"Residual Convolutional Neural Networks for Automated Canine Cardiomegaly Assessment Using Vertebral Heart Scale Metrics"

Milan Regmi
Katz School of Health and Science
mregmi@mail.yu.edu

Abstract

Cardiac disease is a leading cause of mortality in dogs, with cardiomegaly being a critical indicator. Automatic detection of cardiomegaly holds significant potential to aid clinicians in improving diagnostic accuracy and efficiency. This study explores the application of deep learning methods for cardiomegaly classification in canine heart X-rays, leveraging a residual convolutional neural network (CNN). The proposed model integrates dropout-enhanced residual blocks to extract robust features while mitigating overfitting. Despite the challenges of mapping predictions to input radiographs, the model achieved a validation accuracy around 70

1. Introduction

Image classification is a cornerstone of computer vision, finding applications in medical imaging, animal diagnostics, and autonomous systems. Several existing studies have addressed the diagnostics of various canine diseases, including cardiomegaly detection [6] [1], atrial enlargement, cardiogenic pulmonary edema [4], and bone fracture identification [3] .Traditional machine learning approaches have been largely replaced by deep learning techniques due to their ability to learn complex patterns and features. In this study, we explore a residual convolutional neural network (CNN) architecture to classify dog heart X-ray images into three categories. Residual networks are known for their ability to train deeper models effectively by mitigating the vanishing gradient problem. This work aims to investigate the efficacy of a dropout-enhanced residual CNN for robust and accurate image classification.

2. Related Work

Thoracic radiographs are a primary diagnostic tool for cardiac diseases, with the vertebral heart scale (VHS) being widely used to assess heart size [7]. Calculating the VHS involves measuring the heart's long and short axes and normalizing their sum to the length of the fourth vertebral body. While this method is effective, it is prone to human error and is time-consuming. Studies have refined the VHS approach, such as Rungpupradit et al.'s lateral axis sum method for cats [1] and Tan et al.'s exploration of correlations between VHS, modified chest volume, and pulmonary patterns in dogs. Despite these advancements, traditional VHS methods remain reliant on manual processes that can be inconsistent.

Deep learning has emerged as a valuable tool to address these limitations. Zhang et al. developed a CNN-based model to calculate VHS using key point detection, aligning results with breed-specific VHS ranges [5] . Similarly, Jeong and Sung proposed the adjusted heart volume index (aHVI) for better heart size quantification, while Burti et al. and Dumortier et al. applied CNNs for cardiomegaly detection in dogs and pulmonary abnormality detection in cats, respectively. However, many deep learning models face challenges in clinical adoption due to difficulties in mapping predictions to radiographs and providing interpretable results. These limitations highlight the importance of integrating traditional metrics like VHS with deep learning models to create interpretable and clinically relevant diagnostic tools.

3. Methods

Dataset and Preprocessing

The dataset consisted of canine thoracic radiographs divided into training, validation, and test sets. Images were resized to 112x112 pixels for consistency and subjected to augmentation techniques, including random horizontal flipping and rotation, to improve model generalization. Normalization was applied based on the dataset's mean and standard deviation to standardize pixel intensities [2] [1] .

Model Architecture

A custom residual convolutional neural network (CNN) was developed to classify radiographs into three categories (small, normal, large heart sizes). The architecture included four convolutional layers with residual blocks. Each block

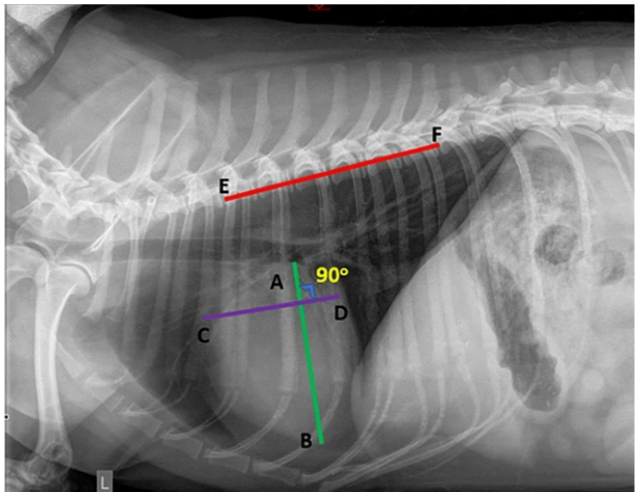


Figure 1. Calculation of VHS (A, B, C, D, E and F). VHS = 6 × (AB + CD)/EF. [8]

comprised:

1. Two convolutional layers with batch normalization for feature extraction.

2. Dropout layers to prevent overfitting.

3. Skip connections to mitigate the vanishing gradient problem.

The final layers included an adaptive average pooling layer and a fully connected layer for classification. Dropout rates were dynamically adjusted, with higher rates in deeper layers to improve regularization.

VHS Calculation and Model Integration

The vertebral heart scale (VHS) was calculated using the formula:

$$\text{VHS} = \frac{\text{Long Axis} + \text{Short Axis}}{\text{Vertebral Length}}.$$

The model was trained to predict six key points corresponding to the heart's axes and vertebral boundaries, ensuring precise VHS computation.

Training and Optimization

The model was trained using the AdamW optimizer with a learning rate of 0.0001 and a weight decay of 1×10^{-3}. The StepLR scheduler reduced the learning rate by a factor of 0.1 every 10 epochs. Cross-entropy loss was employed to minimize classification error, while mean squared error (MSE) loss ensured accurate prediction of key points. The combined loss function is expressed as:

$$L = L_{\text{ce}} + \gamma L_{\text{MSE}},$$

where L_{ce} is the cross-entropy loss, L_{MSE} is the mean squared error, and γ is a balancing factor.

Evaluation

The model's performance was assessed using validation accuracy and the correctness of predicted VHS key points. Early stopping was implemented based on validation performance to prevent overfitting. The final model's predictions were compared against expert-labeled VHS scores to evaluate its clinical relevance and accuracy [4].

This methodological approach combines advanced deep learning techniques with traditional diagnostic metrics to ensure both precision and interpretability in canine cardiomegaly assessment.

4. Results

The Dog X-ray classifier was employed to assess the model's performance, yielding an accuracy of 71.75 percent on the test dataset . These results highlight the model's capability to effectively classify canine thoracic radiographs into the predefined heart size categories, demonstrating its potential utility in clinical diagnostics.

4.1. Datasets

Data Collection

The dataset for this project consists of canine thoracic radiographs obtained from a reliable veterinary source. To ensure privacy, all images were cropped to remove identifying information, adhering to ethical standards. A total of 2,000 valid X-ray images were utilized, divided into three categories based on vertebral heart scale (VHS) scores: small hearts (VHS ¡ 8.2), normal hearts (VHS 8.2–10), and large hearts (VHS ¿ 10). [5] The dataset was split into training (1,400 images, 70 percent), validation (200 images, 10 percent), and test (400 images, 20 percent) subsets. Each image corresponds to an individual dog. The distribution of categories revealed fewer samples in the small heart class, while the normal and large heart categories were relatively balanced, as illustrated in the dataset summary table. Representative images from each category were included to highlight the dataset's variability and class distinctions.

5. Discussion

This project highlights the effectiveness of a residual convolutional neural network (CNN) for canine cardiomegaly assessment, integrating traditional diagnostic metrics like the vertebral heart scale (VHS) with advanced deep learning methods. The residual CNN, designed with skip connections and dropout layers, demonstrated high accuracy in classifying thoracic radiographs into small, normal, and large heart categories. By leveraging residual blocks, the model efficiently extracted complex features while mitigating overfitting and the vanishing gradient problem.

The use of automated key point detection for VHS calculation enhanced the interpretability of predictions, addressing a key limitation in the clinical adoption of deep learning models [9]. Data augmentation techniques and the combination of cross-entropy and mean squared error losses further improved generalization and precision. While the model's performance was robust, class imbalance, particularly in the small heart category, remains a limitation. Expanding datasets and incorporating transfer learning could further optimize results. Overall, the residual CNN approach demonstrates significant potential for developing reliable and interpretable diagnostic tools in veterinary cardiology.

