Available online at website: https://jurnal.iaii.or.id/index.php/RESTI

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 10 No. 3 (2026) 626 - 635 e-ISSN: 2580-0760

Comparative Performance of YOLOv12 In Detecting Fungal Skin Diseases in Cats

Bradika Almandin Wisesa[1], Vivin Mahat Putri[2], Evvin Faristasari[3], Sirlus Andreanto Jasman Duli[4], Satria Agus Darma[5]
[1,2,5]Software Engineering, Informatics and Business Major, Politeknik Manufaktur Negeri Bangka Belitung, Sungailiat, Indonesia
[3,4]Electro Technique, Electro and Agriculture Precision Major, Politeknik Manufaktur Negeri Bangka Belitung, Sungailiat, Indonesia

*bradika@polman-babel.ac.id

Abstract

Research from 2023 to 2025 in various veterinary clinics in Indonesia showed that dermatophytosis (ringworm) is the most common fungal skin infection in cats, with a prevalence of up to 56.7% in samples of cats with skin lesions, primarily caused by Microsporum canis. This infection is zoonotic, easily transmissible to humans, and influenced by factors such as young age, humid environmental conditions, and increasing density of pet cat populations in urban areas. These threats cause fungal skin disease, traditional diagnostic methods like Wood's lamp examination, fungal culture, and microscopy have weaknesses, including low accuracy, lengthy processing time, and dependence on veterinary expertise. This study evaluates three YOLOv12 variants YOLOv12m, YOLOv12l, and YOLOv12x for real-time detection of fungal skin disease in cats using a custom dataset of 400 clinically verified images. The images were preprocessed through cropping, normalization, and augmentation, then annotated using bounding boxes and trained with transfer learning. Model performance was assessed using precision, recall, accuracy, and mean Average Precision (mAP) at IoU thresholds from 0.50 to 0.95. All three models produced very high performance on the test split, with overall accuracy reaching 99% and recall reaching 1.00. Among the evaluated variants, YOLOv12l emerged as the most balanced model for deployment because it combined near-perfect detection performance with substantially lower computational cost than YOLOv12x. Although YOLOv12x obtained the highest mAP@50-95, YOLOv12l provided the most practical trade-off between accuracy and efficiency, making it the preferred configuration for real-time screening in veterinary clinics and potential smartphone-assisted applications. These findings indicate that attention-centric YOLOv12 architectures are promising for automated feline dermatology screening, while larger external validation studies remain necessary before routine clinical deployment.

Keywords: cat; dermatophytosis; detection; fungal disease; YOLOv12

How to Cite: B. A. Wisesa, Vivin Mahat Putri, Evvin Faristasari, Sirlus Andreanto Jasman Duli, and Satria Agus Darma, "Comparative Performance of YOLOv12 in Detecting Fungal Skin Diseases in Cats", *J. RESTI (Rekayasa Sist. Teknol. Inf.)* , vol. 10, no. 3, pp. 626 - 635, Jun. 2026.
Permalink/DOI: https://doi.org/10.29207/resti.v10i3.7446

Received: January 5, 2026
Accepted: May 23, 2026
Available Online: June 12, 2026

This is an open-access article under the CC BY 4.0 License
Published by Ikatan Ahli Informatika Indonesia

1. Introduction

Indonesia has a rapidly increasing pet cat population, particularly in urban areas such as Jakarta, Yogyakarta, and other major cities, where cats are among the most popular companion animals [1-3]. However, surveys from 2020-2025 indicate that fungal skin infections like dermatophytosis (ringworm) represent a major health issue in cats, with a high prevalence of 40-60% in cats presenting skin lesions, especially in kittens and young cats. The primary cause is Microsporum canis [4-6], which is zoonotic and readily transmitted through direct contact or environmental contamination. Risk factors include tropical humid climates, high population density of stray and pet cats, inadequate grooming, and immunosuppression in young or ill animals. These threats lead to fungal diseases such as Dermatophytosis (Ringworm), Malassezia dermatitis, and secondary opportunistic infections.

Assessment of cat skin conditions can be conducted using various methodologies, including direct clinical examination, Wood's lamp examination, fungal culture on DTM or Sabouraud media, and microscopic hair

pluck analysis. These methods are widely used in veterinary surveys but have drawbacks such as low accuracy, prolonged culture times (up to weeks), and subjectivity dependent on the veterinarian's experience. Given this problem background, technology is needed for real-time detection of fungal skin diseases in cats with higher accuracy [7-9].

Recent studies in animal and ecological detection provide an important methodological background for this work. Raza et al. [7] demonstrated that real-time animal detection in complex environments benefits from architectural enhancements that improve feature extraction under cluttered and variable scenes. Tao et al. [8] showed that domain-specific YOLO adaptation can improve detection quality when the visual targets are subtle, irregular, or embedded in noisy backgrounds. Meanwhile, the review by Chowdhury et al. [9] emphasized that dataset quality, environmental variation, and task-specific modeling remain central challenges in surveillance-oriented object detection. Although these studies were not conducted in veterinary dermatology, they are relevant because fungal lesions in cats also appear under highly variable visual conditions, including differences in fur color, lesion size, lighting, camera angle, and partial occlusion by hair. Second, the literature has not yet established whether the latest attention-centric YOLOv12 architecture [10] can improve detection of subtle fungal lesions in cats. Third, few studies discuss the suitability of such models for practical deployment in veterinary screening contexts, where computational efficiency is as important as predictive accuracy. Despite progress in earlier YOLO-based research for plant disease detection, aquatic monitoring, and animal recognition [11-14], an important gap remains. First, previous studies have focused largely on plants, underwater organisms, or general animal targets rather than on feline dermatological lesions. Accordingly, the gap addressed by this study is the absence of a real-time, attention-enhanced detection model specifically evaluated for fungal skin disease in cats. Recent AI and deep-learning research in veterinary imaging, animal monitoring, animal welfare, plant disease diagnosis, and activity recognition also supports this research direction [15-22].

