Using machine learning to detect hanging errors in canine thoracic radiographs

by

Ryan Hennessey

B.S., Pennsylvania State University, 2008

A REPORT

submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Department of Computer Science
Carl R. Ice College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2022

Approved by:

Major Professor
Dr. Lior Shamir

Copyright

Abstract

In medical diagnostic imaging, "hanging" refers to the orientation of a radiographic image (X-ray) made available in the viewing software for interpretation by the radiologist. Industry hanging standards exist to ensure images of a body region are always presented the same way to a radiologist. These hanging standards eliminate the need for radiologists to rotate and flip the images which improves efficiency and accuracy in image interpretation through more effective pattern recognition and reduced distractions. Despite these standards, hanging errors persist in radiology, primarily due to human error in the acquisition process. Machine learning is in its infancy in veterinary medicine, and thus far has been applied to various image interpretation tasks in the field such as detection of pulmonary patterns (Boissady E. , Comble, Zhu, Abbott, & Adrien-Maxence, 2021). However, application of machine learning to veterinary diagnostic image quality control tasks such as image hanging has not yet been investigated.

This investigation employed multiple image preprocessing steps and architectures both pretrained and not, such as SVM using Eigen-Radiographs derived through PCA, EfficientNetB0, AlexNet and a Minimal 6-Layer Self-built Neural Network, in order to find learners that could correctly classify hanging errors in canine ventrodorsal thoracic radiographs with the goal of ultimate application in the imaging workflow process. Hanging errors were introduced into the images and include horizontal flip and rotation (90, 180 and 270 degrees) of the images, representing all possible errors encountered in clinical practice. Radiographic images were acquired and divided into training (800), test (240), and validation (208) sets. The architectures were tested on their ability to identify flip, rotation, and flip and rotation combined. Nearly all models, including non-neural networks, were able to distinguish rotation alone with 80% or greater accuracy. Flip and flip plus rotate were more challenging. The best performing model was a stacked model that categorized rotation, corrected the image, and then categorized flip to achieve an accuracy of 94% on the test set. This investigation serves as an important pilot study into the applicability of machine learning in veterinary diagnostic imaging workflow and quality control, with numerous additional applications warranting future investigation.

Table of Contents

List of Figures

List of Tables

Acknowledgements

To Drs. Erin Hennessey and Matt DiFazio of the KSU Veterinary Health Center Radiology Department for their assistance in collecting the radiographs and providing expert veterinary input.

Introduction

This report details work done to reduce radiographic image hanging error through machine learning detection. Correction of hanging errors was the focus of this study and all images were canine ventrodorsal thoracic radiographs (Figure 1). The dataset was collected from radiographic studies completed at Kansas State University Veterinary Health Center. The machine learning algorithms consisted of a Support Vector Machine using Eigen-Radiographs and numerous neural networks.

Figure 1. Radiograph examples from the dataset (from left to right: image hung correctly, image flipped and rotated 90 degrees, image flipped horizontally)

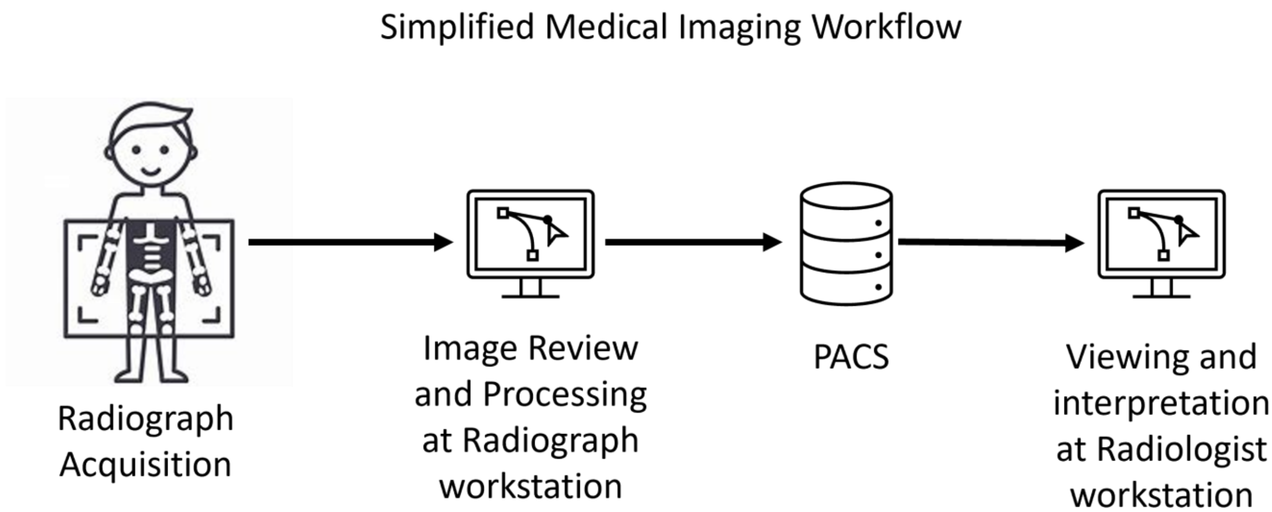
In human medicine, medical errors account for an estimated 250,000 deaths a year across all medical fields. In radiology the error rate is estimated at 4% (interpretive error in radiology) and correct image labeling is identified as one source of such errors and is a Key Safety Metric. Not only are incorrectly labeled images an issue of safety but they interrupt radiologist workflow and are a source of decreased study reading efficiency. Although image labeling errors are not identical to hanging errors, they can have some of the same deleterious outcomes if the hanging error is not detected by the radiologist (Halstead & Froehle, 2008) (Johnson, Krecke, Miranda, Roberts, & Denham, 2009).

