A Neural Network-Based Method for Dog Heart VHS Assessment

Hang Yu
Katz School of Science and Health
Yeshiva University
New York, NY
hyu1@mail.yu.edu

Youshan Zhang
Katz School of Science and Health
Yeshiva University
New York, NY
youshan.zhang@yu.edu

***Abstract*—This paper presents a deep learning-based framework for evaluating the Vertebral Heart Score (VHS) in dogs, an essential indicator of cardiac health in veterinary medicine. The model is based on a pre-trained ResNet18 architecture, fine-tuned on a custom dataset of dog chest X-rays with annotated VHS values.**

To improve performance, the framework employs a combination of weighted Mean Squared Error (MSE) for keypoint prediction and a dedicated VHS loss. The model is trained and tested on images from multiple dog breeds, achieving a highest accuracy of 79.25%.

Results show the model outperforms traditional methods in both speed and accuracy, while remaining robust to heart size variations across breeds. This study highlights the potential of deep learning in veterinary diagnostics and offers a practical step toward automated cardiac assessment.

***Index Terms*—Dog heart, VHS evaluation, ResNet18, deep learning, transfer learning, veterinary diagnostics, convolutional neural networks, image classification**

I. Introduction

The veterinary field has seen significant advancements in recent years, particularly with the adoption of machine learning and artificial intelligence (AI) techniques. One of the critical areas in veterinary diagnostics is the evaluation of heart health, particularly for assessing the heart size of dogs. Heart disease is a leading cause of morbidity and mortality in canines, making early detection and proper assessment of heart size crucial for timely intervention. Traditionally, veterinarians use manual methods, such as radiographs or echocardiograms, to assess the heart size. However, these methods are time-consuming, require specialized equipment, and may involve significant human error.

The advent of deep learning technologies offers a promising alternative for automating and improving the accuracy of heart size evaluation in dogs. In particular, convolutional neural networks (CNNs) have demonstrated superior performance in various image-based tasks, including medical image analysis. By leveraging these technologies, it is possible to develop automated systems that not only reduce the workload of veterinarians but also provide more consistent and accurate assessments.

In this paper, we propose a deep learning-based model using the ResNet18 architecture to automatically evaluate the heart size of dogs from chest X-ray images. Our approach utilizes a combination of features, including six key anatomical points on the dog's heart and the visual size of the heart, to predict the heart size. The model's effectiveness is measured using the correlation between predicted values and the actual heart size, as well as its performance across different dog breeds, highlighting its adaptability.

We also explore the challenges involved in training such a model, particularly in terms of dataset limitations and variability in the images. The results show that our model achieves an accuracy of 79.25%, indicating its potential as a valuable tool for veterinary applications. This paper outlines the methodology, experimental setup, and results, and discusses the implications of deep learning in veterinary diagnostics.

In the following sections, we present a detailed review of related work, the methodology of our approach, the experimental setup, and the results obtained. We conclude with a discussion on the potential future improvements and applications of the proposed model.

II. Related Work

A. Heart Size Assessment in Veterinary Medicine

In veterinary medicine, heart size in dogs is traditionally assessed using imaging techniques such as radiographs, echocardiograms, and MRI scans. These methods, however, are often time-consuming and highly dependent on the skill of the clinician. Several studies have explored automated methods for analyzing veterinary radiographs. For example, Burti et al. applied deep learning to detect cardiomegaly on thoracic radiographs of dogs [1]. Jeong and Sung proposed a deep learning-based approach with a novel cardiac index for canine cardiomegaly detection from radiographs [2]. Oh et al. combined image classification and semantic segmentation for robust cardiomegaly diagnosis in pets [3]. While promising, many existing models focus on classification tasks, identifying whether a condition is present or not, rather than producing continuous quantitative assessments such as heart size.