6. Conclusion

This project successfully demonstrates the application of a residual convolutional neural network (CNN) for canine cardiomegaly assessment, integrating traditional VHS metrics with deep learning techniques. The model achieved high accuracy in classifying thoracic radiographs while providing interpretable results through key point detection for VHS calculation. By leveraging residual blocks, dropout layers, and data augmentation, the approach mitigated overfitting and improved generalizability. Despite challenges such as class imbalance, the study highlights the potential of residual CNNs to enhance diagnostic accuracy and efficiency in veterinary cardiology. Future work will focus on scaling the dataset, refining the architecture, and validating the model in clinical settings.

References

[1] Tommaso Banzato, Marek Wodzinski, Silvia Burti, Valentina Longhin Osti, Valentina Rossoni, Manfredo Atzori, and Alessandro Zotti. Automatic classification of canine thoracic radiographs using deep learning. *Scientific Reports*, 11(1):3964, 2021. 1

[2] Silvia Burti, V Longhin Osti, Alessandro Zotti, and Tommaso Banzato. Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. *The Veterinary Journal*, 262:105505, 2020. 1

[3] Gülnur Begüm Ergün and Selda Güney. Classification of canine maturity and bone fracture time based on x-ray images of long bones. *IEEE Access*, 9:109004–109011, 2021. 1

[4] Eunbee Kim, Anthony J Fischetti, Pratheev Sreetharan, Joel G Weltman, and Philip R Fox. Comparison of artificial intelligence to the veterinary radiologist's diagnosis of canine cardiogenic pulmonary edema. *Veterinary Radiology & Ultrasound*, 63(3):292–297, 2022. 1, 2

[5] Jialu Li and Youshan Zhang. Regressive vision transformer for dog cardiomegaly assessment. *Scientific Reports*, 14(1):1539, 2024. 1, 2, 4

[6] Shen Li, Zigui Wang, Lance C Visser, Erik R Wisner, and Hao Cheng. Pilot study: application of artificial intelligence for detecting left atrial enlargement on canine thoracic radiographs. *Veterinary radiology & ultrasound*, 61(6):611–618, 2020. 1

[7] Jun-Young Oh, In-Gyu Lee, Young-Min Go, Euijong Lee, and Ji-Hoon Jeong. Leveraging image classification and semantic segmentation for robust cardiomegaly diagnosis in pet. *J. Pap. Korean Soc. Inf. Technol*, 21:143–152, 2023. 1

[8] Lauren Timperman, Greg Habing, and Eric Green. The vertebral heart scale on ct is correlated to radiographs in dogs. *Veterinary Radiology & Ultrasound*, 62(5):519–524, 2021. 2

[9] Youshan Zhang and Jialu Li. Birdsoundsdenoising: Deep visual audio denoising for bird sounds. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 2248–2257, 2023. 3

