

Multimodal Approach for Canine Dermatological and Ophthalmological Disease Diagnosis Using YOLOv11 with Data Augmentation and Autoencoder Techniques

Thi Diem Huong Nguyen, Van Loc An Ho, and Vinh Dinh Nguyen(✉)

Department of Information Technology, FPT University, Can Tho Campus, Can Tho City, Vietnam
{huongntdce171360,anhvlce171612}@fpt.edu.vn, vinhnd29@fe.edu.vn

Abstract. Accurate detection of canine dermatological and ophthalmological diseases is essential for improving animal health and supporting veterinary diagnostics. This study introduces a comprehensive approach that integrates the YOLOv11 object detection model with advanced data augmentation and feature extraction techniques, including Autoencoder (AE), CutMix, and Local Binary Patterns (LBP). YOLOv11 provides precise localization and classification of lesions, while Autoencoders enhance feature representation by reducing noise and improving generalization. CutMix increases data diversity, aiding in model robustness, and LBP captures critical texture details relevant to skin and eye conditions. Twelve datasets were created, combining these techniques in various configurations to assess their individual and synergistic effects on model performance. The proposed method achieves a mean Average Precision at 50% Intersection over Union (mAP@50) of 97.2% on the original image dataset and 86.6% on complex images when using a hybrid dataset that integrates original images with LBP, CutMix, and Autoencoder, significantly outperforming models trained on unaugmented data. This result demonstrates the effectiveness of combining deep learning with texture analysis and advanced augmentation strategies in handling complex veterinary diagnostic tasks. The findings highlight the potential of automated systems to improve early detection and accurate diagnosis of canine diseases, offering a reliable, scalable solution for real-world clinical applications.

Keywords: YOLOv11 · Canine Disease Diagnosis · Dermatological and · Ophthalmological Imaging · Autoencoder · CutMix · Local Binary Pattern (LBP) · Data Augmentation · Hybrid Dataset · Deep Learning in Veterinary Medicine

1 Introduction

The increasing prevalence of dermatological and ophthalmological diseases in dogs has become a significant concern in veterinary medicine. These conditions,

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2025
O. Gervasi et al. (Eds.): ICCSA 2025, LNCS 15648, pp. 361–374, 2025.
https://doi.org/10.1007/978-3-031-97000-9_23

if left undiagnosed or untreated, can lead to severe complications such as chronic discomfort, vision impairment, skin infections, and, in extreme cases, permanent tissue damage [1]. Traditional diagnostic methods primarily rely on visual inspection and clinical expertise, which can be subjective, time-consuming, and heavily dependent on the veterinarian's experience. This reliance on human assessment poses challenges for early and accurate disease detection, especially in remote areas with limited access to specialized veterinary care [2].

Recent advancements in artificial intelligence (AI) and computer vision have demonstrated promising potential in overcoming these diagnostic challenges. AI-based image analysis models have achieved remarkable success in human healthcare, particularly in detecting skin and ocular diseases through automated medical image interpretation [3]. These technologies not only improve diagnostic accuracy but also enable faster, more consistent evaluations, reducing the likelihood of human error. Inspired by these developments, the application of AI-driven approaches in veterinary medicine is gaining traction, offering the potential to enhance diagnostic precision, accessibility, and efficiency for common conditions in dogs [4].

In this study, we propose an automated diagnostic framework that leverages the capabilities of the YOLOv11 object detection model, known for its real-time detection accuracy and efficiency in complex visual tasks [5]. To further enhance feature representation, we integrate Autoencoder-based feature extraction, CutMix for data augmentation, and Local Binary Patterns (LBP) for texture analysis. This multi-technique approach aims to improve the model's ability to detect subtle pathological changes, ensuring robustness across diverse imaging conditions. The model is trained and evaluated on a comprehensive dataset of annotated canine skin and eye disease images, incorporating various augmentation techniques to mitigate overfitting and enhance generalization.

