**Bulletin of Electrical Engineering and Informatics**
Vol. 15, No. 3, June 2026, pp. 2263~2274
ISSN: 2302-9285, DOI: 10.11591/eei.v15i3.11181 

# Osprey optimization algorithm for VGG16 hyperparameter optimization in breast cancer detection

**Sabura Banu Urundai Meeran[1], Nafeena Abdul Munaf[2], Vengadeshwaran Velu[3]**
[1]Department of Electrical and Electronics Engineering, Saveetha Engineering College, Chennai, India
[2]Faculty of Engineering and Technology, Villa College, Malé, Maldives
[3]Director of Research Management Centre (RMC), Mila University, Nilai, Malaysia

**Article Info**

***Article history:***

Received Aug 9, 2025
Revised Mar 26, 2026
Accepted Apr 19, 2026

***Keywords:***

Bio-inspired algorithms
Breast cancer detection
Computer-aided diagnosis
Convolutional neural networks
Deep learning
Hyperparameter optimization
osprey optimization algorithm
medical image analysis

**ABSTRACT**

Globally, breast cancer is one of the reason for mortality among women and accurate automated diagnosis remains a critical research challenge. This research is used to improve breast cancer classification performance by optimizing deep learning (DL) model hyperparameters using a bio-inspired optimization technique. The osprey optimization algorithm (OOA) is applied to fine-tune the hyperparameters of the VGG16 convolutional neural network (CNN) for histopathological breast cancer image classification. The optimized model is evaluated using a curated dataset and compared with established DL architectures, including AlexNet, Xception, InceptionV3, and ResNet50. Performance is assessed using standard evaluation metrics such as accuracy, precision, recall, F1-score, specificity, AUC-ROC, Matthews correlation coefficient (MCC), log loss, and inference time. Experimental results indicate that the OOA-optimized VGG16 model achieves superior performance, with an accuracy of 97.7%, precision of 96.71%, recall of 97.79%, AUC-ROC of 99.92%, and MCC of 0.9449, while maintaining competitive computational efficiency. The results demonstrate that bio-inspired hyperparameter optimization significantly enhances classification reliability and diagnostic accuracy. In summary, integrating OOA optimization with the VGG16 architecture yields a dependable framework for breast cancer identification, making it a promising candidate for deployment in automated diagnostic support systems.



***Corresponding Author:***

Sabura Banu Urundai Meeran
Department of Electrical and Electronics Engineering, Saveetha Engineering College
Saveetha Nagar, Thandalam, Chennai: 602105, Tamilnadu, India
Email: saburasec@gmail.com

## 1. INTRODUCTION

Globally, breast cancer remains a critical public health challenge, with high prevalence and a substantial impact on both disease burden and cancer-related deaths among women [1]. Recent global cancer reports emphasize that early diagnosis significantly improves survival rates, reduces treatment complexity, and lowers healthcare burden [2]. However, conventional diagnostic workflows—such as histopathological assessment and radiological interpretation—are often manual, time-intensive, and susceptible to inter-observer variability, especially in large-scale screening programs [3]. To mitigate these drawbacks, researchers have increasingly adopted computer-aided diagnosis systems powered by deep learning (DL) for automated breast cancer detection and classification tasks [4]. Convolutional neural network (CNN), in particular, excel at capturing complex spatial information from medical imaging data, often delivering improved accuracy compared with conventional machine learning techniques that rely on manually designed features [5]. Foundational CNN architectures such as AlexNet [6] and VGG16 [7] established the feasibility of deep hierarchical feature extraction, while subsequent models including InceptionV3 [8], ResNet50 [9], and Xception [10] further improved classification accuracy through architectural innovations such as residual learning and depthwise separable convolutions.

Despite these advancements, CNN performance is strongly influenced by the selection of hyperparameters such as learning rate, batch size, and dropout probability, optimizer selection, and weight initialization strategies [11]. The effects of deep learning assistance on the detection of esophageal cancer on CT were investigated, evaluating performance across both experienced radiologists and radiology residents. The findings highlighted the potential of AI-assisted interpretation in thoracic oncology, demonstrating that diagnostic accuracy could be meaningfully augmented regardless of reader experience level [12]. Conventional tuning strategies such as grid search, random search, and Bayesian optimization become computationally prohibitive as model complexity and parameter dimensionality increase [13]. Consequently, bio-inspired metaheuristic optimization algorithms have gained significant attention for hyperparameter optimization in DL models [14]. Techniques such as genetic algorithms [15], marine predators optimization [16], social ski-driver optimization [17], and manta-ray foraging optimization [18] have demonstrated effectiveness in navigating large and nonlinear search spaces. More recent studies (2022–2024) report that hybrid and nature-inspired optimizers consistently outperform conventional tuning approaches in medical image classification tasks by improving convergence speed and classification robustness [19]–[22]. The osprey optimization algorithm (OOA) is a newly proposed bio-inspired metaheuristic that models the hunting behavior of ospreys, incorporating adaptive exploration and exploitation mechanisms [23]. Preliminary studies have shown that OOA exhibits strong global search capability, avoids premature convergence, and achieves superior optimization performance compared to classical metaheuristics in DL applications [24], [25]. However, its application to breast cancer histopathological image classification remains largely unexplored.