Background

Medical imaging has come a long way since its introduction more than 100 years ago when radiographs were acquired with film screen technology. The advent of digital imaging, picture archiving and communication system (PACS) servers, and electronic viewing and reporting software that integrates with the hospital's radiology information system (RIS) and hospital information system (HIS) has vastly improved both image quality and throughput in the imaging workflow with one study showing up to a 42% reduction in reading times (Halstead & Froehle, 2008).

Briefly, patients will enter the imaging room and be positioned on a table. For the purposes of this study, dogs were positioned lying on their back with the detector plate beneath the table and radiation-producing tube above the table so that the beam passed through them in a ventral to dorsal direction. After exposure, images are initial reviewed at the radiograph workstation. Technologists assess for positioning, hanging, labeling, and exposure errors and ensure all necessary views are obtained, then send the images to the PACS server. Radiologists at their viewing stations then access the images on the server using viewing and reporting software.

Figure 2. Medical Imaging Workflow



Process improvement has already had far reaching effects on patient care quality. The use of artificial intelligence to further improve processes would affect the following areas: include correct image labeling listed as a Key Safety Metric, access and waiting times listed as Key Process Metrics and with further development the reducing in errors would reduce the radiation dose listed as a Method to Assess Professional outcomes by reducing the need to retake images (Johnson, Krecke, Miranda, Roberts, & Denham, 2009).

Methodology

Data Collection

All images used in this study were collected from the Kansas State University Veterinary Health Center. The original images were all in the DICOM (Digital Imaging and Communications in Medicine) format which is a standard in medical imaging. The images were of varying size. The images in the study contained 16-bits available for each pixel but by DICOM standard only allocate the needed bits per pixels. In this case that was 14 or 15 bits per pixel depending on the image with the unallocated bits remaining unused. (NEMA PS3 / ISO 12052)

All collected images were initially correctly hung and all hanging errors were induced. The collection of mis-hung images would be prohibitive due to the length of time it would take to collect incorrectly hung initial images and also to ensure collection before the mis-hanging is corrected.

Image Processing

Images were initially processed into multiple different end formats. The processing output levels were 512x512x1 TIFF images in 16bit floating point format, a 224x224x3 JPEG, and 224x224x3 cropped JPEG. Each of these formats was meant to address a potential concern. The TIFF format was meant to address the concern of loss of information and contrast however, most pretrained models are trained on JPEGs and thus single channel TIFF images are incompatible. The standard JPEG was both an easy format to convert to and has a high compatibility with pretrained algorithms. It is also possible to export images directly as JPEGs (rather than DICOM) from most viewing software. The images were also cropped to reduce superfluous information and ensure only bodily factors were considered by the algorithms. It is, however, important to note that metallic laterality markers (denoting right and left sides of the patient) were retained in many images since these markers are placed close to the patient during image acquisition. In the end the TIFF images were dropped from the study due to their incompatibility with all pretrained algorithms and processing them into the correct format would effectively be the same as the JPEG images already in use.