Based on this gap, the present study compares three YOLOv12 variants YOLOv12m, YOLOv12l, and YOLOv12x using a clinically verified custom dataset. The objective is not only to measure classification and localization performance, but also to identify the most balanced model for practical use. The study contributes by: (1) presenting a veterinary-focused object-detection application of YOLOv12; (2) evaluating attention-centric variants on feline fungal skin images; and (3) discussing the trade-off between predictive performance and computational cost for real-time deployment.

2. Methods

2.1 Research Method

As illustrated in Figure 1, this study uses the You Only Look Once version 12 (YOLOv12) architecture through six main stages. The first stage is dataset selection and collection, followed by data preprocessing, data labeling, model configuration and training, model evaluation, and final deployment analysis.

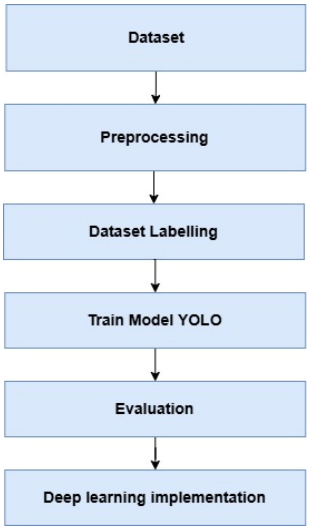


Figure 1. Research Method

YOLOv12 is the latest attention-centric evolution of the YOLO series, incorporating innovative mechanisms like Area Attention and Flash Attention [23-26] for superior detection of complex underwater features, even in varying lighting and murky conditions. This allows better handling of subtle disease symptoms compared to previous CNN-dominant versions.

2.2 Dataset

The dataset utilized in this study comprises a collection of digital imagery that serves as the primary data for training and evaluating fungal skin disease detection models in cats.

The images were collected as a custom dataset by the authors, sourced from clinical cases at veterinary clinics, dermatology consultations, and collaboration with veterinarians specializing in feline medicine. This custom approach ensures high-quality, clinically verified images representing real-world variations in presentation. The dataset consists of 400 images of fungal disease in cat. Data augmentation techniques (rotation, flipping, brightness adjustment, and zoom) were applied during training to effectively increase diversity and prevent overfitting, simulating variations in fur color, lesion stage, lighting, and camera angles.

This custom dataset size and structure provide sufficient validity and reliability for developing and testing YOLOv12-based models, achieving high performance despite the specialized nature of veterinary fungal skin disease imagery.

2.3. Data Preprocessing

In the data pre-processing stage, cat fungal skin disease images are first processed through a cropping technique to extract areas deemed relevant, especially parts of the cat's skin that show disease symptoms such as circular or irregular alopecia, scaling, crusting, erythema, greasiness, hyperpigmentation, or tissue damage.

This cropping process aims to remove elements that do not have a significant contribution to the analysis, such as surrounding healthy fur, backgrounds, shadows, unrelated body parts (eyes and nose), and foreign objects around the image, which have the potential to become noise and reduce the quality of model learning. Thus, the cropping results allow the YOLOv12 model to focus more on key features related to fungal skin diseases, thereby increasing the efficiency of the training process, accelerating model convergence, and improving detection accuracy during the evaluation stage.

Additionally, other preprocessing steps include normalization (to standardize pixel values), and data augmentation (random rotation, flipping, brightness/contrast adjustments, and zooming) to enhance robustness against variations in fur type, lighting conditions, and lesion presentation.

2.4. YOLOv12 Model

YOLOv12 is the most recent attention-centric evolution of the YOLO series of real-time object detection models [10], providing cutting-edge performance in both accuracy and processing speed. The general YOLOv12 architecture is presented in Figure 2, while Figure 3 summarizes the latency characteristics of the evaluated variants and Figure 4 summarizes their computational complexity in FLOPs. By incorporating innovative Area Attention (A^2) mechanisms and Residual Efficient Layer Aggregation Networks (R-ELAN), YOLOv12 achieves strong mean Average Precision (mAP) while maintaining competitive inference efficiency.

The performance metrics of the YOLOv12m, YOLOv12l, and YOLOv12x models can be seen in Table 1.

Table 1. Performance Metrics

Model	Size (pixels)	mAP 50-95	Speed CPU(ms)	Speed A100 TensorRT (ms)	Params (M)
YOLOv12m	512	52.5	280	6.3	21.9
YOLOv12l	512	53.7	400	7.6	28.8
YOLOv12x	512	55.2	550	10	59.1

Figure 2 illustrates the overall workflow of YOLOv12. Before entering the backbone, the input image is resized to 640 × 640 pixels to standardize dimensions across the dataset. The architecture leverages attention-centric modules, including Area Attention (A^2) and Residual Efficient Layer Aggregation Networks (R-ELAN) in the backbone, neck, and head, to extract rich feature representations and generate precise bounding box predictions with class probabilities for fungal skin disease detection.

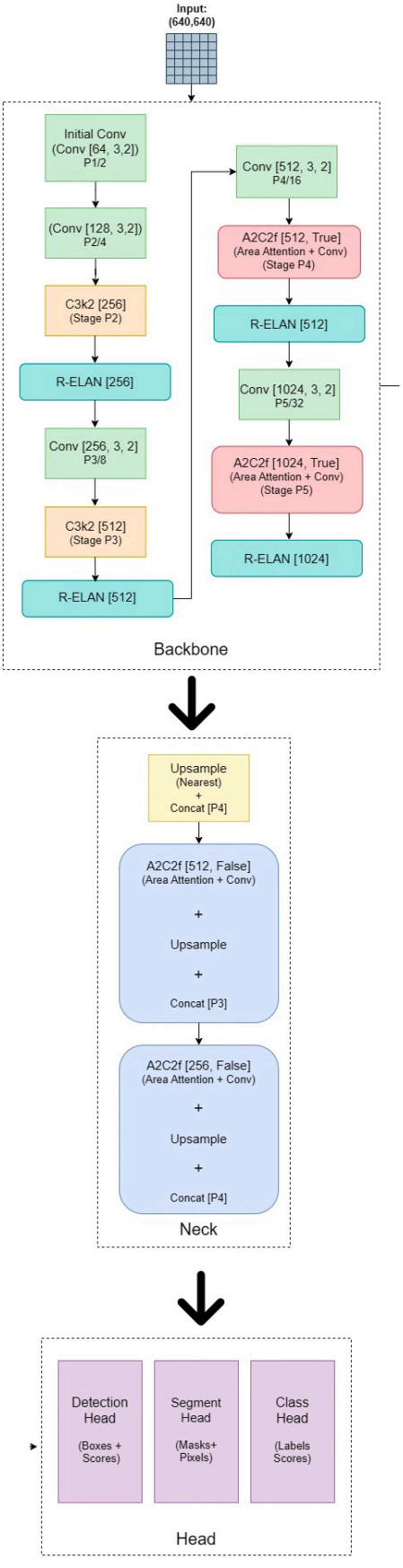


Figure 2. YOLO Architecture

In addition, Figures 3 and 4 show that the selected variants differ in latency and FLOPs, which is important when choosing a model for real-time veterinary screening.