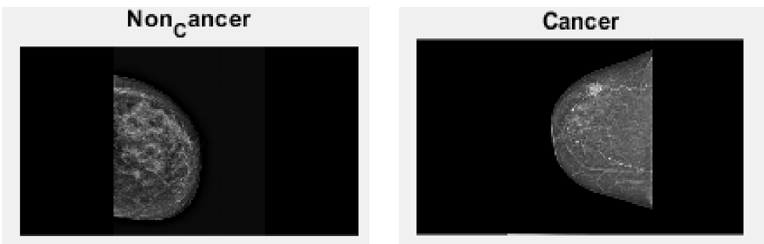
This study proposes the application of the OOA for hyperparameter optimization of the VGG16 network in the context of breast cancer histopathological image classification. The optimized VGG16 model is systematically evaluated against several state-of-the-art CNN architectures, including AlexNet, Xception, InceptionV3, and ResNet50, to assess its relative effectiveness. The contributions of this work include the automated tuning of VGG16 hyperparameters using OOA for a medical imaging application, an extensive comparative benchmarking with widely used CNN models, and a comprehensive performance analysis using multiple evaluation metrics providing a robust and well-rounded assessment of classification performance. The organization of this paper is as follows. Section 2 explains the materials and methodological framework, including data preparation, model choice, and the OOA optimization strategy. Section 3 presents the obtained results along with their discussion. Section 4 summarizes the conclusions and proposes possible future research avenues.

## 2. MATERIALS AND METHOD

This study focuses on optimizing the hyperparameters of VGG16 for breast cancer classification using the OOA. The methodology is structured into several key phases: data acquisition, preprocessing, model selection, optimization, and evaluation.

### 2.1. Data acquisition and preprocessing

The dataset used consists of histopathological breast cancer images, sourced from publicly available repositories such as Mendeley Data and other medical imaging datasets [26]. Figure 1 shows a sample of the breast cancer dataset with Figure 1(a) showing a sample of benign breast image and Figure 1(b) showing malignant breast image. Dataset comprises of 250 cancerous images and 620 healthy images.




(a) (b)

Figure 1. Sample images of breast cancer dataset; (a) benign breast image and (b) malignant breast

All images were scaled to a uniform size to maintain consistency during model training. To enhance variability and reduce the risk of overfitting, data augmentation methods such as rotation, horizontal and vertical flipping, contrast modification, and normalization were employed. The dataset was subsequently divided into training (80%), validation (10%), and testing (10%) subsets to support proper model learning and performance assessment. Although the dataset size is relatively moderate, it is consistent with several prior histopathological breast cancer studies where expert annotation and data availability remain constrained. To mitigate potential bias arising from class imbalance between malignant and benign samples, extensive data augmentation techniques—including rotation, horizontal and vertical flipping, scaling, and intensity normalization—were applied exclusively to the training set. Furthermore, imbalance-sensitive performance metrics such as Matthews correlation coefficient (MCC), area under the ROC curve (AUC-ROC), sensitivity, and specificity were employed to ensure unbiased evaluation. These measures provide a reliable assessment of classifier performance under imbalanced conditions, which is essential for medical diagnostic applications.

### 2.2. Model selection and baseline performance

Several DL architectures were initially assessed for comparison, including AlexNet, Xception, VGG16, InceptionV3, and ResNet50. Among these, VGG16 was chosen for hyperparameter optimization due to its balance between depth and computational efficiency, as well as its strong baseline performance. The initial hyperparameters, such as learning rate were tuned based on default settings before applying optimization.

### 2.3. Hyperparameter optimization using osprey optimization algorithm for breast cancer detection

To enhance VGG16's classification accuracy, we employed the OOA, a bio-inspired metaheuristic technique. This algorithm mimics the hunting behavior of ospreys, dynamically adjusting hyperparameters to achieve optimal convergence. The key parameters optimized include the learning rate, number of neurons in fully connected layers, activation functions, batch size, and dropout rate. To ensure robust learning and improved performance on unseen samples, the training process minimized the cross-entropy loss function. The network was trained over 100 epochs, while an early stopping criterion was incorporated to halt training whenever the validation loss remained unchanged for 10 continuous epochs.

#### 2.3.1. Osprey optimization algorithm

The flowchart shown in Figure 2 (in Appendix) illustrates the OOA, which is designed to optimize hyperparameters for DL models. The process begins with initializing parameters, where the optimization problem is defined, and OOA-specific parameters such as population size, maximum iterations, and stall limits are set. Next, the osprey population is initialized by randomly generating initial candidate solutions within the defined search space. The algorithm then evaluates the fitness function for each osprey using the objective function, which quantifies how well each candidate solution performs. After evaluation, the algorithm proceeds to identify the best osprey, i.e., the candidate with the highest fitness value. In the update osprey positions step, the algorithm employs exploration and exploitation strategies to adjust the positions of the ospreys, aiming to move towards better solutions. This process continues iteratively, with the updated positions undergoing repeated evaluation and best-solution tracking. A decision point checks whether the stall limit or stopping criteria (such as maximum iterations or no further improvement) have been met. If not, the process loops back to the position update step; otherwise, it proceeds to return the best solution, outputting the most optimal set of parameters or solution found during the search. Finally, the algorithm terminates, having identified the optimal solution that balances accuracy and computational efficiency.

#### 2.3.2. Osprey optimization algorithm for VGG16 hyperparameter optimization for breast cancer detection

OOA is proposed as a bio-inspired metaheuristic approach for optimizing hyperparameters in VGG16 for breast cancer detection. Inspired by the hunting strategy of ospreys, the algorithm balances exploration and exploitation to find the most effective hyperparameter configurations, improving classification accuracy and model efficiency. This study evaluates OOA's performance against conventional optimization techniques, suggesting that it can significantly contribute to improving DL systems designed for medical image classification tasks.

The detailed procedural steps of the proposed optimization framework are formally presented in Algorithm 1, which outlines the integration of the OOA with VGG16 hyperparameter tuning.

Algorithm 1. OOA for VGG16 hyperparameter optimization
1. Initialize:
   a. Load breast cancer dataset and preprocess images.
   b. Define search space for hyperparameters (learning rate, batch size, dropout rate).
   c. Initialize a population of hyperparameter sets randomly.