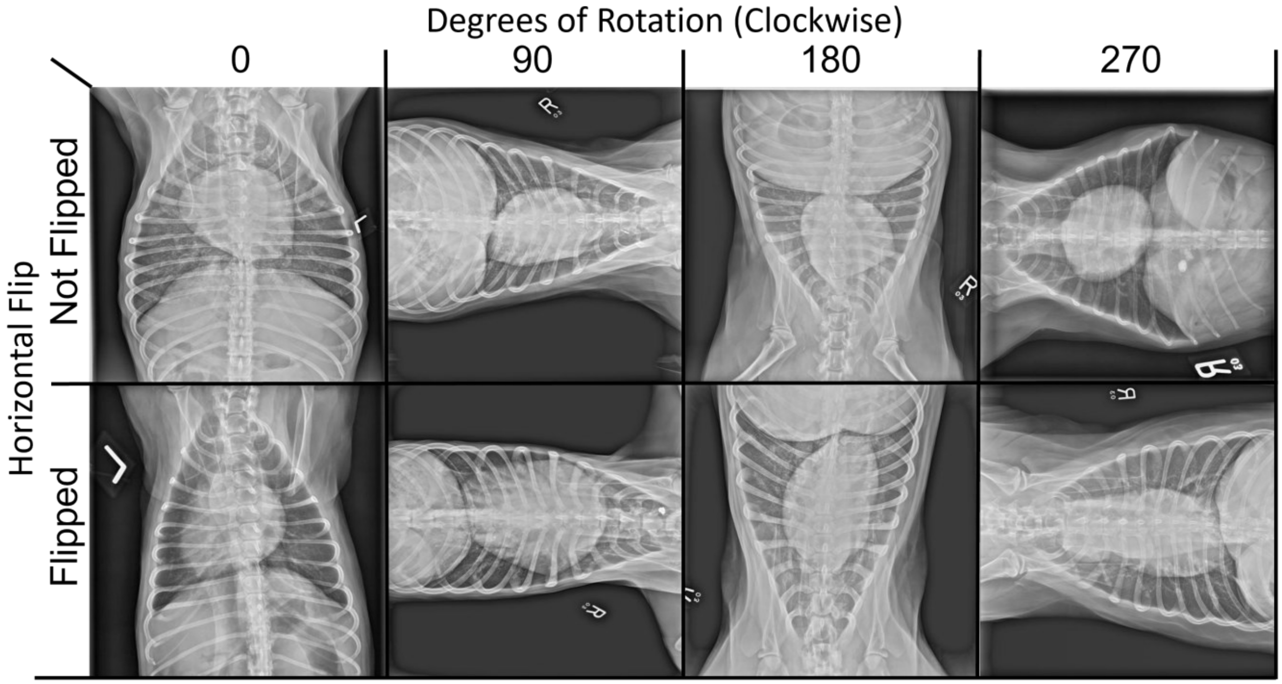
The data was split into three sets: training (800), test (240), and validation (208) sets. This represents an 80/20 split to produce training/validation set and a test set and then another 80/20 split to produce a separate training and validation set from the training/validation set.

Ground Truth Labels

Ground truth labels were initially applied by veterinary radiology residents to manually flipped and rotated images. After refinement of pre-processing steps, only properly hung images were acquired and the hanging errors were introduced programmatically to rotate and flip images into all possible derivations with ground truth labels applied during this process.

The eight classes constitute the eight possible hanging errors of rotation and flip and are pictured below (Figure 3). Images are rotated clockwise 0, 90, 180 and 270 degrees as rotations that are not multiples of 90 degrees would be due to poor alignment of the patient at the time of image acquisition. Zero degrees rotation refers to the patient's head being oriented off the top of the image.   Throughout the study flip refers to a horizontal flipping of the image (about the y axis). Vertical flips were not assessed as a vertical flip is equivalent to a horizontal flip with a 180-degree rotation.

Figure 3. Example image from each class from the non-cropped dataset

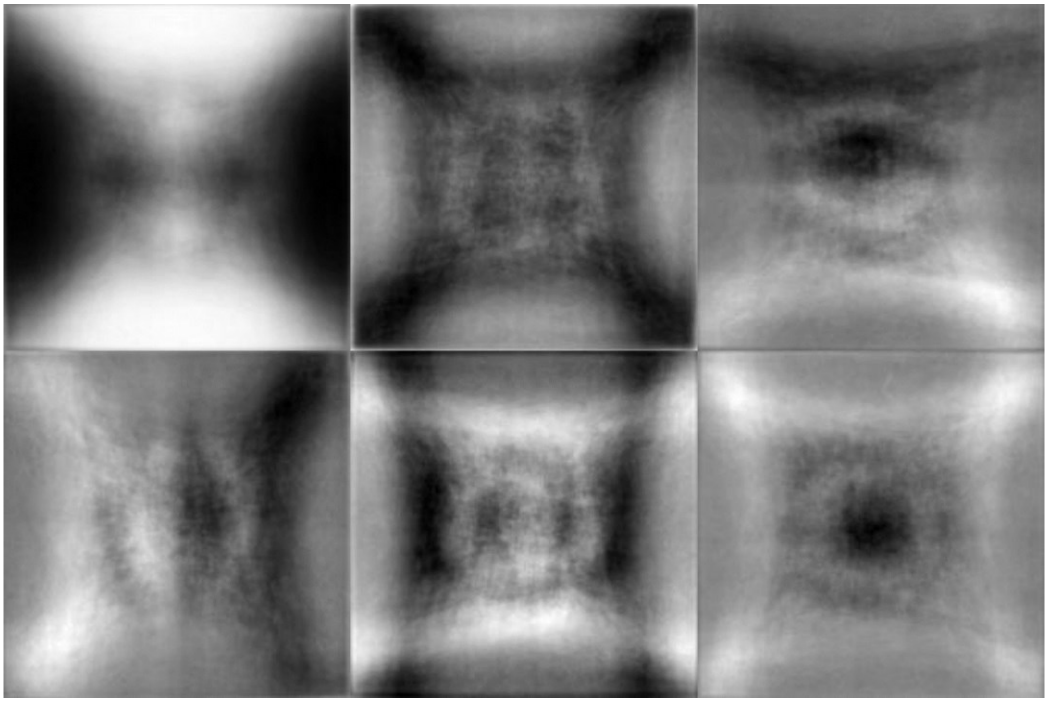


The Models

SVM with Eigen-Radiographs

Principal component analysis (PCA) was used to extract the top eigen-radiographs for each classification problem. For each problem, the PCA algorithm was provided with the features of the training set (with images this means the image is flattened into a 1-dimensional array and each pixel is treated as a feature). The PCA algorithm then works to identify a given number of reduced features. These feature reductions can be shown as eigen-radiographs when the features are remapped to the original image size. These eigen-radiographs are show a reduction that provides the highest variability and thus can be relied upon to provide distinction between images. Training and validation were used to identify the optimal number of eigen-radiographs and a randomized search was used to identify the optimal support vector machine parameters. Once the parameters were identified the model could then be used to predict the test set.

Figure 4. Top 6 Eigen-Radiographs of non-cropped JPEG



The Neural Networks

The neural networks were all trained with a max of 20 epochs using early stopping based on validation loss. Once stopped, the neural network's weights from the epoch with the best validation loss were restored. All networks used softmax as a final activation function, an adam as an optimizer during training, a batch size of 32, and loss was calculated using categorical cross-entropy.

Ten neural networks were utilized for image evaluation including two non-pretrained models: a small-self built Minimal 6-Layer Model and AlexNet (Krizhevsky, Sutskever, & Hinton, 2017). The other eight models were pretrained on ImageNet. Their training was done in two steps. First an added flatten and dense layer were trained while the pretrained model weights were kept frozen. Then, after training the new layers, the whole model was unfrozen and a second training round commenced.

Determining Accuracy

Accuracy for all models was assessed based on a test set that was of 240 images uniformly dispersed among the 8 classes.

Determining Accuracy - All in One Models

For the all-in-one models that assessed the rotation and flip simultaneously and sorted each into one of eight classes, a singular model accuracy measure was calculated.

Determining Accuracy – Stacked Models

For the stacked models, three separate measures of accuracy were calculated.

Rotational prediction accuracy: Rotational classifier accuracy was assessed based on identifying the rotational orientation of the images both flipped and non-flipped.

Flip prediction accuracy: Flip classifier accuracy was assessed based on correctly identifying the flip orientation of the image with the image in the correct rotational orientation (all images for this predictive task were corrected to 0 degrees of rotation based on the true rotation of each image).

Overall prediction accuracy: Overall accuracy was assessed by having the rotational classifier provide predictions. These predictions were used to apply a correction to the rotation of each image. The images were corrected based on the rotational classifier's predictions. The rotationally corrected images were then passed into the flip classifier to obtain flip predictions.

The rotational and flip predictions were combined and then compared to the combined rotational and flip ground truth to provide the overall accuracy score.

Gradient-weighted Class Activation Mapping

In order to allow for insight into what the neural networks were using in their classification decision, Gradient-weighted class activation mapping (Grad CAM) was applied post training on a set of eight images (one of each classification of hanging error) with the exception of the Minimal-6 Layer Network as its lone convolutional layer did not provide any useful heatmapping overlays. Grad CAM allows for the construction of overlays which illustrate which parts of the image are important to the classification. This is important within the medical community as "Explainable AI" is very important to medical providers as they develop trust and confidence in a machine learning system. Providing this type of "explanation" assures users the networks reach the correct classification using the appropriate anatomic indicators versus some image feature that only correlates in the dataset but will not in the world at large.

Results

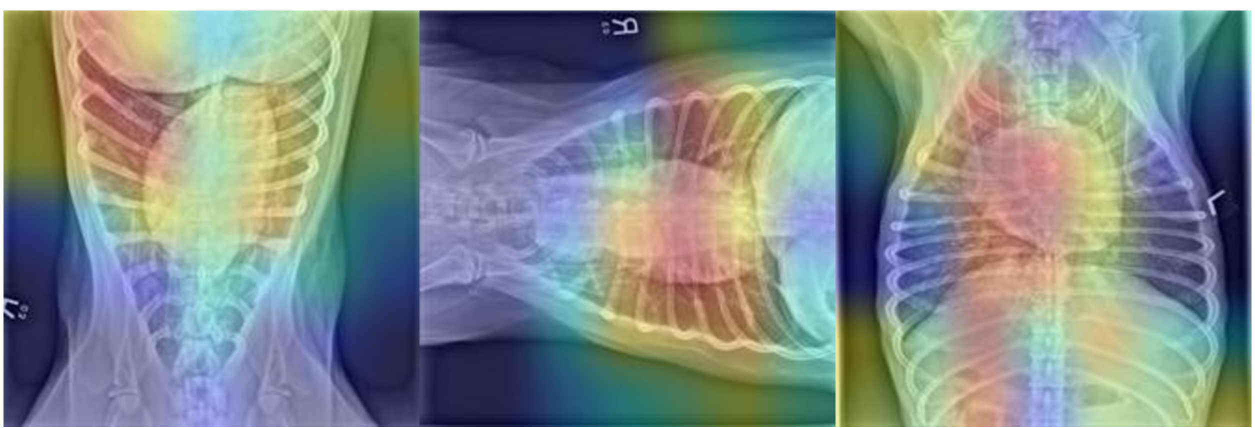
Below is a review of the results. The full table of results can be found in Appendix B.

All-in-One Eight Class Classification

For the first set of classifiers using the simultaneous all-in-one approach with eight separate classes, the classifier performance ranged from a low of 50.83% accuracy by SVC with PCA to a high of 91.67% for pretrained DenseNet121 using non-cropped jpg images.

The Grad CAM for results for DenseNet (figure 5) show that the classifier consistently focused on areas of the radiographs containing the heart and often on the diaphragm, which are two of the major anatomic features used by radiologists to determine whether an image is hung correctly (flipped or not flipped).

Figure 5. Grad CAM results from DenseNet combined predictor on uncropped images All images predicted correctly (Truth from left to right: Flipped_180, Flipped_270, Not Flipped & Not Rotated)



Stacked Models

Rotational Classification

Rotational classification by itself proved the easiest task with the Minimal 5-Layer Network performing the worst at 82.50% and with multiple models reaching accuracy of 99.58%.

Flip Classification

As with rotational classification, the worst performing flip classifier was the Minimal 5-Layer Network with an accuracy of 58.33% and the best performing network being ResNet with an accuracy of 94.17% on cropped JPEGs.

Combined Classification

As it had the two worst separate accuracies the Minimal 5-Layer Network also had the worst combined accuracy at 49.58% and the best combined accuracy was achieved by ResNet50 at 93.75% on regular jpg.

Discussion

The ability to predict and correct hanging and other types of errors before they reach a radiologist will save time and reduce interpretation errors Additionally, integration of hanging error detection and correction could easily be added to the radiologic workflow. Such detection and correction could occur at the radiograph workstation within the acquisition software or at the radiologist workstation in the viewing software.

The use of Grad CAM led to some interesting findings within certain classifiers. In AlexNet, which had a middling accuracy (84.58%) when classifying flips on cropped image, the classifier seemed to hone-in a the right side of each flipped image (patient's left) outside the bounds of the patient's body (Figure 6). This suggests there is a consistent feature on the patient's left during acquisition that allows for a determination of flipped images with decent accuracy. Based on the pattern of activation, a feature of the detector plate such a dead pixel or bias in patient positioning is suspected to be the cause. This area of the image is not something that will be scrutinized by a radiologist, thus highlighting the unique ability of machine learning to see beyond the pattern typically recognized by humans. Further evaluation of these images to determine the cause as well as collection and integration of images from multiple other hospitals would help to train out machine or hospital specific idiosyncrasies and make the classifier more robust.

Figure 6. Grad CAM results from AlexNet flip predictor on cropped images All images predicted Flip (Truth from left to right: Flipped, Flipped, Not Flipped)

While the use of Grad CAM is highly valuable in determining how the classifiers arrive at their predictions Grad CAMs, weaknesses were apparent when in some images no gradient heat map was produced. There has also been research that shows that Grad CAM is prone to highlighting areas that may not actually be important. Future applications should consider the use of HiRes CAM to mitigate this issue (Draelos & Carin).

The success of using SVM after PCA on the images shows that at least for rotational image correction large neural networks may be unnecessarily complex and over-capable since differences between rotated images are quite extensive and very easy to detect. Given these findings, it would be worth pursuing other non-neural network classification models and also a non-AI based solution for rotational radiographic hanging errors. As the cropping algorithm did not rely on AI and was able to successfully crop images regardless of orientation it stands to reason that the programming path followed there might offer some clues as to how to differentiate between the rotational classes.

Detection of horizontal flip-type errors proved much more challenging, and most likely warrants the use of neural networks. When radiologists assess for this error they look at subtle anatomic features such as the margins of the diaphragm and direction in which certain parts of the heart and vessels point. This type of assessment is more akin to image interpretation which has been more thoroughly investigated using neural networks in veterinary medicine (Boissady E. , Comble, Zhu, & Adrien-Maxence, 2020) (Boissady E. , Comble, Zhu, Abbott, & Adrien-Maxence, 2021) (McEvoy & Amigo, 2013) (Biercher, et al., 2021).

A limitation of this investigation is the use of radiographs acquired at an academic teaching hospital. Due to the level of expertise in such a facility, radiographs tend to be of a higher quality with fewer hanging, positioning, and exposure errors. Fortunately hanging errors could be induced in this study. This could be a limitation in further investigation into the use of machine learning in quality control tasks such as labeling (placement of metallic right/left marker) and positioning as images with errors will be difficult to obtain. Collection of a more robust data set from other hospitals and teleradiology sources is suggested to partially mitigate these concerns. Additionally, a more robust dataset will further enhance difference in performance of the tested classifiers.

Conclusion

The findings here act as a proof of the potential of machine learning to improve medical imaging workflow. Research such as this can lead to mutually beneficial improvement for all stakeholders in the imaging field. Hospitals can save time and money by reducing medical imaging errors and increasing patient throughput. Radiologists will benefit by increased individual productivity and efficiency. Additional technologies such as computer aided detection and diagnosis can also improve radiologist accuracy and performance. The author believes implementation of AI in quality control is an overlooked application with much potential benefit. Furthermore, this type of AI application is likely to be more palatable to the medical community as many members have varying levels of concern about AI replacing human jobs. Finally, and perhaps most importantly, radiation dose to imaged patients can be reduced by eliminating the number of re-take radiographs required.

This study should only act as a first step in the development of AI assistance systems which work to improve medical imaging quality control at the source, ensuring that the best images are captured the first time. Immediate follow up investigations may include use AI to detect labeling errors as well as incomplete studies lacking the necessary collection of images. More robust future applications could focus on real time systems monitoring and providing real-time feedback recommendations during rather than after image acquisition. For example, alerting the technologist when the entirety of a body part is not included in the imaging field.

In conclusion, AI is under utilized in veterinary medicine, specifically in diagnostic imaging. Investigations to date have focused on image interpretation rather than quality control, leaving much more for further investigation and application of AI technology in this field.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., . . . et al. (2015). Large-Scale Machine Learning on Heterogeneous Distributed Systems. doi:https://doi.org/10.5281/zenodo.472412

Biercher, A., Meller, S., Wendt, J., N, C., Schmidt-Mosig, J., de Decker, S., & Volk, H. A. (2021). Using Deep Learning to Detect Spinal Cord Diseases on Thoracolumbar Magnetic Resonance Images of Dogs. *8*. doi:doi: 10.3389/fvets.2021.721167

Boissady, E., Comble, A. d., Zhu, X., & Adrien-Maxence, H. (2020). Artificial intelligence evaluating primary thoracic lesions has an overall lower error rate compared to veterinarians or veterinarians in conjunction with the artificial intelligence. *Veterinary Radiology & Ultrasound, 61*, 619-627. doi:DOI:10.1111/vru.12912

Boissady, E., Comble, A. d., Zhu, X., Abbott, J., & Adrien-Maxence, H. (2021). Comparison of a Deep Learning Algorithm vs. Humans for Vertebral Heart Scale Measurements in Cats and Dogs Shows a High Degree of Agreement Among Readers. *Frontiers in Veterinary Science*. doi:https://doi.org/10.3389/fvets.2021.764570

Burti, S., V, L. O., Zotti, A., & T., B. (2020). Use of deep learning to detect cardiomegaly on thoracic radiographs in dogs. *The Veterinary Journal, 262*. doi:https://doi.org/10.1016/j.tvjl.2020.105505

Chollet, F. (n.d.). Keras. GitHub Repository. Retrieved from https://github.com/fchollet/keras

Draelos, R. L., & Carin, L. (n.d.). Use HiResCAM instead of Grad-CAM for faithful explanations of convolutional neural networks. doi:https://doi.org/10.48550/arXiv.2011.08891

Halstead, M. J., & Froehle, C. M. (2008). Design, Implementation, and Assessment of a Radiology Workflow Management System. *American Journal of Roentgenology, 191(2)*, 321-327. doi:https://doi.org/10.2214/ajr.07.3122

Johnson, C. D., Krecke, K. N., Miranda, R., Roberts, C. C., & Denham, C. (2009). Developing a Radiology Quality and Safety Program: A Primer. *RadioGraphics, 29(4)*, 951-959. doi:https://doi.org/10.1148/rg.294095006

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM, 60(6)*, 84-90. doi:https://doi.org/10.1145/3065386

Küstner, T., Gatidis, S., Liebgott, A., Schwartz, M., Mauch, L., Martirosian, P., . . . Schick, F. (2018). A machine-learning framework for automatic reference-free quality assessment in MRI. *Magnetic Resonance Imaging, 53*, 134-147. doi:https://doi.org/10.1016/j.mri.2018.07.003

Lakhani, P., Prater, A. B., Hutson, R. K., Andriole, K. P., Dreyer, K. J., Morey, J., . . . Hawkins, C. M. (2018). Machine Learning in Radiology: Applications Beyond Image Interpretation. *Journal of the American College of Radiology, 15(2)*, 350-359. doi:https://doi.org/10.1016/j.jacr.2017.09.044

