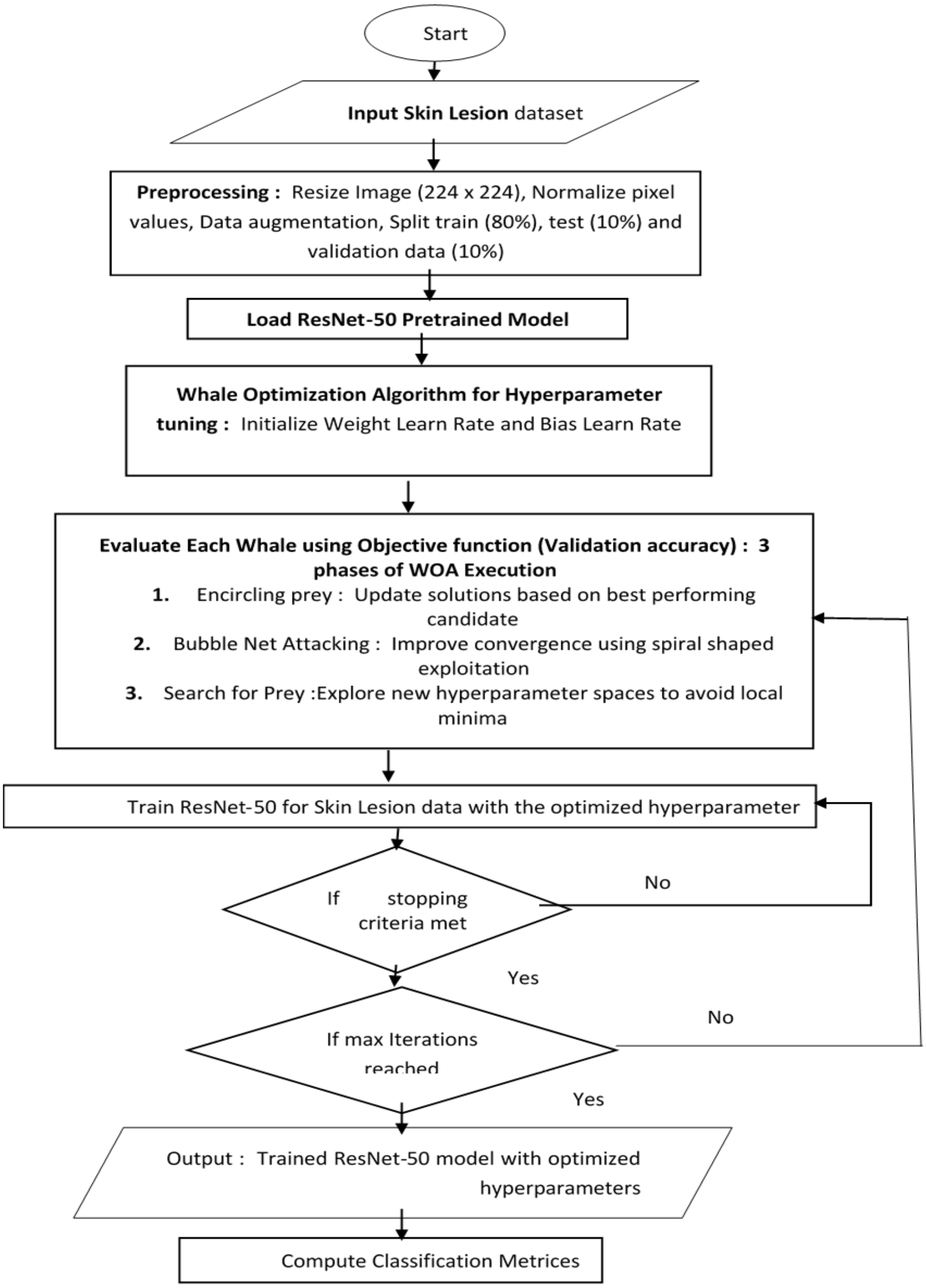

Figure 2. Flowchart of WOA-Optimized ResNet-50

have demonstrated strong potential in medical image classification, yet optimizing their performance re-mains a challenge. This study introduces the Whale Optimization Algorithm (WOA) to fine tune ResNet50, enhancing its classification accuracy, recall, specificity, and computational efficiency.

The following results and discussion section presents a comprehensive evaluation of these models, high-lighting the superior performance of WOA optimized ResNet50 compared to conventional CNN architec tures. The findings underscore the effectiveness of metaheuristic optimization in medical imaging and suggest its potential for further advancements in automated dermatological diagnostics.

Table 1 shows the confusion matrix for the pretrained deep learning networks and the proposed WOA tuned Resnet50 network. Examining the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values: WOA-ResNet-50 achieves 286 TP and 232 TN, showing it correctly identifies most cases, It has the lowest FN (2), proving its exceptional sensitivity. Other models, especially VGG-16 and AlexNet, have significantly higher FN values (40 and 55, respectively), meaning they miss many can-cer cases.

Table 2 presents a comprehensive evaluation of five deep learning models AlexNet, GoogLeNet, VGG16, ResNet50, and the Whale Optimization Algorithm (WOA) optimized ResNet50 for skin cancer classification. Various performance metrics are compared, including accuracy, precision, recall, F1-score, speci-ficity, MCC, log loss, AUC-ROC, inference time, confusion matrix values, and statistical measures. The analysis highlights significant differences in performance, particularly showcasing the superiority of the WOA-optimized ResNet-50.

Among the evaluated models, GoogLeNet demonstrated moderate classification performance with an accuracy of 87.1% and a recall of 89.24%. Its inception modules allowed better feature extraction effi-ciency than AlexNet, yet the relatively shallow depth of the network limited its ability to capture complex patterns in dermoscopic images. This is evident in its false negative count of 31, indicating it failed to correctly identify several malignant cases, which could pose a risk in clinical scenarios. On the other hand, VGG16, while being a deeper network with uniform kernel sizes, underperformed with an accuracy of 81.18% and the highest log loss (0.6078), reflecting lower confidence in its predictions. Its 59 false posi-tives and 40 false negatives reveal a tendency toward both over-diagnosing and under diagnosing lesions.

These limitations likely stem from VGG16's high parameter count, which increases susceptibility to over fitting, especially with a moderate sized dataset. Together, these results highlight the importance of balancing depth, architecture efficiency, and optimization for effective skin lesion classification.

Accuracy Comparison

Accuracy is a crucial metric indicating the proportion of correctly classified cases. Figure 3 shows the comparison of accuracy of the pretrained Deep Learning techniques with the proposed WOA tuned ResNet50. The WOA-optimized ResNet-50 achieves the highest

Table 1. Confusion Matrix of the Pretrained Deep Learning Techniques with the Proposed WOA Tuned ResNet50

Pretrained Deep Learning Technique	Confusion Matrix			
Alexnet			Actual	
			Benign	Malignant
	Predicted	Benign	233	55
		Malignant	16	223
Googlenet				
			Actual	
			Benign	Malignant
	Predicted	Benign	257	31
		Malignant	37	202
ResNet-50				
			Actual	
			Benign	Malignant
	Predicted	Benign	237	51
		Malignant	1	238
VGG16			Actual	
			Benign	Malignant
	Predicted	Benign	248	40
		Malignant	59	180
WOA Tuned RESNET-50			Actual	
			Benign	Malignant
	Predicted	Benign	286	2
		Malignant	7	232