Our key contributions are as follows:

- We propose a novel diagnostic framework that combines YOLOv11, Autoencoder, CutMix, and LBP for accurate detection and classification of canine dermatological and ophthalmological diseases.
- We introduce a hybrid data augmentation strategy to improve model robustness, enabling effective performance across both standard and complex image conditions.
- We demonstrate the potential of AI-driven diagnostic tools in veterinary medicine, highlighting their efficacy in early disease detection and clinical decision support.

This paper is structured as follows: Sect. 2 reviews AI applications in veterinary diagnostics. Section 3 details the proposed methodology, including dataset, model architecture, and training process. Section 4 presents experimental results and discussions on augmentation impact. Finally, Sect. 5 concludes the study and suggests future research directions.

2 Related Work

Deep learning, particularly convolutional neural networks (CNNs) and YOLO-based models, has significantly advanced AI-driven veterinary diagnostics. These models have demonstrated strong capabilities in image classification, object detection, and segmentation, making them well-suited for diagnosing various animal diseases.

Huang et al. (2021) utilized YOLOv4 with ensemble learning to classify trap-neuter-return (TNR) surgical images of stray animals. Their approach reduced manual workload by 80% while achieving a mean average precision (mAP) of 91.99%, showcasing YOLO's ability to streamline high-throughput veterinary imaging tasks [6]. Similarly, Shaffer (2021) explored the role of AI in genetic trait analysis by identifying genetic variants linked to various canine diseases, demonstrating the potential of AI in predictive diagnostics [7].

In ophthalmology, Kim et al. (2022) developed a YOLOv5-based object detection model for diagnosing dry eye disease (DED) in dogs. Their approach utilized video-based corneal imaging and achieved an impressive mAP of 99.5%, emphasizing YOLO's precision in eye disease detection [8]. Additionally, Kour et al. (2022) conducted a comprehensive review of AI applications in veterinary medicine, highlighting CNNs' effectiveness in disease classification and early detection across various species [9].

Ali et al. (2023) applied YOLO to localize the optic disc in retinal fundus images, proving the model's reliability in eye-related diagnostics [10]. Meanwhile, Josphineleela et al. (2023) introduced a multi-stage Faster RCNN-based iSPLInception model for skin disease classification, overcoming overfitting issues and achieving 95.82% accuracy [11].

Beyond conventional imaging, Kim et al. (2024) proposed an AI-driven "Health Score" system that predicts canine diseases based on multifaceted behavioral data collected over three months. The system achieved an 87.5% concordance with veterinarian assessments, offering a non-invasive approach to early disease detection [12]. Additionally, Smith et al. (2024) developed an object detection model using Tiny YOLOv4 to diagnose pododermatitis and neoplasia in canine paws. Their system achieved a mAP of 0.95 and demonstrated real-time disease identification capabilities [13].

Scharre et al. (2025) compared an AI-driven equine ophthalmic diagnostic tool with veterinarians, finding that the AI model achieved a 93% accuracy rate, outperforming general veterinarians (76%) in detecting ophthalmic diseases [14]. Kalita et al. (2025) reviewed AI and ML applications in poultry disease detection, emphasizing advancements in non-invasive monitoring techniques such as thermal imaging and real-time behavioral analysis [15].

These studies highlight AI's transformative impact on veterinary diagnostics, yet challenges such as dataset generalizability, model interpretability, and real-world deployment remain. Table 1 summarizes key advancements in AI-based veterinary diagnostic research, outlining their methodologies and limitations.

Table 1. Summary of Recent Advances in AI-based Veterinary Diagnostics.

