

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

# Machine Learning with Applications

journal homepage: [www.elsevier.com/locate/mlwa](http://www.elsevier.com/locate/mlwa)

## Detecting pulmonary Coccidioidomycosis with deep convolutional neural networks

Jordan Ott <sup>a</sup>, David Bruyette <sup>b,\*</sup>, Cody Arbuckle <sup>b</sup>, Dylan Balsz <sup>b</sup>, Silke Hecht <sup>c</sup>, Lisa Shubitz <sup>d</sup>, Pierre Baldi <sup>a,e,\*</sup>

<sup>a</sup> Department of Computer Science, University of California, Irvine, USA

<sup>b</sup> Anivive Lifesciences, Inc., Long Beach, CA, USA

<sup>c</sup> College of Veterinary Medicine, University of Tennessee, Knoxville, TN, USA

<sup>d</sup> Valley Fever Center for Excellence, The University of Arizona, Tucson AZ, USA

<sup>e</sup> Institute for Genomics and Bioinformatics, University of California, Irvine, USA



### ABSTRACT

Coccidioidomycosis is the most common systemic mycosis in dogs in the southwestern United States. With warming climates, affected areas and number of cases are expected to increase in the coming years, escalating also the chances of transmission to humans. As a result, developing methods for automating the detection of the disease is important, as this will help doctors and veterinarians more easily identify and diagnose positive cases. We apply machine learning models to provide accurate and interpretable predictions of Coccidioidomycosis. We assemble a set of radiographic images and use it to train and test state-of-the-art convolutional neural networks to detect Coccidioidomycosis. These methods are relatively inexpensive to train and very fast at inference time. We demonstrate the successful application of this approach to detect the disease with an Area Under the Curve (AUC) above 0.99 using 10-fold cross-validation. We also use the classification model to identify regions of interest and localize the disease in the radiographic images, as illustrated through visual heatmaps. This proof-of-concept study establishes the feasibility of very accurate and rapid automated detection of Valley Fever in radiographic images.

### 1. Introduction

Coccidioidomycosis, or “Valley Fever”, is a disease found in dogs with the ability to infect humans. Currently found in the southwestern United States, the disease is expected to spread with warming climates in the coming years. The disease can be detected through radiological examination of the lungs. In order to aid in the identification and diagnosis of the disease we leverage deep learning classification models. These models have been applied to medical imaging especially in human clinical settings. In this paper we apply deep learning models to veterinary medicine and radiology images of healthy and infected dogs.

#### 1.1. Background on valley fever

The disease results from infection by the dimorphic, saprophytic fungal organism *Coccidioides immitis* or *Coccidioides posadasii*. Infection by *Coccidioides* spp. occurs in a wide range of host species, including humans, a myriad of domestic and exotic mammals, and,

rarely, reptiles (Shubitz, 2007). Coccidioidomycosis is endemic in semi-arid regions of California, Arizona, New Mexico, Texas, and Northern Mexico and semiarid regions in South America. Across this region, temperature and precipitation influence the extent of the endemic region and number of Valley Fever cases. Climate projections for the western US indicate temperatures will increase and precipitation patterns will shift, which may alter disease dynamics. Using a climate niche model derived from contemporary climate and disease incidence data with projections of climate from Earth system models, researchers have assessed how endemic areas will change during the 21st century. By 2100 in a high warming scenario, the model predicts the area of climate-limited endemicity will more than double, the number of affected states will increase from 12 to 17, and the number of Valley fever cases will increase by 50%. The Valley Fever endemic region will expand north into dry western states, including Idaho, Wyoming, Montana, Nebraska, South Dakota, and North Dakota. Precipitation will limit the disease from spreading into states farther east and along the central and northern Pacific coast (Gorris et al., 2019)

The code (and data) in this article has been certified as Reproducible by Code Ocean: (<https://codeocean.com/>). More information on the Reproducibility Badge Initiative is available at <https://www.elsevier.com/physical-sciences-and-engineering/computer-science/journals>.

\* Corresponding authors.

E-mail addresses: [jott1@uci.edu](mailto:jott1@uci.edu) (J. Ott), [david@anivive.com](mailto:david@anivive.com) (D. Bruyette), [cody@anivive.com](mailto:cody@anivive.com) (C. Arbuckle), [dylan@anivive.com](mailto:dylan@anivive.com) (D. Balsz), [shecht@utk.edu](mailto:shecht@utk.edu) (S. Hecht), [lshubitz@email.arizona.edu](mailto:lshubitz@email.arizona.edu) (L. Shubitz), [pbaldi@ics.uci.edu](mailto:pbaldi@ics.uci.edu) (P. Baldi).