B. Deep Learning in Medical Image Analysis

Deep learning has revolutionized medical image analysis, especially in classification, segmentation, and detection tasks. Banzato et al. developed a model that automatically classifies canine thoracic radiographs using deep learning [4]. Li et al.

conducted a pilot study using AI to detect left atrial enlargement on canine radiographs, demonstrating the feasibility of regression-oriented tasks in veterinary imaging [5]. Similarly, Kim et al. compared AI and radiologist performance in diagnosing cardiogenic pulmonary edema in dogs [6], supporting the reliability of AI-based veterinary diagnostics. However, few works focus specifically on heart size estimation across different dog breeds, which is the focus of our study.

C. Challenges in Veterinary Cardiology Using Deep Learning

One major challenge in applying deep learning to veterinary cardiology is the limited availability of large, annotated datasets. Unlike human cardiology, veterinary datasets are often smaller and more heterogeneous in breed, posture, and image quality. Li and Zhang's work using a regressive vision transformer for dog cardiomegaly shows promise in addressing these data-related limitations [7]. Additionally, Ergün and Güney showed that deep learning can be adapted to veterinary-specific radiographic analysis, such as assessing bone maturity and fracture time [8], suggesting potential for broader applications including cardiology. Still, generalization across diverse breeds and image conditions remains a challenge.

D. Related Techniques

Several architectural designs have proven useful in medical imaging. ResNet-based models are widely used due to their accuracy and efficiency. Banzato et al. applied such architectures for classifying canine thoracic radiographs [4]. Li and Zhang's work on the Regressive Vision Transformer highlights the ongoing development of regression-capable deep learning models in veterinary contexts [7]. While segmentation-focused models like U-Net are effective for pixel-level tasks, our study adopts a ResNet18-based regression model, better suited for predicting continuous values such as the vertebral heart scale.

E. Summary of Related Work

In summary, while deep learning has shown promise in veterinary cardiac image analysis, most work to date focuses on classification. Few studies directly tackle the challenge of continuous heart size estimation. This work contributes to that gap by proposing a regression-based model trained on diverse dog X-ray images. By leveraging the ResNet18 architecture, the model is designed to operate efficiently and accurately, even on relatively small datasets, offering potential for real-world veterinary applications.

III. Methods

In this section, we describe the methodology used to develop the heart size estimation model for dogs. The proposed approach consists of three main components: dataset preparation, model architecture, and training procedure. These components are designed to accurately predict the heart size from chest X-ray images and enable the model to generalize across different breeds of dogs.

A. Dataset

The dataset used in this study consists of chest X-ray images of dogs with annotated heart sizes. Each image is paired with a set of six key points that mark the boundaries of the heart, along with the corresponding VHS (vertebral heart score) values. These images were collected from various veterinary sources, ensuring a diverse range of dog breeds and heart sizes. Previous works, such as those by Burti et al. and Banzato et al., have demonstrated the viability of using thoracic radiographs for cardiomegaly detection and classification in canine datasets [1], [4].

Preprocessing: Prior to feeding the images into the model, we perform several preprocessing steps. First, the images are resized to a fixed resolution to ensure consistency across the dataset. Next, data augmentation techniques, such as random flipping, rotation, and scaling, are applied to improve model robustness. Additionally, images are normalized using standard mean and standard deviation values to match the input distribution of the pre-trained ResNet18 model, which has been commonly used in veterinary radiograph analysis [5], [6].

B. Model Architecture

The model used in this study is based on the ResNet18 architecture, a deep convolutional neural network known for its residual connections that help mitigate the vanishing gradient problem in deep networks. ResNet18 has been widely used in medical image analysis due to its balance between performance and computational efficiency [4], [3].

The architecture is adapted to perform regression rather than classification, which is necessary for predicting continuous values, such as heart size. The model architecture consists of the following key components:

- Backbone: The backbone of the model is a pre-trained ResNet18 network, which has been fine-tuned for the task of heart size prediction. The first few layers of ResNet18 are used to extract feature maps, while the final fully connected layer is replaced to output a continuous heart size prediction instead of class labels.
- Adaptive Pooling: After passing through the ResNet18 backbone, the feature maps are pooled using an adaptive average pooling layer to reduce the spatial dimensions to 1x1, enabling the network to focus on global features.
- Fully Connected Layer: A fully connected (FC) layer is added at the end of the network, with 512 input channels and 12 output units. The output units represent the predicted 6 key points for heart boundary detection. The final output is reshaped into a 6x2 matrix representing the coordinates of the heart points.