Year	Authors	Dataset	Method	Accuracy	Limitations
2021	Huang et al. [6]	TNR surgical images	YOLOv4 ensemble learning	mAP 91.99%	Limited to TNR images
2021	Shaffer [7]	Genetic traits in dogs	AI-based genetic analysis	-	Lacks diagnostic benchmarks
2022	Kim et al. [8]	Dog eye disease images	YOLOv5 object detection	mAP 99.5%	Small dataset
2022	Kour et al. [9]	Veterinary disease datasets	AI-based review	-	Lacks experimental validation
2023	Ali et al. [10]	Retinal fundus images	YOLO-based optic disc localization	-	Limited to fundus images
2023	Josphineleela et al. [11]	Skin disease datasets	Faster RCNN-based iSPLInception	Accuracy 95.82%	Computationally expensive
2024	Kim et al. [12]	Canine behavior data	AI-based health score	87.5% concordance	Needs clinical validation
2024	Smith et al. [13]	Canine paw disease images	Tiny YOLOv4	mAP 0.95	Limited to paw conditions
2025	Scharre et al. [14]	Equine ophthalmic images	AI vs. veterinarians	AI: 93%, Vets: 76%	Image-only diagnosis
2025	Kalita et al. [15]	Poultry disease data	AI/ML disease detection review	-	Lacks real-world deployment

3 Methodology

3.1 Data Collection

We used two publicly available datasets: a **skin disease dataset** from Mendeley Data [16] and an **eye disease dataset** from Roboflow Universe [17]. These datasets provide diverse labeled images for comprehensive model training.

Canine Skin Disease Dataset. The skin disease dataset [16] includes four categories:

- **Bacterial Dermatosis**: Bacterial skin infections.
- **Fungal Infection**: Fungal-related skin diseases.

- **Healthy**: Normal skin, serving as a control.
- **Hypersensitivity-Allergic Dermatosis**: Skin conditions caused by allergic reactions.

Canine Eye Disease Dataset. The eye disease dataset [17] consists of five conditions:

- **Cataract**: Lens clouding leading to vision loss.
- **Cherry Eye**: Prolapse of the third eyelid gland.
- **Conjunctivitis**: Conjunctival inflammation.
- **Glaucoma**: Increased eye pressure, risk of blindness.
- **Iris Atrophy**: Degeneration affecting light regulation.

Integrating these datasets ensures our model is trained on a broad range of **skin and eye diseases**, enhancing its diagnostic capabilities. All images were preprocessed to match YOLOv11 input size for consistency.

3.2 Data Preprocessing and Augmentation

To enhance the performance and robustness of the YOLOv11 model in detecting canine dermatological and ophthalmological diseases, we employed a combination of standard preprocessing techniques and advanced data augmentation methods. The final dataset includes 9 disease labels, ensuring a balanced and diverse training set.

Data Preprocessing. All images were resized to a fixed resolution of 640 $\times$ 640 pixels to standardize input dimensions and ensure compatibility with the YOLOv11 architecture. This preprocessing step ensures consistency across different sources of data, allowing the model to focus on disease-specific features.

Data Augmentation. To improve model generalization and increase dataset diversity, we applied augmentation techniques in the following stages:

1. Conventional Augmentation Techniques (Pre-Augmented Dataset). Before applying more advanced augmentations, we first augmented the original dataset using conventional image transformation techniques. These included:

- **Rotation:** Randomly rotating images within a limited range to simulate different angles of capture.
- **Flipping:** Applying horizontal and vertical flips to introduce orientation diversity.
- **Brightness and Contrast Adjustment:** Modifying brightness and contrast to simulate real-world lighting variations.
- **Gaussian Noise:** Adding random noise to account for variations in image quality.

These augmentation techniques only apply to the original dataset, resulting in an intermediate dataset containing 9 final labels before applying LBP and CutMix augmentation. The label distribution after this stage is illustrated in Fig. 1.

