

Intelligent Decision Support System Based on Deep Learning with the Whale Optimization Algorithm for Oral Cancer

Dasril Aldo¹, Adanti Wido Paramadini²

¹Faculty of Informatics, Telkom University, Purwokerto, Indonesia

²Department of Biomedical Engineering, Telkom University, Purwokerto, Indonesia

Article Info

Article history:

Received Jan 27, 2026

Revised Feb 16, 2026

Accepted Mar 26, 2026

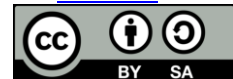
Keywords:

Deep Learning,
Hyperparameter optimization,
Whale Optimization
Algorithm,
Histopathological Images,
Clinical Decision Support
Systems

ABSTRACT

To build an accurate and reliable clinical decision support system, this study seeks to create a classification system using deep learning as a better approach in the analysis of oral cancer histopathological images. The dataset used consisted of 10,002 images, of which the two more balanced classes were normal oral and oral squamous cell carcinoma. Some pre-trained deep learning architectures are taken as baseline models and then optimized using the Whale Optimization Algorithm to obtain the best hyperparameter configuration. Performance evaluation was carried out on test data using accuracy, precision, recall, F1-score, confusion matrix, and operational efficiency metrics as well as evaluation of a trust-based decision support system with the same mechanism as the reject option system. The models are better optimized and all models show improved performance. From the results of the experiments, the model that was best optimized with an F1-score and an accuracy of 98.73%, and also showed the best performance, was the EfficientNet B3 model. This is accompanied by a stable training process and adequate generalization. Therefore, the model shows results with adequate performance in the coverage range of 0.60 - 0.90 and still provides a reasonable inference time for use in the clinic. These results show that this model has high potential to be integrated with clinical decision support systems. Therefore, this model can be used as a diagnostic tool in clinics that is more accurate and ensures consistency in each clinical practice and can also build a better diagnostic decision support system.

This is an open access article under the [CC BY-SA](#) license.



Corresponding Author:

Dasril Aldo

Faculty of Informatics, Telkom University, Purwokerto, Indonesia

Jl. DI Panjaitan No.128, Karangreja, Kec. Purwokerto Sel., Kabupaten Banyumas, Jawa Tengah 53147

Email: dasrilaldo@telkomuniversity.ac.id

1. INTRODUCTION

Oral cancer is a form of cancer that has a fairly high morbidity and mortality rate, especially found in developing countries [1], [2], [3]. Early detection of oral cancer plays an important role in avoiding high mortality rates as well as improving treatment success rates [4], [5]. Histopathological examination is the 'gold standard' for the diagnosis of oral cancer [6], [7]. However, the process and technique of interpreting histopathological images is a very specific field and relies on the experience of the medical personnel concerned [8], [9]. This allows for a very varied diagnosis and diagnostic tool.

Along with technological advancements, especially in the field of artificial intelligence and deep learning, automated and objective medical image analysis is growing rapidly. Several studies have shown that deep learning is able to extract complex features contained in histopathological images and show better results than previously used substandard techniques [10], [11]. However, to implement deep learning, there are several criteria needed, including a high level of accuracy and

precision, consistent prediction, efficiency in calculations, and a system that is able to support decision-making in a clear, consistent, and reliable way [11], [12].

A number of previous studies have explored the use of deep learning to detect oral cancer through histopathological image analysis and have shown significant improvements in classification performance compared to conventional approaches [13], [14]. Other studies have examined the development and variations of deep learning architectures, including explainable AI integration, to improve the sensitivity and specification of oral cancer detection [15], [16]. Furthermore, the application of optimization techniques for model parameter tuning has been reported to improve the performance of machine learning models on medical data [17]. However, most of these studies still focus on the technical evaluation of model performance, while the use of these models as part of a clinical Decision Support System is still lacking in terms of prediction consistency, clear interpretation, and operational efficiency [18]. So, more research is still needed to develop a more analytically applicable approach.

Regarding the dilemma already mentioned, this study proposes the development of an Intelligent Decision Support System for oral cancer detection, which is built on deep learning and integrated with the Whale Optimization Algorithm (WOA). WOA is used as a supporting method to help the deep learning hyperparameter adjustment process be more systematic and adaptive, so that it is expected to improve the stability and reliability of the model [19], [20], [21]. The proposed system would not only make predictions about classes, but also provide predictive confidence information that would be used as the basis for decision-making mechanisms in the DSS.

The novelty of this study lies in the incorporation of deep learning combined with the Whale Optimization Algorithm into an Intelligent Decision Support System for oral cancer, taking into account the classification performance, confidence-based decision stability, and operational efficiency of the model. This approach is expected to make a significant contribution to the development of a more accurate, reliable, and practical medical decision support system, and can help medical personnel in more effective early detection of oral cancer.

2. RESEARCH METHOD

The methodology has been systematically divided into four main components to build a proper, stable, and fully ready decision support system for implementation. The research process includes the data preparation stage, modeling and optimization using the Whale Optimization Algorithm (WOA), evaluation and selection of the best models, and finally, implementation and integration into the decision support system, as shown in Figure 1.