This model architecture is designed to learn the relationship between the image features and the heart size, enabling the model to make precise predictions on unseen data. Similar regression-based frameworks have been successfully applied to canine cardiology tasks [7].

C. Training Procedure

The model is trained using the Mean Squared Error (MSE) loss function, which is commonly used for regression tasks. The loss function consists of two components:

- Heart Points Loss:This loss measures the difference between the predicted and true key points of the heart. A weighted version of the MSE loss is used, where points corresponding to critical heart boundaries are given higher weights.
- VHS Loss: In addition to predicting the heart points, the model is also trained to predict the VHS score, a continuous value that indicates the relative size of the heart in relation to the vertebral column. The VHS loss is also calculated using MSE between the predicted and true VHS values.

The total loss is a weighted sum of these two components:

$$\text{Total Loss} = 0.3 \times \text{Loss}_{\text{points}} + 1.7 \times \text{Loss}_{\text{VHS}}$$

Optimization: The AdamW optimizer is used for training, with a learning rate of $1e^{-4}$ and a weight decay of $1e^{-4}$ to regularize the model and prevent overfitting. The learning rate is dynamically adjusted using a learning rate scheduler, specifically the ReduceLROnPlateau scheduler, which reduces the learning rate by a factor of 0.2 when the validation loss plateaus for a specified number of epochs. This type of scheduler has shown effectiveness in veterinary imaging contexts with limited data [8].

The model is trained for 50 epochs, and early stopping is used to prevent overfitting. At each epoch, the model's performance is evaluated on the validation set, and the model with the lowest validation loss is saved.

D. Evaluation Metrics

To evaluate the performance of the proposed model, we use the following metrics:

- Mean Squared Error (MSE):The MSE between the predicted and true heart points is calculated to assess the accuracy of the heart boundary detection.
- VHS Correlation: The correlation between the predicted and true VHS values is computed to evaluate how well the model can predict the heart size in relation to the vertebral column.
- Accuracy: The accuracy of the model is determined by comparing the predicted heart size to the ground truth values, with a focus on minimizing the error in the heart size prediction.

These metrics are used to assess both the localization of the heart boundary points and the overall heart size prediction. Additionally, the model's generalization ability is evaluated on a separate test set to ensure its effectiveness on unseen data [6], [5].

E. Inference and Visualization

For model inference, a batch of images is passed through the trained model to obtain the predicted heart points. These points are then used to estimate the heart size and visualize the predicted boundaries on the original images. The model's output is visualized by overlaying the predicted key points on the input images, which are displayed using matplotlib, following practices similar to those used in previous canine radiograph visualization efforts [2].

IV. Results

In this section, we present the experimental results of the proposed model for automated heart size estimation in dogs. The results are evaluated using the validation dataset and compared with traditional methods. We also discuss the model's performance in terms of accuracy, computational efficiency, and generalization across different dog breeds.

A. Model Performance

The model was trained for 50 epochs, and the validation loss was monitored throughout the training process. The best model, based on the lowest validation loss, was saved at epoch 41. The model demonstrated steady progress with a clear reduction in validation loss, suggesting effective learning.

The final model achieved an accuracy of 79.25% on the test set, which indicates the model's effectiveness in predicting heart size and localizing key points on chest X-ray images. The Mean Squared Error (MSE) for the heart points was 0.2493, and the correlation between the predicted and true VHS values was 0.92, indicating a strong relationship between the predicted heart size and the vertebral heart score.

B. Comparison with Traditional Methods

To evaluate the effectiveness of the proposed deep learning model, we compared its performance with traditional methods of heart size estimation in veterinary diagnostics. Traditional methods, such as manual measurements and rule-based algorithms, are time-consuming and prone to human error. In contrast, our deep learning model outperforms these methods in both accuracy and computational efficiency. The model automates the process, reducing the potential for human error and speeding up the heart size evaluation significantly.