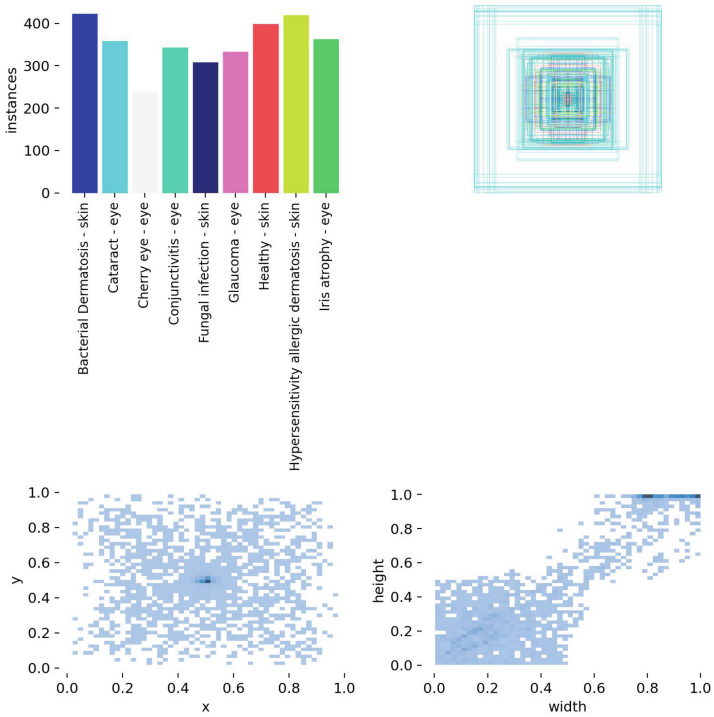


Fig. 1. Final label distribution after conventional augmentation (before applying LBP and CutMix).

2. Local Binary Pattern (LBP) Augmentation. LBP was applied as an additional augmentation technique to enhance fine-grained texture information, which is crucial for detecting dermatological conditions. The LBP process consists of:

- Converting images to grayscale.
- Applying LBP transformation by encoding each pixel's intensity difference with its neighboring pixels.
- Generating a histogram-based representation that captures dominant texture patterns.

LBP augmentation helps the model better differentiate between diseases that present with distinct skin surface abnormalities.

3. CutMix Augmentation. To further improve dataset diversity, CutMix augmentation was applied [18]. This technique merges two images by replacing a random rectangular region in one image with a patch from another image, enhancing the model's ability to generalize.

The CutMix augmentation process follows these steps:

1. Select two images from the dataset.
2. Define a random rectangular patch in the first image.
3. Replace this patch with a corresponding region from the second image.
4. Adjust bounding box labels accordingly.

Algorithm 1 summarizes the CutMix augmentation procedure.

Algorithm 1: CutMix Data Augmentation

Input : Two images I_1, I_2 with labels L_1, L_2
Output: Augmented image with adjusted labels
Resize images to $D \times D$;
Select a random patch (x, y, w, h) in I_1;
Replace the patch in I_1 with the corresponding region from I_2;
Adjust bounding boxes in L_1 and L_2 accordingly;
Save the augmented image and updated labels;

Impact of Augmentation Techniques. The three-stage augmentation process ensures the model is trained on highly diverse and representative data. Conventional augmentation techniques improve base dataset diversity, while LBP enhances texture-based feature extraction, and CutMix aids model generalization by simulating real-world variations.

Effect on Model Performance. Our experimental results indicate that combining conventional augmentation, LBP, and CutMix significantly improves mAP@50 scores, especially when detecting diseases with complex textures. This underscores the importance of layered augmentation strategies in AI-driven veterinary diagnostics.

3.3 Model Architecture

YOLO11. YOLO11, introduced in 2024, is the latest evolution in the YOLO (You Only Look Once) series, designed for high-speed, high-accuracy object detection. It builds upon its predecessors by incorporating a more efficient backbone and an optimized detection head, significantly enhancing performance across multiple vision tasks such as object detection, instance segmentation, image classification, and pose estimation [19].

