

Image Classification on Caltech-256 Dataset

Springboard Data Science Capstone Report



Florian Sestak

Overview

- Approach
- Caltech256
- Models
- Problems
- Conclusion

Approach

- Analyse the dataset
- Train different models on a small part of the dataset
- Take the best performing model and train it on the whole dataset

Caltech256

- http://www.vision.caltech.edu/Image_Datasets/Caltech256/
- 257 Categories
- 30607 images
- 15423 different resolutions

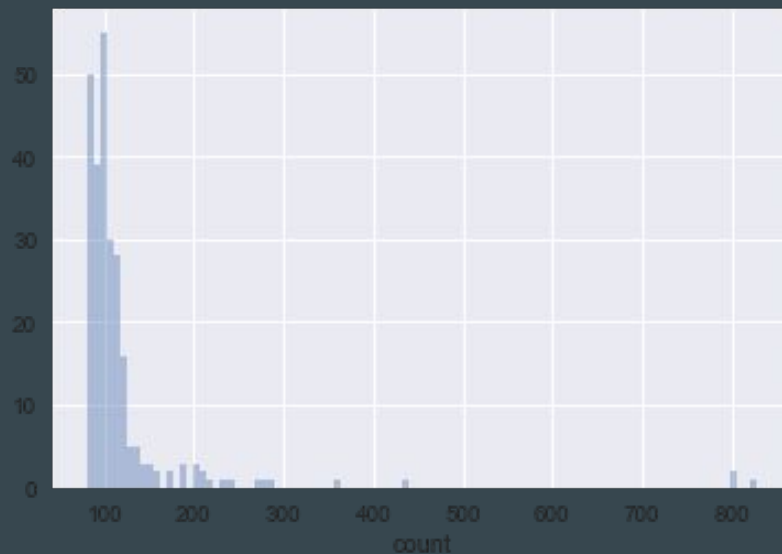
<i>Dataset</i>	<i>Released</i>	<i>Categories</i>	<i>Pictures Total</i>	<i>Pictures Per Category</i>			
				<i>Min</i>	<i>Med</i>	<i>Mean</i>	<i>Max</i>
Caltech-101	2003	102	9144	31	59	90	800
Caltech-256	2006	257	30608	80	100	119	827

http://www.vision.caltech.edu/Image_Datasets/Caltech256/details.html

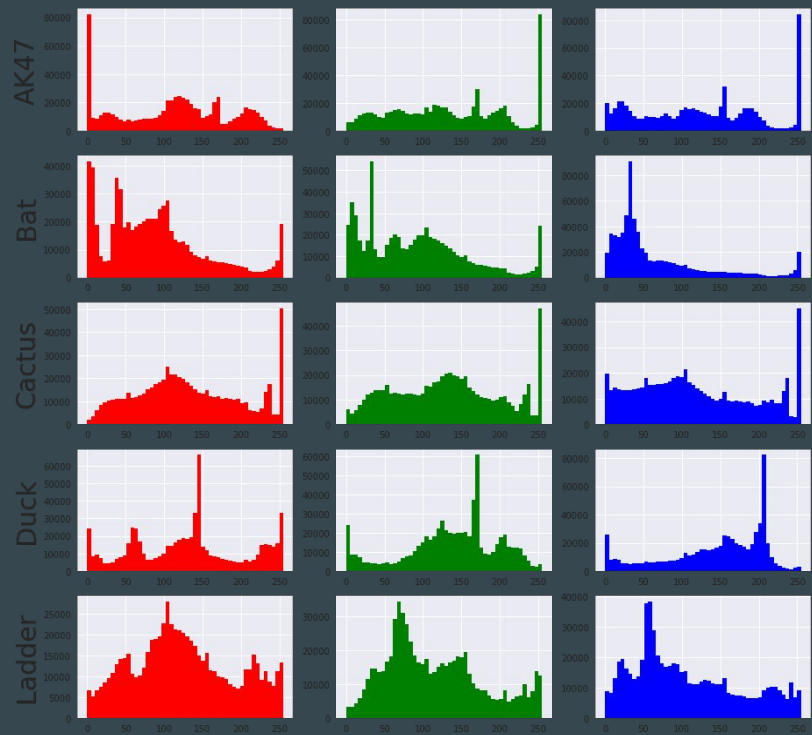
Caltech256 - Samples



Caltech256 - Class distribution



RGB - Distribution of different classes



Models

- Trained on 10 classes
- Comparing the models
- Choose final model and train it on whole data
- K

Models

Library	Model	Train - Accuracy	Test - Accuracy
sklearn	K - Nearest	0.40	0.42
sklearn	SVM	0.19	0.16
Keras	CNN	0.64	0.51

CNN - Final Model

- VGG16 architecture (<https://arxiv.org/pdf/1409.1556.pdf>)
- Pre-trained weights (ImageNet)
- Replaced the last three dense layers
- Freezed the other layers
- Learning Rate, Decay and Momentum based on Paper
- Loading data in batches from directories

