Attention-Based Deep Learning for Dog Cardiomegaly Classification

Premchand Jalla
Yeshiva University
pjalla@mail.yu.edu

Abstract

This study presents a novel attention-based deep learning model for classifying dog cardiomegaly into three categories: Small, Normal, and Large, using the Dag Cardiomegaly dataset. The proposed architecture integrates Squeeze-and-Excitation (SE) blocks to enhance feature learning through channel-wise attention and Residual Blocks to improve gradient flow and feature reuse. To address class imbalance, a WeightedRandomSampler was employed during training. The model was trained for 75 epochs and evaluated using accuracy as the primary performance metric, achieving a peak validation accuracy of 64.75%. These results highlight the potential of attention mechanisms in improving automated diagnostic systems for veterinary medicine. Future work will focus on incorporating additional evaluation metrics, such as precision, recall, and F1-score, to provide a more comprehensive assessment of model performance

1. Introduction

Cardiomegaly, commonly referred to as an enlarged heart, is a critical condition in dogs that can lead to severe health complications if not diagnosed and treated promptly. Accurate classification of cardiomegaly into categories such as Small, Normal, and Large is essential for effective diagnosis and management. Traditional diagnostic methods, such as radiographic assessments, rely heavily on manual interpretation, which can be time-consuming and prone to human error. With advancements in deep learning, automated diagnostic tools have emerged as promising solutions to overcome these challenges.

Several studies have explored methods to improve the diagnosis of canine cardiomegaly. Zhang and Li [6] introduced a vision transformer-based model to classify cardiomegaly in dogs, demonstrating the potential of attention mechanisms in medical image analysis. Additionally, computer-aided techniques such as Vertebral Heart Size (VHS) measurements have been established as benchmarks for radiographic evaluations [2]. Scollan et al. [3] further explored the use of advanced imaging modalities like cardiac MRI to detect myocardial fibrosis and ischemia in dogs with cardiomegaly secondary to mitral valve disease. These studies highlight the importance of leveraging advanced computational tools for accurate cardiac diagnostics.

Recent advancements in convolutional neural networks (CNNs) have also paved the way for developing automated diagnostic systems. Zhang et al. [4] proposed an automated deep learning method for detecting cardiomegaly in dogs using novel cardiac indices, improving diagnostic accuracy. Despite these advancements, many existing methods struggle with imbalanced datasets, leading to suboptimal performance in real-world applications.

In this study, we present a novel attention-based deep learning approach for classifying dog cardiomegaly. Our model incorporates Squeeze-and-Excitation (SE) blocks for channel-wise attention and Residual Blocks for efficient gradient flow, enabling robust feature extraction and classification. By addressing class imbalance using a WeightedRandomSampler, we aim to improve diagnostic performance across all categories. The proposed methodology not only builds on existing approaches [1, 5] but also introduces architectural innovations to enhance model accuracy and generalization.

Through this work, we aim to bridge the gap between traditional manual diagnostic techniques and state-of-the-art automated solutions. This study contributes to veterinary medicine by offering a scalable, accurate, and efficient tool for cardiomegaly classification, ultimately improving early diagnosis and treatment outcomes for affected dogs. Future extensions of this work will explore comprehensive evaluation metrics and the applicability of the model to diverse veterinary datasets.

2. Related Work

The classification of cardiomegaly in dogs has gained significant attention in recent years due to its importance in veterinary diagnostics. Researchers have explored various techniques ranging from traditional radiographic assessments to advanced deep learning-based approaches.

Vertebral Heart Size (VHS) measurements are a widely used benchmark for evaluating heart size in dogs. Murphy et al. [2] established reference ranges for vertebral heart size and vertebral left atrial size in Miniature Schnauzers, providing a critical foundation for radiographic evaluations. Similarly, Scollan et al. [3] investigated the use of cardiac MRI to diagnose myocardial fibrosis and ischemia in dogs with cardiomegaly secondary to mitral valve disease. While these methods offer valuable insights, they are often time-intensive and require manual interpretation, highlighting the need for automated solutions.