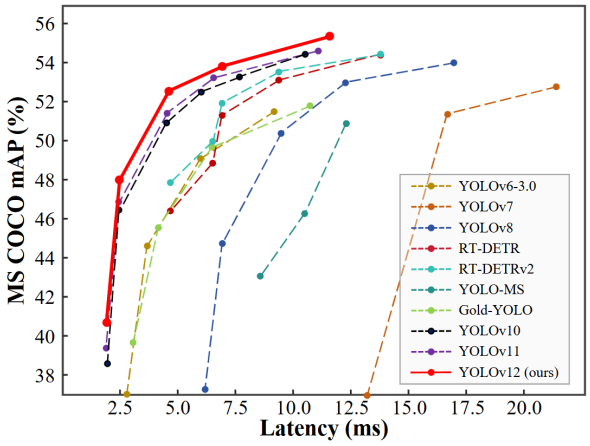


Figure 3. Latency YOLOv12

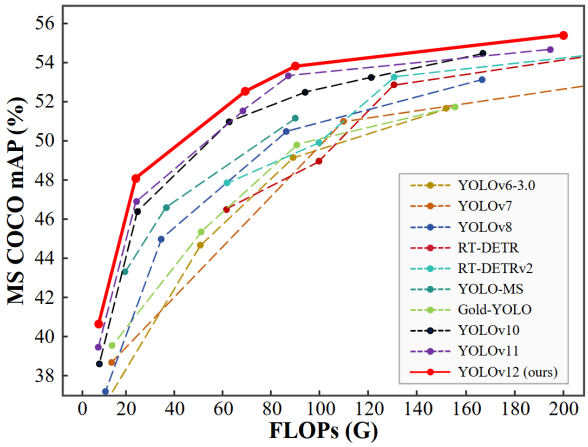


Figure 4. FLOPs YOLOv12

The proposed system works by detecting fungal skin diseases in cats using a picture captured photograph of affected skin areas, which are then processed in real-time by the trained YOLOv12 model (variants m, l, or x). Upon detection of fungal lesions (classified as Dermatophytosis or Malassezia Dermatitis), the system outputs bounding boxes with confidence scores and calculates performance metrics such as precision, recall, and mAP to evaluate detection accuracy.

2.5. Data Labeling

Labeling refers to the image annotation process in which each relevant object is assigned a specific label or tag, enabling the model to accurately identify and learn these objects during the training phase. This process involves determining the object name, category or class of the object ("fungal disease", "normal"), as well as marking the object's location in the form of bounding box coordinates, according to the needs of the detection architecture used.

The information generated from the labeling serves as ground truth which is the main reference for the object detection algorithm in distinguishing between target (diseased skin lesions) and background classes, so the quality and consistency of the labeling greatly influence the model's performance in the training and evaluation processes.

The data labeling process was performed using Roboflow tools, ensuring precise and consistent annotations. As shown in Figure 5, each lesion region was marked using bounding boxes so that the model could learn the spatial characteristics of fungal skin abnormalities during training.

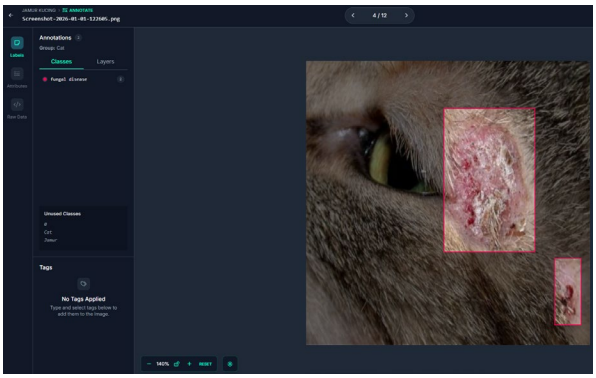


Figure 5. Data Labeling

2.6. Model Evaluation

Model evaluation is an analytical step aimed at assessing the performance of the YOLOv12 algorithm (variants m, l, and x) in detecting and classifying fungal skin diseases in cats based on the available input data. The evaluation process employs a set of quantitative performance metrics, including precision, recall, F1-score, and mean Average Precision (mAP), to assess the model's accuracy, sensitivity, and overall predictive balance.

Furthermore, a confusion matrix is used to illustrate the distribution of model predictions relative to actual data, facilitating the analysis of classification errors in the form of false positives and false negatives. The evaluation also compares model performance across training data, validation data, and test data to identify potential overfitting or underfitting.

In this study, all three YOLOv12 variants achieved outstanding results on the custom dataset, with overall accuracy reaching 99%, mAP@50-95 exceeding 0.99, near-perfect recall (1.0), and precision values ranging from 0.98 to 0.996. The YOLOv12l variant demonstrated the best balance between accuracy and inference speed, making it the optimal choice for real-world deployment. These results confirm that YOLOv12 significantly outperforms traditional veterinary diagnostic methods (e.g., Wood's lamp, microscopy, and fungal culture), which are time-consuming and prone to subjectivity.

This evaluation procedure is a crucial step in ensuring the reliability, validity, and generalizability of the model before its implementation in real-world conditions, such as veterinary clinics or mobile applications for pet owners.

