Classifying Dog Breeds based on images

Watson Whitehurst

ITCS-5156 Fall 2022 – Lee

October 12, 2022

Original paper:

A new dataset of dog breed images and a benchmark for fine-grained classification by Ding-Nan Zo, Song-Hai Zhang, Tai-Jiang M, and Min Zhang

Zou, Ding-Nan, et al. “A New Dataset of Dog Breed Images and a Benchmark for Finegrained Classification.” *Computational Visual Media*, vol. 6, no. 4, 2020, pp. 477–487., https://doi.org/10.1007/s41095-020-0184-6.

Code and Documentation from QuocThangNguyen on GitHub.

https://github.com/QuocThangNguyen/deep-metric-learning-tsinghua-dogs

**Abstract**

Dogs come in many different shapes and sizes. As there are several different breeds of dogs, it may take a person a little while to learn the differences between all of them. As it turns out, we can just classify the breeds of dogs with AI. Ding-Nan Zo et al. [5] created a large dataset of labeled images of dogs corresponding to their breed. User QuocThangNguyen [4] from GitHub shows how they were able to classify that data using deep metric learning. This paper will show how different models classify dog breeds and the noticeable variations from different datasets.

**Introduction**

Dogs are a major species on Earth and are said to be Man’s best friend. They come in several different breeds, and they can all be very different from one another. My goal is to classify dogs into classes based on their breeds using deep learning. I wanted to familiarize myself better with neural networks while creating an AI that has an actual use case. Classifying dog breeds is a real-world application that I thought would be helpful to use AI to classify. One challenge I had was getting my data to be read correctly by the model. I was able to figure this out using FastAI with help from a GitLab tutorial by jhp00 [1]. Another thing I had to learn was which models would work for classifying images. I could not use a simple linear model typically used to classify data points as we have done in previous labs. I compared 3 different models: resnet34, resnet101, and AlexNet. I compared the results of these models to see which one worked best for my data and how they stacked up against the model created by QuocThangNguyen from a different dataset.

**Related Works**

A screenshot of a computer

Description automatically generated with low confidence

Table 1: This compares the Standford Dogs dataset to previous datasets

**Novel Dataset for Fine-Grained Image Categorization: Stanford Dogs:**

In this paper from 2011 by Aditya et al. [2], one of the biggest datasets for classifications of dogs for its time was introduced. The Stanford Dogs dataset includes over 22,000 annotated images of dogs spanning 120 classes of breeds. One element of this dataset compared to past datasets is that images are taken from random environments and sometimes contain humans in the background. This allows for a greater variety and makes the dataset harder to classify. This new dataset allows testing for a greater number of classes and images per class compared to previous datasets as seen in Table 1 above. Each image was individually examined to make sure that the class it was put in matched features from images in the same category. Any unusual or duplicate images were manually removed from the dataset. You can see a sample of the dataset below in Figure 1. This shows the qualities of a dataset very similar to the one we will be using.

**A collage of various animals

Description automatically generated with low confidence**

Figure 1.A sample of

12 of the dog categories

**Plant Disease Detection and Classification by Deep Learning:**

This paper has a similar goal to what I am trying to accomplish here. Instead of classifying dog breeds, they are classifying plants to see if they have diseases or not. In May 2021, Lili et al. [3] presented their work on classifying plants. First, they show how they used K-means clustering, in which features were identified based on a support vector machine. Through this approach, they reached 93% accuracy. Different methods of classification are mentioned until they get to convolutional neural networks (CNNs). By using 3-6 layer CNNs they were able to reach an accuracy of 99.35%. It was mentioned that the datasets had low diversity, but there are no plant databases of greater diversity available currently. The paper notes the history of deep learning algorithms which can be seen in Figure 2 below.

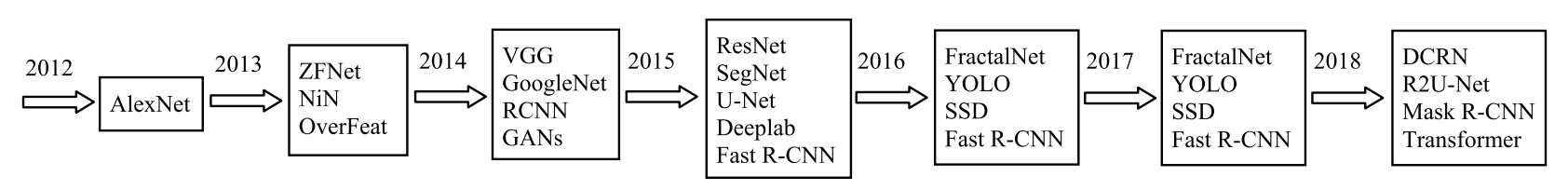


Figure 2. The history of deep learning architectures

The authors also touch on how visualization techniques such as the introduction of heat maps and salient maps help to better understand results from CNNs and other classification methods. The writers mention several different species of plants and what features on them can be studied by deep learning to learn about diseases. This paper is relevant to my work since I will be dealing with CNNs for classifying dog breeds.