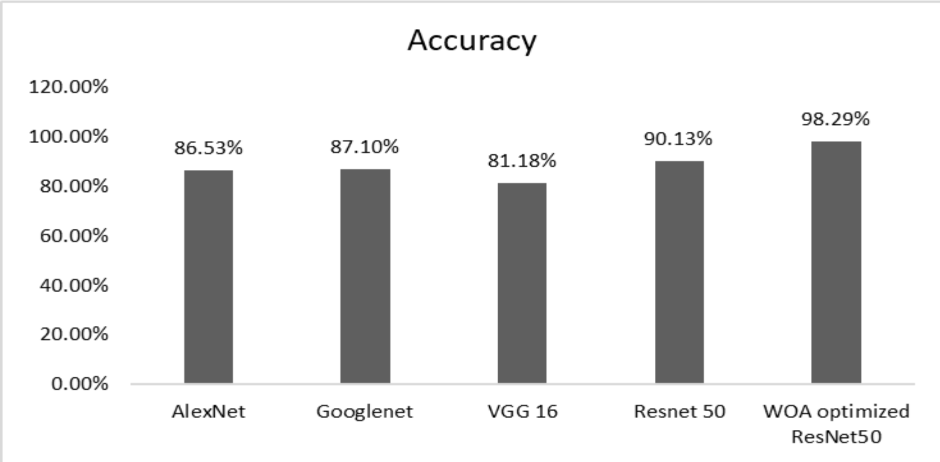

Figure 3. Accuracy Comparison of the Pretrained Deep Learning Techniques with the Proposed WOA Tuned ResNet-50.

Table 2. Performance Metrices of the Pretrained Deep Learning Techniques with the Proposed WOA Tuned ResNet-50.

	AlexNet	Googlenet	VGG 16	Resnet 50	Proposed WOA optimized Resnet 50
Accuracy	86.53%	87.1%	81.18%	90.13%	98.29%
Precision	93.57%	87.41%	80.72%	99.58%	97.61%
Recall	80.90%	89.24%	86.06%	82.29%	99.31%
F1 Score	86.78%	88.32%	83.31%	90.11%	98.45%
Specificity	93.31%	84.52%	75.31%	99.58%	97.07%
Matthews Correlation Coefficient (MCC)	0.7400	0.7393	0.6195	0.8190	0.9657
Log Loss	0.3404	0.2928	0.6078	0.2408	0.0512
AUC-ROC	94.78%	94.73%	90.56%	99.47%	99.84%
Average Inference Time (sec)	0.034	0.1822	0.3571	1.0290	0.1488

accuracy (98.29%), significantly outperform-ing standard ResNet50 (90.13%) and other models such as GoogLeNet (87.1%), AlexNet (86.53%), and VGG16 (81.18%). The optimization algorithm significantly enhances ResNet50's classification ability, making it the most reliable model.

Precision, Recall and F1-Score Analysis

Precision: Measures the proportion of correctly identified positive cases out of all predicted positives. Figure 4 shows the comparison of precision of pretrained Deep Learning techniques with the proposed WOA tuned ResNet50. WOA ResNet50 achieves an outstanding precision of 97.61%, demonstrating its ability to minimize false positives. This is far superior to VGG16 (80.72%) and standard ResNet50 (99.58%, which may suggest an aggressive classification that slightly reduces recall).

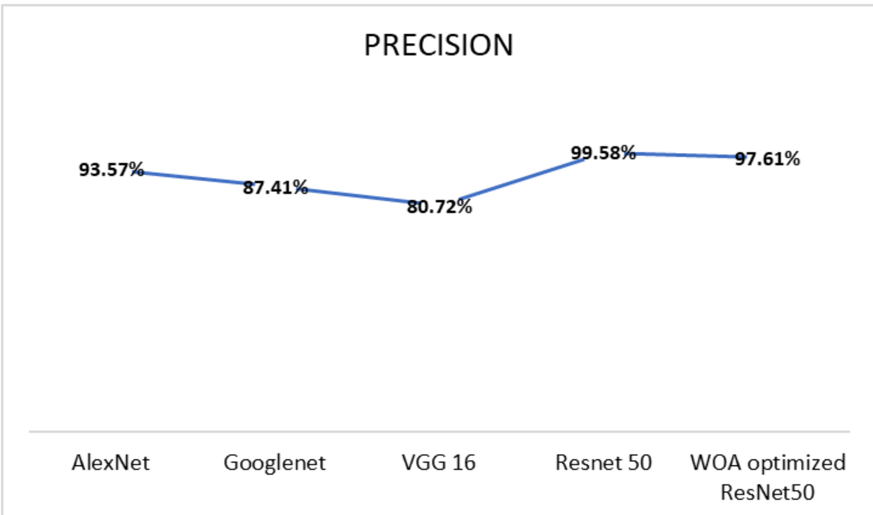

Figure 4. Precision Comparison of the Pretrained Deep Learning Techniques with the Proposed WOA Tuned ResNet50.