<https://doi.org/10.1016/j.mlwa.2021.100040>

Received 19 January 2021; Received in revised form 12 May 2021; Accepted 14 May 2021

Available online 25 May 2021

2666-8270/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1.2. Incidence of valley fever

Inhalation is the most common route of infection in both animals and humans and typically occurs after fungal hyphae have desiccated and matured into arthroconidia that are easily aerosolized. The arthrospores are inhaled and dispersed along the bronchial tree. Here, they undergo structural transformation into spherules, which enlarge and undergo endospore formation. Eventually, new endospores are released into the surrounding tissue and the cycle continues until the host is able to mount an appropriate immune response. Direct cutaneous inoculation of arthrospores has been reported, but these instances are rare and typically result in local granuloma formation (Tortorano et al., 2015)

Infections caused by *Coccidioides* may be subclinical or result in severe illness and death. Dissemination to extrapulmonary organs is a potential sequela in some animals. One study indicated that 80% of dogs develop primary pulmonary infection whereas 20% develop disseminated disease (Davidson & Pappagianis, 1996). Pulmonary Coccidioidomycosis is characterized by clinical signs that include chronic cough, lethargy, and respiratory distress. Radiographic pulmonary lesions may be characterized by interstitial to nodular patterns, with hilar lymphadenopathy (Mehrkens et al., 2016).

Human infection has steadily increased over the last 10 years, and Coccidioidomycosis is considered a reemerging infectious disease. Dogs come into close contact with soil, because of a closer proximity to the ground and digging behavior, and may provide a clearer delineation of broadly defined endemic regions. In this sense, rapid recognition of *Coccidioides* infection in dogs in a clinical setting may benefit human public health as well as minimize morbidity and owner expense.

## 1.3. Deep learning for biomedical imaging

This study can be categorized at three different levels. First, it is an application of deep learning to biomedical imaging, a rapidly expanding field. Second, the proposed method targets a pulmonary disease, narrowing the application to a particular organ. And third is the most distinctive feature of this work as an application of deep learning to veterinary science, where the literature is still very sparse. This categorization serves to connect the current work to the larger field of biomedical imaging while calling attention to its novelty and importance for veterinary medicine.

Biomedical imaging is a large domain of application of deep learning, a rebranding of neural networks, with hundreds of articles already published in the literature (Baldi, 2021). Recent examples from our group alone include applications to colonoscopy screening (Urban et al., 2018), cardiovascular disease detection (Wang et al., 2017a, 2017b), spinal metastasis detection (Wang et al., 2017a, 2017b), genetic mutation classification (Chang et al., 2018), and counting hair follicles (Urban et al., 2021). Many more applications can be found in the literature, including segmentation of retinal blood vessels (Liskowski & Krawiec, 2016), skin cancer detection and classification (Esteve et al., 2017), as well as pediatric bone assessment (Igloukov et al., 2018) just to name a few. Primarily advances in computing power, in particular Graphical Processing Units (GPUs), together with the growing availability of labeled datasets and robust software libraries (e.g. PyTorch, TensorFlow, Keras), are fueling the expansion of deep learning methods in biomedical imaging applications.

Within this expansion, there has been work specifically targeted at pulmonary diseases. For example, convolutional neural networks have been used to evaluate thoracic radiographs in man (Tang et al., 2018), and in lung cancer detection and screening applications (Ardila et al., 2019; Coudray et al., 2018; Hosny et al., 2018).

There has been considerably less work in applying deep learning methods to veterinary medicine in general, and thoracic radiographs in particular. These applications come with their own challenges, not the least of which is the potential for greater variability in anatomy and

dimensions, for instance across dog species. There has been recent work applying deep learning algorithms to classify red blood cell morphology in canines (Pasupa et al., 2020). Likewise, there has been work classifying mammary tumors in canines and humans (Kumar et al., 2020). The current study introduces a novel application of deep learning to thoracic images in order to identify Coccidioidomycosis.

In this study, we apply deep convolutional neural networks in the evaluation of a single well-defined medical condition (pulmonary Coccidioidomycosis) in dogs.

## 2. Materials and methods

### 2.1. Radiographs of normal and abnormal dogs

The challenge study was performed at Colorado State University under an Institutional Animal Care and Use Committee (IACUC) approval. The normal dogs consisted of two groups. One were purpose bred laboratory dogs used as the control group in the development of a challenge model of canine coccidiomycosis. The dogs were equally distributed between males and females and were all of the same age (approximately twelve months) and weight. The second group consisted of normal, healthy pet dogs of various ages and weights with thoracic radiographs reviewed by a veterinary board-certified radiologist.

The diseased dogs consisted of two groups. One were purpose bred laboratory dogs that were infected with varying doses of live spores of *Coccidioides posadasii* as part of the aforementioned challenge study. The dogs were all of the same age (approximately twelve months) and weight. Infection was confirmed by histopathology and fungal lung and lymph node cultures. The second group consisted of pet dogs living in the endemic area with a confirmed diagnosis of Coccidioidomycosis. All cases were evaluated at a single veterinary specialty hospital in Phoenix, Arizona.

