Can We Spot The Right Spot

Ridley Arp

Machine Learning – Metis

April 20th, 2022

Better Together

- Prior Classification Project
  - Breed
  - Tumour Type
  - Location
- Breed information was considerably sparse
  - More Affordable
  - Better Preventative

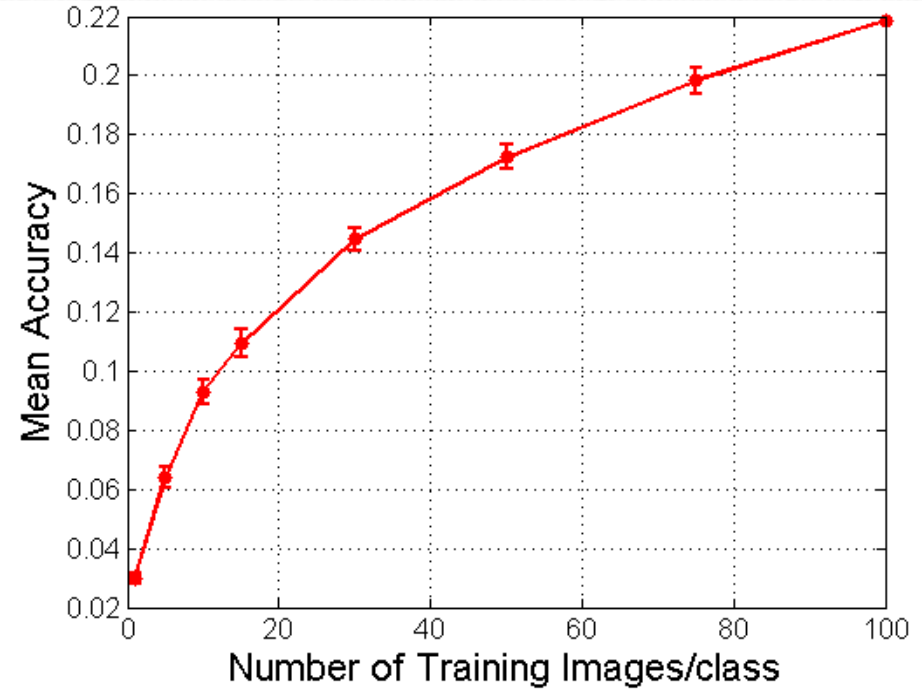
Primary Goal

- Train a CNN to identify individual breeds among various dogs
- Eventually provide a spectrum output to provide probable breed mixes for mutts

Dataset & Methods

- **Stanford Dogs Dataset**
  - 20,580 Images
  - 120 Labeled Breeds
  - 150+ Images/Breed

Stanford Start



- Prior work on Dataset offered some starting help on base metrics
- Data had a test/train split inbuilt

*Aditya Khosla, Nityananda Jayadevaprakash, Bangpeng Yao and Li Fei-Fei. **Novel dataset for Fine-Grained Image Categorization.** *First Workshop on Fine-Grained Visual Categorization (FGVC), IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011. [pdf] [poster] [BibTex]

Model Selection

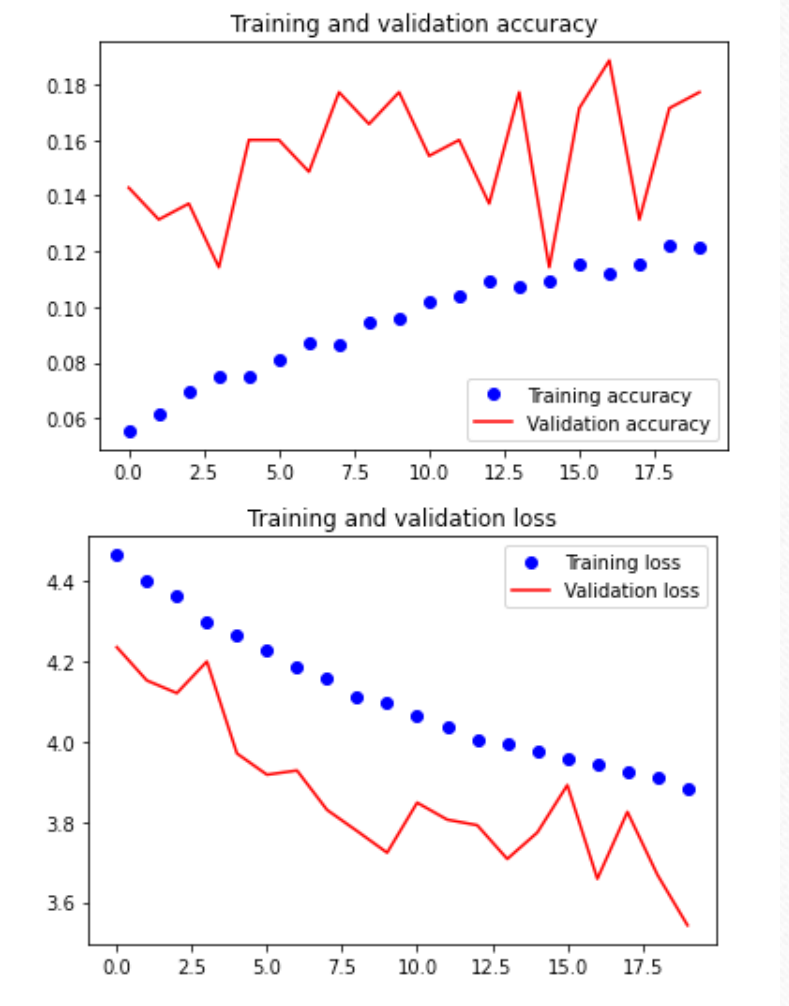
- VGG16 ImageNet weights transferred in as the convolutional base
  - Images originated from the ImageNet Dataset
  - Ease of implementation with transfer learning
- Optimizers
  - RMSprop
  - Adam
- Image Augmentation

Results and Reflections

The Good, The Bad & The Better

- Initial Models were incredibly inaccurate
  - Accuracy never exceeded 5%
  - Convergence after 4 Epochs
  - Lack of data required many hyperparameter tunings

Later Models show significant promise:



Final Results

Despite significant augmentation and tuning, the VGG16 base is only able to meet the Stanford at just over 20%

Layer (type)	Output Shape	Param #
vgg16 (Functional)	(None, 3, 3, 512)	14714688
flatten_10 (Flatten)	(None, 4608)	0
dense_23 (Dense)	(None, 576)	2654784
dropout_9 (Dropout)	(None, 576)	0
dense_24 (Dense)	(None, 120)	69240

Total params: 17,438,712
Trainable params: 17,438,712
Non-trainable params: 0

Final Stanford Accuracy: .22
Final VGG16 Accuracy: .208

Next Steps

- The limitations of computational availability as well a additional data may have rendered VGG16's convolutional base less effective. Next steps would be Resnet50 with less input image reduction.
- Look at the effect of including known mixed-breeds (e.g.Puggle:Beagle/Pug)
- Build interface to receive an input image and output the soft probability predictions for that image