

Dog Breed Identification

Wenting Shi, Jiaquan Chen, Fangyu Liu, Muyun Liu
University of California, San Diego

UC San Diego

Abstract

Pattern recognition(PR) is realized as a human recognition process which can be completed by computer technology. Our project is based on PR which is to identify the dog's breed. In our project, based on 10,000+ images of dogs, we use 4 methods to do the identification. Based on our model, we also make some improvements to increase our identification accuracy. After our comparisons, we find that the DenseNet model is the best, and we take it as our prime model. Our best accuracy can be up to 85.14%.

Introduction

Our project is to identify different dogs from different breeds. Our dataset including 10,000+ images of 120 breeds of dogs, is downloaded from ImageNet. Also, we have a label document, which has 2 columns(ID;Breed), and 10,000+ rows(the number of images). We have tried 4 methods to train our data, and make some improvements on some of them to increase our identification accuracy.

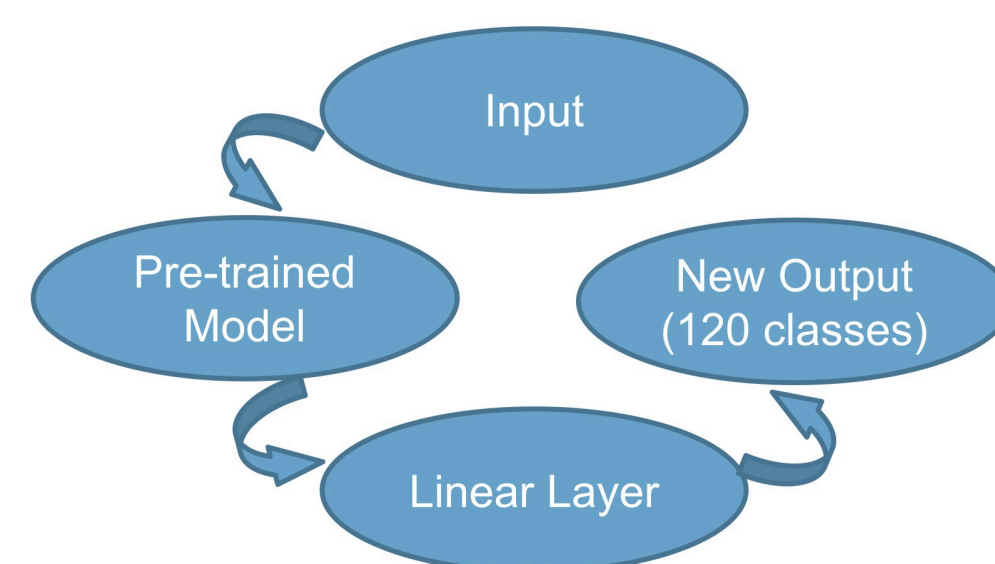


Figure 1: Block

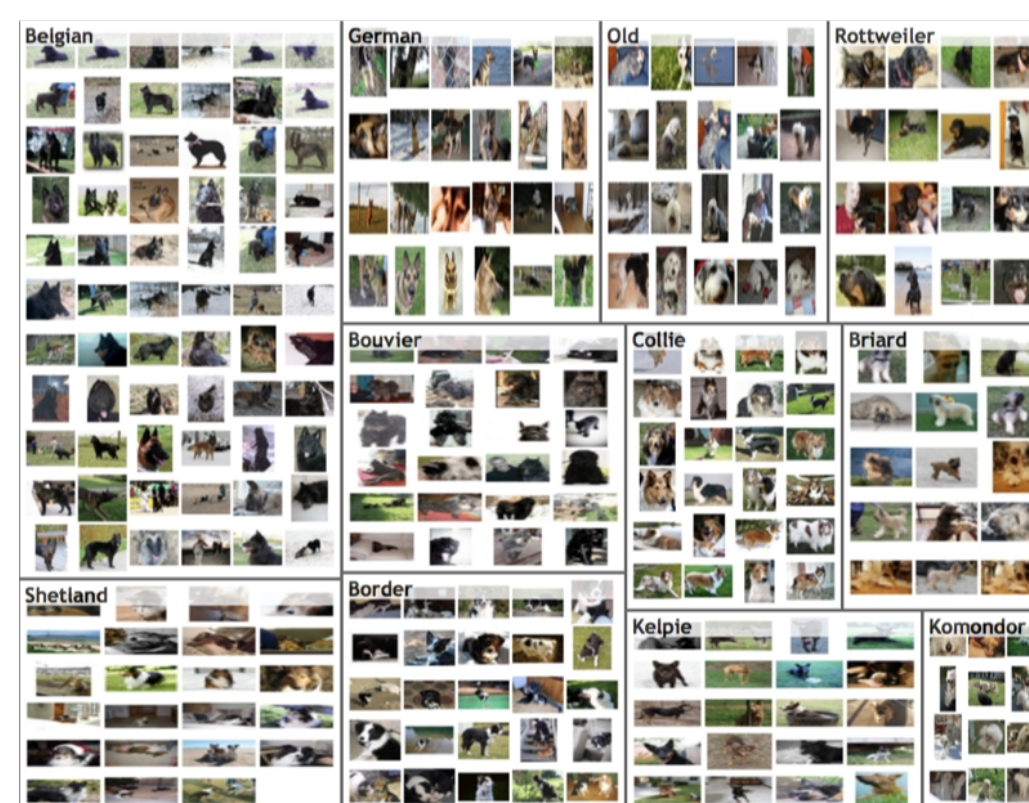


Figure 2: Dataset

Methods

In our experiment, we chose pytorch as our framework. We applied data augmentation on preprocessing our dataset and transfer learning in order to avoid overfitting. Because there are 120 breeds of dogs and more than 10,000 images, which is not big enough for models. Our exact methods of augmentation are cropping, flipping, randomly changing the brightness, contrast and saturation of each image. Transfer learning we used is to choose a pre-trained model and apply it to a different but related problem. In our experiment, we downloaded different models, such as ResNet, VGG16, AlexNet, DenseNet161, and added two fully-connected layers to replace the last layer of each pre-trained model.

Comparison of Models

We'd like to compare several different pre-trained models, and choose the one with the best performance to do our classification.

- ResNet18. ResNet means deep residual network. The residual learning framework eases the training of deep networks. ResNet is much deeper than their plain counterparts, yet it requires a similar number of parameters.
- VGG16. VGG is a convolutional neural network model proposed by K. Simonyan and A. Zisserman. It increases depth using an architecture with very small (3x3) convolution filters.
- AlexNet. AlexNet is a convolutional neural network that won the ImageNet Large Scale Visual Recognition Challenge.
- DenseNet161. DenseNet is a convolutional neural network which has following advantages: It alleviates the vanishing-gradient problem, strengthens feature propagation, encourages feature reuse, and substantially reduces the number of parameters.

