

# CS229 Final Project

Eden Grown-Haerberli  
edengh@stanford.edu

Sophia Andrikopoulos  
sophiala@stanford.edu

June 2021

## 1 Introduction

Fungal infections caused by dermatophytes are scientifically known as dermatophytosis and commonly known as ringworm for the round, red ring marking the inflammation of animals infected with the disease. Ringworm is often found in dogs and cats. Because of the visual nature of the infection, most veterinarians are easily able to suspect the presence of dermatophytes, but official diagnosis waits until the infected animal can be inspected using a fluorescent lamp - which is able to diagnose most, but not all, cases of *M. canis*, or until fungal culture of the infected area confirms the presence of dermatophytes. Unfortunately, fungal culture can take multiple days, and the dermatophytes causing ringworm are both zoonotic and highly contagious, meaning that humans are likely to contract the disease if their pets are afflicted.

As a result, a classification system which can return quick, accurate identification of images containing ringworm would be an important contribution to the veterinary space. Although machine learning has been successfully applied to the human dermatological diagnoses, using machine learning approaches to diagnose clinical veterinary diseases is much less common. Therefore, we have worked to develop a trained classifier for diagnosing canine and feline dermatophytosis. The input to our algorithm is an image of a cat or a dog. We explore an SVM and a convolutional neural network to output whether that animal is healthy or is infected with ringworm.

## 2 Related work

In recent years, machine learning has been applied to several dermatological problems. Skin disease classification using clinical images or images taken on mobile devices has become increasingly common. Support-vector machines (SVMs), k-nearest neighbors, and convolutional neural networks (CNNs) are commonly used for dermatological classification, tackling diseases such as melanoma, psoriasis, and atopic dermatitis.

The veterinary field has also seen an up-tick in the use of machine learning for diagnosis. The most frequently used methods are CNNs and other neural networks, logistic regression, linear regression, multiple regression, PCA, and k-means. However, the use of machine learning in the veterinary field is still quite new, and is focused around the use of clinical data.

In the Phillipines, a research group at University of Mindanao trained an Android-based application which used an artificial neural network to identify ringworm and yeast infections, but the group did not distinguish between the two types of dermatologic afflictions, and in addition did not include cats in their classification (Mellores).

### 3 Dataset and Features

75 unique images of dogs and cats with ringworm were downloaded from a variety of veterinary practice websites and websites directed towards pet owners. We received no response from over 40 veterinarians and veterinary dermatologists we contacted to further expand the dataset. Images of healthy dogs and cats were sampled from the Kaggle "Stanford Dogs" and "Cat" datasets. Both datasets contain images of a variety of cat and dog breeds captured with different angles, zoom, and environment. Due to the low availability of images of cats and dogs infected with ringworm, transformations were applied to the images to enlarge the dataset. The images were flipped, rotated, skewed, and zoomed using Augmentor, resulting in a final set of 500 images in the "ringworm" class. Due to variance in image sizing, after augmentation, the images were standardized to the same size (220x220) and resolution. (Figure 1)

At this point, the data was processed differently for both learning approaches.

In order to perform SVM, features were extracted from the entire dataset. First, using sklearn's StandardScaler package, the mean was subtracted from the entire dataset and the dataset was scaled to unit variance. Then, sklearn's PCA package was run in order to extract the top 500 principal components. These components were then used to train and predict the SVM model. For the CNN, the preprocessed images were normalized but kept otherwise the same.

### 4 Methods

In order to classify ringworm using images of dogs and cats, we selected two machine learning models to examine. The first and likely most straightforward model is a Support Vector Machine (SVM). SVMs, which are a supervised method, have been shown to classify images with a high percentage accuracy. However, using SVM requires separate feature extraction, which may not focus on the important features. The second model is a Convolutional Neural Network (CNN), which is a deep learning technique that aids in feature extraction. CNNs require less computation than normal neural networks because of the filtering which is part of the algorithm, but are still able to provide accurate

image classification results.

With respect to the SVM, we aim to maximize the geometric margin. This means that we perform the following optimization:

$$\begin{aligned} \min_{w,b,\epsilon} & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \epsilon_i \\ \text{s.t.} & y^{(i)}(w^T x^{(i)} + b) \geq 1 - \epsilon_i, i = 1, \dots, n \\ & \epsilon_i \geq 0 \end{aligned}$$

. This optimization essentially maximizes the geometric margin when we constrain the functional margin of  $w, b$  to be 1. Using this maximization will result in some number of support vectors which will provide non-zero Lagrangian multipliers  $\alpha_i$ . In order to maximize the geometric margin and develop an effective SVM, we ran the libsvm python package. We were able to run the SVM both with a regularization term (as a C-SVC) and without, as a basic SVM.

Convolutional Neural Networks (CNNs) consist of convolutional layers, pooling layers, optional drop-out layers, and fully-connected layers. The convolutional layers are used for feature extraction and function by sliding a filter across each image channel and taking the dot product between the filter and the parts of the input with respect to the size of the filter. Oftentimes multiple filters are used in one convolutional layer, resulting in several feature maps. We use the ReLu activation function,  $f(x) = \max(0, x)$ , to increase the non-linearity of the output of such convolutional layers. The pooling layers reduce the size of the feature maps, decreasing connections by pooling the values across a 2x2 window using the MaxPool function, which takes the maximum value in the window. Dropout layers drop neurons from with a pre-determined probability  $p$ . The outputs from the feature extraction layers are flattened and fed into the fully-connected layer, resulting in classification. For our CNN we use a ReLu layer followed by SoftMax layer ( $\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$ ) for classification. We implement our CNN in Keras.

