Dog Breed Identification
- A Convolutional Neural Network Approach via Transfer Learning

Cheng Zhong, Ze Zheng, Ye Zhang
Georgetown University, Analytics – Data Science Program

ABSTRACT

To determine breed characteristics in dogs is difficult due to the fact that many dogs come in similar shapes and sizes. This project aims to identify different dogs from different breeds. Convolutional Neural Networks (CNNs) is the main method in species classification. A baseline model utilized random forest method is set up to compare the performance of the CNNs model. Due to the fact that the dataset contains roughly 10000 images which led to small number of images per breed, we utilize data augmentation and transfer learning methods to avoid the overfitting on training set.

DATASET

The dataset is downloaded from Kaggle. The dataset is originally from a strictly canine subset of ImageNet website, which is composed of 120 different breeds with 10222 images in total. The dataset is separated into training set, test set and a completely new dataset is created as a validation dataset.

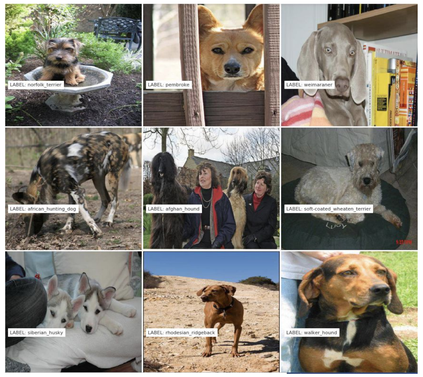

Figure 1. Dataset Overview

METHODS

Since we only have about 10000 images but 120 different breed of dogs. We want to use data augmentation when processing our data and use transfer learning to avoid overfitting. So we chose Keras as our framework. The data augmentation method (Figure 2) we mainly use here is by Flip, Crop, and Scale. Also, by applying the pretrained models like Xception, VGG16, VGG 19, and ResNet-50, we can conduct the transfer learning on the dataset by freezing the previous weights in the pretrained models and train the last layer.

MODELS

Baseline Model: Since the data set is a little bit unbalanced, the baseline model here we choose is random forest with bagging. Thus a classifier like random forest with bagging will perform well in reduce variance. However, the result of the baseline model is not as good as we think.

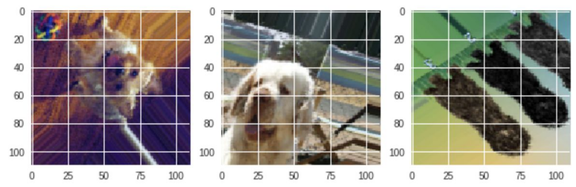

Figure 2. Data Augmentation Example

Transfer Learning: Transfer learning method is applied to avoid overfitting. Transfer learning method can be helpful to reduce computational time and deal with the problems caused by small dataset. We compared the result from the following models:

- *Xception*: An extension of the inception architecture which replace the standard inception modules with depth-wise separable convolutions.
- *VGG16*: A widely used CNNs model proposed by K. Simonyan and A. Zisserman.
- *VGG 19*: VGG19 is a network with 19 layers used by Visual Geometry Group at ImageNet ILSVRC-2014.
- *ResNet-50*: ResNet is the deep residual network. The ResNet-50 model has 50 layers.

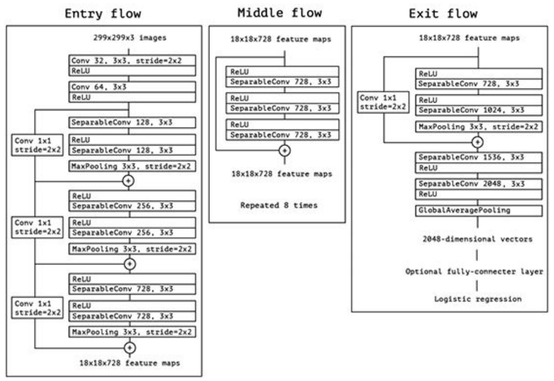

Figure 3. Xception Model Structure

RESULTS

Exploratory Data Analysis:
The dog breed appeared most in data is scottish_deerhound.