2. Evaluate Initial Population:
   a. Train VGG16 with each hyperparameter set.
   b. Compute fitness function (e.g., accuracy, loss, AUC-ROC).
3. Iterative Search Process:
   Repeat for maximum iterations or until convergence:
   a. Select top-performing hyperparameter sets based on fitness scores.
   b. Apply Osprey-inspired adaptive search:
      i. Exploitation: Fine-tune hyperparameters near high-performing values.
      ii. Exploration: Introduce variations to avoid local optima.
   c. Update population with new hyperparameter sets.
4. Check Convergence:
   a. If performance improvement stagnates for ‘n’ iterations, stop optimization.
   b. Else, continue searching for optimal hyperparameters.
5. Select Best Hyperparameters:
   a. Choose the hyperparameter set with the highest fitness score.
   b. Train the final VGG16 model using optimized parameters.
6. Final Evaluation:
   a. Compute classification metrics.
   b. Compare results with baseline models.
7. Output optimized model for breast cancer classification

For reproducibility, all experiments were conducted using a fixed number of training epochs, a predefined population size for the optimization algorithm, and consistent early stopping criteria based on validation loss. The fitness function was defined as a weighted combination of classification accuracy and validation loss, ensuring stable convergence. These settings, along with the algorithmic steps outlined above, enable straightforward replication of the proposed optimization framework. This study presents the OOA for VGG16 hyperparameter optimization, specifically designed for breast cancer detection using DL. The algorithm follows a structured approach, starting with the initialization of parameters and the generation of an initial population of candidate solutions. Each solution is evaluated based on a fitness function that considers accuracy, precision, recall, F1-score, and AUC-ROC. The algorithm dynamically updates hyperparameters through adaptive search mechanisms, ensuring efficient exploration of the solution space while preventing local optima stagnation.

Experimental results demonstrate that the OOA-optimized VGG16 model achieves superior classification performance compared to standard CNN architectures and other optimization techniques. The study highlights the efficacy of bio-inspired optimization algorithms in DL-based medical diagnostics, paving the way for further research in hyperparameter tuning for medical image classification. By integrating OOA, the proposed approach significantly enhances accuracy, specificity, and computational efficiency, contributing to improved breast cancer detection and patient outcomes.

### 2.4. Evaluation metrics and performance analysis

The optimized VGG16 model was assessed using standard classification metrics. Additionally, inference time was recorded to measure computational efficiency. The confusion matrix was used to analyze the model’s ability to distinguish between malignant and benign cases, ensuring a balance between sensitivity and specificity. The results were compared with other state-of-the-art architectures and optimization techniques, highlighting the effectiveness of osprey optimization in improving breast cancer classification performance.

## 3. RESULTS AND DISCUSSION

Breast cancer detection using DL has significantly improved with advancements in CNNs. This study evaluates the performance of several CNN architectures, including AlexNet, Xception, VGG16, Inception V3, and ResNet50, in comparison to an optimized VGG16 model enhanced using the osprey optimization algorithm (OOA). The results demonstrate that hyperparameter tuning via the OOA considerably improves classification performance, as evidenced by various evaluation metrics.

Figure 3 provides insights of the confusion matrices of various pretrained models Figure 3(a) AlexNet, Figure 3(b) Xception, Figure 3(c) VGG 16, Figure 3(d) Inception V3, Figure 3(e) Resnet50, and Figure 3(f) osprey VGG16, highlighting the number of correctly and incorrectly classified cancerous and non-cancerous cases. AlexNet demonstrates moderate performance, correctly classifying 41 cancerous cases but misclassifying 9 non-cancerous samples as cancer (false positives) and missing 6 actual cancer cases (false negatives). Xception shows a similar pattern, with 39 true positives, 11 false positives, and 5 false negatives, indicating slightly lower sensitivity. VGG16 improves upon these results with 44 correctly identified cancer cases, fewer false positives (6), but slightly more false negatives (7). Inception V3 excels in detecting non-cancerous cases, with 0 false negatives, meaning all actual non-cancer cases were classified correctly; however, it struggles with cancer detection, misclassifying 11 cases. ResNet50 performs even better, reducing false positives to just 3 and maintaining a low false-negative count of 6. Finally, the osprey optimized VGG16 model significantly outperforms all others, with 49 true positives, only 1 false positive, and just 3 false negatives, demonstrating its superior ability to distinguish between cancerous and non-cancerous cases. These results highlight that the osprey-optimized model provides the most reliable and balanced performance, achieving high sensitivity and specificity, which are crucial for accurate breast cancer diagnosis. From a clinical perspective, minimizing false-negative predictions is of paramount importance in breast cancer diagnosis, as missed malignant cases may lead to delayed treatment and adverse patient outcomes. The proposed OOA-optimized VGG16 model demonstrates high sensitivity and AUC-ROC values, indicating its strong capability to correctly identify malignant cases. Moreover, the superior MCC achieved by the proposed approach confirms balanced predictive performance across both classes, making it particularly suitable for real-world clinical screening scenarios where class imbalance is prevalent.