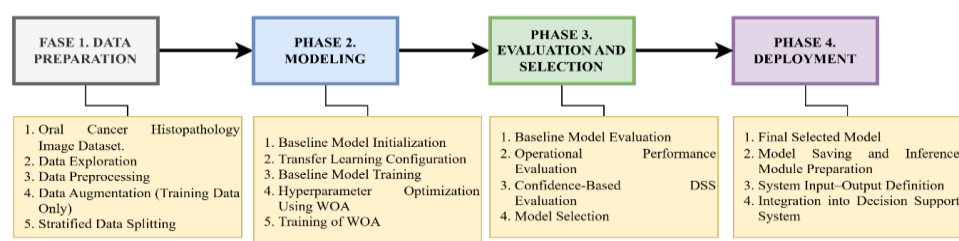


Figure 1. Research Framework

PHASE 1. Data Preparation

This analysis has been conducted on a dataset comprising 10,002 oral cancer histopathological images, which can be publicly accessed through the Kaggle Multi Cancer Dataset. The dataset includes two classes: Oral Normal, and Oral Squamous Cell Carcinoma (Oral SCC), which has 5,001 images per class. The class labels accompanying the images originate from the dataset's histopathological diagnosis, which acts as the ground truth across previous studies utilizing the dataset. A review of class distribution reveals a maximum-to-minimum ratio of 1.0, which shows that the dataset is balanced and does not need class imbalance correction strategies. The dataset was split employing stratified sampling. This method preserves class ratios in all the resulting subsets, thus 70%, 15%, and 15% of the dataset was assigned for training, validation, and testing respectively. Consequently, the training set consisted of 7,001 images, 3,501 of which are

for Oral SCC, and 3,500 for Oral Normal. The validation set consisted of 1,500 images, with 750 for each class, and the testing set has 1,501 images, 751 for Oral Normal and 750 for Oral SCC. To guarantee the consistency of the original dataset and the generated splits, thorough data integrity checks were performed to ascertain that no images were missing or damaged in any subset. In Figure 2, the representative examples of each class's histopathological images used in the training and testing processes are displayed. In addition, the dataset's credibility and strength are also supported by the community's wide usage. For example, over 124 publicly accessible notebooks and scripts on Kaggle that apply this dataset to classify medical images. Such broad reusing demonstrates that the dataset has been analyzed and evaluated many times with various deep learning models, indicating that this dataset is fit to be used as a benchmark, even though this study lacked outside cross-institutional validation.

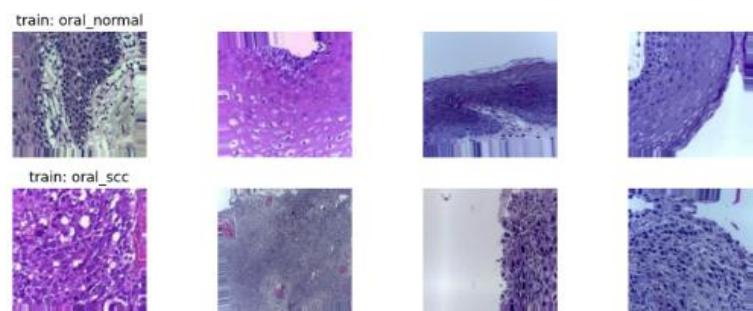


Figure 2. Example of a Data Train

All images are pre-processed by resizing to 224×224 pixels and normalizing using the standard mean values and deviations from ImageNet. Class labels are numerically coded into two classes, oralnrmal (0) and oralcc (1). To reduce the risk of overfitting, data augmentation is applied moderately, only for training data, including horizontal flip, light rotation, and light color jitter. After the augmentation and pre-processing steps, a data loader is created, and a health check is performed, confirming the suitability of the data dimensions to the batch size [32, 3, 224, 224]. The check also confirms successful data processing in the GPU (CUDA), so the dataset is considered ready for use in the modeling phase.

PHASE 2. Modeling

The modeling phase begins by initializing the base model with some pre-trained deep learning architectures commonly used for medical image classification. Each architecture was chosen to represent different model characteristics in terms of depth, complexity, and computational efficiency, as summarized in Figure 1. Here is an explanation of each architecture.

1. ResNet50

ResNet50 is a Convolutional Neural Network (CNN) that combines residual learning concepts to solve the problem of fading gradients in deep networks. With 50 layers and residual joints, this architecture effectively extracts high-level features (HLFs) and is often used as a robust baseline in a wide range of medical image classification tasks [22], [23], [24]. In the methodological workflow, ResNet50 is located as part of the base model, as illustrated in Figure 3.

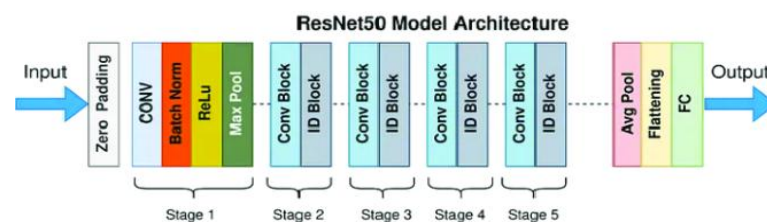


Figure 3. ResNet50 Architecture [25]

2. EfficientNet-B0

EfficientNet-B0 adalah varian dasar dari keluarga EfficientNet, dan menerapkan Composite scaling approach, which involves scaling the depth, width, and resolution of the input uniformly. This architecture is built to achieve an optimal balance between accuracy and compute efficiency, which means it is suitable for systems with limited resources [26], [27], [28]. Figure 4 illustrates the structure and role of EfficientNet-B0 in this study.

