Classification of Dog's Breed using CNN and Transfer Learning.

Md. Monoarul Islam Bhuiyan

Abstract

Computability is where the computer shines. Give it a big chunk of computational instructions and it will solve them within a moment, eventually a lot faster than humans but when it comes to solving fuzzy problems like reasoning and cognition humans outperform them with ease. Classifying images of the breed of a dog is thus quite challenging for computers. Experiments in this project show that by leveraging convolutional neural network architecture and transfer learning, computers can also be able to classify images at a near human-level performance.

1 Introduction

Dogs have been a great companion for human beings for a while and there are quite a few types of dogs around the world depending on the location they reside, their characteristics, and most importantly breeds. Classifying breeds of dogs with the help of traditional expert systems are becoming obsolete as the number of dog breed is increasing and their characteristics differ as well.

In this project, by leveraging CNN and transfer learning, I implemented a considerable amount of models to classify breeds of dogs. The Base CNN model outperforms the ANN model with almost less than one-third of trainable parameters which makes it understand why CNN is specifically great for computer vision problems. Although the base model performs well on training data it fails to produce remarkable outputs on unseen data for which another model is introduced which uses regularization to reduce overfitting from the base CNN model. Lastly, two more models are created with transfer learning's feature extraction and fine-tuning approaches where in feature extraction, we start with a pre-trained model and only update the final layer weights from which we derive predictions whereas, in fine-tuning, we start with a pre-trained model and update some of the model's convolutional layer's parameters for our task.

2 Problem Definition and Algorithm

Classifying images is fairly a simple task for humans while computers continuously struggle to achieve near-human-level performance on this task. Since the paradigm shifted with the advancement of deep neural networks, the computer seems to achieve this feat also.

2.1 Convolutional Neural Network

CNN is used heavily in computer vision problems such as image classification, object detection, Self-driving cars, image segmentation, and many more. With the help of convolution and pooling layers, it gradually extracts the most important features from the images where filters, padding (same or valid), and stride play imperative roles. Flatten layer converts the extracted features into a single-dimensional vector and passed them consequently onto the fully connected layer(s) and output layer. In the backpropagation stage, it tries to learn the trainable parameters by using optimization algorithms.

2.2 Transfer Learning

Custom models that I created fail to generalize greatly on the unseen data due to several reasons. Firstly, because of insufficient data (less than 1000 images). On the other hand building complex models with a lot of layers requires significant computational power. As a result, I leverage transfer learning. Transfer learning is a technique where instead of reinventing the wheel, we can leverage the already existing pre-trained model on large datasets. I use VGG16[1] architecture which was already trained on huge amounts of data and produced remarkable results on my project's unseen data.

3 Dataset

In this image classification project, a Kaggle dataset namely "Dog's Breed Dataset" with 1030 image files of 5 different types of dog breeds (french bulldog, german shepherd, golden retriever, poodle, yorkshire terrier) is used.

4 Experimental Evaluation

4.1 Methodology

Project's dataset is split into training and validation sets with a ratio 0f 80-20 and the image shape is set to (224, 224, 3). Later, the data is augmented so that it could generalize better on the unseen data and reduce overfitting at the same time. Using different versions of activation functions for example ReLU, Leaky ReLU and regularization techniques such as dropout[2] and ridge regression makes the generalization much more better as they introduce randomness to the models but requires longer training time and much more complex model to become highly efficient for unseen data. When creating models with transfer learning, it does wonder to our validation accuracy.

4.2 Results

Table 1 shows the training and validation accuracy with respect to the models created whereas the figures depict the loss and accuracy curves of the best custom and transfer learning model.

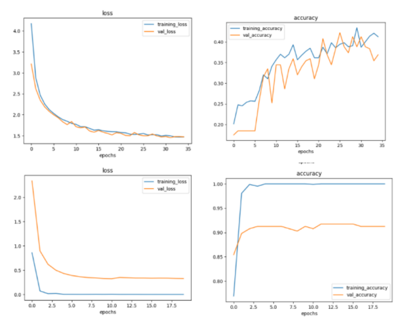

Figure 1: Loss and Accuracy Curves

Models	Epochs	Train Acc.	Val Acc.
ANN	20	0.9854	0.3689
CNN	20	1.0000	0.3689
CNN Regularized	20	0.3871	0.3447
CNN Regularized 2	35	0.4126	0.3689
VGG 16 FE	20	1.0000	0.9126
VGG 16 FT	20	0.98	0.6942

Table 1: results of different models

4.3 Discussions

When testing the models with completely unknown images of different dog breeds customize transfer learning models perform comparatively better than custom models and give a better probability of a dog being a particular breed.

5 Conclusion

Although the base models and regularized model perform less effectively on the unseen data due to resource limitations, they perform incredibly well while leveraging transfer learning.

References

[1] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.

[2] Nitish Srivastava, Geoffrey E. Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov.