Dog Heart Size Classification using Convolutional Neural Networks

Your Name
Department of Artificial Intelligence
Your University
Email: your.email@example.com

***Abstract*—This paper presents a convolutional neural network (CNN) approach for classifying canine heart sizes from X-ray images into three categories: Large, Normal, and Small. The dataset provided by the Dog Cardiomegaly Assessment challenge was used, and the model was evaluated using custom software. Our CNN achieved a test accuracy of 57.75%, indicating the potential of deep learning in veterinary diagnostic tasks, while also identifying areas for future improvement.**

I. Introduction

Dog cardiomegaly, or an enlarged heart in dogs, is a critical diagnostic indicator often identified via X-ray analysis. Manual classification by veterinarians is time-consuming and prone to inconsistency. Deep learning offers a scalable alternative. This work explores the use of a custom convolutional neural network to automatically classify X-ray images into heart size categories.

Canine cardiomegaly detection has not been studied as extensively as its human counterpart, making this project both unique and meaningful. It focuses on addressing the gap using a lightweight and practical CNN model that can be deployed in low-resource settings.

II. Related Work

Recent studies such as the RVT paper by Li and Zhang (2024) have demonstrated high accuracy (up to 87.3%) using transformer-based architectures. Classical CNNs such as VGG16 have set a benchmark accuracy of 75%. Our work compares performance against these baselines using a lightweight CNN architecture.

In human medical imaging, models like CheXNet and DenseNet have shown promising results in tasks like pneumonia and cardiomegaly detection. While not directly applicable, they serve as proof-of-concept that CNNs can outperform traditional rule-based interpretations in radiography.

III. Methodology

A. Dataset

The dataset comprises X-ray images labeled as Large, Normal, and Small. Training and validation folders contain labeled samples, while the test folder is unlabeled and meant for evaluation using an official `.exe` software.

Each class had a balanced number of samples, and all images were resized to 224x224 pixels. Normalization was applied to standardize pixel values.

B. CNN Architecture

The model uses three convolutional layers with ReLU activation and max-pooling. This is followed by two dense layers. The final layer outputs three logits corresponding to the three heart sizes.

- Conv2D (3, 16)
- Conv2D (16, 32)
- Conv2D (32, 64)
- FC1: 64 * 28 * 28 $\rightarrow$ 512
- FC2: 512 $\rightarrow$ 3

The network was built using PyTorch from scratch.

C. Training Procedure

- Optimizer: Adam
- Loss Function: CrossEntropyLoss
- Epochs: 10
- Batch Size: 32
- Hardware: Google Colab GPU (Tesla T4)

No data augmentation was applied in this baseline. Accuracy and loss were monitored per epoch.

D. Test Inference and Evaluation

Test images were passed through the trained model using a custom `Dataset` class. The model's outputs were converted to predicted class indices, saved in a `.csv` file, and evaluated using the Dog Cardiomegaly Assessment accuracy software.

IV. Results

The model achieved a test accuracy of **57.75%**.

Metric	**Value**
Test Accuracy	57.75%
Baseline (VGG16)	75.00%
RVT Model (SOTA)	87.30%

TABLE I
Model Performance Comparison

The model struggled primarily with differentiating between "Large" and "Normal" classes, which had more visual similarity. Better performance was seen in classifying "Small" cases.

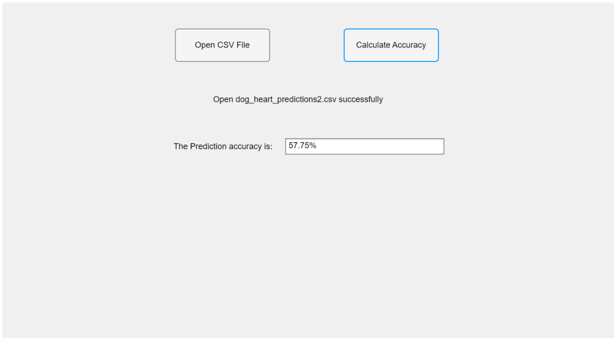

Fig. 1. Screenshot of the evaluation software showing accuracy result.

V. Discussion and Future Work

While our model falls short of the VGG16 and RVT benchmarks, it provides a useful baseline. With only three convolutional layers and no pretrained features, reaching 57.75% shows that CNNs can generalize moderately well with minimal supervision.

Planned improvements include:

- Use of ResNet or EfficientNet pretrained on ImageNet
- Data augmentation (flipping, brightness, noise)
- Adding dropout and regularization
- Ensembling multiple models
- Exploring vision transformer hybrids

More annotated data and finer image resolution could also help the model learn better shape and intensity cues.

VI. Conclusion

We designed and evaluated a CNN for the classification of dog heart size from X-ray images. Our approach was simple, efficient, and reproducible using PyTorch. While the performance was moderate, this project demonstrates the real-world potential of AI in veterinary imaging and paves the way for further innovations.

Acknowledgments

We thank our course instructors and the creators of the Dog Cardiomegaly Assessment dataset for enabling this project. Special thanks to the developers of the evaluation tool.

References

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

[2] P. Rajpurkar *et al.*, "CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning," 2017.

[3] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," 2014.