The radiographic images were captured using three different commercially available veterinary digital imaging systems (Sound technology DR System 1417E; Philips Diagnost Digital RF Tri-mode 12"/9"/7"; Philips Diagnost Digital RF Tri-mode 15"/12"/9"). A brief description of each system can be found in the Supplementary Materials section. Recorded images were de-identified and given a unique id number indicating normal or diseased, prior to being evaluated in the study. Right lateral (RL), left lateral (LL), and ventrodorsal (VD) views were obtained and the images annotated with the dogs unique ID number, date and view. The dataset consists of 106 infected dogs and 240 healthy dogs. The digital captures of these subjects produced 1,186 total images, 12 of which were marked as irrelevant and as a result discarded. The discarded images captured other body regions such as the front or hind legs, and critically, did not include the lungs or chest region, making the identification of the disease impossible in these images. After filtering irrelevant images, the remaining images were of 769 healthy subjects and 405 infected subjects, for a total of 1,174 images. All images were resized to 500 by 400 pixels for computational efficiency and to produce uniform image sizes across the dataset. A constant image size is required for the classification models employed in this study, which are described in the following section.

All images underwent a manual preprocessing phase, where each image was subject to cropping by a human expert. This process ensured the removal of many spurious features such as textual and label markings. Examples of selected crops are shown in Fig. 1.

### 2.2. Neural networks

Fully-connected, feedforward neural networks (NNs) receive a single vector representing the raw data at the input layer. This input is then transformed through a series of weighted connections to hidden layers that perform non-linear operations, before being routed to an output layer for the purposes of classification, regression, and other tasks. In these architectures, each node in a layer is connected to

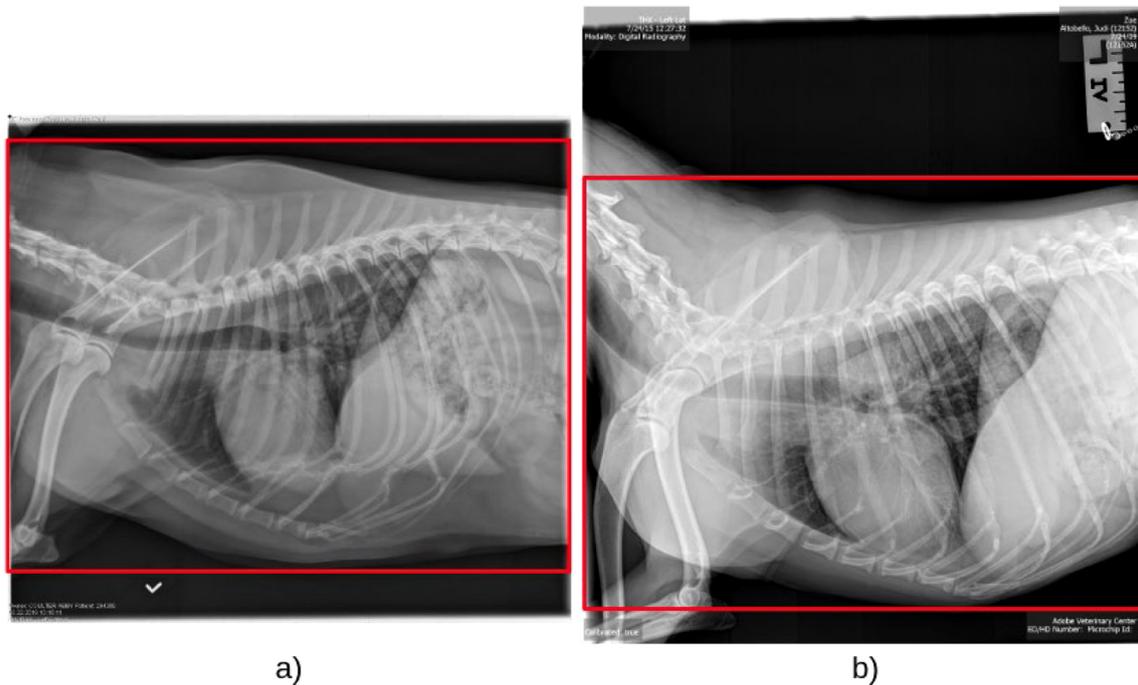


Fig. 1. Images of canine subjects. Red bounding boxes represent cropped regions during preprocessing. (a) Healthy subject; (b) Subject with Coccidioidomycosis.

every node in the following layer. Though powerful, fully-connected feedforward networks can require a large number of parameters and are not ideally suited for processing images and recognize features in ways that are invariant under translations and other deformations. For computer vision applications, most of the time Convolutional Neural Networks (CNNs) (Baldi & Chauvin, 1993; Fukushima & Miyake, 1982; LeCun et al., 1998) are a better choice. CNNs use weight sharing to reduce the number of parameters and convolve the image with arrays of local filters that are learnt from the data, providing a basis for building equivariance/invariance to translation and other geometric transformations in the overall response.

In this work five different CNN architectures are implemented, trained, and compared to classify the disease status of canine subjects: Inception (Szegedy et al., 2016), MobileNet (Sandler et al., 2018), ResNet (He et al., 2016), VGG (Simonyan & Zisserman, 2014), as well as a relatively shallow four-layer network. The first four of these architectures have been applied to a variety of image based problems in the literature.