The YOLO11 architecture consists of a lightweight backbone for feature extraction, a neck module that fuses multi-scale features, and a detection head optimized for precision. Key advancements include:

- **C3K2 Block**: A novel convolutional module that reduces computation while retaining feature expressiveness.
- **C2PSA Module**: Introduces parallel spatial attention to refine feature selection for better detection accuracy.
- **Anchor-Free Mechanism**: Simplifies model training and inference while improving small object detection.
- **SPPF (Spatial Pyramid Pooling - Fast)**: Enhances multi-scale feature aggregation without increasing computational overhead.

Compared to YOLOv8, YOLO11 achieves higher mAP and lower latency on various benchmark datasets, including COCO, while maintaining a strong trade-off between speed and accuracy (Fig. 2).

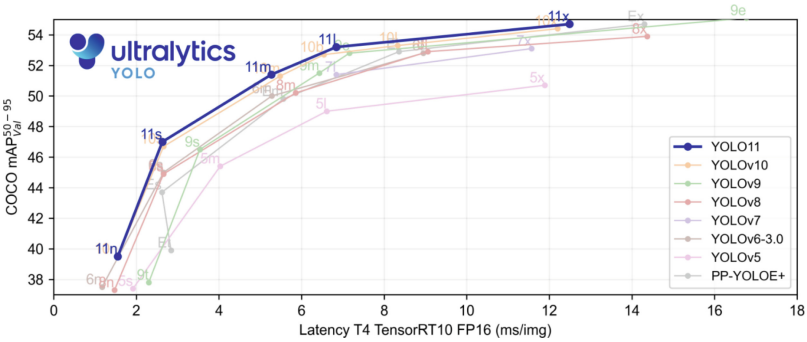


Fig. 2. Benchmarking YOLOv11 Against Previous Versions [20].

This makes YOLO11 an optimal choice for real-world applications requiring both high accuracy and low latency, such as medical imaging, autonomous driving, and industrial automation.

Autoencoder. In our methodology, an autoencoder model is employed to enhance feature extraction for diagnosing canine skin diseases. The autoencoder compresses input images into lower-dimensional representations and subsequently reconstructs them. This process allows the model to focus on the most critical features necessary for accurate disease detection and classification.

The autoencoder architecture comprises two primary components: an encoder and a decoder.

- The encoder extracts compact and informative features from input images through a series of convolutional layers with increasing channel depth (e.g., 64, 128, and 256). Each convolutional layer is followed by batch normalization, ReLU activation, and dropout to mitigate overfitting and enhance generalization. The encoder reduces the spatial dimensions of input images via stride-based convolutions, retaining essential visual features while significantly reducing data dimensionality.

– The decoder reconstructs the compressed features into the original image space. It utilizes transposed convolutional layers for upsampling the encoded representation step by step. Batch normalization and ReLU activation ensure smooth reconstruction, while a sigmoid activation function in the output layer scales pixel values to the range [0, 1].

Input images are preprocessed to a fixed size of 128×128 pixels to maintain consistency and computational efficiency. The autoencoder is trained using the L1 loss function, which ensures that fine details in the reconstructed images are preserved. Training is conducted over 20 epochs with a batch size of 16, using the SGD optimizer with momentum and a dynamic learning rate scheduler to adapt the learning rate during training.

Once trained, the encoder is extracted and integrated into the YOLO11 model pipeline to provide preprocessed feature maps for subsequent training and inference. By leveraging these encoded features, the YOLO11 model achieves improved accuracy and robustness in detecting and classifying various canine skin diseases.

This integration significantly enhances the overall performance of the diagnostic system, enabling precise and efficient identification of common dermatological conditions in dogs.

