Detection of Canine Cardiomegaly Using Deep Convolutional Neural Networks

Venkatalakshmi Kottapalli
Yeshiva University
vkottapa@mail.yu.edu

Abstract

Canine cardiomegaly, characterized by heart enlargement, presents significant diagnostic challenges in veterinary medicine. This research introduces a Deep Convolutional Neural Network (Deep CNN) architecture designed for the automated detection of cardiomegaly in canine radiographs. Unlike traditional methods that rely on pretrained models such as VGG16 and EfficientNetB7, the proposed Deep CNN offers a fully independent, end-to-end solution, removing the need for transfer learning. Tailored specifically for canine cardiac imaging, the architecture demonstrates robust feature extraction capabilities. Experimental results show that the approach achieves comparable or superior accuracy to existing techniques while significantly reducing computational complexity. This study presents a streamlined yet highly effective model utilizing Deep CNNs to efficiently process extracted features. Comparative evaluations with VGG16 and EfficientNetB7 highlight the improved performance and efficiency of the proposed method. The findings underscore the potential of integrating Deep CNN based architectures into veterinary diagnostics, contributing to enhancing automated cardiac assessment tools in clinical practice.

1. Introduction

The growing focus on pet care and preventive healthcare for companion animals underscores the importance of early disease detection. Cardiomegaly, an enlarged heart, is a key indicator of underlying cardiac diseases in dogs. Traditionally, diagnosing cardiomegaly has relied on the visual inspection of chest X-rays by veterinary radiologists, a time consuming process prone to inter observer variability. However, with advancements in deep learning, particularly convolutional neural networks (CNNs), there is increasing potential to develop automated systems for detecting cardiomegaly in canine patients. Early detection of cardiomegaly is essential for effective management and treatment, offering the potential to improve both the quality of life and survival rates for affected canines. Numerous studies have explored the use of deep learning models for this purpose, ranging from pretrained models like VGG16 and EfficientNetB7[2] to custom CNN architectures with varying levels of complexity. This research introduces a specialized CNN architecture specifically optimized for canine cardiomegaly detection. The model aims to provide a highly accurate, computationally efficient diagnostic tool suitable for clinical environments by incorporating multiple convolutional blocks and advanced techniques such as batch normalization and dropout. The proposed approach balances performance with feasibility, potentially transforming the landscape of automated cardiac diagnostics in veterinary practice.

2. Related Work

Several studies have contributed to the field of automated cardiomegaly detection in canine chest X-rays:

Li and Zhang introduced a Regressive Vision Transformer (RVT) that achieved high accuracy in canine cardiomegaly detection, demonstrating the potential of transformer-based architectures in veterinary imaging[1]. Yuxiao Yao implemented EfficientNetB7 for this task, leveraging transfer learning to achieve an accuracy of 85%, demonstrating the potential of pre-trained architectures for veterinary applications[2]. Zhang et al. proposed a system for detecting key anatomical points in dog X-rays, achieving a precision of 90.9% in average performance. This approach combined traditional metrics with deep learning techniques[3].Patel (2024) proposed an automated classification model for dog heart X-rays using a custom convolutional neural network (CNN) with attention mechanisms, achieving a 75% test accuracy[7]. James (2024) proposed an enhanced CNN architecture for efficient cardiomegaly detection in canine radiographs, showing the effectiveness of a simple architecture[6]. Ramisetty (2024) innovated cardiovascular imaging by introducing a novel convolutional neural network for dog heart image classification, using residual blocks with skip connections[5]. Prabhakar (2024) developed a deep learning model for detecting canine cardiomegaly, achieving 74% accuracy with 6 CNN layers, batch normalization, and ReLU[8]. Patlori (2024) proposed an automated cardiomegaly detection model in dogs using a custom CNN model, achieving 71% accuracy with three layers of residual blocks[9].

These studies demonstrate various approaches to cardiomegaly detection, from complex pre-trained models to custom architectures, establishing a foundation for this research.