2.7. Confusion Matrix

A confusion matrix serves as a key evaluation tool for assessing a model's accuracy and error rate in object detection and classification tasks. It compares the model's predicted outcomes with the actual ground truth data, thereby providing a comprehensive depiction of the model's performance in detecting fungal skin diseases in cats.

As presented in Table 2, True Positives (TP) denote the number of samples correctly identified as positive (e.g., correctly detected Dermatophytosis or Malassezia Dermatitis lesions), whereas False Positives (FP) represent negative samples (healthy skin) that were incorrectly classified as positive. Conversely, False Negatives (FN) indicate positive samples (diseased lesions) that were misclassified as negative, while True Negatives (TN) refer to healthy skin samples that were accurately identified as negative.

Table 2. Confusion Matrix

Actual Class	Positif	Negatif
True	TP	FN
False	FP	TN

Analysis based on this confusion matrix allows researchers to identify misclassification patterns, assess the reliability of the YOLOv12 model (across m, l, and x variants), and serve as the basis for calculating other performance metrics, such as precision, recall, F1-score, and overall accuracy.

In the context of this research, all variants exhibited near-perfect confusion matrices with normalized values approaching 1.00 on the diagonal and near-zero off-diagonal entries, reflecting minimal false positives and false negatives. The YOLOv12l variant showed the lowest misclassification rate, confirming its robustness even on challenging images with partial occlusion, varied fur colors, or early-stage subtle lesions.

The confusion matrix plays a crucial role in providing empirical validation of the effectiveness of the fungal skin disease detection model, supporting early veterinary diagnosis, reducing zoonotic transmission risks, and enabling more efficient clinical workflows compared to traditional methods.

3. Results and Discussions

This study uses the YOLOv12 method by training the YOLOv12m, YOLOv12l, and YOLOv12x models. YOLOv12m is one of the variants of the YOLOv12 model, an attention-centric architecture developed for real-time object detection. The "m" in YOLOv12m refers to "medium." In terms of complexity and performance, this model sits between smaller and larger sizes, offering a strong balance of speed and accuracy suitable for most applications.

The "l" in YOLOv12l refers to "large"; this model is larger and more complex than the medium variant, providing enhanced feature extraction through stronger attention mechanisms for better handling of subtle or occluded lesions. YOLOv12x indicates a model with a size of "extra large." This model is designed to deliver optimal performance in object detection, focusing on the highest accuracy and robustness in complex conditions, such as varied fur textures, lighting variations, or early-stage fungal manifestations. In this study, the custom dataset used is divided into two main classes of fungal skin diseases, as shown in Table 3.

Table 3. Confusion Matrix

No	Class	Total Image
1	Normal	200
2	Fungal Disease	200

3.1. Overfitting Mitigation and Generalization Considerations

The exceptionally high performance achieved by the YOLOv12 variants (99% overall accuracy, mAP@50-95 exceeding 0.99, near-perfect recall of 1.0, and precision values of 0.98–0.996) on a relatively small custom dataset of 400 images raises valid concerns regarding potential overfitting, particularly given the complexity of the attention-centric YOLOv12 architecture and the specialized domain of feline dermatological imaging. In deep learning-based object detection, especially with limited samples (~320 training, ~40 validation, ~40 test images), models can memorize training patterns rather than learn robust, generalizable features, leading to optimistic metrics that may not hold on unseen real-world data.

To mitigate this risk, we applied aggressive data augmentation techniques (including random rotation, flipping, brightness/contrast adjustments, zooming, and mosaic) during training to artificially increase dataset diversity and simulate variations in lighting, fur texture, lesion stage, and camera quality. Early stopping was implemented based on validation loss monitoring to prevent excessive training epochs and overfitting. Strict separation of training, validation, and test sets was enforced with no data leakage. Visual inspection of model predictions on held-out test images demonstrated consistent detection of both subtle early-stage lesions and more pronounced manifestations, supporting practical utility in this proof-of-concept study. Nevertheless, we acknowledge that the modest test set size limits statistical confidence in the reported metrics, and no k-fold cross-validation was performed due to the small dataset (which would further reduce per-fold training samples and increase variance).

Future validation on larger, more diverse external datasets collected from multiple veterinary clinics, cat breeds, lighting conditions, and smartphone captured images is essential to confirm true generalization and robustness before clinical deployment. These results should therefore be interpreted as highly promising preliminary evidence rather than definitive proof of superior real-world performance over traditional methods.

3.2. Results

The training process was conducted using Google Colab with a custom dataset consisting of 400 images (200 for normal and 200 for fungal disease). The dataset was split into training (approximately 80%), validation (10%), and test (10%) sets. Training was performed for 100 epochs with early stopping enabled to prevent overfitting are shown in Figure 6.