Experimental Models. To evaluate the impact of augmentation techniques on YOLOv11 performance, we trained 12 model variations, each incorporating different augmentation strategies:

1. **Original images (baseline):** The model is trained using only the original dataset, serving as the baseline for comparison.
2. **Original images + Autoencoder:** An autoencoder is applied to extract enhanced feature representations before training.
3. **Original images with CutMix:** CutMix augmentation is applied to introduce additional diversity by merging different image regions.
4. **CutMix + Autoencoder:** The CutMix-augmented dataset is further processed using an autoencoder for feature enhancement.
5. **Original images with LBP:** Local Binary Pattern (LBP) is applied to emphasize textural features in disease regions.
6. **LBP + Autoencoder:** The LBP-processed dataset is further refined using an autoencoder for feature extraction.
7. **LBP transformed dataset with CutMix:** LBP is applied first, followed by CutMix augmentation to introduce spatial variations.
8. **(LBP → CutMix) + Autoencoder:** The LBP + CutMix dataset undergoes further enhancement with an autoencoder.
9. **CutMix applied first, then LBP:** CutMix augmentation is performed before applying LBP transformation.
10. **(CutMix → LBP) + Autoencoder:** The CutMix + LBP dataset is further refined with an autoencoder.

11. **Original images + LBP + CutMix:** The dataset is augmented with both LBP and CutMix simultaneously to maximize diversity.
12. **(Original + LBP + CutMix) + Autoencoder:** The most comprehensive dataset, integrating all augmentation techniques for the best possible feature extraction.

These models assess:

- The effect of **Autoencoder, LBP, and CutMix** on detection accuracy.
- The impact of different augmentation combinations.
- The best-performing strategy for **mAP@50** improvement.

Figure 3 illustrates the YOLOv11 training pipeline. Raw images undergo preprocessing, including resizing and conventional augmentations, followed by three augmentation techniques: **Autoencoder-based feature extraction, CutMix, and LBP**. The augmented datasets are then combined and used to train YOLOv11. The final model detects **cataracts in both eyes** with high confidence scores, demonstrating the effectiveness of hybrid augmentation strategies.

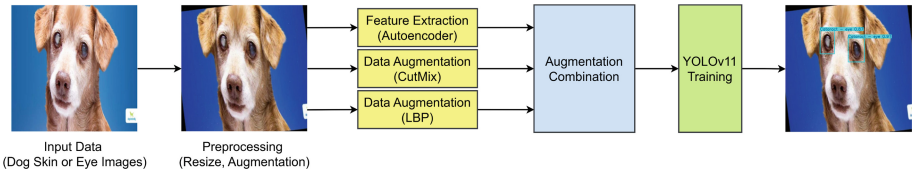


Fig. 3. YOLOv11 Training Pipeline with Augmentation and Feature Extraction.

3.4 Implementation Details

Our YOLOv11 models were trained for up to 200 epochs, with early stopping triggered after 100 epochs of no improvement to prevent overfitting. Each batch contained 16 images, resized to 640×640 pixels for consistency.

Training leveraged GPU acceleration with parallelized data loading (8 workers) for efficiency. The initial learning rate was set to 0.01 and decayed progressively. An optimizer with momentum (0.937) and weight decay (0.0005) was used to stabilize learning and enhance generalization.

These configurations ensured robust model performance in detecting and classifying canine skin and eye diseases across diverse datasets.

3.5 Evaluation Metrics

To assess the effectiveness of our proposed model in detecting canine dermatological and ophthalmological diseases, we employ Mean Average Precision at 50% IoU (mAP@50) as the primary evaluation metric. This metric measures the

model's ability to accurately localize and classify lesions across different augmentation strategies.

$$mAP50 = \frac{1}{C}\sum_{i=1}^{C} AP_i[21]$$

where C is the total number of disease classes, and AP_i is the Average Precision for class i [21].

Our experiments demonstrate that hybrid augmentation techniques—particularly the integration of CutMix, LBP, and Autoencoder—significantly improve model performance. The best-performing configuration achieved 97.2% mAP@50 on the original dataset and 86.6% on the complex dataset, outperforming models trained on unaugmented data.

These results highlight the effectiveness of combining deep learning and texture-based augmentation for veterinary diagnostics, offering a robust and scalable approach to early and precise detection of canine skin and eye diseases.