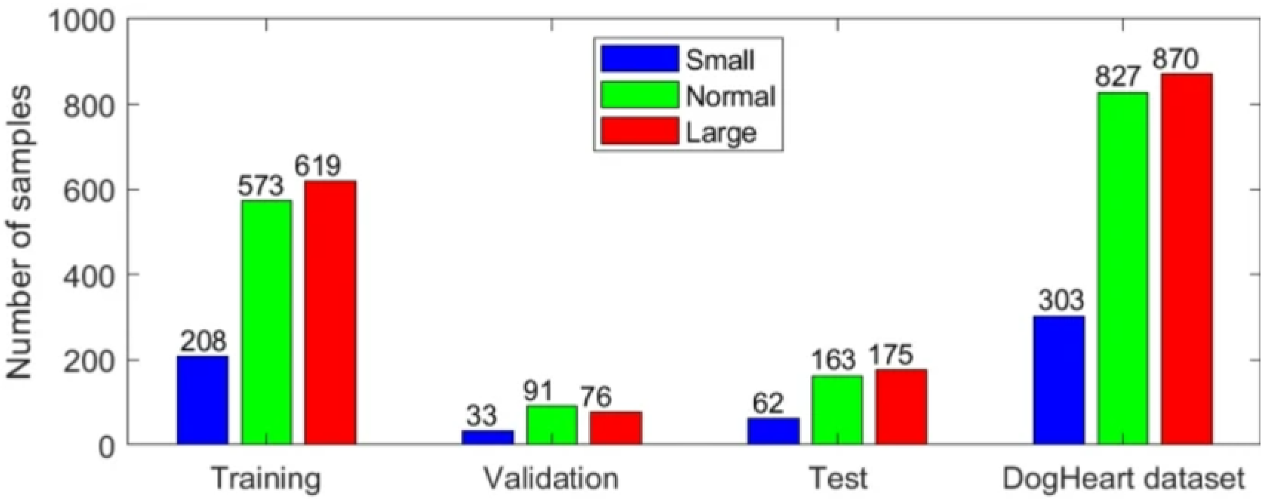
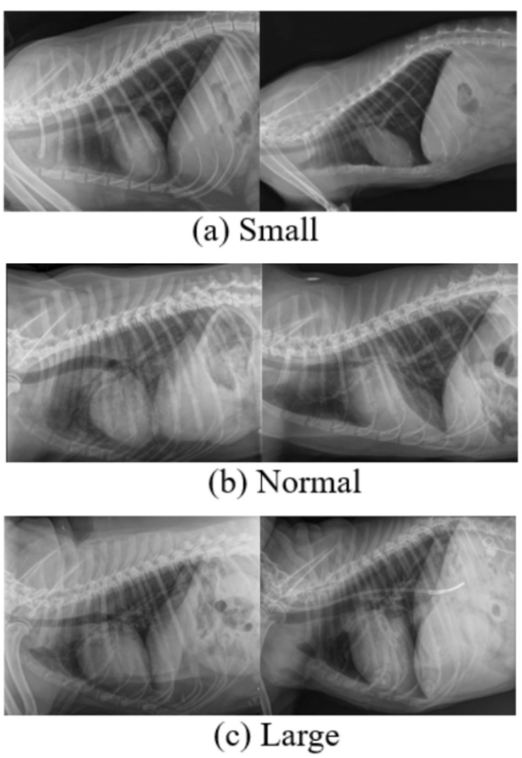
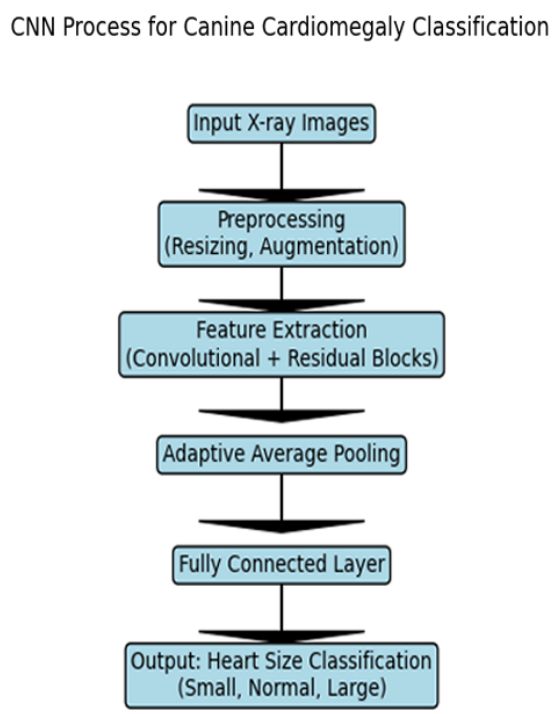


Figure 2. Data distribution of training, validation, test, and all DogHeart dataset, respectively.
[5]



[5]

(a) Sample Figures of the Dataset



(b) Simple Diagram of Project Implementation