Liew, C. (2018). The future of radiology augmented with Artificial Intelligence: A strategy for success. *European Journal of Radiology, 102*, 152-156. doi:https://doi.org/10.1016/j.ejrad.2018.03.019

Lustgarten, J. L., Zehnder, A., Shipman, W., Gancher, E., & Webb, T. L. (2020). Veterinary informatics: forging the future between veterinary medicine, human medicine, and One Health initiatives—a joint paper by the Association for Veterinary Informatics (AVI) and the CTSA One Health Alliance (COHA). *JAMIA Open, 3(2)*, 306-317. doi:https://doi.org/10.1093/jamiaopen/ooaa005

Mason, D., & Lemaitre, G. (2021). PyDicom v2.2.2. Retrieved from https://github.com/pydicom/pydicom

McEvoy, F. J., & Amigo, J. M. (2013). Using Machine Learning To Classify Image Features from Canine Pelvic Radiographs: Evaluation of Partial Least Squares Discriminant Analysis and Artificial Neural Network Models. *Veterinary Radiology & Ultrasound, 54(2)*, 122-126.

NEMA PS3 / ISO 12052. (n.d.). Digital Imaging and Communications in Medicine (DICOM) Standard,. Rosslyn, VA, USA: National Electrical Manufacturers. Retrieved from (available free at http://www.dicomstandard.org/)

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, E. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research, 12*, 2825-2830.

Schemmel, A., Lee, M., Hanley, T., Pooler, B. D., Kennedy, T., Field, A., . . . Yu, J. P. (2016). Radiology Workflow Disruptors: A Detailed Analysis. *Journal of the American College of Radiology, 13(10)*, 1210-1214. doi:https://doi.org/10.1016/j.jacr.2016.04.009

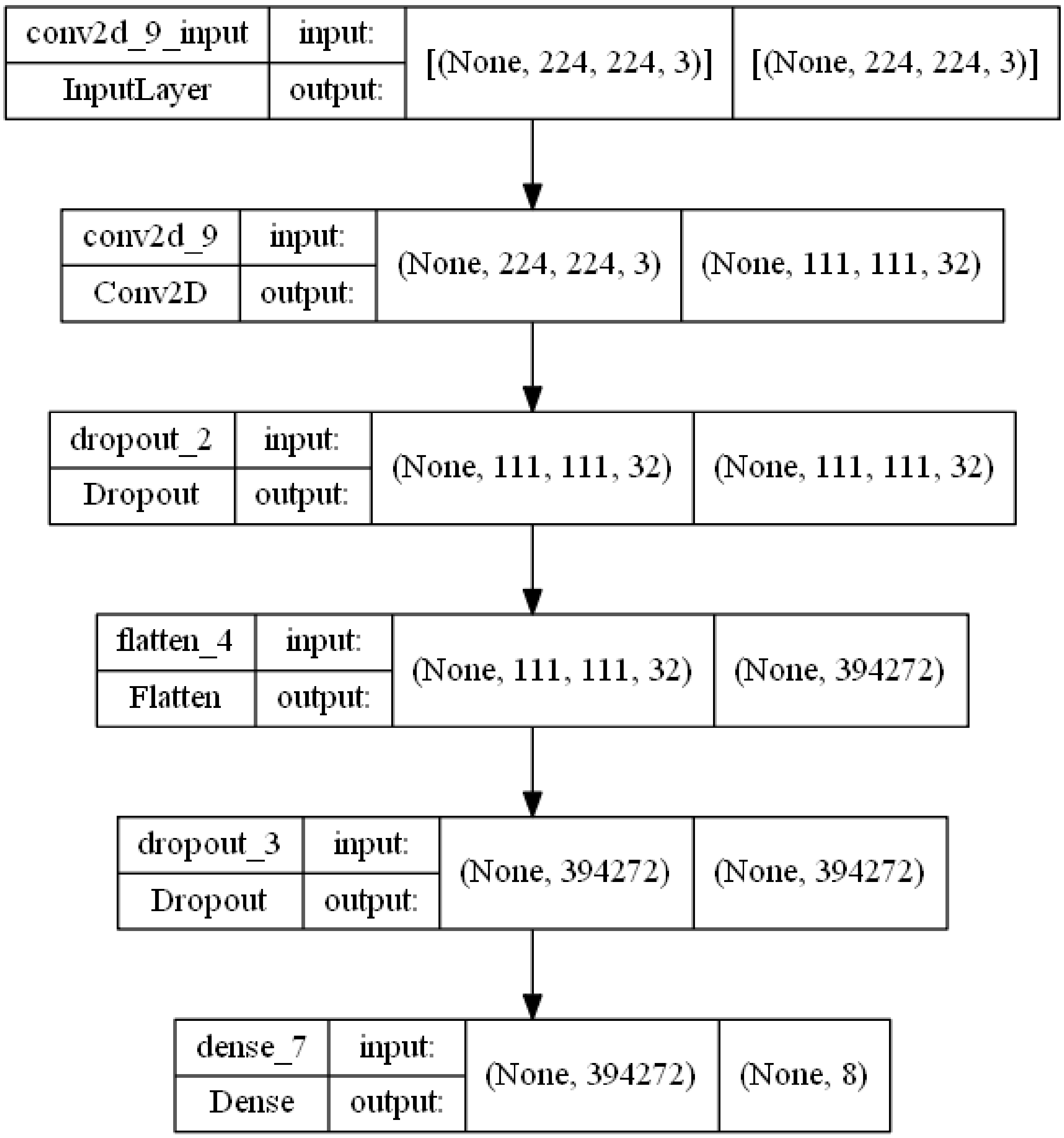
Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2019). Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *International Journal of Computer Vision, 128(2)*, 336-359. doi:https://doi.org/10.1007/s11263-019-01228-7