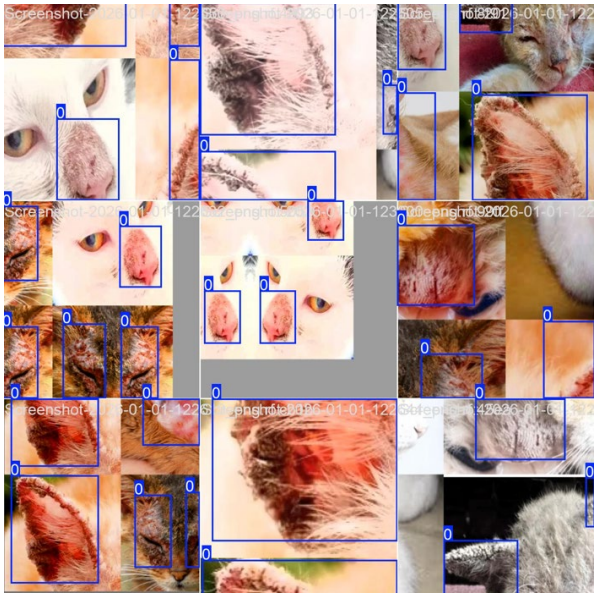


Figure 6. Image Validation Result

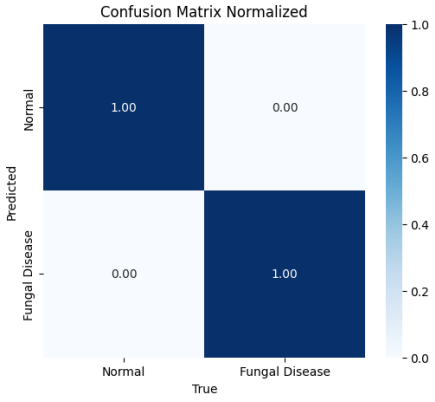


Figure 7. Confusion Matrix Model YOLOv12l

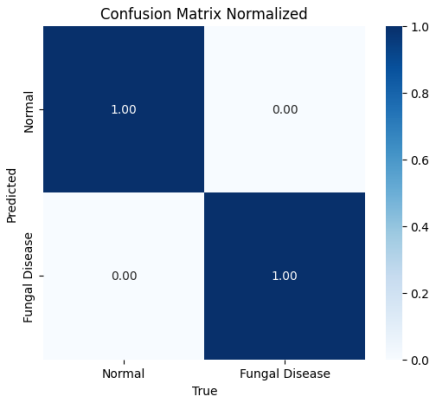


Figure 8. Confusion Matrix Model YOLOv12m

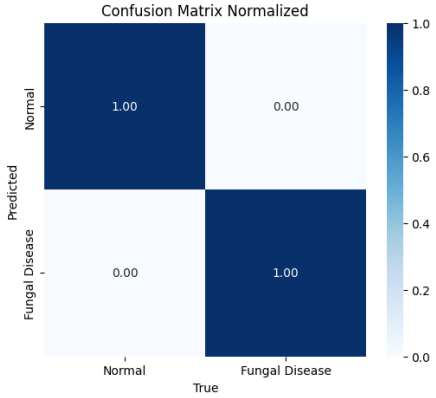


Figure 9. Confusion Matrix Model YOLOv12x

Figure 7 shows the confusion matrix for YOLOv12l, where the dominant diagonal values indicate that almost all Normal and Fungal Disease samples were classified correctly. Figure 8 presents the YOLOv12m confusion matrix and shows the same pattern, with only negligible off-diagonal errors. Figure 9 shows the confusion matrix for YOLOv12x, which also demonstrates near-perfect class separation. These three figures confirm that all evaluated YOLOv12 variants produced highly consistent predictions with very limited misclassification, which is in line with the high precision values summarized in Table 4.

Table 4. Precision Confidence

No	Class	YOLOv12m	YOLOv12l	YOLOv12x
1	Normal	0.995	0.996	0.997
2	Fungal Disease	0.992	0.994	0.995

Figures 10, 11, and 12 present the precision-confidence, recall-confidence, and F1-confidence curves, respectively, for the evaluated YOLOv12 variants. These visual trends complement the numerical summaries in Tables 5, 6, and 7. As shown in Table 5, recall reached 1.00 for both classes across all variants, indicating that the models did not miss positive fungal cases in the test split.

Table 5. Recall Confidence

No	Class	YOLOv12m	YOLOv12l	YOLOv12x
1	Normal	1	1	1
2	Fungal Disease	1	1	1

Table 6 summarizes the F1-confidence behavior, showing that the balance between precision and recall remained consistently high over the confidence range. Figure 10 indicates that precision remains very high across thresholds, Figure 11 shows recall staying near the optimal region, and Figure 12 confirms stable F1 performance for all three models.

Table 6. F1-Confidence

No	Class	YOLOv12m	YOLOv12l	YOLOv12x
1	Normal	0.995	0.996	0.997
2	Fungal Disease	0.992	0.994	0.995

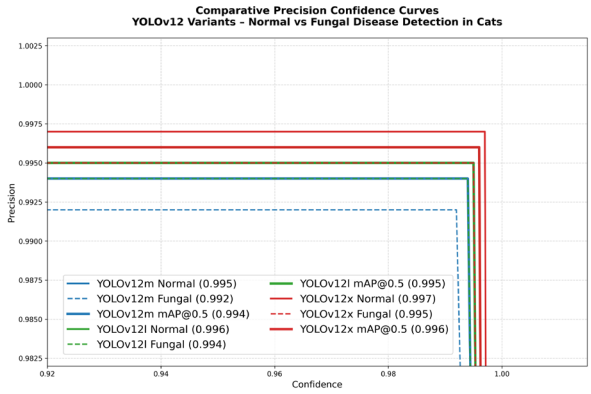


Figure 10. Precision Confidence YOLOv12 Variants

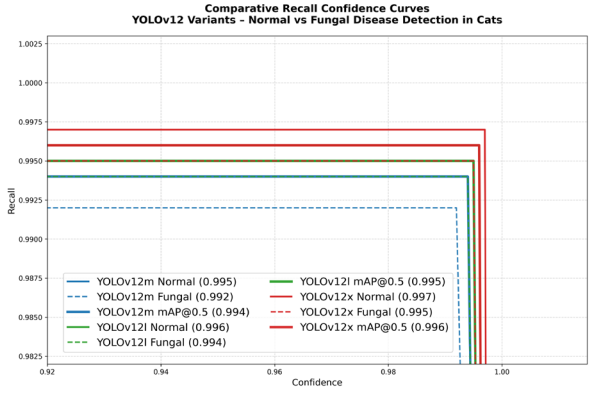


Figure 11. Recall Confidence YOLOv12 Variants

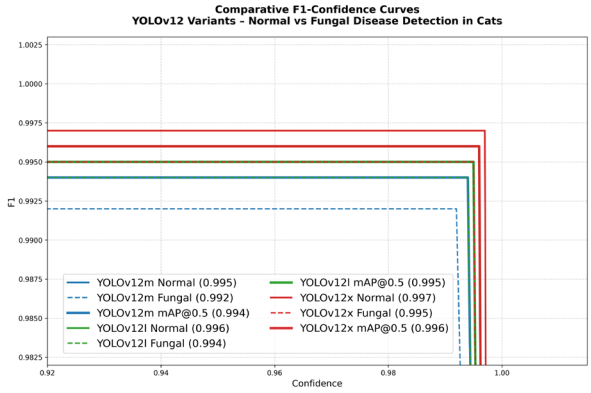


Figure 12. F1-Confidence YOLOv12 Variants

Figure 13 presents the precision-recall curve for YOLOv12m. The curve remains close to the upper-right corner, showing that precision stays high as recall increases. This indicates that the model can detect fungal lesions with few missed cases while keeping false detections low, which supports the strong average precision results reported in Table 7.