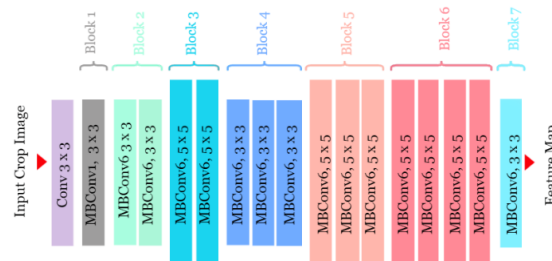


Figure 4. ResNet50 Architecture [29]

3. EfficientNet-B3

Among the EfficientNet variants, the EfficientNet-B3 has a larger model capacity than the B0. This is achieved by increasing the depth and resolution of the input. This architecture is capable of capturing more complex morphological patterns in histopathological images, which is likely to improve classification performance [30], [31]. Figure 5 illustrates the use of EfficientNet-B3 as a key candidate in the optimization and model selection process.

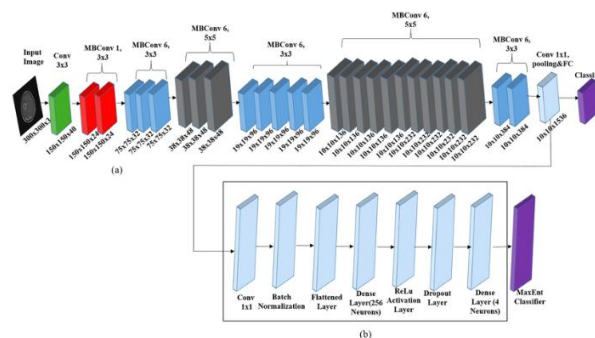


Figure 5. EfficientNet-B3 Architecture [32]

4. MobileNetV2

MobileNetV2 is a lightweight CNN architecture that uses separate convolutions by depth, and inverse residual blocks, which reduce parameters and computational loads. This model is designed for real-time applications and devices with limited resources. Therefore, it is relevant to assess the operational efficiency in the Decision support system [33], [34]. Figure 6 illustrates the integration of MobileNetV2 in the modeling phase.

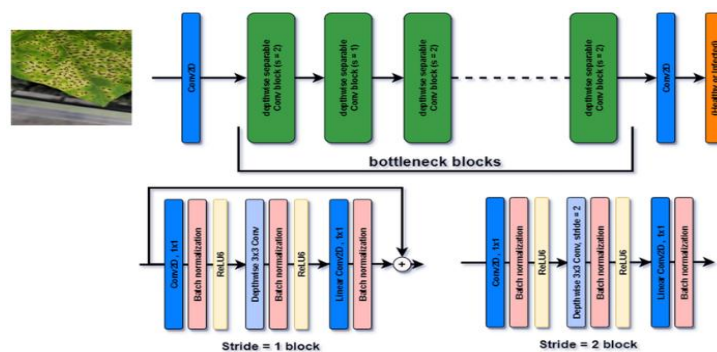


Figure 6. MobileNetV2 Architecture [35]

The complete model was adapted for the oral cancer classification task involving two classes (oral normal and oral SCC) by altering the final fully connected layer of each pre-trained network. As for the histopathological images, they were treated as image patches. The initial resolution was set to 224×224 pixels, a size widely adopted in deep learning-based histopathology research and a size that aligns with the pre-trained ImageNet backbones. All images were resized to this resolution without the use of explicit multi-scale magnification modeling, this was possible because the dataset consists of pre-extracted patches and not whole slide images. Background tissue exclusion was addressed by the dataset curation process, which avoided non-informative areas, and no background removal or tissue segmentation was applied during preprocessing. Each image was an individual patch, and no multiple patch sampling was done from the same slide, guaranteeing that each patch was regarded as an independent sample during both training and evaluation. A transfer learning strategy was utilized in which the backbone layers were frozen, and only the classification head was trained, this came as a way for the model to concentrate on the histopathology specific discriminative features. The entire training procedure using the baseline duration of 10 epochs and a weight decay of 0.0001 was performed on a GPU (CUDA) and saved uniformly with a learning rate of 0.0001. The baseline model with the least validation loss was chosen and set aside for additional optimization. For the purpose of enhancing performance and stability, the model was further optimized through hyperparameter tuning using the Whale Optimization Algorithm (WOA), with learning rate, batch size, dropout rate, number of epochs, and depth of fine-tuning as the parameters of concern. The WOA for each architecture was separately optimized. The deep learning model that was created through this process was used to evaluate and select the decision support system in the final phase.

PHASE 3. Evaluation and Selection

Model performance evaluation is performed on test data to provide an objective comparison between the base model and the model optimized by WOA. Each model is evaluated using standard classification metrics that include Accuracy. Used to measure the accuracy of the model's prediction, i.e. the correct proportion of the predicted values of the entire test dataset. These metrics provide an overview of the model's performance, however, it does not necessarily reflect the balance of performance across the class.

1. Precision

Measure the accuracy of positive predictions, i.e. the extent to which images predicted to be oral cancer (oral SCC) are actually cases of cancer. Precision is essential for assessing the risk of false positives, which, in a clinical context, can result in unnecessary further examinations [36].