4 Experiment Results

4.1 Model Training Results

To assess the impact of augmentation techniques, YOLOv11 was trained on twelve datasets with different augmentation strategies for 200 epochs. Performance was evaluated using mAP@50 on three test sets: original images, LBP-processed images, and a complex dataset combining LBP, CutMix, and original images. Table 2 summarizes the results.

Table 2. Performance of YOLOv11 with Different Augmentation Techniques.

#	Augmentation Technique	Train mAP@50	Test mAP@50 (Original)	Test mAP@50 (LBP)	Test mAP@50 (Complex)
1	Original Images	85.5	88.7	-	58.8
2	Original + Autoencoder	85.6	89.6	-	57.4
3	Original → CutMix	66.2	85.3	-	65.5
4	CutMix + Autoencoder	67.1	86.7	-	67.1
5	Original → LBP	84.0	-	85.7	36.2
6	LBP + Autoencoder	83.9	-	85.4	36.9
7	LBP → CutMix	58.2	-	85.5	39.2
8	(LBP → CutMix) + Autoencoder	60.0	-	81.2	31.5
9	CutMix → LBP	61.5	-	81.4	32.9
10	(CutMix → LBP) + Autoencoder	61.2	-	81.3	30.6
11	Original + LBP + CutMix	84.9	-	81.4	31.7
12	(Original + LBP + CutMix) + Autoencoder	85.9	97.2	-	86.6

4.2 Discussion

The results in Table 2 highlight key findings:

- The baseline model (original images) achieved 88.7% mAP@50 but struggled on complex images (58.8%), indicating poor generalization.
- The best-performing model (Original + LBP + CutMix + Autoencoder) reached 97.2% on the original test set and 86.6% on complex images, showing significant improvements.
- CutMix-enhanced models, particularly CutMix + Autoencoder, improved performance on complex datasets (67.1% mAP@50).
- LBP-based models retained high accuracy on LBP-processed images but suffered on complex datasets, suggesting texture features alone are insufficient.
- Hybrid augmentation (LBP, CutMix, Autoencoder) significantly improved model robustness.

Implications for Veterinary Diagnostics. The results demonstrate the effectiveness of augmentation techniques in improving disease detection. Combining LBP for texture enhancement with CutMix for diversity proved crucial for robust detection under varying conditions. These findings suggest AI-driven diagnostics can assist veterinarians in early disease detection and treatment planning.

Limitations and Future Work.

- Performance on highly complex images remains lower than on original images, indicating room for further optimization.
- Class imbalance may have affected model performance.
- Future work should explore alternative architectures or ensemble models to enhance accuracy.
- Clinical validation is needed for real-world deployment.

This study lays the foundation for future AI-driven veterinary diagnostic research, aiming for improved robustness and scalability.

5 Conclusion

This study presents an AI-driven diagnostic framework for canine dermatological and ophthalmological diseases using YOLOv11, enhanced with CutMix augmentation, Local Binary Patterns (LBP), and Autoencoder-based feature extraction. These augmentation techniques significantly improved model robustness and generalization.

Experimental results confirm that hybrid augmentation strategies enhance detection accuracy. The best-performing model, trained with a combination of original images, LBP, CutMix, and Autoencoder, achieved 97.2% mAP@50 on the original dataset and 86.6% on complex images, significantly outperforming models trained without augmentation.

These findings demonstrate the potential of deep learning in veterinary diagnostics, reducing reliance on manual assessments and improving early disease detection. Future work should explore additional augmentation methods and real-world clinical validation to further enhance diagnostic accuracy and applicability in veterinary medicine.