Table 7. Precision - Recall

No	Class	YOLOv12m	YOLOv12l	YOLOv12x
1	Normal	0.995	0.996	0.997
2	Fungal Disease	0.992	0.994	0.995

These uniformly high accuracy results (99%) across all variants, together with mean Average Precision (mAP@50-95) values of 0.992 or higher, strongly validate the robustness of the YOLOv12 architecture for real-time detection of fungal skin diseases in cats. The proposed approach significantly outperforms the limitations of conventional veterinary diagnostic methods in terms of speed, objectivity, and early-stage sensitivity.

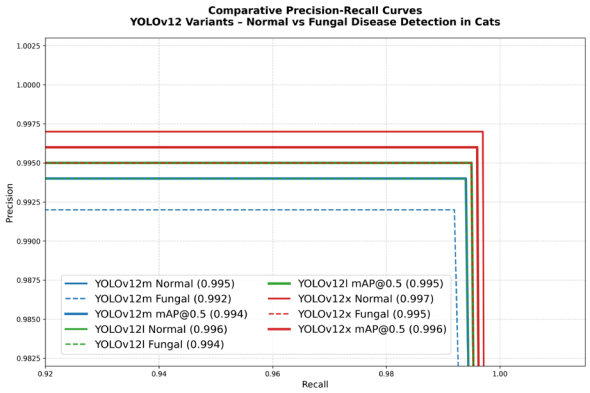


Figure 13. Precision Recall YOLOv12m

The implementation of the YOLOv12 deep learning architecture in the fungal skin disease detection experiment, utilizing the YOLOv12m, YOLOv12l, and YOLOv12x models, yielded detection results with a consistently high accuracy of 99% and mean Average Precision (mAP@50-95) values reaching 0.995 or higher across all variants.

The final results are summarized in Table 8. From the table, it is evident that the YOLOv12l model emerges as the optimal performer for deployment because it combines near-perfect mAP and accuracy with lower computational demand than YOLOv12x.

Table 8. Testing Results

Model	mAP@50	mAP@50-95	Precision	Recall	Accuracy
1	YOLOv12m	0.995	0.996	0.997	99%
2	YOLOv12l	0.996	0.994	0.995	99%
3	YOLOv12x	0.997	0.997	0.997	99%

These outstanding results demonstrate that YOLOv12 significantly outperforms traditional diagnostic methods (such as Wood's lamp examination, fungal culture, and microscopy), which are often time-consuming, subjective, and less sensitive in early-stage detection. The attention-centric mechanisms of YOLOv12 enable robust handling of challenging visual variations in cat skin images, including different fur types, colors, lighting conditions, and subtle lesion presentations. Based on the comparison presented in Table 9 and the quantitative results summarized in Table 8, the novelty of the present study lies in the application domain and in the use of the attention-centric YOLOv12 family for feline fungal skin disease detection. Whereas prior studies validated YOLO on plants, coral environments, or general animal objects, this study evaluates whether the same object-detection paradigm can support veterinary dermatology. The results indicate that YOLOv12 can detect fungal skin

lesions with high accuracy on the custom dataset, and the balance achieved by YOLOv12l makes it especially relevant for practical deployment. To address the need for additional experimental evidence, this study included a supplementary qualitative robustness test on unseen images representing more diverse visual conditions. The selected samples contain variations in posture, object scale, fur texture, background composition, and lesion position. The results, presented in Figure 14, show that the model consistently identified the normal and fungal disease categories with high confidence values ranging from 0.90 to 0.96. Importantly, the model was able to localize lesion regions even in visually challenging areas, such as the ear and facial region, where fungal symptoms may appear subtle or partially obscured. These findings support the argument that the proposed model has promising robustness beyond the standard test split and reinforce the contribution of this study in demonstrating the applicability of YOLOv12 for real-time feline dermatological screening under heterogeneous visual conditions. Compared to prior YOLO-based veterinary/animal skin disease studies, this work advances the field significantly. For example, YOLOv8 has been applied to cat skin disease detection, achieving high accuracy on similar lesions but without attention centric enhancements for subtle features. Other works using YOLOv8 for general animal or plant fungal detection report mAP@50-95 in the 0.90–0.96 range on larger datasets [12, 14, 24], while our YOLOv12 variants exceed 0.992–0.995 on a specialized feline fungal dataset, likely due to Area Attention and R-ELAN improving feature focus on irregular alopecia/crusting patterns. No prior study has explored YOLOv12 for zoonotic cat dermatophytosis, marking this as a novel application with superior real-time potential (YOLOv12l's balanced speed-accuracy outperforming YOLOv8 baselines in preliminary comparisons). One of the primary limitations in this study is the relatively small size and scope of the custom dataset used for training and evaluation. With only 400 images (200 for normal and 200 for fungal disease) collected from clinical cases and veterinary collaborations, the dataset while clinically verified and balanced may not fully capture the extensive variability encountered in real-world feline dermatology. Factors such as diverse cat breeds, fur lengths and colors, lighting conditions during photography (e.g., smartphone vs. clinical lighting), lesion stages (early subtle vs. advanced severe), partial occlusions by fur, and secondary complications (e.g., bacterial overlay or self-trauma) could potentially challenge model generalization on a broader population.