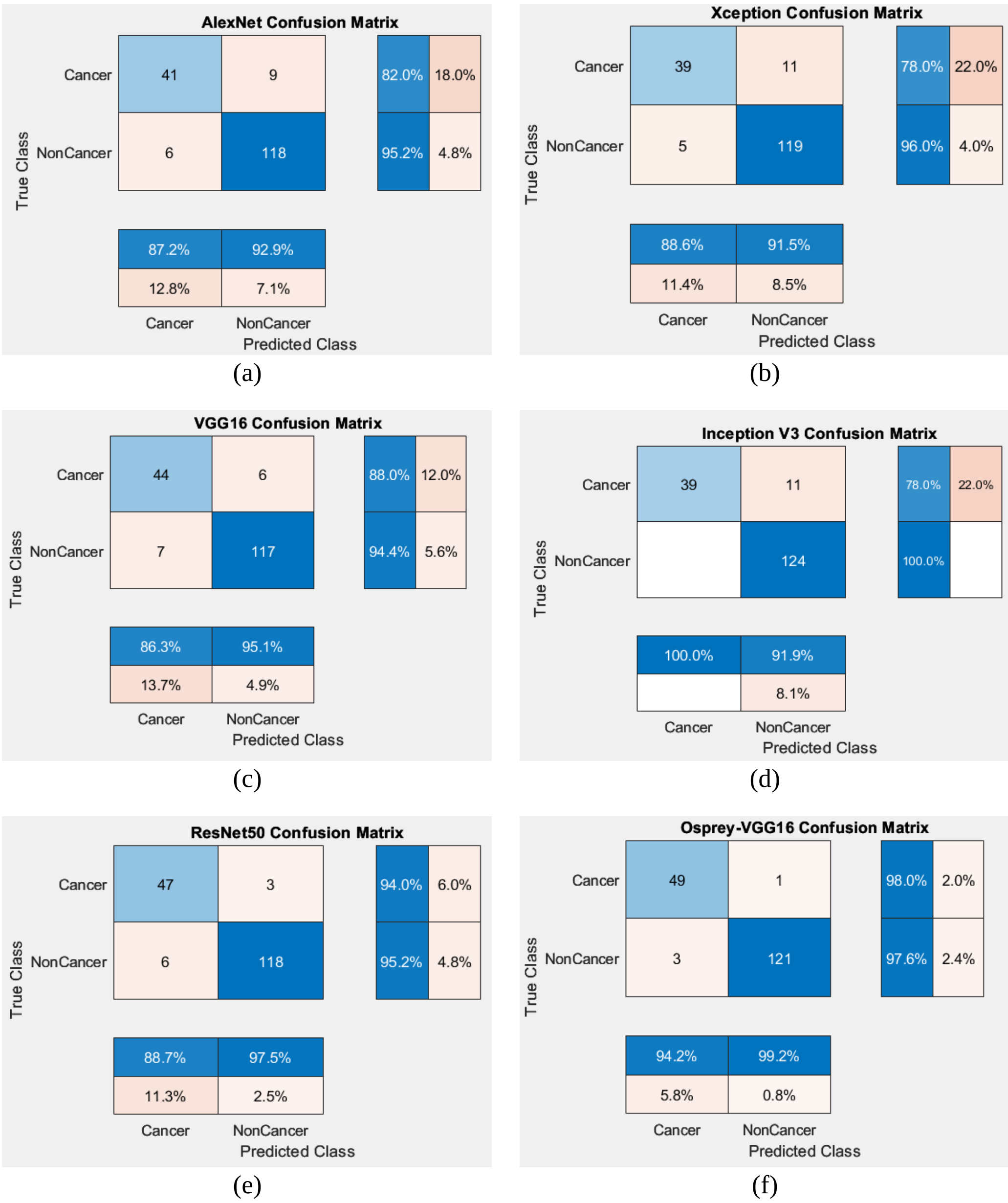



Figure 3. Confusion matrix of the DL models and proposed technique; (a) AlexNet confusion matrix, (b) Xception confusion matrix, (c) VGG16 confusion matrix, (d) Inception V3 confusion matrix, (e) ResNet50 confusion matrix, and (f) osprey-VGG16 confusion matrix

Table 1 presents a comparative analysis of multiple DL architectures for breast cancer detection. Among the models, osprey optimized VGG16 demonstrates the highest accuracy (97.7%) and superior performance across all evaluation metrics, including precision (96.71%), recall (97.79%), F1-score (97.23%), and AUC-ROC (99.92%), indicating its ability to achieve highly reliable predictions. In contrast, other pretrained DL models show moderate performance, with accuracy ranging from 90.80% to 92.53%, and relatively lower precision and recall values. Inception V3 and ResNet50 perform better than these models, with ResNet50 achieving 94.83% accuracy and an MCC of 0.8767, reflecting its strong predictive capability. However, osprey-optimized VGG16 surpasses them, as indicated by its highest MCC (0.9449) and lowest log loss (0.0542), signifying a well-calibrated model with minimal classification errors.

Table 1. Performance comparison of different DL models for breast cancer detection

| | AlexNet | Xception | VGG16 | Inception V3 | ResNet50 | Osprey optimized VGG16 |
|---|---|---|---|---|---|---|
| Accuracy | 91.38 | 90.80 | 92.53 | 93.68 | 94.83 | 97.7 |
| Precision | 90.07 | 90.09 | 90.70 | 95.93 | 93.10 | 96.71 |
| Recall | 88.58 | 86.98 | 91.18 | 89 | 94.58 | 97.79 |
| F1-score | 89.28 | 88.34 | 90.93 | 91.7 | 93.78 | 97.23 |
| Specificity | 88.58 | 86.98 | 91.18 | 89 | 94.58 | 97.79 |
| AUC-ROC | 95.77 | 94.84 | 98.73 | 97.6 | 98.16 | 99.92 |
| MCC | 0.7864 | 0.7701 | 0.8187 | 0.0846 | 0.8767 | 0.9449 |
| Log loss | 0.2015 | 0.3453 | 0.1953 | 0.1987 | 0.1761 | 0.0542 |
| Average inference time (sec) | 0.067 | 0.667 | 0.9546 | 0.3149 | 0.1359 | 0.1488 |

In terms of inference time, AlexNet (0.067 sec) and ResNet50 (0.1359 sec) are the fastest models, making them more suitable for real-time applications. However, Xception and VGG16 have significantly higher inference times (0.667 sec and 0.9546 sec, respectively), which could limit their practical deployment. The osprey-optimized VGG16 achieves a balanced inference time of 0.1488 sec, making it both efficient and accurate. This trade-off is clinically acceptable, as marginal increases in inference time are outweighed by the reduction in false-negative predictions, which is critical for early breast cancer detection. Overall, the results highlight the effectiveness of OOA in improving VGG16's hyperparameters, resulting in a highly optimized model that significantly outperforms traditional DL architectures for breast cancer detection. Recent studies (2023–2025) have explored a variety of bio-inspired and hybrid optimization techniques for DL hyperparameter tuning in medical image analysis. While these methods report incremental performance improvements, they often suffer from premature convergence or increased computational overhead. In contrast, the OOA demonstrates a balanced exploration–exploitation mechanism, enabling faster convergence and improved generalization. The superior MCC, reduced log loss, and consistent classification accuracy observed in this study highlight the effectiveness of OOA when compared to both conventional CNN training strategies and recently reported optimization-based frameworks.