2. Recall

Demonstrate the model's ability to detect all true positive cases, i.e. the proportion of images of oral cancer that have been correctly identified. Recalls play an important role in clinical applications because they are directly linked to the risk of false negatives or missed cancer cases [37].

3. F1-score

It is a harmonious average between precision and recall, reflecting the balance between accuracy and completeness of detection. F1-score is used as the primary metric in this study, as it is the best at representing the model's performance against two classes [38].

4. Confusion Matrix

Applied to analyze the distribution of true and false predictions for each class, including true positive, true negative, false positive, and false negative. This analysis helps in understanding model error patterns and their implications in clinical diagnosis. The results of this evaluation were used to compare the performance of the base model and the optimized WOA model, as well as to assess the performance improvement obtained after the optimization process [38].

In addition to the evaluation of classification performance, operational performance evaluation is also carried out to assess the feasibility of the model in the context of the decision support system. This evaluation revolves around the measurement of model size, the number of

parameters, and the time required for inference per image, all of which are indicators of computational efficiency. Furthermore, a trust-based evaluation of the DSS is used by setting a confidence threshold that provides a decision coverage between 0.60 and 0.90. Within that threshold range, the trade-off between the coverage and prediction accuracy of the covered cases is analyzed. From all these evaluations, the selection of models is carried out through normalization and bottling of metrics, resulting in one optimal model that is balanced between classification performance, operational efficiency, and decision stability to be implemented in the decision support system.

PHASE 4. Deployment

According to the results of the selection from the previous phase, the best model is chosen as the final model that will be applied in the decision support system. This final model is stored in a system-compatible format for inference and continuous use. Next, an inference module is prepared. The module is designed to load models, process input images according to the training configuration, and automatically predict classes and provide confidence scores. At this stage, the definition of the input and output of the system has also been determined, with the input in the form of histopathological images of oral cancer and the output in the form of classification results from the class (oralnormal or oralcc) with the appropriate confidence level of the prediction. The inference module is then integrated into an Intelligent Decision Support System (IDSS) based on a web or mobile application. The developed system is designed to provide informative and reliable decision support, so it is ready to be used as a tool for medical staff in supporting the detection and decision-making process related to oral cancer.

3. RESULTS AND DISCUSSION

This section describes the results of experiments and discussions that focus on performance comparison between baseline deep learning models and models that have been optimized using the Whale Optimization Algorithm (WOA) for oral cancer classification. Table 1 shows a comparison of baseline deep learning models and deep learning models that have been optimized by WOA in terms of classification, training time, and reasoning efficiency.

Table 1. Comparison of Baseline and WOA Model Performance in Oral Cancer Classification

Model	Accuracy	Precision	Recall	F1	Training_time	Inference_time
EfficientNet-B0	0.8441	0.8442	0.8441	0.8441	395.29	0.3321
EfficientNet-B0_WOA	0.9793	0.9794	0.9793	0.9793	153.44	0.3324
EfficientNet-B3	0.8081	0.8113	0.8081	0.8076	441.81	0.6816
EfficientNet-B3_WOA	0.9873	0.9873	0.9873	0.9873	665.00	0.6813
MobileNetV2	0.8321	0.8321	0.8321	0.8321	395.27	0.2667
MobileNetV2_WOA	0.9807	0.9807	0.9807	0.9807	429.64	0.2693
ResNet50	0.8201	0.8209	0.8201	0.8200	480.05	0.7323
ResNet50_WOA	0.9827	0.9827	0.9827	0.9827	444.77	0.7327

Table 1 reflects a performance comparison showing that at baseline conditions, all deep learning models can perform oral cancer classification with sufficient performance, but there are still shortcomings. EfficientNet-B0 provided the best baseline performance with an F1-score of 0.8441, followed by MobileNetV2 (0.8321), ResNet50 (0.8200), and EfficientNet-B3 (0.8076). This suggests that although transfer learning can still function well for the initial model, the default hyperparameter is not good enough to represent the level of complexity of histopathological images. By optimizing the Whale Optimization Algorithm (WOA), all models showed a very significant improvement in performance, with the F1-score value increasing by more than 0.97. A significant improvement is shown by EfficientNet-B3, which increases the F1-score from 0.8076 to 0.9873, thus establishing it as the final model for decision support systems.

As shown in Figure 1, the training curve versus accuracy validation and training against WOA EfficientNet-B3 optimization validated against accuracy loss indicate that the learning process is stable and convergent. The significant increase in accuracy and the loss of consistently validated data suggest that this model has a fairly good generalization. This curve pattern confirms that the WOA optimization hyperparameter not only serves to increase evaluation values, but is also efficient to improve the performance of clinical decision support systems.

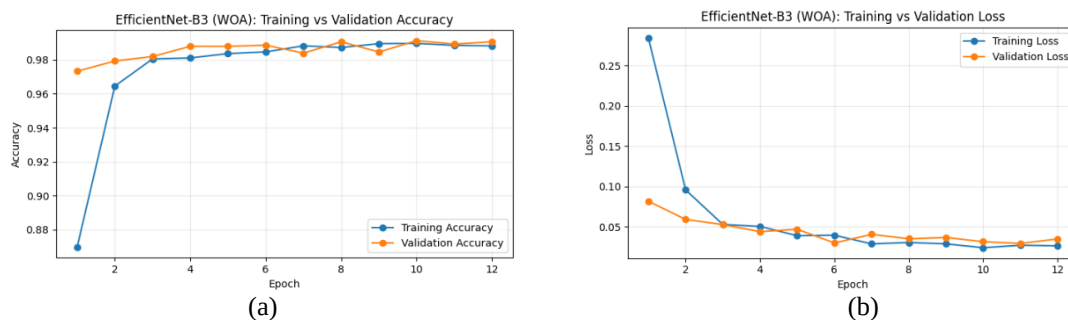


Figure 7. Training and validation performance of the WOA-optimized EfficientNet-B3 model, (a) accuracy, (b) loss.