### 3. Experiments

#### 3.1. Data augmentation

CNNs require large amounts of training data, specifically when the input space is large and the model has a high number of parameters. Insufficient amounts of data can lead the network to overfit, where it performs very well on the training set but fails to generalize properly to outside examples.

In order to combat overfitting, one basic approach is data augmentation where the original data is used to create additional training examples through various transformations. Data augmentation in this study was performed using rotations, translations, reflections, noise addition, and zooming. Each batch presented to the neural network during training was first normalized by color channel, by subtracting the mean value and dividing by the standard deviation. Then each normalized image was rotated by a random amount chosen uniformly in the 0–15 degrees range, translated left or right, and up or down, by a random amount chosen uniformly in the 0%–10% range of the original width and height, horizontally flipped, perturbed by adding Gaussian noise

with mean 0 and standard deviation 0.1 to each pixel, and cropped by zooming in on the center of the image by a random uniform amount in the 0%–10% range. Starting from the small dataset of 1,174 images, we randomly augmented each batch of training examples, effectively increasing the number of available examples by roughly two orders of magnitude.

#### 3.2. Neural network training

Of the architectures used in this study, Inception, MobileNet, ResNet, and VGG were selected based on their popularity and performance on benchmark datasets. The fifth model, the four layer shallow network, was selected to serve as a baseline architecture for comparison purposes. Inception, MobileNet, ResNet, and VGG have the option to use a set of weights that were previously trained on the ImageNet dataset (Deng et al., 2009). This process is known as transfer learning or pre-training, where weights trained on one task are then reused for a different task. When using the ImageNet weights, the primary convolutional layers can be frozen, and only new output layers are trained. The option of using pre-trained weights was treated as a hyperparameter to be optimized, and thus we trained both networks that were initialized randomly and networks that were initialized with pre-trained weights.

All networks were trained for 100 epochs using an early stopping condition if the validation accuracy did not improve in the most recent 15 epochs. All models were trained using 10-fold cross-validation. This process partitions the data evenly into 10 distinct subsets and each fold trains a model using 9 subsets and holds one subset out for testing purposes. The partitioning was stratified to ensure a constant ratio of representation amongst healthy and infected examples — roughly 66% healthy and 34% diseased, in keeping with the overall image ratio. As the classes are unbalanced the network may learn to favor the majority class thus increasing its accuracy simply by chance. Class weighted penalties in the loss function can alleviate this issue. The use of such losses are explored in the hyperparameter options in the next section. Ten-fold cross-validation requires each model to be trained ten distinct times (re-initializing network parameters each time) and ensures different subsets of the data are used for training and testing. The dataset was randomly shuffled and partitioned into 10 distinct folds

for cross-validation. During a given fold, images in the training set never appear in the validation set and images in the validation set are not present in the training set. All models were implemented in Keras with a Tensorflow backend.

### 3.3. Hyperparameter optimization

When implementing neural networks there are a number of choices – hyperparameters – that must be set prior to training. Choosing hyperparameters arbitrarily can lead to suboptimal results. To address this, we conducted neural network optimization via a random hyperparameter search using SHERPA (Hertel et al., 2020), a Python library for hyperparameter tuning. The random search algorithm has the advantage of making no assumptions about the structure of the hyperparameter search problem and is ideal for exploring a variety of settings.

We detail the hyperparameters of interest in Table 1, as well as the range of available options during the search. The hyperparameters of interest consisted of the amount of dropout, learning rate, learning rate decay, nodes per layer, the number of dense layers added at the end, a class weighted loss penalty, and the optimizer. As mentioned in the previous section, additional hyperparameters related to the use of pre-trained weights were also considered. The Inception, MobileNet, ResNet, and VGG networks had the option to use weights pre-trained on the ImageNet dataset or a standard random initialization. Finally, the option of training the base convolutional layers was tested as well. The use of ImageNet weights and training the base layers are not mutually exclusive. For example, one model may start with the pre-trained ImageNet weights, as an initialization, and training may proceed for both the top layers and the base layers.

Twenty-five different hyperparameter settings were tried for each network architecture, for a total of 125 models. Each trial was trained on a single split of the data, then following the hyperparameter search the best trials from each model were retrained in full using the 10 fold cross-validation method. The hyperparameters from the best performing models are displayed in the corresponding columns of Table 1.

## 4. Results

The five architectures examined in this study were each run with twenty-five different combinations of hyperparameters. The best hyperparameter configuration from each architecture is shown in Table 1. Following the hyperparameter search, the best performing variant (determined according to validation accuracy) of each architecture was selected for further analysis. The validation accuracies during training are shown in Fig. 2. All architectures demonstrated success reaching classification performance well above chance, represented by the dashed line. From the accuracy curves over time only MobileNet shows degradation at the later stages of training. The solid lines in Fig. 2 show the mean validation accuracies while the shaded region represents one standard deviation of the 10 separate cross-validation folds.

Next, we analyzed the relationship between the model's false positive and true positive rate in Fig. 3. All models perform well above chance, with the worst performing model, MobileNet, achieving a mean AUC of 0.92. Examining the 10 fold cross-validation results shows the ResNet architecture as the dominant model. It achieves the highest mean AUC and lowest standard deviation over the 10 folds at 0.991 and 0.006 respectively.