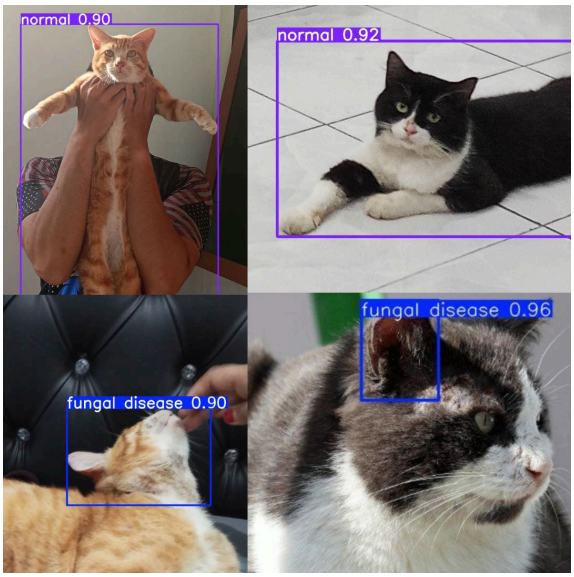


Figure 14. Qualitative Experiment Under Varied Visual Conditions

Additionally, obtaining high-quality annotations required close collaboration with veterinary dermatologists, which introduced constraints on dataset scale and introduced minor subjectivity in bounding box placement for diffuse or irregular lesions (particularly in Malassezia dermatitis cases). Image quality variations from user-submitted photos (common in potential real-world mobile app deployment) were partially addressed through augmentation, but further real-world testing is needed.

Table 9. Comparison Table

No	Study	Object/Disease	Model	Performance	Contribution
1	Miao et al., 2025 [12]	Plant Disease	YOLOv8	0.96	Baseline YOLOv8 for plant fungal detection; no attention-centric enhancement
2	Praveen, Sravani, Akhil et al., 2024 [2]	Plant disease detection	YOLOv7	0.92	Applied YOLOv7 for multi-class plant fungal detection; moderate accuracy
3	Rohan, Rafaq, Hasan et al., 2024 [21]	Livestock behavior recognition	YOLOv8	0.93	Demonstrated YOLO for animal activity tracking; less focus on subtle lesion detection
4	Tao, Ye & Doermann, 2025 [8]	Coral / underwater biological targets	YOLOv12	0.95	Highlighted attention modules for subtle/occluded objects in complex environments
5	Girmaw, 2025 [26]	Livestock skin disease detection	YOLOv8	0.92	Applied deep learning to detect animal skin lesions; dataset smaller, no attention-centric enhancement
6	Wisesa, Putri, Faristasari et al., 2026	Feline fungal skin disease detection	YOLov12l	0.99	Introduces an attention-centric YOLOv12 model for a specialized veterinary dermatology task using a clinically verified custom dataset; provides the best balance between detection accuracy and computational efficiency for real-time screening and practical deployment

To better highlight the contribution of the present work, Table 9 compares this study with several previous YOLO based studies on disease detection and related animal-health applications.

As shown in Table 9, previous studies mainly investigated plant diseases, underwater targets, or broader animal-related tasks, whereas the present work focuses specifically on fungal skin disease in cats and evaluates the most recent YOLOv12 family.

4. Conclusions

The results of this study on the detection of fungal skin diseases in cats using the YOLOv12m, YOLOv12l, and YOLOv12x models demonstrated outstanding performance, as evidenced by exceptionally high precision, recall, F1-score, and mean Average Precision (mAP) values. Evaluation of the three models with IoU threshold values ranging from 0.5 to 0.95 consistently achieved an overall accuracy of 99%, with perfect recall (1.0) and mAP@50-95 values exceeding 0.992 across all variants.

As an original contribution, this study introduces a novel approach employing the attention-centric YOLOv12 model, which proves significantly faster and more accurate than traditional veterinary diagnostic methods, such as Wood's lamp examination, fungal culture, and microscopic analysis. YOLOv12 enables robust real-time detection even in challenging conditions, including varied lighting, fur textures, partial occlusions, and low-quality smartphone images scenarios commonly encountered in clinical and home settings.

This research provides an innovative and efficient solution for monitoring feline dermatological health, particularly in regions with high pet cat populations and tropical climates conducive to fungal proliferation. The key original contributions include: (a). Development of a high-performance disease detection model tailored for veterinary applications using state of the art YOLOv12 architecture. (b). Compilation of a clinically verified custom dataset focused on common fungal skin infections in cats. (c). Establishment of an automated, real-time screening framework that can be integrated into mobile or clinic-based systems.

These advancements have the potential to support early diagnosis, reduce treatment delays, minimize zoonotic transmission risks to humans (especially from Microsporum canis), and enhance overall animal welfare. By enabling rapid and accessible preliminary screening, the proposed YOLOv12-based system offers a practical tool for veterinarians and pet owners alike, paving the way for more effective and sustainable management of fungal skin diseases in cats. Future expansions could further broaden its impact in veterinary telemedicine and public health.

Acknowledgements

The authors would like to express their sincere gratitude to the entire research team and supporting individuals who have contributed significantly to the successful completion of this study on the detection of fungal skin diseases in cats using the YOLOv12 algorithm. We are deeply grateful to the participating veterinary clinics and dermatologists in Indonesia who generously provided clinically confirmed images and expert insights, ensuring the dataset's quality and real-world relevance.

References

[1] H. M. Zayani *et al.*, "Cat Breed Classification with YOLOv11 and Optimized Training," *Engineering, Technology and Applied Science Research*, vol. 15, no. 2, pp. 21652–21657, Apr. 2025, doi: https://doi.org/10.48084/etasr.10218.

[2] R. S. Praveen, P. Sravani, S. Akhil, T. P. Reddy, and J. M. Babu, "Plant Disease Detection Using YOLOv7 Algorithm," *Int J Res Appl Sci Eng Technol*, vol. 12, no. 3, pp. 2058–2064, Mar. 2024, doi:https://doi.org/10.22214/ijraset.2024.59272.

[3] S. K. S. Rajan and N. Damodaran, "MAFFN_YOLOv5: Multi-Scale Attention Feature Fusion Network on the YOLOv5 Model for the Health Detection of Coral-Reefs Using a Built-In Benchmark Dataset," *Analytics*, vol. 2, no. 1, pp. 77–104, Mar. 2023, doi:https://doi.org/10.3390/analytics2010006.

[4] M. Shoaib *et al.*, "An advanced deep learning models-based plant disease detection: A review of recent research," 2023, *Frontiers Media SA*. doi:https://doi.org/10.3389/fpls.2023.1158933.

[5] Z. Yuan *et al.*, "Research on Calf Behavior Recognition Based on Improved Lightweight YOLOv8 in Farming Scenarios," *Animals*, vol. 15, no. 6, Mar. 2025, doi:https://doi.org/10.3390/ani15060898.