As shown in figure 2, the WOA-optimized EfficientNet-B3 training accuracy and training accuracy curve tends to increase and is stable at 98-99%, and the loss curve, decreases sharply and steadily at a loss of <0.05 . On both the loss and accuracy, training and validation curves, small differences indicate a stable and convergent training process. This shows that the results of WOA optimization for hyperparameters not only improve the final evaluation of the model, but also for the model training process, in this case it is not overfitting, and has good generalizations. To produce a quantitative evaluation of the classification over a deeper range of classes, it is further shown in Figure 8 of the confusion matrix, which shows the distribution of true and false predictions in the tst data between the normal and oral classes of the SCC.

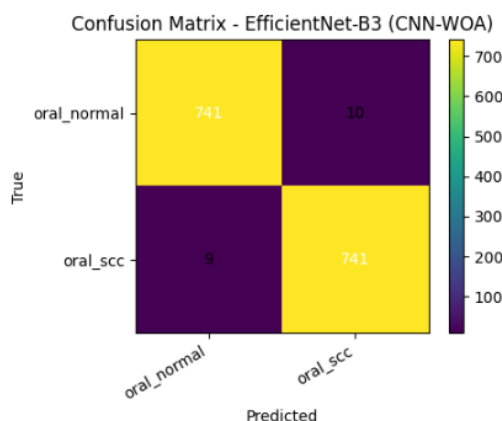


Figure 8. Confusion Matrix

The confusion matrix of the EfficientNet-B3 model optimized using WOA, shown in Figure 8, provides a detailed breakdown of classification performance on the test set consisting of 1,501 images. Of the 751 oral normal images, 741 were correctly classified, while 10 were misclassified as oral SCC (false positives). Similarly, out of 750 oral SCC images, 741 were correctly identified, with 9 images misclassified as oral normal (false negatives). This results in a total of 1,482 correctly classified images and only 19 misclassifications, corresponding to an overall accuracy of 98.73%, with balanced recall and F1-score values across both classes. From a clinical perspective, the low number of false negatives is particularly important, as it indicates a reduced risk of missed oral cancer cases, supporting the reliability of the proposed model as a decision support aid for early diagnosis.

In addition to classification accuracy, the potential for implementation in clinical practice also depends on practical operational performance and stable decision-making. Further evaluation revealed that although EfficientNet-B3 has a comparatively larger model size and number of parameters, the inference time is still at an acceptable level of 0.6816 ms per image on the GPU. This suggests that the increase in model complexity is proportional to the increase in prediction accuracy and reliability.

Figure 9 shows the results of qualitative diagnosis from five test images using the EfficientNet-B3 model as a result of WOA's optimization. All images are classified with a confidence level of >99%, and all predictions have an ACCEPT status based on DSS. This example shows that the model is not only superior in aggregate, but also consistent across all predictions, which in the context of clinical use is important.

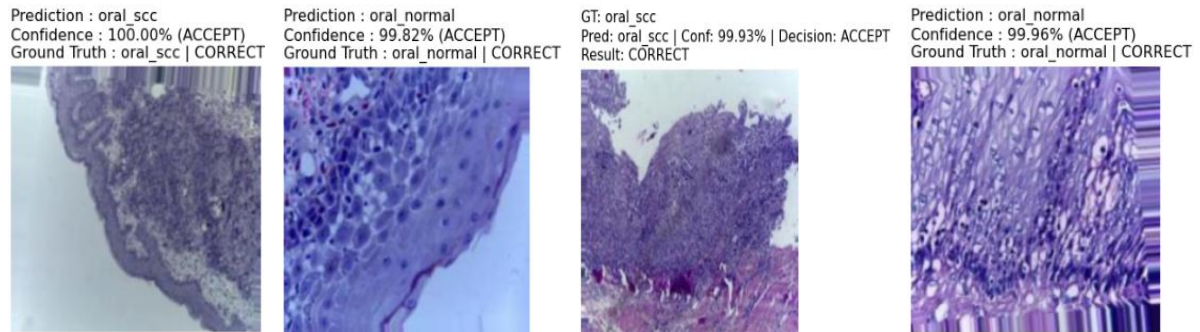


Figure 9. Test Results

Furthermore, the evaluation of a trust-based decision support system (DSS) with a rejection option mechanism explained that the model maintained consistent performance in the coverage range of 0.60 to 0.90, with a tendency to increase accuracy and F1-score in accepted cases. This strategy allows the system to reject predictions at low confidence levels, thereby reducing the risk of not making worse decisions. Overall, the addition of the DSS evaluation provides positive insights into the model, this explains that the EfficientNet-B3 Modified WOA is more than feasible to be integrated as a clinical decision support system.