These results indicate the CNN models employed in this study can successfully solve this classification task. Having identified ResNet as the best performing model, we next analyzed its ability to locate important features related to Coccidioidomycosis through the use of class activation maps. This exploration will ensure the model learns features relating to Coccidioidomycosis and not spurious artifacts of the radiographic images (see Table 2).

### 4.1. Visualizations

Neural networks are generally considered “black-box methods”, which makes model interpretability an issue. However, certain methods like Class Activation Maps (CAM) project information back to the input space, allowing a visual understanding of the models decision (Zhou et al., 2016). CAM gives convolutional networks tremendous interpretability despite being trained on image-level labels. CAM requires the use of a global average pooling layer (Lin et al., 2013), which is added to the last convolutional layer of the models used in this study. The Keras Visualization Toolkit (Kotikalapudi, 2017) is used to produce CAM results. Using CAM, we are able to visualize what regions of input images the network attends to when making its classification prediction. This allows one to ensure the network is learning features directly related to coccidioidomycosis and not other circumstantial features contained in the images (e.g. borders or image annotations).

Shown in Fig. 4a and b are correctly classified diseased and healthy subjects, respectively. The right panel shows the original images and the left panel displays the CAM results from the ResNet models classification. The CAM images are interpreted like heat maps, where the more red the region, the higher weight the network associates with those features when making its prediction. From the CAM images shown in Fig. 4, one can verify the network weights features in the chest and lungs region more heavily than abdominal regions or the X-ray background. These heatmaps confirm that not only can CNNs correctly classify coccidioidomycosis at a high skill level but they also recognize correct features when making their predictions.

## 5. Discussion

A dataset of 1,174 images is small for deep learning applications. In such a regime, multiple techniques are available to avoid overfitting including (1) using pre-trained weights; (2) performing various forms of data augmentation; (3) adding various forms of regularization; (4) early stopping as well as using additional randomness in the training procedure, such as dropout (Baldi & Sadowski, 2014; Srivastava et al., 2014).

The results from all CNN models demonstrate a high level of skill classifying Coccidioidomycosis in canine subjects. The high skill level achieved by the models on the unseen test sets demonstrates the models did not overfit and in fact generalized well to new data. Despite the dataset containing canine subjects at different orientations and projections (right lateral, left lateral, and ventral–dorsal) the models learned to correctly classify Coccidioidomycosis in these differing scenarios.

For CNN models to be really useful in a medical capacity they must demonstrate not only high accuracy but also a good degree of interpretability. The ResNet model probed here demonstrates a high skill level while providing essential interpretability through the use of heatmaps. Coccidioidomycosis is primarily found in the lung fields, the hilar lymph nodes, or both. Examining Fig. 4, the lung field, located dorsal and caudal of the heart, receives high importance both when predicting the presence of cocci as well as its absence. The lung field cranial of the heart is indicated as important as well.

A critical aspect of the models implemented in this study is the speed at which they compute predictions. The training took place on NVIDIA Titan V GPUs with 12 GB of memory. Each epoch took roughly two minutes to complete. Training for a typical number of fifty epochs took about sixteen hours for the full 10-fold cross-validation. However, at production time, an image can be classified in less than 50 ms, and a heatmap produced in less than 1 s. The low latency offered by these models is further evidence of their potential for being deployed and used alongside medical professionals.

The evidence provided supports the robustness of our approach along four different dimensions. First, the model learns to identify Coccidioidomycosis across many different dogs, with variations in the

**Table 1**

Hyperparameter space. The hyperparameters from the best performing architecture variants are shown in their respective columns.

Name	Range	Parameter type	Inception	MobileNet	ResNet	Shallow	VGG
Dropout	0., 0.25	Continuous	0.28	0.29	0.31	0.44	0.21
Learning rate	0.00001, 0.01	Continuous	0.0006	0.0096	0.0071	0.0025	0.0056
Learning rate decay	0.5, 1.	Continuous	0.93	0.82	0.98	0.93	0.86
Loss penalty	Yes, No	Choice	No	No	No	No	Yes
Number of layers	0, 2	Discrete	0	0	1	1	0
Number of nodes	128, 512	Discrete	–	–	152	180	–
Optimizer	Adam, SGD, RMSProp	Choice	Adam	SGD	SGD	RMSProp	RMSProp
Weights	ImageNet, Random	Choice	Random	ImageNet	ImageNet	Random	ImageNet
Train base	Yes, No	Choice	Yes	Yes	Yes	No	No

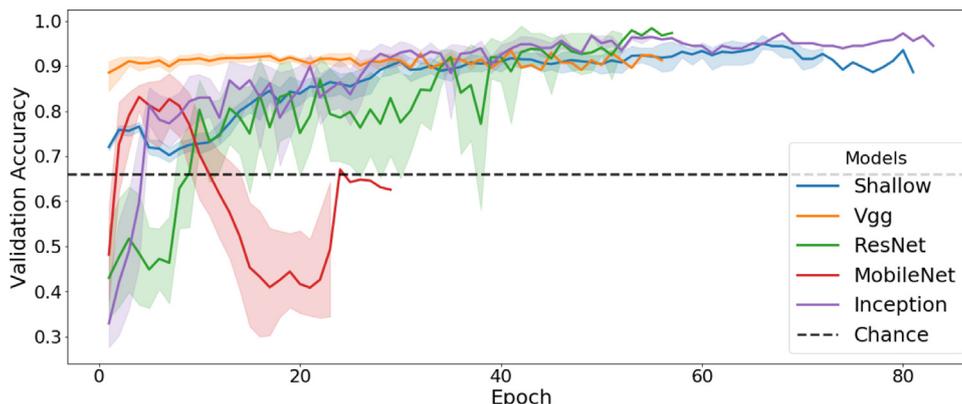


Fig. 2. Validation accuracy during training. The legend indicates the corresponding model and the dashed line represents the accuracy from chance (around 66%).

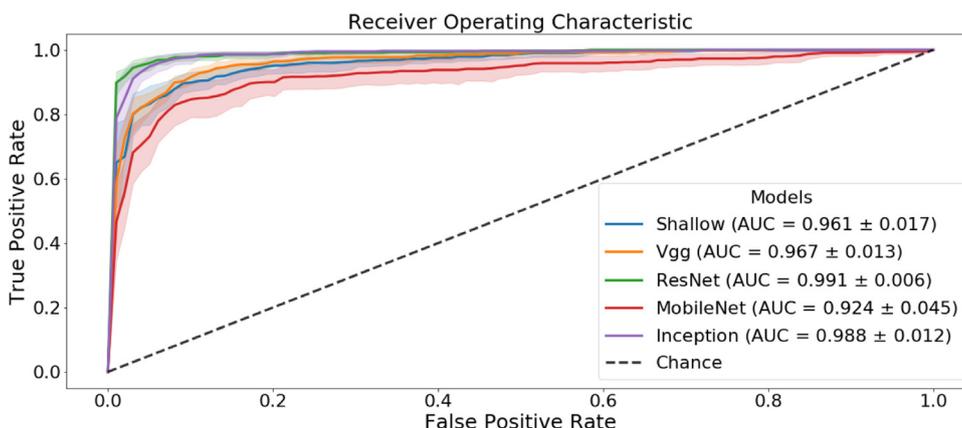


Fig. 3. Receiver Operating Characteristic (ROC) curves displaying the relationship between the model’s false positive rate and true positive rate. The legend indicates the corresponding model as well as the mean and standard deviation of the AUC (Area Under the Curve) obtained by 10-fold cross validation.

**Table 2**

Result metrics showing sensitivity, specificity, positive predictive value, and negative predictive value.

	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Inception	0.906	0.975	0.949	0.960
VGG	<b>0.925</b>	0.885	0.784	<b>0.965</b>
ResNet	0.846	<b>0.996</b>	<b>0.991</b>	0.936
Shallow	0.802	0.970	0.922	0.918
MobileNet	0.769	0.871	0.800	0.901

height, width, and size of the dogs. Second, the model learns to identify Coccidioidomycosis across different radiography imaging systems and quality of scans. Third, the model learns to identify Coccidioidomycosis across different projections. Right, left, and Ventral–dorsal projections all occur in the dataset. Fourth, ten-fold cross-validation results demonstrate robust generalization properties of the approach. In addition, a high level of performance is achieved by multiple models, further confirming the viability and robustness of the overall approach.

Finally, there is some room to further improve the results, for instance by acquiring more data. As the dataset and diversity of samples increases, one can expect the performance of CNN models to likewise increase. With more training examples the models will make more accurate predictions and provide more informative localization heatmaps. It may also be useful to leverage transfer learning and, for instance, develop CNNs that can diagnose and classify multiple lung diseases simultaneously.