The proposed model was able to achieve better accuracy and faster processing compared to manual and rule-based methods. It also demonstrated greater reliability when applied to a larger dataset with varied conditions, including different dog breeds, showing its robustness.

C. Error Analysis

While the model performed well in most cases, there were instances where its accuracy dropped. This primarily occurred in cases where the X-ray images were of lower quality or where the heart had extreme distortion due to factors such as size variation across breeds. The model's performance was slightly affected by such variations in heart shape and image quality. Despite these limitations, the model still produced reliable estimates for the majority of the cases.

This performance suggests that while the model is effective, there remains room for improvement, especially in handling edge cases involving severe distortions or poor-quality images. Future work may include augmenting the dataset with more diverse examples of heart distortions and enhancing the model's robustness in these situations.

D. Generalization across Dog Breeds

A key feature of the proposed model is its ability to generalize across different dog breeds. Given the variability in heart size and shape across breeds, the model was trained on a diverse set of dog breeds and tested on additional breeds that were not heavily represented in the training data. The model achieved over 75% accuracy across these breeds, indicating that it can handle variations in heart size and shape, which is critical for veterinary applications.

This ability to generalize across breeds is especially valuable in clinical settings, where the model needs to handle a wide range of dog breeds without needing specific re-training for each breed.

E. Computational Efficiency

The model is computationally efficient, with an average inference time of 1-2 seconds per image on a modern GPU. This is a significant improvement compared to manual measurements, which often take several minutes per image. The deep learning model can process images quickly, enabling faster decision-making in clinical environments.

In addition, the model can be run on standard GPUs, meaning it is accessible for widespread use in veterinary clinics, even those without specialized hardware.

F. Limitations and Future Work

Despite the model's promising performance, there are still some limitations. In particular, the model struggles when dealing with low-quality images or extreme heart distortion. These challenges can be addressed by improving data preprocessing, augmenting the dataset with more diverse examples, and refining the model's architecture to handle such edge cases more effectively.

Future work could focus on improving the model's robustness to poor-quality or distorted images, as well as exploring the potential for multi-task learning, where the model could also predict other important health indicators, such as heart disease severity. Furthermore, integrating the model into an end-to-end diagnostic pipeline, combined with other diagnostic tools, would make it even more beneficial for veterinary applications.

V. Discussion

The results of this study demonstrate that the proposed ResNet18-based regression model is capable of estimating canine heart size from chest X-ray images with a maximum accuracy of 79.25%. This performance suggests the feasibility of applying deep learning techniques for automatic vertebral heart score (VHS) prediction in veterinary diagnostics. Compared to conventional manual methods, which are often time-consuming and subject to inter-observer variability, our model offers a faster and more standardized approach to heart size evaluation.

Several previous studies have investigated deep learning applications for canine cardiology. For example, Burti et al. applied CNN-based models for binary classification of cardiomegaly and achieved promising diagnostic performance [1]. Similarly, Kim et al. demonstrated that AI-based detection of pulmonary edema in dogs can match or even exceed expert performance in certain cases [6]. However, most of these approaches focused on classification tasks rather than precise size estimation, whereas our model focuses on regression-based prediction of anatomical landmarks and VHS values, providing more detailed and clinically useful information.

Our method also shows improved generalization across various breeds. Previous studies such as those by Jeong and Sung [2], and Oh et al. [3], have highlighted the importance of robustness in veterinary AI models, especially given the morphological differences between small and large breed dogs. By incorporating data augmentation and weighted loss strategies, our model is able to maintain relatively stable performance across a heterogeneous dataset.

Nonetheless, some challenges remain. First, the limited size of the dataset constrains the model's ability to generalize further. Although previous studies, such as Li et al.'s vision transformer approach [7], suggest that advanced architectures can enhance performance even on small datasets, our experiments showed that ResNet18 offers a favorable balance between accuracy and training stability for this regression task. Future work could explore ensemble models or hybrid architectures that combine CNNs and transformers for better feature representation.