Recall

Represents the proportion of correctly identified actual positive cases. Figure 5 shows the comparison of recall of pretrained Deep Learning techniques with the proposed WOA tuned ResNet50. WOA ResNet50 outperforms all models with an excellent 99.31%, ensuring that nearly all true cancer cases are identified. In contrast, standard ResNet-50 achieves 82.29%, and AlexNet falls behind at 80.90%.

F1-Score

A balance between precision and recall, WOA ResNet50 again leads with 98.45%, confirming its robustness. Other models struggle, such as VGG16 with 83.31%, which suggests a weaker balance be-tween precision and recall. Supplementary Figure 1 shows the comparison of F1-Score of pretrained Deep Learning tech-niques with the proposed WOA tuned ResNet50.

Specificity

Specificity indicates the model's ability to correctly identify negative cases (non-cancerous). Supplementary Figure 2 shows the comparison of specificity of pretrained Deep Learning techniques with the proposed WOA tuned ResNet50. WOA ResNet50 scores 97.07%, second only to standard ResNet50 (99.58%), demon-strating minimal false positives.

False Positives (FP) and False Negatives (FN)

Table 3 show TP, TN, FP and FN comparison of the pretrained Deep Learning techniques with the proposed WOA tuned ResNet50. WOA ResNet50 has only 7 false positives and 2 false negatives, the lowest among all models. In contrast, VGG16 has 59 false positives and 40 false negatives, indicating a high misclassification rate. Google Net and AlexNet also show moderate misclassification rates, with false negatives of 31 and 55, respectively. One of the most significant outcomes observed was the remarkably low false negative rate of the WOA optimized ResNet50 model, with only 2 false negatives compared to 55 for AlexNet and 51 for standard ResNet50. This indicates the model's exceptional sensitivity in identifying malignant cases, which is critical in a clinical setting where missing a cancer diagnosis can have severe consequences. Furthermore, while the standard ResNet50 model achieved a high specificity of 99.58%, it came at the cost of a much higher inference time (1.029 seconds). The WOA optimized model, in con-trast, struck a balanced trade-off achieving 97.07% specificity with a drastically reduced inference time of 0.1488 seconds. This makes it not only highly accurate but also suitable for real-time diagnostic applications. Additionally, the high Matthews Correlation Coefficient (0.9657) of the WOA ResNet50 supports the reliability of its predictions across both classes, affirming its robustness even in imbalanced datasets.

Mathews Correlation Coefficient (MCC) and Log Loss

MCC, which measures overall classification performance (where 1 indicates perfect classification), shows that WOA-ResNet-50 (0.9657) significantly outperforms ResNet-50 (0.8190) and other models, proving its superior reliability. Supplementary Figure 3 shows the comparison of MCC of pretrained Deep Learning techniques with the proposed WOA tuned ResNet-50.

Log Loss, a measure of classification confidence, is

Table 3. TP, TN, FP and FN Comparison of the Pretrained Deep Learning Techniques with the Proposed WOA Tuned ResNet-50.

	AlexNet	Googlenet	VGG 16	Resnet 50	Proposed WOA optimized Resnet 50
True Positive	233	257	247	237	286
True Negative	223	202	180	238	232
False Positive	16	37	59	1	7
False Negative	55	31	40	51	2