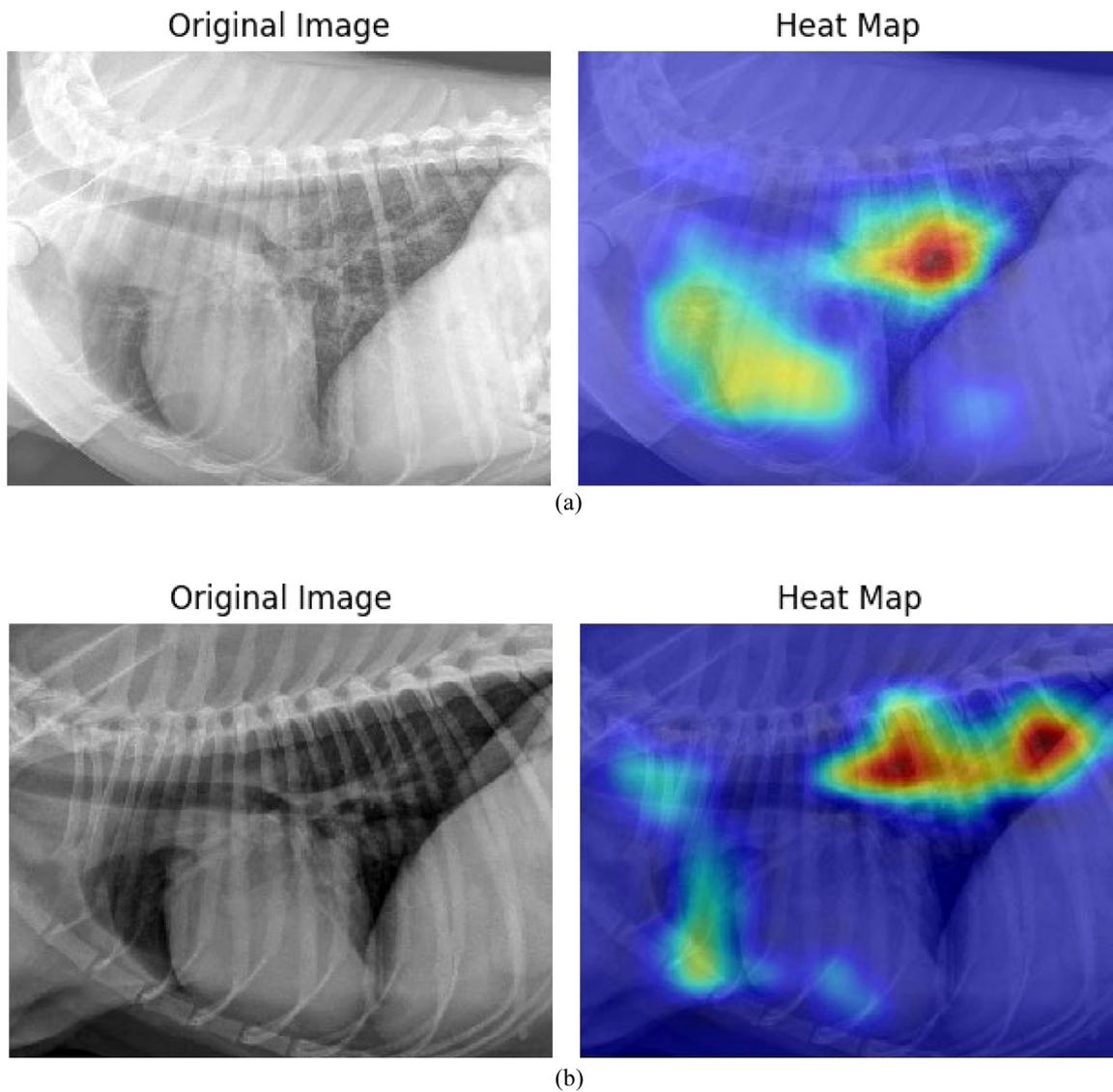


Fig. 4. Original image fed to convolutional network and the class activation map produced by the ResNet’s classification prediction. (a) A canine subject with Coccidioidomycosis correctly classified by the network. (b) A healthy canine subject correctly classified by the network.

## 6. Conclusion

In this paper we presented a novel application of convolutional networks to identify the presence of pulmonary Coccidioidomycosis in canine subjects. After assembling a data set of annotated radiographic images, we have trained and compared several deep learning models and demonstrated that successful classification and localization of the disease in radiographic images is feasible. The best model achieves a high skill level and is interpretable via the use of heat maps that can be understood by human experts. While training can be time-consuming, at production time the model is fast and can produce useful results in less than a second. As more data is gathered, an automated system driven by machine learning could further improve itself. The same approach could be extended to other organs, diseases, and animal species in veterinary medicine. As is the case for human medicine, the major remaining challenges are in overcoming the numerous barriers to data collection and aggregation, and the actual deployment of deep learning systems in clinical settings where the concerns of multiple stakeholders must be addressed.

## CRediT authorship contribution statement

**Jordan Ott:** Designed, Trained, Tested all the machine learning models, Writing and editing the manuscript. **David Bruyette:** Designed the overall study, Data annotation, Additional veterinary expertise, Writing and editing the manuscript. **Cody Arbuckle:** Designed the overall study, Writing and editing the manuscript. **Dylan Balsz:** Designed the overall study, Writing and editing the manuscript. **Silke Hecht:** Provided the original data, Writing and editing the manuscript. **Lisa Shubitz:** Provided the original data, Writing and editing the manuscript. **Pierre Baldi:** Designed the overall study, Designed, Trained, Tested all the machine learning models, Writing and editing the manuscript.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Supplementary data

Supplementary material related to this article can be found online at <https://doi.org/10.1016/j.mlwa.2021.100040>.