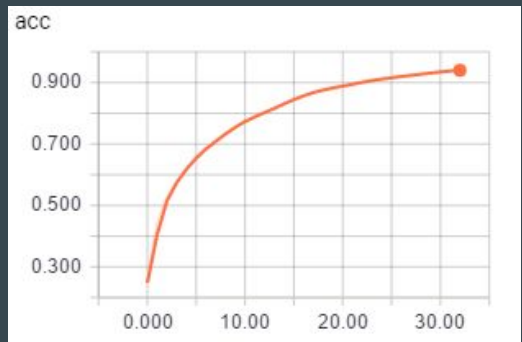
CNN - Final Model - Parameters

- Epochs 32
- Batch_size: 50
- Steps_per_epoch: $\text{training_samples} / \text{batch_size} = 492$
- Stochastic Gradient Descent
 - Learning rate: 0.01
 - Decay: 0.0005
 - Momentum: 0.9
 - Nesterov

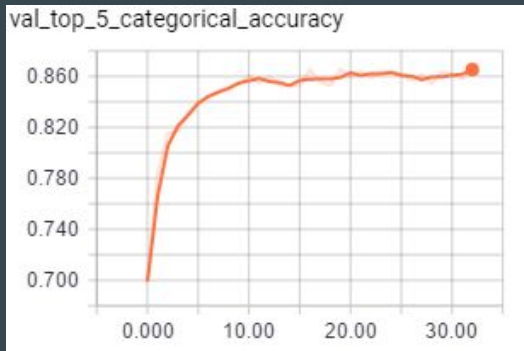
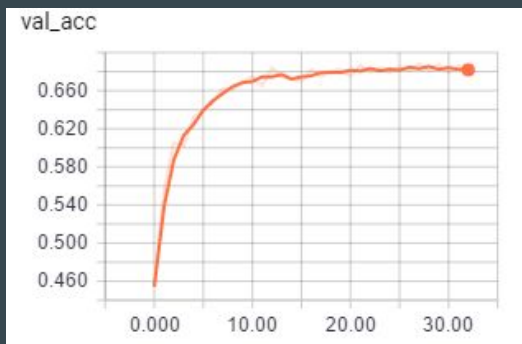
CNN - Final Model - Image Augmentation

- `rotation_range=40,`
- `width_shift_range=0.2,`
- `height_shift_range=0.2,`
- `rescale=1./255,`
- `shear_range=0.2,`
- `zoom_range=0.2,`
- `horizontal_flip=True,`
- `fill_mode='nearest'`

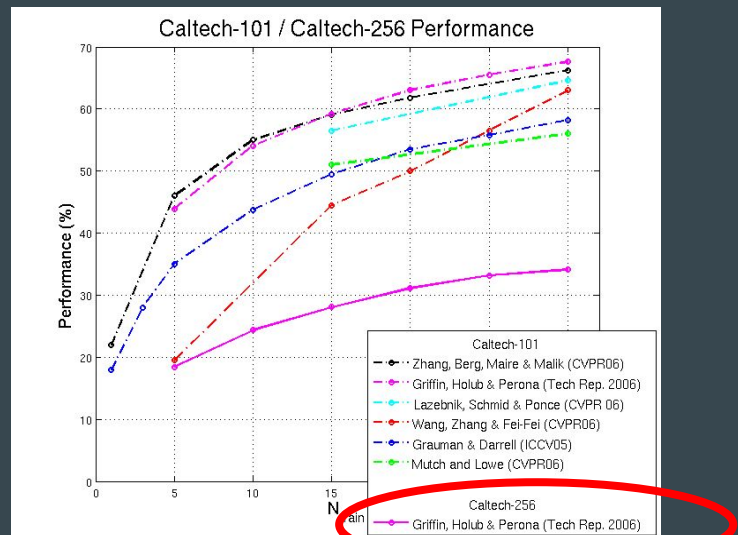
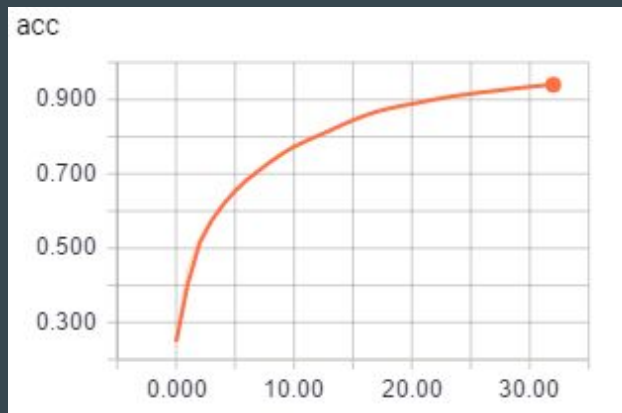
CNN - Final Model Result



Accuracy:	0.93
Validation Accuracy	0.68
Top 5 Validation Accuracy	0.87



CNN - Final Model - Compared with benchmark



http://www.vision.caltech.edu/Image_Datasets/Caltech256/performance/performance.png

Problems

AK47



People



Problems

Billiards



Conclusion

- Beat Caltech256 benchmark
- Improvement with cleaner data
- More data augmentation
- Very easy with Keras
- Big training time improvement on GPU

Thank you