From various evaluations of the performance and stability of DSS as a confidence-based decision support system, the next step is dedicated to implementing the WOA-optimized EfficientNet-B3 model, at the interface of the clinical decision support system. Figure 10 shows that the implementation translates the output of the model in the form of class predictions, confidence levels, accept/reject decision status, and clinical recommendations, into a structured and intuitive interface. This supports practical, transparent, and real-time decision-making for its users in the clinical realm.

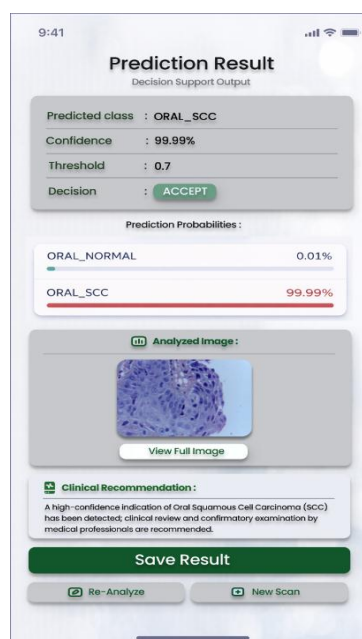


Figure 10. Results of EfficientNet-B3+WOA Implementation

This study proves that hyperparameter optimization using the Whale Optimization Algorithm (WOA) has a very significant influence in improving the performance of the baseline deep learning model in the classification of medical images. All model architectures tested, especially in the baseline model, provided optimization that induced an increase in the evaluation score, including in the accuracy metric and F1-score. This shows that WOA implements hyperparameters optimally. Pipelines with WOA in advanced performance-enhancing deep learning models contribute to improving performance from baseline to moderate to quite high with poor performance.

These findings are in line with several previous studies that emphasized the effectiveness of optimization algorithms in improving the stability and generalization of deep learning models. Albahli et al. report that the integration of optimization algorithms into CNN training is able to improve classification performance while reducing the risk of overfitting medical image data. Sharma et al. show that the EfficientNet architecture is heavily influenced by the selection of hyperparameters, thus, proper optimization can result in substantial improvements in performance. In addition, Mirjalili and Lewis emphasized that the Whale Optimization Algorithm is very effective given the excellent balance of exploration and exploitation to solve optimization problems in high dimensions, such as adjusting deep learning hyperparameters. From the perspective of decision support systems, Esteva et al. emphasize that improvements in evaluation metrics, in order to be clinically applicable, must be supported by stable training behaviors and reliable predictions from the model. Therefore, the findings of this study show not only numerically superior results, but also reinforce the methodological validity of the proposed approach.

The study found that the integration of Whale Optimization Algorithm-based hyperparameter optimization and proper deep learning architecture can result in models with high performance, good training stability, and can be implemented for image-based decision support systems, thereby making a relevant methodological contribution to the development of reliable artificial intelligence systems.

4. CONCLUSION

This study successfully proves that the integration of deep learning with WOA hyperparameter optimization significantly improves the classification performance and reliability of decision support systems. The results of the experiment corresponded to the research objectives stated in the introduction where all models improved after optimization as indicated by the F1 score which increased from baseline 0.80–0.84 to above 0.97 after optimization for all models. The best model, the EfficientNet-B3 WOA, achieved an accuracy of 98.73% and an F1 score on test data with stable training behavior that converged without significant signs of overfitting.

In terms of implementation, EfficientNet-B3 WOA creates a good balance between accuracy and efficiency with an inference time of 0.6816 ms per image on the GPU. Trust-based decision-making from DSS with stable confidence in the optimal coverage range of 0.60–0.90. The rejection option mechanism allows the system to ignore predictions with low confidence which reduces the likelihood of making unreliable decisions. In light of the above, this study confirms that the proposed method is not only quantitatively superior but also ready for practical use as an accurate, reliable, and stable DSS. For the future, further development can focus on multi-class classification scenarios, explainable artificial intelligence integration, and system implementation on larger-scale clinical platforms.

ACKNOWLEDGEMENTS

The author expresses his appreciation and gratitude to Telkom University for the institutional support, research facilities, and conducive academic environment so that this research can be carried out properly. Thank you are also expressed to all parties who have contributed directly or indirectly to the provision of computing resources, academic discussions, and technical support during the research process and preparation of this article.