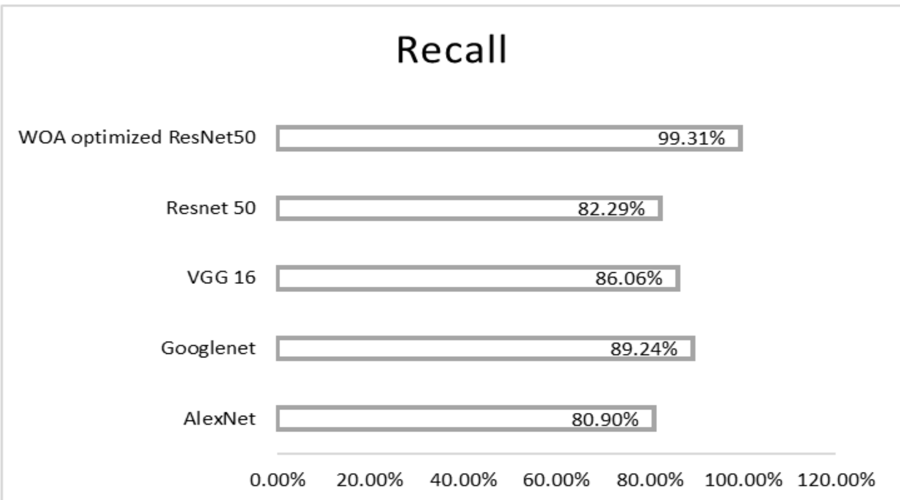

Figure 5. Recall Comparison of the Pretrained Deep Learning Techniques with the Proposed WOA Tuned ResNet50.

lowest for WOA ResNet50 (0.0512), signifying fewer misclassified cases. VGG16 has the highest log loss (0.6078), suggesting poor confidence in its predic-tions. Supplementary Figure 4 shows the comparison of Log Loss of pretrained Deep Learning techniques with the proposed WOA tuned ResNet50.

AUC-ROC Analysis

AUC-ROC, which measures the model's ability to distinguish between classes, is highest for WOA ResNet50 (99.84%), confirming its superior discriminative power. Standard ResNet50 follows closely at 99.47%, while AlexNet and GoogLeNet perform at 94.78% and 94.73%, respectively. VGG16 scores 90.56%, reflecting its struggles in classification.

Supplementary Figure 5 shows the comparison of AUC-ROC of pretrained Deep Learning techniques with the proposed WOA tuned ResNet50. Supplementary Figure 6 shows the ROC curve for benign and malignant class with the proposed WOA tuned ResNet50. The ROC curve indicates that the classification model demonstrates excellent performance in distinguishing between benign and malignant skin lesions, as both curves are near the top-left corner, signifying high true positive rates and low false positive rates.

Average Inference Time (Computation Efficiency)

Inference time is essential for real-time medical applications. Supplementary Figure 7 shows the comparison of Average Inference Time of pretrained Deep Learning techniques with the proposed WOA tuned ResNet50. The fastest model is AlexNet (0.034 sec), but at the cost of lower accuracy. WOA ResNet50 takes 0.1488 sec, demonstrating efficient classification without compromising accuracy. ResNet-50 (1.029 sec) is the slowest, suggesting that WOA significantly enhances computation speed.

The spider plot (radar chart) displayed in the Supplementary Figure 8 provides a comparative visualization of multiple performance metrics for five deep learning models: AlexNet, GoogLeNet, VGG16, ResNet50, and WOA Optimized ResNet50. The WOA-Optimized ResNet50 (purple) shows the highest values for key metrics like Accuracy, Precision, Recall, F1 Score, AUC-ROC, and MCC, making it the best-performing model overall. It maintains high specificity and a low log loss, further reinforcing its robustness. ResNet50 and WOA-Optimized ResNet50 outperform the other models in terms of Accuracy. GoogleNet and AlexNet have relatively good Precision and Recall, but not as high as ResNet50 and WOA-Optimized ResNet50. VGG16 lags behind in Accuracy and Specificity, indicating weaker classification performance. ResNet50 has the highest inference time (~1.029 sec), making it computationally expensive. The WOA-Optimized ResNet50 significantly reduces inference time, making it more efficient than standard ResNet50 while maintaining better performance. AlexNet has the lowest inference time (~0.034 sec), but at the cost of reduced accu-racy. Lower log loss values indicate better model confidence in predictions. WOA-Optimized ResNet50 has the lowest log loss (0.0512), while VGG16 has the highest log loss (0.6078), indicating a higher uncer-tainty in VGG16's predictions. WOA-Optimized ResNet50 has the highest AUC-ROC (0.9984), meaning it achieves the best trade-off between sensitivity and specificity. Specificity is highest for ResNet50 and WOA-Optimized ResNet50, suggesting that they minimize false positives effectively. VGG16 has poor specificity, indicating a higher rate of false positives.