Results

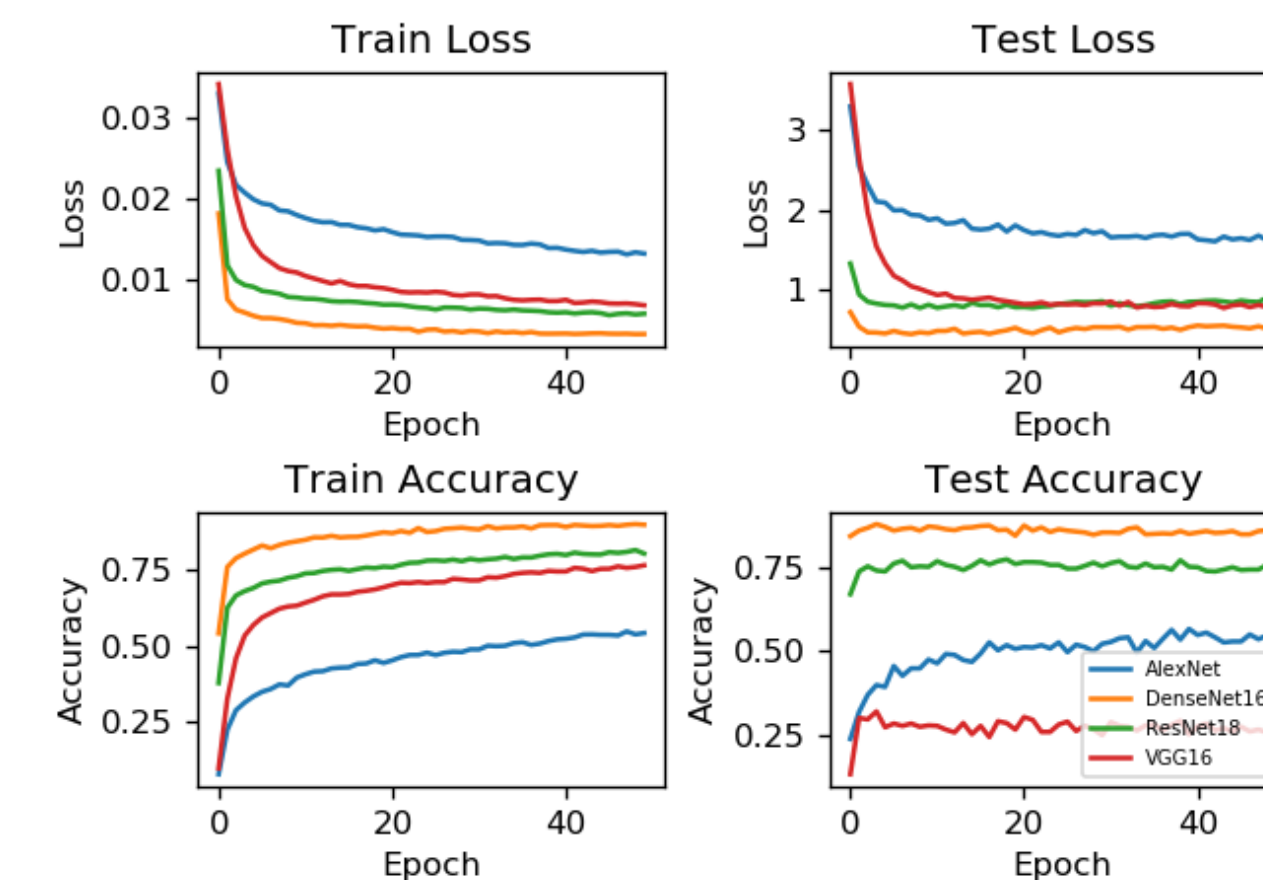


Figure 3: Train Loss Comparison

Table 1: Comparison of Different Models(50 epochs)

Model	Loss (Test/Train)	Accuracy (Test/Train)
ResNet18	0.0056/0.8782	0.8013/0.7542
DenseNet161	0.0031/0.5522	0.8954/0.8514
AlexNet	0.0131/1.6958	0.5419/0.5556
VGG16	0.0067/0.7371	0.7630/0.2667



Breed	Probability
maltese_dog	87.82%
sealyham_terrier	11.12%
west_highland_white_terrier	0.214%
shih-tzu	0.153%
malinois	0.055%

Figure 4: A Sample of Output

Discussion

In this project:

- We use transfer learning to deal with the problem that the dataset is not large enough;
- We use multi methods (like dropout and weight decay) to avoid overfitting;
- After comparing different pre-trained models and optimization algorithms, we decide to choose: DenseNet161+2 FC layers as pre-trained model and SGD+mini-batch size 128 as a optimizer;
- After 50 epochs, the accuracy of our model can reach to 85.14%. The accuracy is not high enough because there still exists overfitting and the layers we add after the pre-trained model is not enough (we only add 2 in our project).

Future Work

In the future, there are three directions we can go: First, we can compare CNN which we used in our project with other train models like SVM and PCA. Second, we can try to add more layers after Pre-trained models. Third, we can try other methods to avoid overfitting (like stop early).

References

- [1] Liu, Jiongxin, et al. "Dog breed classification using part localization." European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2012.
- [2] Sun, Yi, Xiaogang Wang, and Xiaoou Tang. "Deep learning face representation from predicting 10,000 classes." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.
- [3] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.

Contact Information

{wes053,mul046,jic496,fal053}@eng.ucsd.edu