REFERENCES

- [1] I. Perera *et al.*, “An overview of the burden of oral cancer in Sri Lanka and its inequalities in the face of contemporary economic and social malaise,” *Community Dent. Oral Epidemiol.*, vol. 51, no. 4, pp. 680–696, Aug. 2023, doi: 10.1111/cdoe.12888.
- [2] M. D. L. G. Aguiñaga *et al.*, “Tongue cancer: An overview and current update,” *International Journal of Applied Dental Sciences*, vol. 8, no. 3, pp. 356–360, Jul. 2022, doi: 10.22271/oral.2022.v8.i3d.1612.
- [3] Ms. M. Kudale, “Epidermiology, Risk Factor and Treatment of Oral Cancer,” *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 13, no. 4, pp. 4394–4399, Apr. 2025, doi: 10.22214/ijraset.2025.69232.
- [4] A. K. Jain, “Oral Cancer Screening: Insights into Epidemiology, Risk Factors, and Screening Programs for Improved Early Detection,” *Cancer Screening and Prevention*, vol. 3, no. 2, pp. 97–105, Jun. 2024, doi: 10.14218/CSP.2023.00029S.
- [5] H. Seok, “Role of oral and maxillofacial surgeons in treating oral cancer,” *J. Korean Assoc. Oral Maxillofac. Surg.*, vol. 48, no. 6, pp. 329–330, Dec. 2022, doi: 10.5125/jkaoms.2022.48.6.329.
- [6] U. Tomić *et al.*, “Comparison of transepithelial cytology and histopathology in the diagnosis of potentially malignant and malignant lesions of the oral mucosa,” *Balkan Journal of Dental Medicine*, vol. 26, no. 2, pp. 113–117, 2022, doi: 10.5937/bjdm20220531-001.
- [7] C. Gupta, A. Anjum, S. Sharma, and M. Jamwal, “Histopathological spectrum of lesions of oral cavity and oropharynx in a tertiary care center,” *Panacea Journal of Medical Sciences*, vol. 15, no. 1, pp. 14–17, Mar. 2025, doi: 10.18231/pjms.v.15.i.1.14-17.
- [8] T. Mezei, M. Kolcsár, A. Joó, and S. Gurzu, “Image Analysis in Histopathology and Cytopathology: From Early Days to Current Perspectives,” *J. Imaging*, vol. 10, no. 10, p. 252, Oct. 2024, doi: 10.3390/jimaging10100252.
- [9] F. Ghezloo *et al.*, “An analysis of pathologists’ viewing processes as they diagnose whole slide digital images,” *J. Pathol. Inform.*, vol. 13, p. 100104, 2022, doi: 10.1016/j.jpi.2022.100104.
- [10] M. Unger and J. N. Kather, “Deep learning in cancer genomics and histopathology,” *Genome Med.*, vol. 16, no. 1, p. 44, Mar. 2024, doi: 10.1186/s13073-024-01315-6.
- [11] V. Ramachandra, “Histopathology: Deep machine learning based semantic segmentation features predict patient survival,” Jan. 14, 2023, doi: 10.1101/2023.01.14.23284554.
- [12] J. K. Sandhu, C. Sharma, A. Kaur, S. K. Pandey, A. Sinha, and J. Shreyas, “Development of a clinical decision support system for breast cancer detection using ensemble deep learning,” *Sci. Rep.*, vol. 15, no. 1, p. 26098, Jul. 2025, doi: 10.1038/s41598-025-06784-2.
- [13] H. T. N. PHAN and A. AKTER, “HYBRID MACHINE LEARNING APPROACH FOR ORAL CANCER DIAGNOSIS AND CLASSIFICATION USING HISTOPATHOLOGICAL IMAGES,” *International Journal of Medical Science and Dental Health*, vol. 11, no. 01, pp. 63–76, Jan. 2025, doi: 10.55640/ijmsdh-11-01-08.
- [14] Z. Yang, H. Pan, J. Shang, J. Zhang, and Y. Liang, “Deep-Learning-Based Automated Identification and Visualization of Oral Cancer in Optical Coherence Tomography Images,” *Biomedicines*, vol. 11, no. 3, p. 802, Mar. 2023, doi: 10.3390/biomedicines11030802.
- [15] K. C. Figueroa *et al.*, “Interpretable deep learning approach for oral cancer classification using guided attention inference network,” *J. Biomed. Opt.*, vol. 27, no. 01, Jan. 2022, doi: 10.1117/1.JBO.27.1.015001.
- [16] V. P. G. Pushparathi, S. V. NarayanS.R, P. R.S., and V. Naveen, “Histopathological image analysis and enhanced diagnostic accuracy explainability for oral cancer detection,” *Polish Journal of Pathology*, vol. 76, no. 2, pp. 120–130, 2025, doi: 10.5114/pjp.2025.153973.
- [17] V. N. Thatha *et al.*, “Histopathological image based breast cancer diagnosis using deep learning and bio inspired optimization,” *Sci. Rep.*, vol. 15, no. 1, p. 19034, May 2025, doi: 10.1038/s41598-025-04136-8.
- [18] M. K. Karuppan Perumal, R. Rajan Renuka, S. Kumar Subbiah, and P. Manickam Natarajan, “Artificial intelligence-driven clinical decision support systems for early detection and