Turk, M. A., & A, P. (1991). Face Recognition using Eigenfaces. . *In Proc. of Computer Vision and Pattern*, 586-591.

Waite, S., Scott, J., Gale, B., Fuchs, T., Kolla, S., & Reede, D. (2017). Interpretive Error in Radiology. *American Journal of Roentgenology, 208(4)*, 739-749. doi:https://doi.org/10.2214/ajr.16.16963

Appendix A - Minimal 6 Layer Architecture

Appendix A Figure 7 – Minimal 6-Layer Neural Network Architecture



Appendix B - Model Accuracy Tables

Appendix B Table 1 – All-in-One Model Classification Accuracy

(Green and Red highlighting denote three highest and lowest accuracy ratings)

Classifier	Image Type	Flip and Rotational Accuracy
EfficientNetB0	jpg	88.75%
	Cropped jpg	83.75%
AlexNet	jpg	77.92%
	Cropped jpg	53.75%
Xception	jpg	77.92%
	Cropped jpg	69.17%
MobileNetV3Small	jpg	76.25%
	Cropped jpg	79.58%
Inception V3	jpg	77.50%
	Cropped jpg	62.08%
SVM with PCA	jpg	50.83%
	Cropped jpg	66.66%
ResNet50	jpg	86.67%
	Cropped jpg	85.42%
VGG16	jpg	87.50%
	Cropped jpg	87.92%
DenseNet121	jpg	91.67%
	Cropped jpg	79.85%
NasNetMobile	jpg	76.25%
	Cropped jpg	67.92%
Minimal 6-Layers	jpg	54.17%
	Cropped jpg	45.83%

Appendix B Table 2 - Stacked Model Classification Accuracy

(Green and Red highlighting denote three highest and lowest accuracy ratings)

Classifier	Image Type	Rotational Accuracy	Flip Accuracy	Total Accuracy
EfficientNetB0	jpg	99.58%	93.33%	93.34%
	Cropped jpg	99.58%	90.42%	90.00%
AlexNet	jpg	90.42%	72.08%	65.42%
	Cropped jpg	98.33%	84.58%	83.75%
Xception	jpg	99.58%	90.42%	90.00%
	Cropped jpg	99.58%	82.92%	82.92%
MobileNetV3Small	jpg	99.58%	91.25%	90.83%
	Cropped jpg	99.58%	93.33%	93.34%
Inception V3	jpg	99.17%	83.33%	82.92%
	Cropped jpg	99.17%	82.08%	81.25%
SVM with PCA	jpg	96.67%	62.92%	61.25%
	Cropped jpg	99.16%	79.16%	78.75%
ResNet50	jpg	99.58%	93.75%	93.75%
	Cropped jpg	99.17%	94.17%	93.34%
VGG16	jpg	99.17%	93.33%	92.92%
	Cropped jpg	99.58%	94.58%	94.16%
DenseNet121	jpg	99.58%	94.58%	94.16%
	Cropped jpg	99.58%	91.67%	91.67%
NasNetMobile	jpg	99.58%	87.92%	87.50%
	Cropped jpg	99.17%	70.00%	69.53%
Minimal 6-Layers	jpg	82.50%	58.33%	49.58%
	Cropped jpg	87.92%	46.67%	40.83%

Glossary of Terms

PACS – Picture Archiving and Communications System – a system for storing and accessing images captured by the various imaging modalities used in the medical community.

DICOM - Digital Imaging and Communications in Medicine – The international standard medical imaging format for all modalities as outlined in the ISO 12052 standard.

Grad CAM - Gradient-weighted Class Activation Mapping- A method for creating overlays to show the areas of an image which were important for the class prediction provided by a Neural Network.

RIS - Radiology Information System – The base system for imaging departments that handles scheduling, examination performance tracking, reporting, distribution of results, etc.

HIS - Hospital Information System- Hospital computer system that deals with the administrative needs of the hospital.