Dog Breed Classification via Convolutional Neural Network

Yizhou Chen
Electrical and Computer Engineering
UCSD
yic244@ucsd.edu

Xiaotong Chen
Mechanical and Aerospace Engineering
UCSD
xic363@ucsd.edu

Xuanzhen Xu
Electrical and Computer Engineering
UCSD
xux021@ucsd.edu

Abstract

In this project, a convolutional neural network which can classify 40 breeds of dogs with more than 80% accuracy is built by transfer learning. Existing results with different pretrained model were compared and VGG19 network is finally chosen in this project. Also, an investigation of the category number's effect on test accuracy is done and the largest breeds number with tolerable accuracy is found.

I. Introduction

A. Motivation

The breed of dogs carries a lot of information, a lot more than just their appearance. Dogs of different breeds may have different temperament, different space and climate requirement, different effort and time it takes the owner to train them, different price and feeding cost, and so on. So, it might be helpful if families can get all the information when choosing their pets instead of just noticing how the candidates look like. This motivates the idea of dog breed classification from pictures, which falls into a fine-grained category recognition problem.

B. Problem statement

The input of the algorithm is an image of dog. We than use a pre-trained VGG19 network and logistic regression to output a predicted breed of the dog in the image.

II. Related work

A. Results without pre-trained model before 2014

In Angelova et al [1], a comparison of several methods where no pre-trained model included, and a benchmark of 'state-of-the-art' at that time is performed.

Table 1. The benchmarking result of Angelova et al [1].

Method	120 Dogs
Khosla et al [16]	22.0
Yang et al [28]	38.0
Chai et al [5]	26.0
Ours (from Table 1)	**48.26**
Ours, improvement over the best state-of-the-art	+10.26

This benchmarking result made us believe our first test accuracy for all 120 breeds (48.6%) with some mistakes in our code was correct and caused lots of worry till the last week.

B. Results by transfer learning

There are also many results including different pre-trained models as shown in table 2.

Table 2. Results out of different pre-trained model

Model	test accuracy	reference
VGG-19 and Resnet-50	82.30%	[2]
VGG16BN	76.53%	[3]
VGG16	40.23%	[3]
ResNet50	85.50%	[4]
ResNet50	84.00%	[5]
Inception v3	74%	[5]

The algorithms with transfer learning works better than the previous one. One possible reason is that the pre-trained models from ImageNet are based on much larger dataset.

Among different pre-trained models, VGG-19 and Resnet-50 has the best performance. A test in [2] has verified this.

III. Dataset

The dataset we use is Stanford Dogs Dataset[6] which is taken as training and test data in this project. It covers 120 breeds of dogs in 20580 images in total. This dataset is built upon images and annotation from ImageNet. Most of the pictures are life photo instead of close-ups with clean background.

IV. Method

VGG19

VGG19 is a network with 19 layers used by Visual Geometry Group at ImageNet ILSVRC-2014[7]. Given an image, the VGG network will output probabilities of different classes that an image could potentially belong to. In VGG19, 224*224 RGB images are passed through five blocks of convolutional layers and blocks are separated by max pooling layers. For the convolutional layer, the filter size is 3*3, the stride is 1. The final layer is a logistic regression layer that outputs the probability of each classes. The architecture of VGG19 is shown in figure 1. We use transfer learning on this project by implement our model on Keras. We load the pretrained VGG19 model from the library and use the ImageNet weights (all layers frozen).

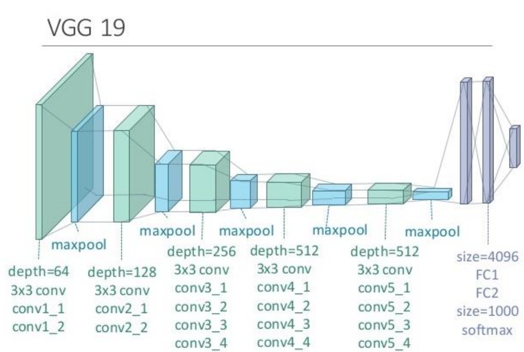

Figure 1. VGG19 architecture[7]

V. Result & Discussion

A. Results

We change the number of categories for our model and get their corresponding accuracy in table 3.

Table 3. Accuracy versus the categories selected

Category number	Test accuracy
20	84.73%
40	81.41%
60	79.02%
80	75.03%
100	73.35%
120	68.63%

And the relation between accuracy and number of categories selected is shown in figure 2.

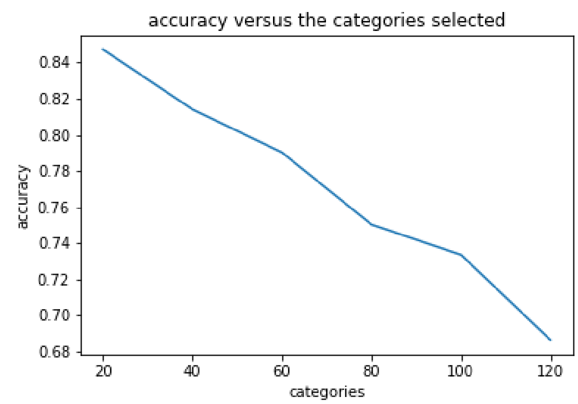

Figure 2. Plot of accuracy versus the categories selected

To achieve the goal of the project that our tolerant error rate is 20%，we select top 40 categories of dogs with highest frequency in the dataset.