## References

- Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G., & Naidich, D. P. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature Medicine*, 25(6), 954–961.
- Baldi, P. (2021). *Deep Learning in Science*. Cambridge University Press.
- Baldi, P., & Chauvin, Y. (1993). Neural networks for fingerprint recognition. *Neural Computation*, 5(3), 402–418.
- Baldi, P., & Sadowski, P. (2014). The dropout learning algorithm. *Artificial Intelligence*, 210, 78–122.
- Chang, P., Grinband, J., Weinberg, B. D., Bardis, M., Khy, M., Cadena, G., Su, M. Y., Cha, S., Filippi, C. G., Bota, D., & Baldi, P. (2018). Deep-learning convolutional neural networks accurately classify genetic mutations in gliomas. *American Journal of Neuroradiology*, 39(7), 1201–1207.
- Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N., & Tsirigos, A. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nature Medicine*, 24(10), 1559–1567.
- Davidson, A. P., & Pappagianis, D. (1996). Canine coccidioidomycosis: 1970 to 1993. In *Proceedings of the 5th international conference on coccidioidomycosis* (pp. 155–162). Bethesda, MD: National Foundation for Infectious Disease.
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248–255). Ieee.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118.
- Fukushima, K., & Miyake, S. (1982). Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and cooperation in neural nets* (pp. 267–285). Berlin, Heidelberg: Springer.
- Gorris, M. E., Treseder, K. K., Zender, C. S., & Randerson, J. T. (2019). Expansion of coccidioidomycosis endemic regions in the United States in response to climate change. *GeoHealth*, 3(10), 308–327.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778).
- Hertel, L., Collado, J., Sadowski, P., Ott, J., & Baldi, P. (2020). Sherpa: Robust hyperparameter optimization for machine learning. *SoftwareX*, Also: arXiv:2005.04048.
- Hosny, A., Parmar, C., Coroller, T. P., Grossmann, P., Zeleznik, R., Kumar, A., Bussink, J., Gillies, R. J., Mak, R. H., & Aerts, H. J. (2018). Deep learning for lung cancer prognostication: a retrospective multi-cohort radiomics study. *PLoS Medicine*, 15(11), Article e1002711.
- Iglovikov, V. I., Rakhlin, A., Kalinin, A. A., & Shvets, A. A. (2018). Paediatric bone age assessment using deep convolutional neural networks. In *Deep learning in medical image analysis and multimodal learning for clinical decision support* (pp. 300–308). Springer, Cham.
- Kotikalapudi, R. (2017). Contributors, keras-vis.
- Kumar, A., Singh, S. K., Saxena, S., Lakshmanan, K., Sangaiah, A. K., Chauhan, H., Shrivastava, S., & Singh, R. K. (2020). Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer. *Information Sciences*, 508, 405–421.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.
- Lin, M., Chen, Q., & Yan, S. (2013). Network in network. arXiv preprint arXiv:1312.4400.
- Liskowski, P., & Krawiec, K. (2016). Segmenting retinal blood vessels with deep neural networks. *IEEE Transactions on Medical Imaging*, 35(11), 2369–2380.
- Mehrkens, L. R., Mohr, F. C., & Sykes, J. E. (2016). Clinicopathologic and histopathologic renal abnormalities in dogs with coccidioidomycosis. *Journal of Veterinary Internal Medicine*, 30(5), 1667–1671.
- Pasupa, K., Tungitnob, S., & Vatathanavaro, S. (2020). Semi-supervised learning with deep convolutional generative adversarial networks for canine red blood cells morphology classification. *Multimedia Tools and Applications*, 79(45), 34209–34226.
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4510–4520).
- Shubitz, L. F. (2007). Comparative aspects of coccidioidomycosis in animals and humans. *Annals of the New York Academy of Sciences*, 1111(1), 395–403.
- Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929–1958.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. 2016. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818–2826).
- Tang, H., Kim, D. R., & Xie, X. (2018). Automated pulmonary nodule detection using 3d deep convolutional neural networks. In *2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018)* (pp. 523–526). IEEE.
- Tortorano, A. M., Carminati, G., Tosoni, A., & Tintelnor, K. (2015). Primary cutaneous coccidioidomycosis in an Italian nun working in South America and review of published literature. *Mycopathologia*, 180(3–4), 229–235.
- Urban, G., Feil, N., Csuka, E., Hashemi, K., Ekelem, C., Choi, F., Mesinkovska, N. A., & Baldi, P. (2021). Combining deep learning with optical coherence tomography imaging to determine scalp hair and follicle counts. *Lasers in Surgery and Medicine*, 53(1), 171–178.
- Urban, G., Tripathi, P., Alkayali, T., Mittal, M., Jalali, F., Karnes, W., & Baldi, P. (2018). Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology*, 155(4), 1069–1078.
- Wang, J., Ding, H., Bidgoli, F. A., Zhou, B., Iribarren, C., Molloy, S., & Baldi, P. (2017a). Detecting cardiovascular disease from mammograms with deep learning. *IEEE Transactions on Medical Imaging*, 36(5), 1172–1181.
- Wang, J., Fang, Z., Lang, N., Yuan, H., Su, M. Y., & Baldi, P. (2017b). A multi-resolution approach for spinal metastasis detection using deep siamese neural networks. *Computers in Biology and Medicine*, 84, 137–146.
- Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., & Torralba, A. (2016). Learning deep features for discriminative localization. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2921–2929).