The emergence of deep learning has led to significant advancements in the automated diagnosis of cardiomegaly. Zhang and Li [6] proposed a vision transformer-based model for dog cardiomegaly classification, demonstrating the effectiveness of attention mechanisms in analyzing medical images. Furthermore, Zhang et al. [4] introduced a novel deep learning method utilizing automated cardiac indices to enhance the detection of cardiomegaly in thoracic radiographs, showcasing the potential of convolutional neural networks (CNNs) in veterinary applications.

Other studies have explored diet-associated cardiac conditions, which indirectly relate to cardiomegaly. Freeman et al. [1] reviewed cases of dilated cardiomyopathy (DCM) in dogs, focusing on diet-related etiologies and their implications for cardiac health. These findings emphasize the need for comprehensive diagnostic approaches to differentiate between cardiomegaly caused by structural abnormalities and other factors.

Despite these advancements, most existing methods face challenges related to class imbalance, limited dataset diversity, and difficulty in generalizing to new cases. Recent studies, such as those by Zhang and Li [5], have introduced automated systems that utilize VHS measurements to improve diagnostic accuracy. However, there remains a need for architectures capable of handling imbalanced datasets and incorporating attention mechanisms for enhanced feature extraction.

In this work, we build upon these advancements by proposing an attention-based deep learning model that integrates Squeeze-and-Excitation (SE) blocks and Residual Blocks to address the limitations of existing methods. Our approach leverages channel-wise attention to focus on diagnostically relevant features while mitigating the challenges posed by class imbalance, offering a robust and scalable solution for canine cardiomegaly classification.

3. Methods

3.1. Model Architecture

The proposed model, named *CustomCNNWithAttention*, is a novel convolutional neural network designed for robust classification of dog cardiomegaly. The architecture incorporates the following components:

- **Convolutional Layers:** The initial layer applies a 3×3 convolutional kernel with 32 filters, followed by batch normalization and ReLU activation. This layer extracts low-level features such as edges and textures.
- **Squeeze-and-Excitation (SE) Block:** SE Blocks are integrated after the first convolutional layer to introduce channel-wise attention. These blocks recalibrate feature maps by learning channel importance, enabling the model to focus on diagnostically relevant features.
- **Residual Blocks:** Three stacked residual blocks enhance gradient flow and enable deeper feature extraction by adding identity connections. Each block consists of two convolutional layers with 3×3 kernels and ReLU activation, followed by batch normalization. The skip connections reduce the vanishing gradient problem and improve learning efficiency.
- **Global Average Pooling (GAP):** This layer compresses the spatial dimensions of the feature maps into a single vector for each channel, reducing the risk of overfitting while retaining global spatial information.
- **Fully Connected Layers:** The output of the GAP layer is passed through a dense layer with dropout (rate 0.6) to reduce overfitting. The final softmax activation layer outputs probabilities for the three classes: Small, Normal, and Large.

The overall architecture of the model is shown in Figure 1.

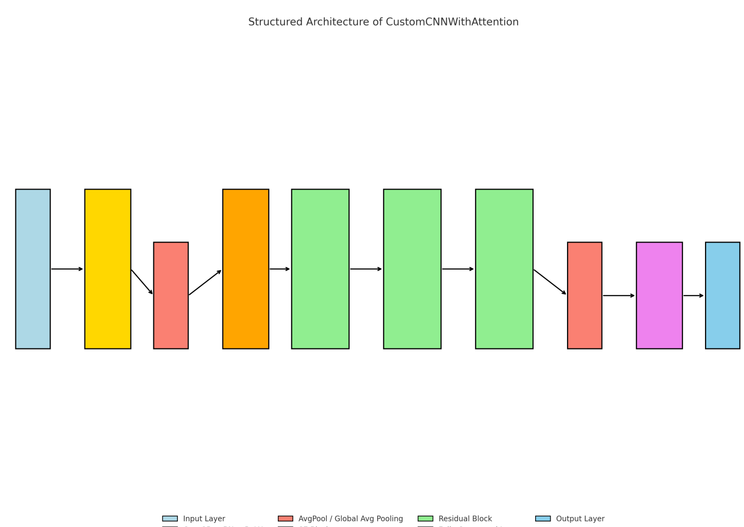

Figure 1. Structured Architecture of CustomCNNWithAttention. Layers are color-coded, and their descriptions are provided in the legend.