**Method**

There are many different models we can use to classify images of dogs. First, let’s start with the dataset that was used. I used a smaller variation of the Stanford Dogs dataset from FastAI. This dataset includes 37 classes and approximately 1300 images for training. The reasoning behind using a smaller dataset is so that I can run the models on the data in a timely manner. Also, using the dataset provided by FastAI allows our data to already be preprocessed as needed and formatted into the dataset. The model by QuocThangNguyen I am basing my testing on used the much larger Tsinghua Dogs dataset from 2020. This dataset is comprised of 65,000 training images and 130 classes. For my testing, I used three different models. The models used were resnet34, resnet101, and AlexNet**.** I did not preprocess the data, as FastAI does this for me. For each model I used 3 epochs for the sake of time, however, this can be easily scaled up for more accurate results. Each model was run three times to make sure of no anomalies. The accuracy between runs varied by about .5% to 1%.

**Text

Description automatically generated Graphical user interface, text

Description automatically generated Graphical user interface, text, chat or text message

Description automatically generated**

Figure 2. Includes results from the training of each model. 3 epochs were used each time for testing purposes and the metrics were set to error rate. The models are presets given by FastAI and this is my code implementing them.

**Experiments**

We can see above in Figure 2 how the models performed. AlexNet was one of the earlier vision deep-learning models. From the AlexNet model, we saw an accuracy of 81%. With the more relevant models, Res-Net-34 and Res-Net-101, we saw results of 92.5% and 93.5% accuracy respectively. Remember that these results have small variations between runs. If we go back to the original paper using the Tsinghua Dog dataset. The model they used was Res-Net-50 with an embedding size of 128, batch size of 48, and used 100 epochs. Using this model, we can see that they received an accuracy of about 76%. Now we see that this is a fair bit lower than our accuracy, but we must remember the Tsinghua Dog dataset has 130 different classes while ours only has 37. It is somewhat surprising to see that my results were better when I only used 3 epochs compared to 100 even with the AlexNet model. It is important to note how much of a difference using a bigger dataset affects the results. Through these results, we learn that having to classify data into a larger number of classes is much more demanding than a smaller number of classes, even with a larger amount of training data. There also may be a few differences in the datasets that are not mentioned that are affecting these results. If I were to increase the number of epochs used in my training to 100 as was done in the paper, I believe that my results would have higher accuracy.

**Conclusion**

I believe that by spending more time I could increase the accuracy of the models simply by increasing the number of epochs. I could also look into creating my own models and see how they stack up against AlexNet or even Res-Net. I could also try using a larger dataset myself such as the full Stanford Dogs dataset and use data preprocessing before training my model. That was a challenge I had earlier with learning how to correctly import the dataset, but FastAI was able to help me with this. I would have to work further with this project to allow the models to work with a different dataset. I think the project was very helpful for me as I was able to use multiple models for training a classification model. It was nice that I was able to use a real-life scenario for deep learning.

**References**

[1] Jhp00, “fastbook” GitHub Repository, 2022, https://github.com/fastai/fastbook/blob/master/05\_pet\_breeds.ipynb

[2] Khosla, Aditya, et al. “Novel Dataset for Fine-Grained Image Categorization: Stanford Dogs.” Stanford Dogs Dataset, 2011, https://doi.org/10.1109/cvpr.2011.5995368.

[3] Li, Lili, et al. “Plant Disease Detection and Classification by Deep Learning-A Review.” IEEE Access, vol. 9, 2021, pp. 56683–98, https://doi.org/10.1109/ACCESS.2021.3069646.

[4] Nguyen, Quoc, “Deep metric learning tsinghua dogs.” GitHub Repository, 2021, https://github.com/QuocThangNguyen/deep-metric-learning-tsinghua-dogs

[5] Zou, Ding-Nan, et al. “A New Dataset of Dog Breed Images and a Benchmark for Finegrained Classification.” *Computational Visual Media*, vol. 6, no. 4, 2020, pp. 477–487., https://doi.org/10.1007/s41095-020-0184-6.

My GitHub Repo: <https://github.com/0rangeB0y/animalbreeds>