The graph shown in Figure 4 visually compares the performance of various CNN architectures used for breast cancer detection. Osprey optimized VGG16 consistently achieves the highest values across all metrics, demonstrating its superior classification capability. The accuracy curve (light blue) shows a significant increase, peaking at 97.7% for the optimized VGG16. Similarly, the AUC-ROC metric follows a steady upward trend, reaching 99.92%, indicating an excellent ability to distinguish between cancerous and non-cancerous cases. Other models, such as ResNet50 and Inception V3, show competitive performance but with slightly lower recall and F1-scores. Inception V3 exhibits high precision but a drop in recall, suggesting it is more conservative in predicting cancer cases. Xception and AlexNet demonstrate lower overall performance, with AlexNet showing a decline in specificity and recall. The specificity (green curve) fluctuates, with a noticeable dip for Xception before increasing for ResNet50 and osprey VGG16. The results emphasize the effectiveness of OOA in fine-tuning VGG16 hyperparameters, resulting in a highly balanced model with minimal misclassification errors. The significant performance boost makes it a promising approach for enhancing DL-based medical diagnosis systems.

The 3D surface plot shown in Figure 5 provides a comprehensive visualization of the performance of different CNN architectures across multiple evaluation metrics. The osprey optimized VGG16 achieves the highest values across all performance indicators, as represented by the uppermost region of the plot, with accuracy, AUC-ROC, and recall values close to 98–100%. ResNet50 and Inception V3 also perform well but exhibit slight variations in recall and specificity. VGG16, Xception, and AlexNet demonstrate relatively lower performance, particularly in recall and F1-score, as indicated by the deeper regions of the plot. The color gradient, ranging from blue (lower performance) to yellow (higher performance), further emphasizes the superior classification ability of osprey VGG16, with AlexNet and Xception showing the weakest performance among the models tested. This visualization reinforces the effectiveness of the OOA in enhancing CNN-based breast cancer detection by fine-tuning hyperparameters for optimal accuracy and reliability.

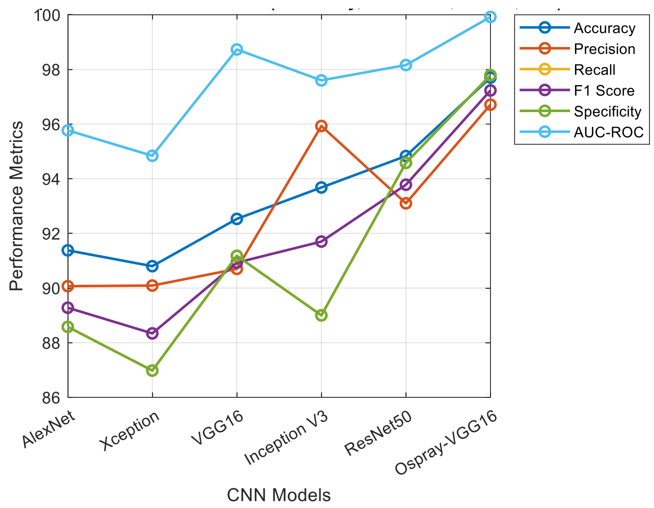



Figure 4. Line plot of performance comparison of different CNN models for breast cancer detection across multiple evaluation metrics

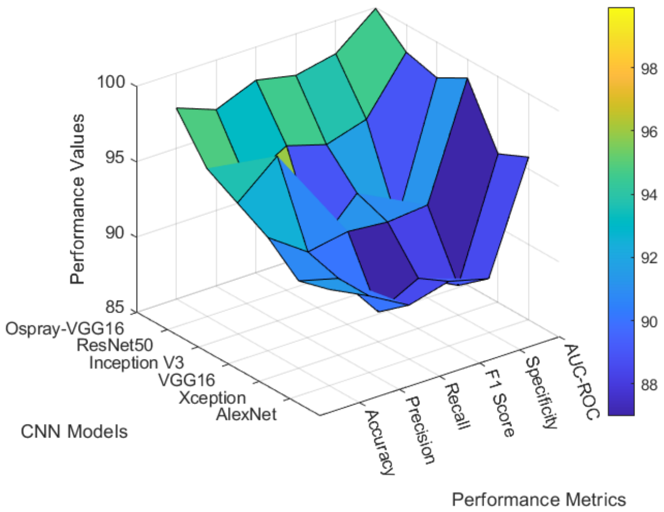



Figure 5. 3D surface plot depicting the performance metrics of various CNN models for breast cancer detection, including accuracy, precision, recall, F1-score, specificity, and AUC-ROC

The bar chart shown in Figure 6 presents a comparison of CNN models based on MCC (blue), log loss (red), and inference time (yellow), providing deeper insights into model reliability and computational efficiency. Osprey optimized VGG16 achieves the highest MCC (~0.9449), indicating superior predictive performance and a strong correlation between actual and predicted classifications. ResNet50 follows closely, while Inception V3 exhibits a notably lower MCC, highlighting its inconsistency in classification. In terms of log loss, osprey VGG16 records the lowest value (~0.0542), suggesting the highest confidence in its probability-based predictions, whereas Xception and Inception V3 show relatively higher values, indicating greater uncertainty in classification. Inference time, a crucial metric for real-time applications, is the lowest for AlexNet and ResNet50, making them computationally efficient choices. However, Xception and baseline VGG16 exhibit significantly higher inference times, which could limit their practical deployment. The results reaffirm the effectiveness of OOA, which not only improves classification accuracy but also optimizes computational efficiency by reducing log loss and maintaining a balanced inference time. This makes the optimized VGG16 model the most reliable choice for breast cancer detection in clinical applications.