## 5 Experiments/Results/Discussion

For both of these algorithms, we chose to use a confusion matrix, allowing us to calculate accuracy, as well as a ROC AUC score to rate the performance of the algorithm. The confusion matrix gave us insight into the explicit performance of each algorithm, including true positives, false positives, true negatives and false negatives. The ROC curve represents the discrimination ability of a classifier, and the AUC is a quantitative measure of whether the algorithm can distinguish between two classes. With regard to the SVM, we tuned two SVM hyperparameters, namely the kernel type and the cost used in the optimization function.

Before tuning the hyperparameters, we experimented with multiple SVM types, particularly the C-SVC and a basic SVM. The basic SVM, which does

not use a regularization term, returned a 0% accuracy rate, whereas the C-SVC originally returned a 95% accuracy rate. This led us to focus on the C-SVC. Within the C-SVC model, we aimed to tune the regularization term as well as the kernel used.

First, we ran the C-SVC with a linear kernel. This experiment resulted in the following confusion matrix:

		Ground Truth	
		Healthy	Ringworm
Predicted	Healthy	12190	0
	Ringworm	302	0

Unfortunately, this C-SVC returned an accuracy of 97.582% and an ROC-AUC of 0.5. Although the accuracy percentage looks considerably high, we can see that the classifier achieved this by classifying all images as healthy. The large number of healthy images in our dataset skewed accuracy as a performance evaluator. Next, we aimed to improve performance by tuning the regularization term using the RBF kernel. Running the C-SVC with both  $C = 2$  and  $C = 4$ , we obtained the following confusion matrix:

		Ground Truth	
		Healthy	Ringworm
Predicted	Healthy	12190	0
	Ringworm	244	58

Both of these iterations resulted in an accuracy of 98.0467%. Although this accuracy sounds high, examining the confusion matrix leads to the conclusion that this SVM is not performing as well as it should. This can be confirmed by looking at the ROC-AUC, which is 0.5960 for both iterations. At this point, we determined that the SVM was not performing well enough to use this as a classifier of ringworm in images of dogs and cats. The next algorithm to focus on was the CNN.

Due to the large amount of training data and the resulting training duration, we were unable to fine-tune each hyperparameter for the CNN. Instead, we focused on optimizing our learning rate, dropout, and training epochs. We began with a sequence of three pairs of 2D convolutional layers with ReLu activation followed by MaxPooling with a 2x2 window. Following the third block, we implemented a dropout layer with a  $p = 0.6$ , a dense layer with a ReLu activation function, and a final dense 2-class-output SoftMax layer. We trained with a learning rate of 0.001 for 10 epochs, resulting in 99% and 98% accuracy for the training and test sets, respectively:

		Ground Truth	
		Healthy	Ringworm
Predicted	Healthy	4492	21
	Ringworm	70	130

The above model overfit to the data with high loss and 65% recall for positive cases in the validation set. In order to partially resolve overfitting and improve recall with our limited dataset, we experimented with the number and probabilities of our dropout layers, as well as our learning rate and the number of training epochs. A similar model as the one above with a learning rate of 0.0001 trained for 50 epochs performed similarly to the above model, while learning rates greater than 0.001 performed worse overall. Our final model consists of a sequence of three blocks of one 2D convolutional layer with ReLu activation, one MaxPooling layer with a 2x2 window, and a  $p = 0.4$  dropout layer. Following the third block, we implemented a flattening layer, a dense layer with a ReLu activation function, and a final dense 2-class output SoftMax layer. We implemented our neural network with the Adam optimizer and cross-entropy loss. Adding additional dropout layers helped reduce overfitting and increase recall (74%) at the cost of a slight decrease in accuracy (98% and 97% for training and test sets respectively):

		Ground Truth	
		Healthy	Ringworm
Predicted	Healthy	4462	51
	Ringworm	51	149

## 6 Conclusion/Future Work

Although the SVM algorithms that we experimented with were able to identify some small number of ringworm cases, the overall performance in terms of the AUC score was somewhat disappointing. Using this particular C-SVC to diagnose ringworm in veterinary offices would be suboptimal, as more than half of the ringworm cases were misdiagnosed in the best case scenario. While the CNNs performed better than the SVM algorithms, they still only caught 65% of the ringworm cases. In the future, in order to allow accuracy to act as a better representation of the performance of these algorithms, decreasing the number of images in the test set will be an important step.

With an improved dataset, both algorithms would likely perform better. Two specific aspects of the dataset might improve SVM performance in the future: an increase in the number and variety of images of patients with ringworm, as well as an improved feature extraction pipeline. Using PCA to extract features is not the only approach with respect to training an SVM. Other methods to try include using the grayscale pixel values themselves, filtering the image in some way, or even performing a Fourier or a Laplacian transform on the image. CNN performance might be improved by increasing the dataset as described above, as well as further optimizing the layers in the network by adding and removing layers, exploring the effects of different non-linear activation functions such as tanh and Swish, and exploring different optimization techniques such as SGD, Adagrad, RMSProp, and Adadelata.