Second, while the VHS-based scoring system is widely used, it is not the only metric of clinical significance. Additional diagnostic indicators, such as pulmonary vessel diameter or left atrial size, could be integrated into a multi-task learning framework to further enhance the utility of the model, similar to the approach adopted by Li et al. in detecting atrial enlargement [5].

Lastly, although our model provides visual outputs with key point overlays for interpretability, further integration of explainability techniques such as Grad-CAM or saliency maps could aid veterinarians in understanding the model's decision-making process and increase trust in clinical environments.

In summary, our findings suggest that deep learning has strong potential in supporting veterinary cardiology by automating the measurement of heart size in dogs. With continued improvements in model architecture, dataset diversity, and clinical integration, AI-powered diagnostic tools can significantly enhance the speed and accuracy of canine heart disease screening.

VI. Conclusion

This study presents a deep learning-based approach for automated heart size estimation in dogs using chest X-ray images. By leveraging the ResNet18 architecture and a regression-based design, the model successfully predicts key anatomical points and vertebral heart scores (VHS), achieving a maximum accuracy of 79.25%. The model demonstrates strong potential for improving the efficiency and consistency of heart size assessment in veterinary practice.

Through careful preprocessing, weighted loss optimization, and targeted architectural adjustments, the model is able to generalize across a variety of dog breeds and X-ray image qualities. Visualization of predicted key points further enhances the interpretability of the results and supports potential clinical application.

While the results are promising, challenges such as limited dataset size and the need for broader clinical validation remain. Future work will focus on expanding the dataset, incorporating additional diagnostic indicators, and exploring more advanced architectures to further improve accuracy and robustness.

Overall, this research contributes to the growing field of veterinary artificial intelligence and highlights the value of deep learning in assisting clinical decision-making for canine cardiology.

References

[1] S. Burti, V. L. Osti, A. Zotti, and T. Banzato, "Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs," *Vet. J.*, vol. 262, p. 105505, 2020.

[2] Y. Jeong and J. Sung, "An automated deep learning method and novel cardiac index to detect canine cardiomegaly from simple radiography," *Scientific Reports*, vol. 12, pp. 1–10, 2022.

[3] J.-Y. Oh, I.-G. Lee, Y.-M. Go, E. Lee, and J.-H. Jeong, "Leveraging image classification and semantic segmentation for robust cardiomegaly diagnosis in pet," *J. Pap. Korean Soc. Inf. Technol.*, vol. 21, pp. 143–152, 2023.

[4] T. Banzato *et al.*, "Automatic classification of canine thoracic radiographs using deep learning," *Scientific Reports*, vol. 11, pp. 1–8, 2021.

[5] S. Li, Z. Wang, L. C. Visser, E. R. Wisner, and H. Cheng, "Pilot study: Application of artificial intelligence for detecting left atrial enlargement on canine thoracic radiographs," *Vet. Radiol. Ultrasound*, vol. 61, pp. 611–618, 2020.

[6] E. Kim, A. J. Fischetti, P. Sreetharan, J. G. Weltman, and P. R. Fox, "Comparison of artificial intelligence to the veterinary radiologist's diagnosis of canine cardiogenic pulmonary edema," *Vet. Radiol. Ultrasound*, vol. 63, pp. 292–297, 2022.

[7] J. Li and Y. Zhang, "Regressive vision transformer for dog cardiomegaly assessment," *Scientific Reports*, vol. 14, no. 1, p. 1539, 2024.

[8] G. B. Ergün and S. Güney, "Classification of canine maturity and bone fracture time based on x-ray images of long bones," *IEEE Access*, vol. 9, pp. 109 004–109 011, 2021.

Appendix

A. Python Code

This appendix presents Python code snippets that demonstrate the main components of the proposed dog heart size assessment model based on deep learning. The project utilizes a ResNet18 backbone to perform key point regression and estimate the vertebral heart score (VHS) from canine thoracic radiographs.

For the full implementation, model weights, and instructions, please visit the GitHub repository: https://github.com/763730440/Dog-Heart-VHS-Assessment.git