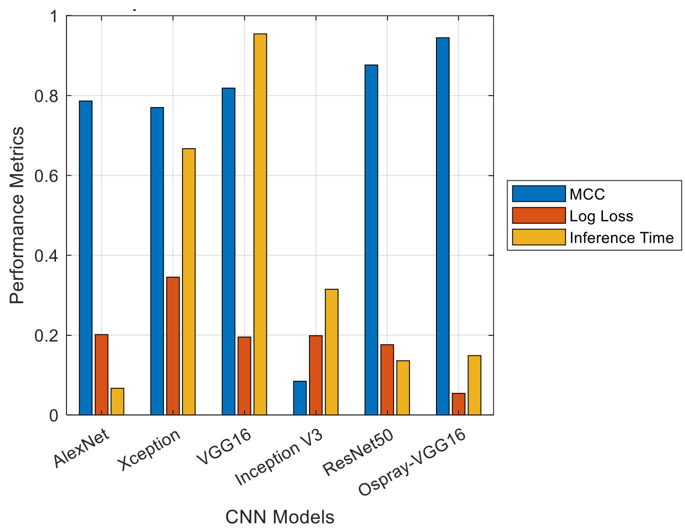



Figure 6. Comparative analysis of CNN models for breast cancer detection based on MCC, log loss, and inference time

Figure 7 presents a three-dimensional surface comparison of the evaluated CNN models based on MCC, log loss, and average inference time. These metrics collectively reflect classification reliability, prediction confidence, and computational efficiency. The osprey-optimized VGG16 model occupies the most favorable region of the surface, exhibiting the highest MCC and the lowest log loss while maintaining a balanced inference time. This indicates strong agreement between predicted and actual classifications with minimal uncertainty. In contrast, baseline architectures such as AlexNet, Xception, and standard VGG16 demonstrate lower MCC values and higher log loss, reflecting reduced classification reliability. Although ResNet50 and InceptionV3 show competitive performance, their overall trade-off between reliability and efficiency remains inferior to the optimized VGG16 model. The visualization highlights the effectiveness of the OOA in simultaneously enhancing predictive accuracy and computational efficiency, confirming its suitability for breast cancer detection applications. These results suggest that hyperparameter optimization through the OOA significantly enhances the performance of CNN-based breast cancer detection models.

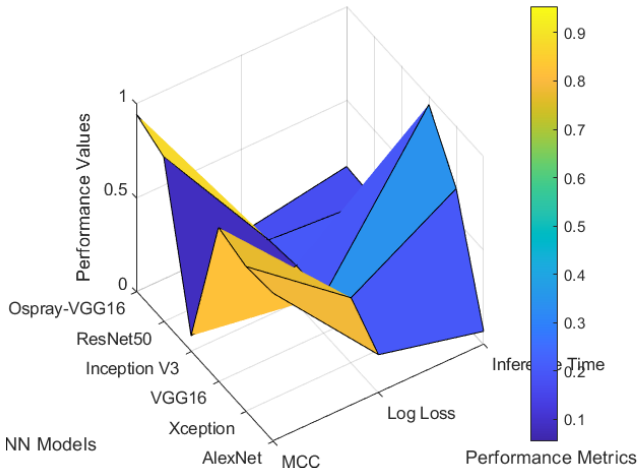



Figure 7. 3D surface plot comparing CNN models based on MCC, log loss, and inference time for breast cancer detection

The findings highlight the potential for improved early detection, leading to better patient outcomes and clinical decision-making. Future work may explore further refinements and real-world deployment to validate its effectiveness in diverse medical imaging datasets. To ensure robustness and reliability of the proposed optimization framework, the model training incorporated early stopping based on validation loss convergence, preventing overfitting and unstable learning behavior. The use of multiple complementary evaluation metrics, including MCC and log loss, further ensures that performance gains are not attributed solely to class dominance. Additionally, confusion matrix-based analysis was conducted to validate consistency in classification outcomes across both malignant and benign categories, thereby reinforcing the stability of the optimized model.

## 4. CONCLUSION

The experimental results demonstrate the superior performance of the osprey-optimized VGG16 model in breast cancer detection. With an accuracy of 97.7%, it significantly outperforms other DL architectures, including ResNet50 (94.83%) and Inception V3 (93.68%). Its precision (96.71%), recall (97.79%), and F1-score (97.23%) further highlight its robustness in classification, ensuring minimal false positives and negatives. The specificity (97.79%) and AUC-ROC (99.92%) values confirm its excellent discriminatory power, making it a highly reliable model for medical diagnosis. Additionally, the MCC of 0.9449 underscores its strong classification reliability. Despite a slightly higher inference time (0.1488 sec) compared to AlexNet (0.067 sec), it remains computationally efficient compared to Xception (0.667 sec) and standard VGG16 (0.9546 sec). The lowest log loss (0.0542) further supports the model's high confidence in its predictions. The confusion matrix analysis reaffirms its diagnostic effectiveness, with only 1 false positive and 3 false negatives, correctly identifying 49 cancerous and 121 non-cancerous cases. This minimal misclassification rate highlights the model's potential for clinical implementation. Despite the promising results achieved, this study has certain limitations. The evaluation is conducted on a single histopathological dataset, which may limit generalizability across diverse imaging conditions and acquisition protocols.

