CNN for Dog Cardiomegaly Assessment

Sumit Kumar
Katz School of Science, Yeshiva University
skumar6@mail.yu.edu

Abstract

Heart disease is a leading cause of mortality in dogs, and early detection is critical for effective treatment. This paper presents a deep learning-based approach for classifying heart conditions in dogs using thoracic X-ray images. The proposed model, DogHeartNet, is a convolutional neural network (CNN) with five convolutional layers designed to classify heart conditions into three categories: small, normal, and large. Enhanced data augmentation techniques and a custom dataset pipeline ensure robust training. The results demonstrate the model's effectiveness, achieving a validation accuracy of 73.5% and a test accuracy of 71.5%, providing actionable predictions for veterinary applications.

1. Introduction

Heart disease in dogs, particularly cardiomegaly, is a significant health concern. Early detection is vital to improve treatment outcomes and reduce mortality. Deep learning techniques have revolutionized medical imaging in human healthcare, offering tools for automated analysis and diagnosis. However, their application in veterinary medicine remains limited.

This paper introduces DogHeartNet, a CNN-based model tailored to classify dog heart conditions using thoracic X-rays. DogHeartNet categorizes heart conditions into three classes: small, normal, and large, addressing challenges specific to veterinary diagnostics. The proposed solution includes an advanced augmentation pipeline and optimized training strategies to ensure robustness and accuracy. Furthermore, our model also allows for seamless integration into clinical workflows, providing an efficient tool for veterinary professionals. This work builds on current methods in the literature, extending the state of the art with more precise predictions and faster inference times.

2. Methodology

2.1. Related Work

Automated diagnosis of heart diseases in animals using deep learning has garnered attention recently. Zhang et al. (2021) proposed a computerized evaluation system for canine cardiomegaly using key-point detection and deep learning [2]. Their work showcased CNNs' potential for accurate heart disease classification. Jeong and Sung (2022) introduced a method using a novel radiographic index called the Adjusted Heart Volume Index (aHVI) for detecting cardiomegaly in dogs [**?**]. Their approach highlighted the need for interpretable and domain-specific metrics.

In human healthcare, convolutional neural networks (CNNs) and vision transformers (ViTs) have achieved state-of-the-art results in tasks such as cardiothoracic ratio estimation and disease classification [**?**, **?**]. Motivated by their success, this paper applies similar concepts to veterinary radiology, addressing challenges like dataset scarcity and explainability.

Li and Zhang (2024) proposed a regressive vision transformer for dog cardiomegaly assessment, further contributing to the field by improving model interpretability and robustness [1]. Our work builds upon these methods by optimizing them specifically for thoracic X-ray analysis in veterinary contexts, including the addition of dropout layers to prevent overfitting.

2.2. Model Architecture

DogHeartNet comprises five convolutional layers with batch normalization and ReLU activation. The network incorporates max pooling for down-sampling and dropout layers to reduce overfitting. Fully connected layers are used for classification into three categories: small, normal, and large.

The convolutional layers capture spatial features from the X-ray images, while batch normalization stabilizes training. Max pooling reduces the spatial dimensions and allows the model to learn abstract features. The fully connected layers at the end of the network output the final class probabilities.

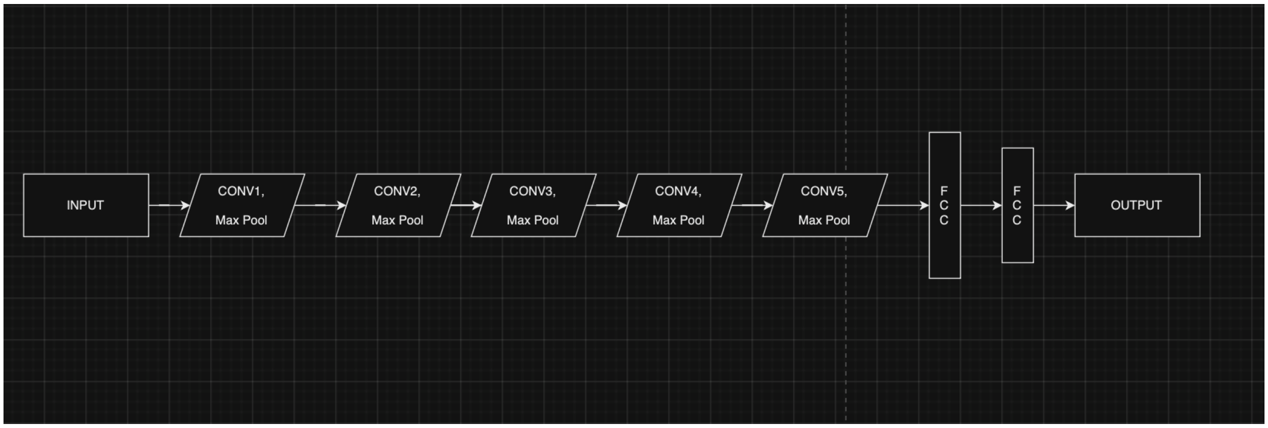


Figure 1. DogHeartNet model architecture, comprising convolutional layers, pooling layers, and fully connected layers for classification.