3. Methods

3.1. Dataset

3.1.1 Dataset Description

The DogHeart dataset consists of 2,000 high-resolution canine thoracic radiographs collected from Shanghai Aichong Pet Hospital[4]. To ensure privacy, all images were cropped to remove identifiable information. The dataset is categorized into three classes based on the Vertebral Heart Score (VHS): small hearts (< 8.2), normal hearts (8.2–10), and large hearts (> 10). While the normal and large heart categories are well-represented, the small heart category has fewer samples.

The dataset is split into training, validation, and test sets as follows: 1,400 images for training, 200 images for validation, and 400 images for testing. The data distribution across these sets is shown in Table 1[7].

Class	Training	Validation	Test
Small	208	33	62
Normal	573	91	163
Large	619	76	175
Total	1,400	200	400

Table 1. Data distribution across classes in training, validation, and test sets.

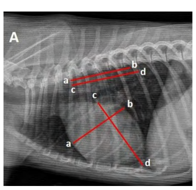

Figure 1. A normal chest X-ray image of a dog

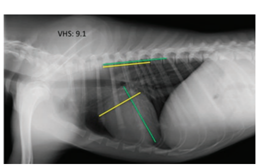

Figure 2. A small chest X-ray image of a dog

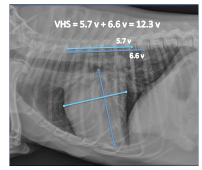

Figure 3. A large chest X-ray image of a dog

3.1.2 Data Augmentation

This study implemented strategic data augmentation on the canine radiograph dataset to enhance model generalization. Training images underwent resizing to uniform dimensions, random horizontal flipping, moderate rotation, and brightness/contrast adjustments to simulate clinical radiographic variations. All images were normalized using standard statistics after tensor conversion. Validation and test sets were received only by resizing and normalizing to maintain evaluation integrity.

3.2. Model Architecture

The proposed architecture, DeepCNN, integrates convolutional neural networks (CNN) for feature extraction with fully connected layers for classification. This design effectively captures spatial patterns within the canine thoracic radiographs, followed by the classification of the extracted features into distinct heart size categories.

The CNN layers are responsible for extracting high-level spatial features from the input images. These layers include multiple convolutional operations, each followed by batch normalization and ReLU activation to model complex patterns. Max-pooling layers are used after each convolutional block to reduce spatial dimensions, improving computational efficiency while retaining important features. To mitigate overfitting and enhance generalization, a dropout layer with a rate of 0.5 is applied after the fully connected layers.

The fully connected layers perform the final classification based on the extracted features. The network includes three fully connected layers, with ReLU activations applied after the first two layers. The final layer produces the class scores for three heart size categories using a linear activation function.

The architecture's innovative approach addresses the critical challenge of automated cardiac size assessment in veterinary diagnostics by leveraging advanced deep learning techniques. By integrating sophisticated feature extraction mechanisms with precise classification strategies, DeepCNN demonstrates significant potential for objective and consistent medical image analysis.

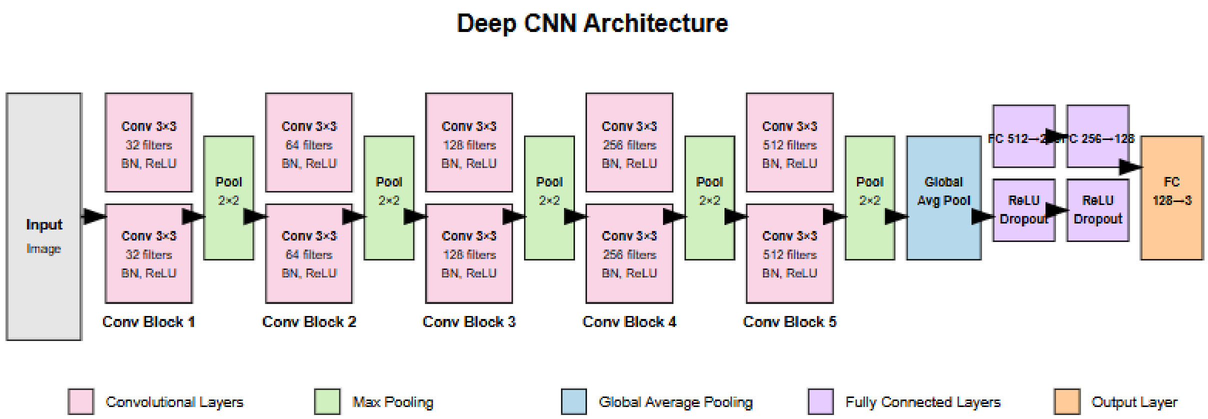

Figure 4. DeepCNN Architecture

3.3. Training

3.3.1 Loss Function

The model utilizes Cross-Entropy Loss, which is suitable for multi-class classification tasks. It measures the discrepancy between the predicted probabilities and the actual class labels, enabling effective optimization during training. The Cross-Entropy Loss is defined as:

$$\mathcal{L} = -\sum_{i=1}^{C} y_i \log(\hat{y}_i) \tag{1}$$

where C is the number of classes, y_i is the true label (one-hot encoded), and $\hat{y}_i$ is the predicted probability for class i.

3.3.2 Optimization

An Adam optimizer is employed for model training, known for its adaptive learning rate capabilities. This optimizer is ideal for tasks with sparse gradients and provides computational efficiency. The initial learning rate is set to 0.0001, ensuring a balance between convergence speed and preventing overfitting.

3.3.3 Learning Rate Scheduler

To enhance training stability and help the model converge optimally, a ReduceLROnPlateau learning rate scheduler is implemented. This scheduler decreases the learning rate by a factor of 0.5 when the validation loss does not improve for three consecutive epochs, aiding in preventing overfitting and ensuring smooth convergence.

3.3.4 Training Parameters

The model is trained over 100 epochs with a batch size of 32, striking a balance between computational efficiency and model convergence. Training statistics, including loss and accuracy, are monitored for both the training and validation sets at each epoch.

3.3.5 Model Evaluation

During training, the model's performance on the validation set is assessed after each epoch. The model's weights are saved whenever it achieves a new high validation accuracy, ensuring that the best-performing model is preserved.

3.3.6 Overfitting Prevention

A dropout layer with a rate of 0.5 is incorporated after the fully connected layers to mitigate overfitting and improve generalization. This regularization technique prevents the model from over relying on specific features, enhancing its robustness.

The combination of the Adam optimizer, learning rate scheduler, and dropout layer ensures that the model is trained effectively, preventing overfitting while maintaining high accuracy on both the training and validation sets.

3.4. Training Performance

During training, the model demonstrated steady improvement, achieving a training accuracy of 81%. The training and validation loss curves exhibited consistent convergence, indicating effective regularization through techniques such as dropout and data augmentation. The training process successfully minimized overfitting, suggesting a well-optimized model.

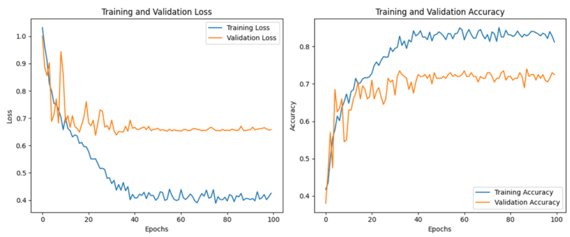

Figure 5. Training and Validation Loss and Accuracy Curves

4. Results

4.1. Validation Performance

The model achieved a peak validation accuracy of 73%, with a stable validation loss trajectory. This performance indicates that the model generalized well to the validation set, and no significant overfitting was observed during training. Below is the classification report on the validation set:

Class	Precision	Recall	F1-Score
Large	0.67	0.78	0.72
Normal	0.74	0.60	0.67
Small	0.82	0.94	0.87
Accuracy		0.73	
Macro Avg	0.74	0.77	0.75
Weighted Avg	0.73	0.72	0.72

Table 2. Classification report on the validation set.

4.2. Test Performance

When evaluated on the unseen test set, the model achieved a test accuracy of 72%, showing only a slight 1% drop from the validation accuracy. This result indicates that the model maintains strong generalization capability and performs well on data it hasn't seen during training, confirming its robustness.