Additionally, although data augmentation partially addresses dataset size constraints, larger multi-institutional datasets could further enhance model robustness. These limitations do not undermine the validity of the proposed approach but instead provide meaningful directions for future research. Future research can focus on real-world deployment of the osprey-optimized VGG16 model in clinical settings, integrating it with radiology workflows for automated breast cancer screening. Additionally, exploring transfer learning with larger, diverse datasets could further enhance generalization across different demographics. The incorporation of explainable AI (XAI) techniques can improve interpretability, making the model more trustworthy for medical professionals. Moreover, hardware acceleration techniques, such as TensorRT and FPGA optimizations, can be explored to reduce inference time, ensuring real-time processing. Finally, extending the osprey optimization framework to other medical imaging tasks, such as lung cancer detection and brain tumor classification, can broaden its impact in healthcare AI applications.

## ACKNOWLEDGMENTS

The authors declare that no individuals contributed to the work in a manner that requires acknowledgment.

## FUNDING INFORMATION

Authors state no funding involved.

## AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

| Name of Author | C | M | So | Va | Fo | I | R | D | O | E | Vi | Su | P | Fu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sabura Banu Urundai Meeran | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Nafeena Abdul Munaf | | | | | ✓ | | | | | ✓ | | | | |
| Vengadeshwaran Velu | | | | | | ✓ | | | | ✓ | | | | |

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

## CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

## DATA AVAILABILITY

Data availability is not applicable to this paper as no new data were created or analyzed in this study.

## REFERENCES


[1] R. L. Siegel, A. N. Giaquinto, and A. Jemal, "Cancer statistics, 2024," *CA: A Cancer Journal for Clinicians*, vol. 74, no. 1, pp. 12–49, Jan. 2024, doi: 10.3322/caac.21820.

[2] H. Sung *et al.*, "Global Cancer Statistics 2023: GLOBOCAN Estimates of Incidence and Mortality Worldwide," *CA: A Cancer Journal for Clinicians*, vol. 73, no. 5, pp. 209–249, 2023.

[3] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[4] A. Esteva *et al.*, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, no. 1, pp. 24–29, Jan. 2019, doi: 10.1038/s41591-018-0316-z.

[5] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 2, pp. 1097–1105, 2012.

[7] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations*, 2015.

[8] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 2818–2826, doi: 10.1109/CVPR.2016.308.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[10] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 1800–1807, doi: 10.1109/CVPR.2017.195.

[11] J. Bergstra and Y. Bengio, "Random Search for Hyper-Parameter Optimization," *Journal of Machine Learning Research*, vol. 13, no. 2, pp. 281–305, 2012, doi: 10.5555/2188385.2188395.

[12] K. Yasaka *et al.*, "Effects of deep learning on radiologists' and radiology residents' performance in identifying esophageal cancer on CT," *British Journal of Radiology*, vol. 96, no. 1150, 2023, doi: 10.1259/bjr.20220685.

[13] J. Snoek, H. Larochelle, and R. P. Adams, "Practical Bayesian optimization of machine learning algorithms," *Advances in Neural Information Processing Systems*, vol. 4, pp. 2951–2959, 2012.

[14] A. Darwish, A. E. Hassanien, and S. Das, "A survey of swarm and evolutionary computing approaches for deep learning," *Artificial Intelligence Review*, vol. 53, no. 1, pp. 1767–1812, 2020, doi: 10.1007/s10462-019-09719-2.

[15] K. Davoudi and P. Thulasiraman, "Evolving convolutional neural network parameters through the genetic algorithm for the breast cancer classification problem," *Simulation*, vol. 97, no. 8, pp. 511-527, 2021, doi: 10.1177/0037549721996031.

[16] E. H. Houssein, M. M. Emam, and A. A. Ali, "An optimized deep learning architecture for breast cancer diagnosis based on improved marine predators algorithm," *Neural Computing and Applications*, vol. 34, no. 20, pp. 18015-18033, 2022, doi: 10.1007/s00521-022-07445-5.

[17] P. Pramanik, S. Mukhopadhyay, S. Mirjalili, and R. Sarkar, "Deep feature selection using local search embedded social ski-driver optimization algorithm for breast cancer detection in mammograms," *Neural Computing and Applications,* vol. 35, no. 7, pp. 5479-5499, 2023, doi: 10.1007/s00521-022-07895-x.

[18] N. A. Baghdadi, A. Malki, H. M. Balaha, Y. AbdulAzeem, M. Badawy, and M. Elhosseini, "Classification of breast cancer using a manta-ray foraging optimized transfer learning framework," *PeerJ Computer Science*, 2022, doi: 10.7717/peerj-cs.1054.

[19] S. Thirumalaisamy *et al.*, "Breast Cancer Classification Using Synthesized Deep Learning Model with Metaheuristic Optimization Algorithm," *Diagnostics (Basel)*, vol. 13, no. 18, pp. 1-21, 2023, doi: 10.3390/diagnostics13182925.

[20] E. H. Houssein, M. M. Emam, A. A. Ali, and P. N. Suganthan, "Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review," *Expert Systems with Applications*, vol. 167, doi: 10.1016/j.eswa.2020.114161.

[21] A. A. Asiri, A. Shaf, T. Ali, M. Aamir, M. Irfan, and S. Alqahtani, "Enhancing brain tumor diagnosis: an optimized CNN hyperparameter model for improved accuracy and reliability," *PeerJ Computer Science*, 2024, doi: 10.7717/peerj-cs.1878.

[22] M. Obayya *et al.*, "Hyperparameter Optimizer with Deep Learning-Based Decision-Support Systems for Histopathological Breast Cancer Diagnosis," *Cancers*, vol. 15, no. 3, pp. 1-19, 2023, doi: 10.3390/cancers15030885.