There are 4029 and 4051data in training and test set, respectively. We get 0.03% training error and 18.59% test error by our model as shown in table 4.

Table 4 Result of top 40 selected categories

Training set	Test set	Training error	Test error
4029	4051	0.03%	18.59%

After the training, we get some correctly classified and misclassified figures as shown in figure 3 and figure 4 respectively.

Figure 3. Some correctly classified dog images

Figure 4. Some misclassified dog images

B. *Discussion*

Overall accuracy

The test accuracy of our algorithm is 68.63% if all 120 dog breeds are included, which is higher than the result of methods without pre-trained model but lower than the ones by transfer learning except the one using VGG16. One reason might come from the fact that we set all layers of the pre-trained model frozen. Another main reason for this may be that some of the images in the testing and training dataset have poor quality. The two misclassified images in figure 4 are just good examples. The dogs in the two images are both wet all over their bodies and thus many features of a fluffy breed like them is changed thoroughly. And what make things worse is the background. The color of the background is so much close to the dog especially at the contour. We also notice that in the training test there exist some images contains more than one dog and same image appears twice in two different breeds as shown in figure 5 and figure 6, respectively.

Figure 5. example of multi breed in one image (image ID: n02085620_2981)

Figure 6. repeated image in different breeds

The last two problems, say, multiple breeds in one figure and repeated test data, seem not solvable via programming. It is also a great challenge to teach the algorithm to classify dogs not in their normal state, which is the solution for the wet dog problem. But there is some method can be applied to improve the performance against bad background. Many algorithms we found implement pre-processing of the images like dog face detection and color histogram. This is discussed thoroughly in Liu [8] where they successfully reach 67% recognition rate from a relatively small dataset without any pre-trained model.

Category number effect on test accuracy

As shown in table 4 and figure 3, the test accuracy drops when the breed number included in a linear pattern.

We expected to see the decreasing but the fact that it decreases linearly is not expected. For each additional breed, we add around 150 test data, so the linearly decreasing pattern means that if we keep adding new breeds (assume available), for some large enough breed number, our algorithm won't work at all.

This observation astonishes us but for linear regression, only one group of tests is not enough. Same test needed to be performed on different problems and finer grid to reach a solid conclusion which is a meaningful work to be done in the future.

VI. CONCLUSION

Overall, we are satisfied with the results which achieves the goals we set in the proposal. An algorithm able to classify 40 dog breeds at an accuracy of >80% is built. With the classification among the number of dog breeds ranging from 20 to 60 with a step of 20, the testing accuracy drops not so fast and maintain a minimum of 79.02%. Also, an interesting linear pattern is observed for test accuracy – category number relationship.

Some future work is left. The first one is designing image preprocess algorithm to improve performance against bad testing data. And the second one is investigating test accuracy – category number relationship with finer grid and for different problems.

References

[1] A. Angelova and M. Philip Long, "Benchmarking large-scale fine-grained categorization," Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on. IEEE, 2014.

[2] K. Nikolay, “Dog breed image classification with Keras,” machinememos.com, July. 11, 2017. [Online]. Avalible: http://machinememos.com/python/keras/artificial%20intelligence/machine%20learning/transfer%20learning/dog%20breed/neural%20networks/convolutional%20neural%20network/tensorflow/image%20classification/imagenet/2017/07/11/dog-breed-image-classification.html. [Accessed: Jun. 16, 2018].

[3] saksham789, “An example of using CNN for classification of dogs,” github.com, Jan. 14, 2018. [Online]. Avalible: https://github.com/saksham789/DOG-BREED-CLASSIFICATION-STANFORD-DOG-DATASET. [Accessed: Jun. 16, 2018].

[4] Y. Adrian, “Dog Breed Classifier,” adrianyi.com, July, 2017. [Online]. Avalible: https://adrianyi.com/2017/07/DogBreedClassifier.html. [Accessed: Jun. 16, 2018].

[5] P. Jacob, “An example of using CNN for classification of dogs,” github.com, Jun. 26, 2017. [Online]. Avalible: https://github.com/JacobPolloreno/dog-breed-web-classifier. [Accessed: Jun. 16, 2018].

[6] K. Aditya et al, “Stanford Dogs Dataset,” Stanford University, [Online]. Avalible: http://vision.stanford.edu/aditya86/ImageNetDogs/. [Accessed: Jun. 16, 2018].

[7] S. Karen and Z. Andrew, “Very Deep Convolutional Networks for Large-Scale Visual Recognition,” University of Oxford, http://www.robots.ox.ac.uk/~vgg/research/very_deep/

[8] J. Liu et al, "Dog breed classification using part localization," European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2012.