- precision therapy in oral cancer: a mini review,” *Frontiers in Oral Health*, vol. 6, Apr. 2025, doi: 10.3389/froh.2025.1592428.
- [19] O. F. Altal *et al.*, “WOAENet: a whale optimization-guided ensemble deep learning with soft voting for uterine cancer diagnosis based on MRI images,” *Front. Artif. Intell.*, vol. 8, Oct. 2025, doi: 10.3389/frai.2025.1664201.
- [20] S. B. Urundai Meeran, “AI-Powered Skin Lesion Diagnosis using Whale Optimization Algorithm Enhanced ResNet 50 for Cancer Prediction,” *Asian Pacific Journal of Cancer Prevention*, vol. 26, no. 8, pp. 2919–2928, Aug. 2025, doi: 10.31557/APJCP.2025.26.8.2919.
- [21] E. al. Vijay Arumugam R., “Modelling of Whale Optimization Algorithm with Deep Learning for Skin Cancer Diagnosis on Dermoscopic Images,” *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 11, no. 10, pp. 2018–2026, Nov. 2023, doi: 10.17762/ijritcc.v11i10.8885.
- [22] H. S. Ali, A. I. Ismail, E. M. El-Rabaie, and F. E. Abd El-Samie, “Deep residual architectures and ensemble learning for efficient brain tumour classification,” *Expert Syst.*, vol. 40, no. 6, Jul. 2023, doi: 10.1111/exsy.13226.
- [23] M. H. Abid, R. Ashraf, T. Mahmood, and C. M. N. Faisal, “Multi-modal medical image classification using deep residual network and genetic algorithm,” *PLoS One*, vol. 18, no. 6, p. e0287786, Jun. 2023, doi: 10.1371/journal.pone.0287786.
- [24] T. Zhang, “Performance Analysis of Residual Networks for Pneumonia Diagnosis on Chest X-Rays,” *Applied and Computational Engineering*, vol. 138, no. 1, pp. 104–109, Mar. 2025, doi: 10.54254/2755-2721/2025.21363.
- [25] T. Kanimozhi *et al.*, “RWHO: A hybrid of CNN architecture and optimization algorithm to predict basal cell carcinoma skin cancer in dermoscopic images,” *The Scientific Temper*, vol. 15, no. 02, pp. 2138–2142, Jun. 2024, doi: 10.58414/SCIENTIFICTEMPER.2024.15.2.25.
- [26] A. E. Minarno, L. R. Wandani, and Y. Azhar, “Classification of Breast Cancer Based on Histopathological Image Using EfficientNet-B0 on Convolutional Neural Network,” *International Journal of Emerging Technology and Advanced Engineering*, vol. 12, no. 8, pp. 70–77, Aug. 2022, doi: 10.46338/ijetae0822_09.
- [27] B. Jacob, R. S. V. Kumar, and S. S. Kumar, “Multiple Scaling Based EfficientNet Modelling for Liver Tumor Classification on CT Images,” *International Journal Bioautomation*, vol. 28, no. 3, pp. 151–160, Sep. 2024, doi: 10.7546/ijba.2024.28.3.001001.
- [28] A. Bruno, G. Ignesti, O. Salvetti, D. Moroni, and M. Martinelli, “Efficient Lung Ultrasound Classification,” *Bioengineering*, vol. 10, no. 5, p. 555, May 2023, doi: 10.3390/bioengineering10050555.
- [29] M. Naeeni Davarani *et al.*, “Efficient segmentation of active and inactive plaques in FLAIR-images using DeepLabV3Plus SE with efficientnetb0 backbone in multiple sclerosis,” *Sci. Rep.*, vol. 14, no. 1, p. 16304, Jul. 2024, doi: 10.1038/s41598-024-67130-6.
- [30] T.-B.-T. Nguyen, M.-V. Ngo, and V.-P. Nguyen, “Histopathological Imaging Classification of Breast Tissue for Cancer Diagnosis Support Using Deep Learning Models,” 2022, pp. 152–164. doi: 10.1007/978-3-031-08878-0_11.
- [31] S. Anjum *et al.*, “Lung Cancer Classification in Histopathology Images Using Multiresolution Efficient Nets,” *Comput. Intell. Neurosci.*, vol. 2023, no. 1, Jan. 2023, doi: 10.1155/2023/7282944.
- [32] D. Beaulah Princiba, P. Ezhilarasi, and S. Rajeshkannan, “CSMEC-based deep learning model for detection and classification of brain tumours in MR images,” *Neural Comput. Appl.*, vol. 36, no. 29, pp. 18479–18498, Oct. 2024, doi: 10.1007/s00521-024-10168-4.
- [33] K. Ding, Z. Niu, J. Hui, X. Zhou, and F. T. S. Chan, “A Weld Surface Defect Recognition Method Based on Improved MobileNetV2 Algorithm,” *Mathematics*, vol. 10, no. 19, p. 3678, Oct. 2022, doi: 10.3390/math10193678.
- [34] D. Bağcı Daş, “An Intelligent and Lightweight Approach Based on MobilenetV2 Architecture for Identifying Brain Tumors,” *Sakarya University Journal of Computer and Information Sciences*, vol. 8, no. 3, pp. 392–399, Sep. 2025, doi: 10.35377/saucis...1590213.

- [35] F. Shahoveisi, H. Taheri Gorji, S. Shahabi, S. Hosseinirad, S. Markell, and F. Vasefi, "Application of image processing and transfer learning for the detection of rust disease," *Sci. Rep.*, vol. 13, no. 1, p. 5133, Mar. 2023, doi: 10.1038/s41598-023-31942-9.
- [36] H. M. Whitney, K. Drukker, and M. L. Giger, "Performance metric curve analysis framework to assess impact of the decision variable threshold, disease prevalence, and dataset variability in two-class classification," *Journal of Medical Imaging*, vol. 9, no. 03, May 2022, doi: 10.1117/1.JMI.9.3.035502.
- [37] Y. J. Hong *et al.*, "Deep learning-based classification of pleural malignancy using medical thoracoscopic images," *Transl. Lung Cancer Res.*, vol. 14, no. 10, pp. 4475–4484, Oct. 2025, doi: 10.21037/tlcr-2025-588.
- [38] H.-S. Ham, H.-S. Lee, J.-W. Chae, H. C. Cho, and H.-C. Cho, "Improvement of Gastroscopy Classification Performance Through Image Augmentation Using a Gradient-Weighted Class Activation Map," *IEEE Access*, vol. 10, pp. 99361–99369, 2022, doi: 10.1109/ACCESS.2022.3207839.