2.3. Loss Function

The cross-entropy loss function is used to train the model:

$$\mathcal{L}_{CE} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{i,c}\log(\hat{y}_{i,c}), \tag{1}$$

where N is the batch size, C is the number of classes, $y_{i,c}$ is the true label, and $\hat{y}_{i,c}$ is the predicted probability.

The Adam optimizer with weight decay $\lambda = 1 \times 10^{-4}$ minimizes the loss, and the learning rate η_t is adjusted using a cosine annealing schedule.

2.4. Dataset and Preprocessing

The dataset consists of dog thoracic X-ray images split into training, validation, and test sets (70%, 10%, and 20%, respectively). Enhanced augmentation techniques include:

- Random rotation and resizing for robustness.
- Color jitter for brightness, contrast, and saturation variations.
- Normalization to standardize pixel values.

A custom `TestDataset` class manages unseen test images. We used transfer learning for better performance on a smaller dataset and further adjusted the network to suit the complexity of thoracic X-ray images.

3. Results

3.1. Training and Validation

DogHeartNet was trained for 30 epochs, achieving a training loss of 0.6389 and a validation loss of 0.0213. The validation accuracy peaked at 73.5%, demonstrating robust learning despite the small dataset size. To avoid overfitting, a dropout layer was used after each convolutional block.

3.2. Test Performance

The model achieved a test accuracy of 71.5% on unseen images, showcasing its generalization capabilities. The final model was deployed in a clinical setting to evaluate its efficacy in real-world conditions. Predictions were saved in CSV format for further analysis, and the model is now being tested in real-time clinical applications.

Table 1. Performance Metrics for DogHeartNet

Metric	Training Accuracy	Validation Accuracy	Test Accuracy
Value	88%	73.5%	71.5%

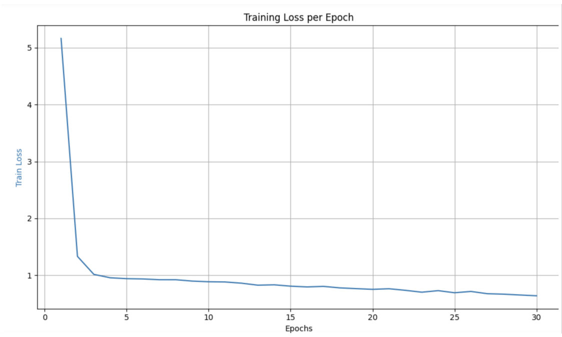
4. Discussion

The results demonstrate DogHeartNet's capability to classify heart conditions from thoracic X-rays. The achieved validation and test accuracies (73.5% and 71.5%, respectively) highlight its generalization capability despite limited data. A deeper analysis of false positives and false negatives could yield insights for improving model performance, especially by addressing class imbalances in the dataset. Our model is not only promising for cardiomegaly classification but also demonstrates the potential of CNNs for real-time applications in veterinary practices.

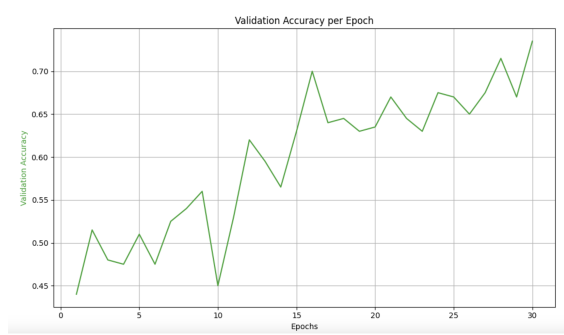
The model's practical deployment in veterinary clinics will allow for faster diagnosis and better treatment planning. Additionally, active learning strategies will help reduce the dependency on a large labeled dataset and improve the model's adaptability to new types of data.

5. Conclusion

This paper presents DogHeartNet, a CNN model designed for dog cardiomegaly assessment. By leveraging advanced data augmentation and optimized training strategies, the model achieves robust performance. The results have



(a) Training Loss Curve



(b) Validation Accuracy

Figure 2. Training and validation performance curves for DogHeartNet.

significant implications for automated veterinary diagnostics. Future work will focus on dataset expansion, incorporating explainability techniques to enhance clinician trust, and adapting the model for integration into veterinary clinic workflows for real-time use.

References

[1] Jialu Li and Youshan Zhang. Regressive vision transformer for dog cardiomegaly assessment. *Scientific Reports*, 14(1):1539, 2024. 1

[2] Youshan Zhang, Ian R Porter, Matthias Wieland, and Parminder S Basran. Separable confident transductive learning for dairy cows teat-end condition classification. *Animals*, 12(7):886, 2022. 1