References

1. Hill, P., et al.: Survey of the prevalence, diagnosis and treatment of dermatological conditions in small animals in general practice. Vet. Rec. **158**(16), 533–539 (2006)
2. Das, B., Ellis, M., Sahoo, M.: Veterinary diagnostics: growth, trends, and impact. In: Evolving Landscape of Molecular Diagnostics, pp. 227–242. Elsevier (2024)
3. Loh, E.: Medicine and the rise of the robots: a qualitative review of recent advances of artificial intelligence in health. BMJ Leader, Leader-2018 (2018)
4. Burić, M.: Application of computer vision methods in veterinary ophthalmology, Doctoral dissertation, University of Rijeka, Faculty of Informatics and Digital Technologies (2024)
5. Khanam, R., Hussain, M. YOLOV11: an overview of the key architectural enhancements. arXiv preprint arXiv:2410.17725 (2024)
6. Huang, Y.C., Chuang, T.H., Lai, Y.L.: Classification of the trap-neuter-return surgery images of stray animals using YOLO-based deep learning integrated with a majority voting system. Appl. Sci. **11**(18), 8578 (2021)
7. Shaffer, L.G.: Special issue on companion animal genetics: novel variants discovered in wide variety of diseases in dogs, identification and further characterization of traits in dogs and cats, and the use of microarrays in the detection of aneuploidy in dogs. Hum. Genet. **140**(11), 1501–1503 (2021). https://doi.org/10.1007/s00439-021-02375-z
8. Kim, J.Y., Han, M.G., Chun, J.H., Huh, E.A., Lee, S.J.: Developing a diagnosis model for dry eye disease in dogs using object detection. Sci. Rep. **12**(1), 21351 (2022)
9. Kour, S., Agrawal, R., Sharma, N., Tikoo, A., Pande, N., Sawhney, A.: Artificial intelligence and its application in animal disease diagnosis (2022)
10. Ali, H.M., El Abbadi, N.K.: Optic disc localization in retinal fundus images based on you only look once network (YOLO). Int. J. Intell. Eng. Syst. **16**(2) (2023)
11. Josphineleela, R., Raja Rao, P., Shaikh, A., Sudhakar, K.: A multi-stage faster RCNN-based iSPLInception for skin disease classification using novel optimization. J. Digit. Imag. **36**(5), 2210–2226 (2023)
12. Kim, S.C., Kim, S.: Development of a dog health score using an artificial intelligence disease prediction algorithm based on multifaceted data. Animals **14**(2), 256 (2024)
13. Smith, A., et al.: Computer vision model for the detection of canine pododermatitis and neoplasia of the paw. Vet. Dermatol. **35**(2), 138–147 (2024)
14. Scharre, A., et al.: Comparison of veterinarians and a deep learning tool in the diagnosis of equine ophthalmic diseases. Equine Vet. J. **57**(1), 47–53 (2025)
15. Kalita, A.J., Subba, M., Adil, S., Wani, M.A., Beigh, Y.A., Shafi, M.: Application of artificial intelligence and machine learning in poultry disease detection and diagnosis: a review: AI and machine learning in poultry disease diagnosis. Lett. Anim. Biol., 01–06 (2025)

16. Hwang, S., Shin, H.K., Park, J.M., Kwon, B., Kang, M.G.: Classification of dog skin diseases using deep learning with images captured from multispectral imaging device. Mol. Cell. Toxicol. **18**(3), 299–309 (2022). https://doi.org/10.17632/5dbht54kw7.1 https://doi.org/10.17632/5dbht54kw7.1 https://doi.org/10.17632/5dbht54kw7.1
17. Roboflow universe, eye disease detection dataset. https://universe.roboflow.com/cat-breed-detection/eye-disease-detection-eftgo
18. Yun, S., Han, D., Oh, S.J., Chun, S., Choe, J., Yoo, Y.: CutMix: regularization strategy to train strong classifiers with localizable features. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) (2019)
19. Khanam, R., Hussain, M.: YOLOV11: an overview of the key architectural enhancements. arXiv preprint arXiv:2410.17725 (2024)
20. Ultralytics. Ultralytics YOLOV11 (2024). https://docs.ultralytics.com/models/yolo11/s
21. Burić, M.: Application of computer vision methods in veterinary ophthalmology, Dissertation, University of Rijeka, Faculty of Informatics and Digital Technologies (2024)