Discussion

WOA-Optimized ResNet50 outperforms all models in nearly every metric while maintaining lower infer-ence time and log loss. ResNet50 performs well but suffers from high inference time. AlexNet and GoogleNet are decent but not as strong in overall performance. VGG16 struggles the most, with lower accuracy, specificity, and higher log loss. This analysis highlights the effectiveness of the Whale Optimization Algorithm (WOA) in optimizing ResNet50, making it a strong candidate for practical applications requiring high classification accuracy and efficiency.

The findings of this study highlight the effectiveness of the Whale Optimization Algorithm (WOA) in fine-tuning deep learning models for skin lesion classification. By optimizing ResNet50's hyperparameters, WOA significantly enhanced its accuracy, recall, specificity, and overall predictive reliability, outperform-ing traditional CNN architectures such as AlexNet, GoogleNet, VGG16, and standard ResNet50. The WOA-optimized ResNet50 not only achieved the highest classification performance with an accuracy of 98.29% and an AUC-ROC of 99.84% but also maintained computational efficiency with a reduced inference time of 0.1488 seconds. These improvements suggest that metaheuristic optimization techniques can play a crucial role in advancing deep learning-based medical diagnostics. The model's superior performance in minimizing false positives and false negatives further demonstrates its reliability for early skin cancer detection, which is vital for timely and effective treatment.

Given these promising results, future research could explore WOA optimization for multi-class skin lesion classification and integration with real-time medical imaging systems, potentially revolutionizing auto-mated dermatological diagnostics and clinical decision-making. While the current study successfully demonstrates the effectiveness of the WOA optimized ResNet50 for binary classification of skin lesions (benign vs. malignant), it opens up promising avenues for future work. Given the model's strong performance in accurately differentiating between two classes, its underlying architecture and optimization strategy suggest potential for adaptation to more complex classification tasks. In particular, extending this approach to multi-class skin lesion classification encompassing various lesion types such as melanoma, basal cell carcinoma, and squamous cell carcinoma could significantly enhance its clinical utility. This transition would require curated multi-class datasets and further fine-tuning but remains a feasible and valuable direction for advancing automated dermatological diagnostics.

Author Contribution Statement

As a single author manuscript, all components of the study-including conceptualization, methodology, implementation, analysis, interpretation, and manuscript writing were solely undertaken by the author.

Acknowledgements

General
The author sincerely acknowledges the facilities and support provided by Saveetha Engineering College during the course of this research.

Scientific Approval
This work was conducted as part of an approved faculty-led research project within Saveetha Engineering College. It is not associated with a student thesis submission or review by any external scientific body.

Ethical Declaration
This study involved retrospective analysis of openly available and anonymized medical imaging datasets and did not require patient consent or direct ethical clearance. The dataset used complies with public data usage standards under Creative Commons Attribution 4.0 International License. Therefore, ethical approval from a specific institutional review board was not required.

Data Availability
The dataset used in this study, titled “Medical Imaging (CT scan, MRI, X-ray, and Microscopic Imagery) Data”, is publicly available at Mendeley Data [https://data.mendeley.com/datasets/5kbjrgsncf/3], under a CC BY 4.0 license.

Study Registration
Not applicable. This study does not involve clinical trials or patient registration and was not submitted to a clinical study registry.

Conflict of Interest
The author declares no conflict of interest.