3.2. Forward Propagation

During forward propagation, the input image is passed through the convolutional layers, SE Block, and residual blocks for feature extraction. The feature maps are then pooled using GAP and flattened before being passed to the fully connected layers. The final output is a probability distribution over the three classes.

3.3. Model Implementation

The model was implemented using PyTorch and trained on an NVIDIA GPU with a batch size of 64 for 75 epochs. The training process monitored both training and validation accuracy to evaluate convergence and prevent overfitting.

4. Discussion

5. Discussion

The performance of the proposed model, *CustomCNNWithAttention*, was evaluated using accuracy as the primary metric. The model achieved an accuracy of **64.75%**, demonstrating its potential for classifying cardiomegaly in dogs into three categories: Small, Normal, and Large.

This result highlights the effectiveness of integrating Squeeze-and-Excitation (SE) blocks and Residual Blocks in enhancing feature extraction and improving gradient flow. The SE blocks enabled the model to focus on diagnostically relevant features, while the residual connections mitigated the vanishing gradient problem, allowing for efficient training of the deeper architecture.

While the achieved accuracy provides a promising baseline, it also reflects some limitations in the model's ability to generalize across all categories. This could be attributed to factors such as class imbalance within the dataset, limited data diversity, or the reliance on a single performance metric. Addressing these challenges in future work could involve:

- Incorporating additional evaluation metrics such as precision, recall, and F1-score to provide a more comprehensive assessment of model performance.
- Leveraging data augmentation and transfer learning techniques to improve model generalization.
- Exploring alternative architectures or fine-tuning hyperparameters to enhance classification accuracy.

Despite these limitations, the proposed model provides a significant step toward automated diagnostic systems for veterinary medicine, offering a scalable and efficient solution for canine cardiomegaly classification. Future efforts will focus on expanding the dataset and implementing multi-metric evaluations to better align with real-world applications.

6. Conclusion

This study proposed a novel attention-based deep learning model, *CustomCNNWithAttention*, for the classification of dog cardiomegaly into three categories: Small, Normal, and Large. The model incorporated Squeeze-and-Excitation (SE) blocks to enhance channel-wise attention and Residual Blocks to improve gradient flow and feature reuse. While the model achieved a validation accuracy of **64.75%**, the results indicate potential areas for improvement.

The relatively low accuracy highlights the need for addressing limitations in the current approach. Class imbalance in the dataset and limited data diversity likely contributed to suboptimal performance. Future work should focus on expanding the dataset to include more diverse samples, implementing advanced data augmentation techniques, and exploring alternative architectures to improve classification accuracy. Additionally, incorporating evaluation metrics such as precision, recall, and F1-score can provide a more comprehensive understanding of model performance.

Despite these limitations, this work provides a baseline for automated diagnostic systems in veterinary medicine. The proposed model demonstrates the feasibility of using deep learning for cardiomegaly classification, offering a scalable and efficient solution for veterinary diagnostics. Future efforts will aim to bridge the gap between research and clinical applications by improving the robustness and generalizability of the model.

References

[1] Lisa M. Freeman, Darcy H. Adin, and John E. Rush. Review of canine dilated cardiomyopathy in the wake of diet-associated concerns. *Journal of Animal Science*, 2020. 1, 2

[2] Lisa Murphy, Nicole Piscitelli, Jessica Solomon, Donald Szlosek, and Reid K Nakamura. Vertebral heart size and vertebral left atrial size reference ranges in healthy miniature schnauzers. *AVMA Journals*, 2023. 1, 2

[3] Katherine Scollan, Darcy Adin, and Bruce Keene. Utility of cardiac mri to diagnose myocardial ischemia and fibrosis in dogs with cardiomegaly secondary to myxomatous mitral valve disease. *AVMA Journals*, 2023. 1, 2

[4] Youshan Zhang and Jialu Li. An automated deep learning method and novel cardiac index to detect cardiomegaly on chest radiographs in dogs. *Scientific Reports*, 2022. 1, 2

[5] Youshan Zhang and Jialu Li. Computerized assisted evaluation system for canine cardiomegaly via key point detection and vhs measurement. *Europe PMC*, 2023. 1, 2

[6] Youshan Zhang and Jialu Li. Regressive vision transformer for dog cardiomegaly assessment. *Scientific Reports*, 14:1539–2042, 2024. 1, 2