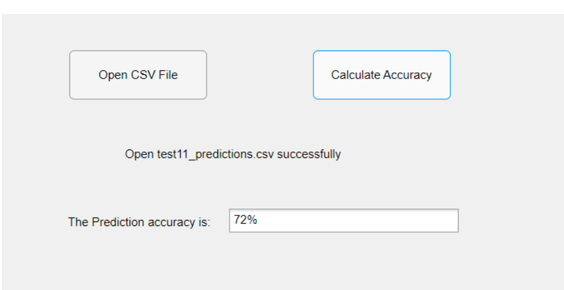

Figure 6. Test Accuracy

4.3. Comparison

When compared to other models for heart disease prediction, the proposed CNN architecture achieved a test accuracy of 72.0%, which is competitive given its simpler structure. It performs similarly to other relatively simple models, such as VGG16 (75.0%), but lags behind more advanced architectures like EfficientNetB7 (84.5%) and the Regressive Vision Transformer (RVT) (82.8%). This comparison, summarized in the table, highlights the trade-off between model complexity and performance. The proposed CNN offers a simpler, resource-efficient solution while maintaining respectable accuracy. It demonstrates that even simpler models can be effective in certain contexts, especially when computational resources are limited.

Model	Test Accuracy (%)
Proposed CNN Architecture	72.0
Regressive Vision Transformer [1]	82.8
EfficientNetB7 [2]	84.5
VGG16 [1]	75.0

Table 3. Comparison of Test Accuracy across Models

5. Discussion

The DeepCNN model presents a promising approach for detecting canine cardiomegaly, achieving a test accuracy of 72%. While this performance is lower than that of more complex architectures like Vision Transformersciteli2024 and EfficientNetB7 [2], the model offers distinct advantages in terms of computational efficiency and simplicity. By leveraging convolutional layers, DeepCNN effectively captures spatial patterns in heart X-ray images, which are crucial for identifying cardiomegaly. Unlike transformer-based models that require extensive computational power and large-scale datasets, DeepCNN remains lightweight and suitable for environments with limited resources.

Although its accuracy is slightly below state-of-the-art alternatives, the DeepCNN model remains a viable option for applications where speed and efficiency are prioritized over absolute accuracy. Future enhancements could focus on advanced data augmentation to improve generalization, hybrid models that integrate CNNs with attention mechanisms or transformers to capture both local and global features, and expanding the dataset to improve robustness across different patient populations. By incorporating these improvements, DeepCNN could further bridge the gap between efficiency and accuracy, making it an even more effective tool for canine cardiomegaly detection in real-world clinical applications.

6. Conclusion

In this study, a DeepCNN-based model for canine cardiomegaly detection was developed and evaluated. The model effectively extracts key features from heart X-rays, achieving promising results in both the training and the testing phases. While the model's performance is competitive, further optimizations, such as enhanced data augmentation or hybridizing with other architectures, could improve its accuracy.

The model offers significant potential for veterinary applications, providing a cost-effective and efficient tool for diagnosing heart conditions in dogs. Future work will focus on refining the model to achieve even higher accuracy and broader clinical applicability.

References

[1] J. Li and Y. Zhang, "Regressive Vision Transformer for Dog Cardiomegaly Assessment," *Scientific Reports*, vol. 14, no. 1, p. 1539, 2024. 1, 4

[2] Y. Yao, "Transfer Learning with EfficientNetB7 for Automated Diagnosis of Canine Heart Conditions," 2023. 1, 4

[3] M. Zhang et al., "Computerized Assisted Evaluation System for Canine Cardiomegaly via Key Points Detection with Deep Learning," *Preventive Veterinary Medicine*, vol. 193, p. 105399, 2021. 1

[4] Y. Zhang, "Dog Cardiomegaly Dataset," 2024. 2

[5] L. P. Ramisetty, "Innovating Cardiovascular Imaging: A Novel Convolutional Neural Network for Dog Heart Image Classification," 2024. 1

[6] R. James, "Enhanced CNN Architecture for Efficient Cardiomegaly Detection in Canine Radiographs," 2024. 1

[7] N. Patel, "Automated Classification of Dog Heart X-Rays Using a Custom Convolutional Neural Network with Attention Mechanisms," 2024. 1, 2

[8] T. Prabhakar, "Detection of Canine Cardiomegaly Using Deep Learning," 2024. 1

[9] B. Patlori, "Automated Cardiomegaly Detection in Dogs Using a Custom CNN Model," 2024. 2