References

1. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural net-works. Advances in neural information processing systems. 2012;25.
2. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. InProceedings of the IEEE conference on computer vision and pattern recognition 2015; pp. 1-9.
3. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014 Sep 4.
4. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. InProceedings of the IEEE conference on computer vision and pattern recogni-tion 2016; pp. 2818-2826.
5. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. 2014 Dec 22.
6. Tieleman T, Hinton G. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural networks for machine learning. 2012;4(2):26.
7. Bottou L. Stochastic gradient descent tricks. In: Montavon G, Orr GB, Müller KR, editors. Neural Networks: Tricks of the Trade. Berlin: Springer; 2012. p. 421–36.
8. Holland JH. Adaptation in natural and artificial systems: an introductory analysis with applica-tions to biology, control, and artificial intelligence. MIT press; 1992 Apr 29.
9. Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of the IEEE International Con-ference on Neural Networks. 1995.
10. Yang XS. Firefly algorithm for optimization. In: Nature-Inspired Metaheuristic Algorithms. Berlin: Springer; 2010. p. 79–90.
11. Dorigo M, Di Caro G. Ant colony optimization: a new meta-heuristic. IEEE Transactions on Evolu-tionary Computation. 1999;1(1):53–66.
12. Karaboga D. An idea based on honey bee swarm for numerical optimization. Technical Report. 2005.
13. Mirjalili S, Lewis A. The whale optimization algorithm. Adv Eng Softw. 2016;95:51-67. https://doi.org/10.1016/j.advengsoft.2016.01.008
14. Fathollahi-Fard M, Hajiaghaei-Keshteli M, Tavakkoli-Moghaddam R, Safaeian M. A whale optimi-zation algorithm for supply chain network design. Appl.Soft Comput. 2018;74:326–46.
15. Sharma A, Reddy G, Jha D, Anand V, Singh G. Whale optimization algorithm-based deep learning approach for brain tumor detection. IEEE Access. 2020;8:115769–81.
16. Heidari AA, Mirjalili S, Faris H, Aljarah I, Mafarja MM, Chen H. Enhanced whale optimization algo-rithm for global optimization. Expert Syst Appl. 2019;114:259–77.
17. Zhang Z, Li W, Yang X, Wang J. Whale optimization algorithm-based feature selection for medical image analysis. Neural Comput Appl. 2021;33(18):11329–44.
18. Mirzaei MR, Salimi M, Abbasi S, Zolfaghari S. Improved deep learning using whale optimization algorithm for melanoma detection. Sensors (Basel). 2021;21(21):7032.
19. Tizhoosh H, Rahnamayan S, Faez K. Whale optimization algorithm-based hyperparameter tuning for convolutional neural networks. Pattern Recognit Lett. 2022;157:78–85.
20. Kanchana K, Kavitha S, Anoop KJ, Chinthamani B. Enhancing skin cancer classification using effi-cient net b0-b7 through convolutional neural networks and transfer learning with patient-specific data. Asian Pac J Cancer Prev. 2024;25(5):1795-802. https://doi.org/10.31557/apjcp.2024.25.5.1795.
21. Deepika R, Sivakumar R, Manikandan R. Hybrid deep learning model for melanoma detection using ensemble of convolutional neural networks and transformer networks. Biomed Signal Pro-cess Control. 2024;87:105387.
22. Saranya P, Prabha S, Jayalakshmi V. Optimized MobileNetV3 architecture for real-time skin lesion classification using augmented dermoscopy images. Comput Biol Med. 2024;165:107630.
23. Vidhya R, Priya M, Sakthi S. Improved skin cancer prediction using metaheuristic-tuned DenseNet model and hybrid feature fusion strategy. Healthc Anal. 2023;5:100157.
24. Meenakshi S, Arun S, Maheswari K. Firefly algorithm-based multi-class skin lesion detection using ResNet and attention modules. Expert Syst Appl. 2023;226:119590.
25. Lakshmi R, Raj Kumar N, Banu US. Comparative analysis of metaheuristic algorithms in optimizing deep learning models for dermatological diagnostics. J Imaging and Biomed Inform. 2024;3(2):75–83.
26. Syed S, Ahmed R, Iqbal A. Medical imaging (CT scan, MRI,

X-ray, and microscopic imagery) data [Internet]. Mendeley Data, Version 3. 2024 [cited 2024 Feb 10]. Available from: https://data.mendeley.com/datasets/5kbjrgsncf/3.