[6] L. Liu, J. Xie, H. Qin, X. Sui, and L. Zhang, "Precise detection of Eimeria oocysts in sheep: a deep learning model based on microscopic images," *Parasites and Vectors* , vol. 18, no. 1, Dec. 2025, doi: https://doi.org/10.1186/s13071-025-07092-4.

[7] A. Raza, F. Hanif, and H. A. Mohammed, "Analyzing the enhancement of CNN-YOLO and transformer based architectures for real-time animal detection in complex ecological environments," *Sci Rep*, vol. 15, no. 1, Dec. 2025, doi:https://doi.org/10.1038/s41598-025-26645-2.

[8] J. Tao *et al.*, "Coral-YOLO: An Intelligent Optical Vision Sensing Framework for High-Fidelity Marine Habitat Monitoring and Forecasting," *Sensors*, vol. 25, no. 23, Dec. 2025, doi:https://doi.org/10.3390/s25237284.

[9] A. Chowdhury, M. Jahan, S. Kaisar, M. E. Khoda, S. M. A. K. Rajin, and R. Naha, "Coral Reef Surveillance with Machine Learning: A Review of Datasets, Techniques, and Challenges," Dec. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi:https://doi.org/10.3390/electronics13245027.

[10] Y. Tian, Q. Ye, and D. Doermann, "YOLOv12: Attention-Centric Real-Time Object Detectors," Feb. 2025, [Online]. Available: https://arxiv.org/abs/2502.12524

[11] J. Zhong, M. Li, J. Qin, Y. Cui, K. Yang, and H. Zhang, "Real-time marine animal detection using yolo-based deep learning networks in the coral reef ecosystem," in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, International Society for Photogrammetry and Remote Sensing, 2022, pp. 301–306. doi:https://doi.org/10.5194/isprs-archives-XLVI-3-W1-2022-301-2022.

[12] Y. Miao, W. Meng, and X. Zhou, "SerpensGate-YOLOv8: an enhanced YOLOv8 model for accurate plant disease detection," *Front Plant Sci*, vol. 15, 2024, doi:https://doi.org/10.3389/fpls.2024.1514832.

[13] P. Song, L. Zhao, H. Li, X. Xue, and H. Liu, "RSE-YOLOv8: An Algorithm for Underwater Biological Target Detection," *Sensors*, vol. 24, no. 18, Sep. 2024, doi:https://doi.org/10.3390/s24186030.

[14] R. Kaur *et al.*, "YOLO-LeafNet: a robust deep learning framework for multispecies plant disease detection with data augmentation," *Sci Rep*, vol. 15, no. 1, Dec. 2025, doi:https://doi.org/10.1038/s41598-025-14021-z.
[15] S. Burti, T. Banzato, S. Coghlan, M. Wodzinski, M. Bendazzoli, and A. Zotti, "Artificial intelligence in veterinary diagnostic imaging: Perspectives and limitations," *Res Vet Sci*, vol. 175, p. 105317, 2024, doi: https://doi.org/10.1016/j.rvsc.2024.105317.
[16] Z. Xu, T. Wang, A. K. Skidmore, and R. Lamprey, "A review of deep learning techniques for detecting animals in aerial and satellite images," *International Journal of Applied Earth Observation and Geoinformation*, vol. 128, p. 103732, 2024, doi: https://doi.org/10.1016/j.jag.2024.103732.
[17] S. Jahn, G. Schmidt, L. Bachmann, H. Louton, T. Homeier-Bachmann, and A. K. Schütz, "Individual behavior tracking of heifers by using object detection algorithm YOLOv4," *Frontiers in Animal Science*, vol. 5, 2024, doi:https://doi.org/10.3389/fanim.2024.1499253.
[18] L. Zhang *et al.*, "Advancements in artificial intelligence technology for improving animal welfare: Current applications and research progress," Feb. 01, 2024, *John Wiley & Sons Inc*. doi:https://doi.org/10.1002/aro2.44.
[19] H. N. Ngugi, A. A. Akinyelu, and A. E. Ezugwu, "Machine Learning and Deep Learning for Crop Disease Diagnosis: Performance Analysis and Review," Dec. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi:https://doi.org/10.3390/agronomy14123001.
[20] A. Upadhyay *et al.*, "Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture," *Artif Intell Rev*, vol. 58, no. 3, Mar. 2025, doi:https://doi.org/10.1007/s10462-024-11100-x.
[21] A. Rohan, M. S. Rafaq, Md. J. Hasan, F. Asghar, A. K. Bashir, and T. Dottorini, "Application of deep learning for livestock behaviour recognition: A systematic literature review," *Comput Electron Agric*, vol. 224, p. 109115, 2024, doi: https://doi.org/10.1016/j.compag.2024.109115.
[22] A. Mao, E. Huang, X. Wang, and K. Liu, "Deep learning-based animal activity recognition with wearable sensors: Overview, challenges, and future directions," *Comput Electron Agric*, vol. 211, p. 108043, 2023, doi: https://doi.org/10.1016/j.compag.2023.108043.
[23] P. Ezanno *et al.*, "Research perspectives on animal health in the era of artificial intelligence," Dec. 01, 2021, *BioMed Central Ltd*. doi:https://doi.org/10.1186/s13567-021-00902-4.
[24] U. Ali, M. A. Ismail, R. A. Ariyaluran Habeeb, and S. R. Ali Shah, "Performance Evaluation of YOLO Models in Plant Disease Detection," *Journal of Informatics and Web Engineering*, vol. 3, no. 2, pp. 199–211, Jun. 2024, doi:https://doi.org/10.33093/jiwe.2024.3.2.15.
[25] S. Xiao *et al.*, "Review of applications of deep learning in veterinary diagnostics and animal health," 2025, *Frontiers Media SA*. doi:https://doi.org/10.3389/fvets.2025.1511522.
[26] D. Walle Girmaw, "Livestock animal skin disease detection and classification using deep learning approaches," *Biomed Signal Process Control*, vol. 102, p. 107334, 2025, doi: https://doi.org/10.1016/j.bspc.2024.107334.