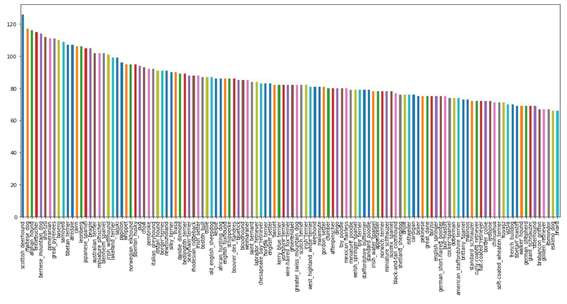

Figure 4. Dog Breed Distribution

Baseline Model Performance:
The Baseline model performance is not good. The random forest model performed on the test data got a **0.0488** accuracy. The reason might be the dimensional of the data is too high since we have over 100 classes here.

Pretrained Model Performance:

Model	Train Accuracy	Test Accuracy
Xception	0.8502	0.8450
VGG 16	0.8052	0.8026
VGG 19	0.8108	0.7943
ResNet-50	0.8371	0.8266

Sample Predict Result:

dingo	0.974487
dhole	0.011417
kelpie	0.003814
basenji	0.001791
redbone	0.000578
eskimo_dog	0.000574
malinois	0.000321
appenzeller	0.000249
cardigan	0.000242
boxer	0.000227
schipperke	0.000225
welsh_springer_spaniel	0.000175
siberian_husky	0.000172
golden_retriever	0.000172
chow	0.000147
pembroke	0.000146
soft-coated_wheaten_terrier	0.000145
saluki	0.000133
african_hunting_dog	0.000130
norwich_terrier	0.000129
cocker_spaniel	0.000118
norwegian_elkhound	0.000115

Figure 5. Sample Output

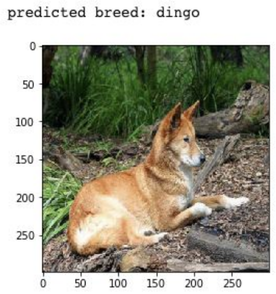

Figure 6. Sample Prediction Output

CONCLUSION

In this Project, we tried to solve the dog breed identification problem using a small dataset with only several images per breed. First of all, a random forest model with bagging was built as a baseline model, the model does not performed well and got a 4.88% accuracy on the test dataset. The reason lead to this is because the high dimensional of the data.

Next, a data augmentation and transfer learning was applied to the dataset. By applying flipping, crop, and scale to data, we got a more robust data. The Xception pretrained transfer learning model achieved a 84.5% accuracy on the test set which is the highest accuracy among all the transfer learning models. Following models ResNet-50, VGG 16, and VGG 19 achieved accuracy 82.66%, 80.26%, 79.43% on test data respectively.

The results shows that although the dog breed identification seems like a challenging task with 120 dog breeds, the CNN based transfer learning model with data augmentation is very powerful and has a decent performance on the dataset.

FUTURE WORKS

Since the size of the dataset is small, overfitting can be a continuing concern. There are many methods can be applied to avoid model overfitting. Approaches like adding more images to dataset or a more aggressive data augmentation might help solve this problem. The customized architecture that we used only have one flatten layer and one hidden layer. For the future works, the performance of the pre-trained model can be improved by adding few more layers such as add a more aggressive dropout layer might also help the model with less chance to overfit and more accurate.

REFERENCES

[1] LaRow, Whitney, et al. *Dog Breed Identification*. Stanford University,2016,web.stanford.edu/class/cs231a/prev_projects_2016/output%20(1).pdf.
[2] Chollet, Francois. “The Keras Blog.” *The Keras Blog ATOM*, 5 June 2016,blog.keras.io/building-powerful-image-classification-models-using-very-little-data.html.
[3] Chollet, Francois. *Xception: Deep Learning with Depthwise Separable Convolutions*. Google, Inc., 7 Oct. 2016,arxiv.org/abs/1610.02357.

Contact Info

Cheng Zhong cz220@georgetown.edu
Ze Zheng zz186@georgetown.edu
Ye Zhang yz597@georgetown.edu