[23] M. Dehghani and P. Trojovský, "Osprey optimization algorithm: A new bio-inspired metaheuristic algorithm for solving engineering optimization problems," *Frontiers in Mechanical Engineering*, vol. 8, 2022, doi: 10.3389/fmech.2022.1126450.

[24] B. Yao, L. Chao, M. Asadi, and K.A. Alnowibet, "Modified osprey algorithm for optimizing capsule neural network in leukemia image recognition," *Scientific Reports,* vol. 14, 2024, doi: 10.1038/s41598-024-66187-7.

[25] A. Heikal, A. El-Ghamry, S. Elmougy, and M. Z. Rashad, "Fine tuning deep learning models for breast tumor classification," *Scientific Reports,* vol. 14, 2024, doi: 10.1038/s41598-024-60245-w.

[26] M. Pereira, "BreakHis - Breast Cancer Histopathological Database," Mendeley Data, V1, 2023, doi: 10.17632/jxwvdwhpc2.1.

## APPENDIX

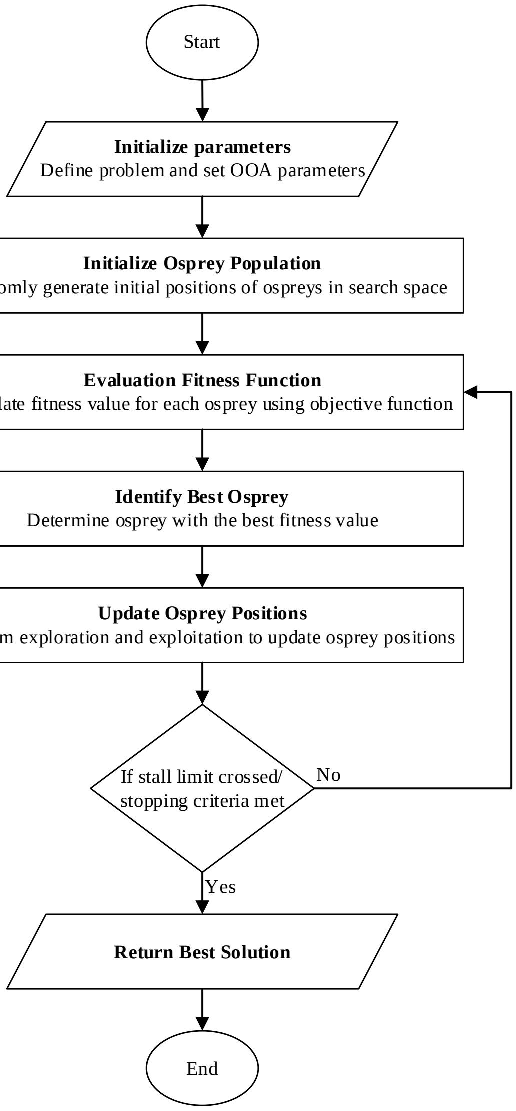



Figure 2. Flowchart of OOA

## BIOGRAPHIES OF AUTHORS

**Dr. Sabura Banu Urundai Meeran** was born in Chennai, Tamilnadu, India on May 8, 1975. She is presently working as a Professor in the Department of Electrical and Electronics Engineering, Saveetha Engineering College, Chennai, Tamilnadu, India. She has an illustrious academic and professional background in engineering, innovation, and consultancy. She completed her Ph.D. in Intelligent Modeling and Control at College of Engineering, Guindy, Anna University (2006–2009) after earning her M.S. by research in the same institution and B.E. in Instrumentation and Control Engineering. She also holds a Diploma in Computer Science Engineering and a B.Sc. in Mathematics. With over two decades of experience, she has served in various leadership roles, including Director at Arcres Techsol Pvt. Ltd., Founder & CEO, Intellectz Tech for All and Head of Departments at esteemed institutions like SRM Institute of Science and Technology and B.S. Abdur Rahman University. She has mentored numerous student projects that garnered national and international accolades, including awards, patents, and startup funding. Her notable achievements include 34 international journal publications in leading Journals like Taylor and Francis, Applied Soft Computing, ACTA Press, AMSE, Inderscience, contributions to five book chapters, and 132 conference presentations. She can be contacted at email: sabura_banu@rediffmail.com and saburasec@gmail.com.

**Dr. Nafeena Abdul Munaf** is Assistant Professor and Coordinator for Electrical and Electronics Engineering program. Villa College, Maldives. Besides that, she has published many papers in Scopus, Web of Science, Science Indexed and UGC journals. Her recent published paper in IEEE got 49 citations on the title' An ML approach for Household, power consumption. Currently she is a reviewer in reputed international and national journals and editor in special issues of conference publications. Also, she has given many keynote lectures in reputed colleges and universities in Smart Grid Challenges and Towards Sustainable University Residences: Comprehensive Energy Audit, Retrofit Techniques and Financial Implications Sustainability Issues in Environment. She is interested in writing articles and publishing on Life Cycle Analysis of various products and Smart Energy management systems for smart grid. Also, she has received grants on SDG based researches. She can be contacted at email: nafeena.abdul@villacollege.edu.mv.

**Associate Professor Vengadeshwaran Velu** is the Director of Research Management Centre (RMC) at Mila University, Malaysia. He is a Professional Engineer registered with the Board of Engineers Malaysia. He is a Chartered Engineer and the Fellow of The Institute of Engineers, Scotland, UK. He is also a Certified Energy Manager, a senior member of the IEEE and a member of the IET, UK. He has published number of research papers in reputed ISI Journals and conferences. He also served as a reviewer for number of reputed peer reviewed ISI journals and conferences, and written Engineering articles for students' magazines. He can be contacted at email: vengadeshwaran.velu@mila.edu.my.