# Who's a good dog?

Using Convolutional Neural Networks to classify dog breeds from images


Garrick Chu (May 2018)

Springboard Data Science Career Intensive Track


# Executive Summary

- Advancements in CNNs and Transfer Learning allow social media platforms to gather more information about its users.
- Opportunities to capitalize on existing image data by serving better ads, improved content suggestions, deeper depth of user demographics.
- Thanks to such advancements, models can be built and deployed with relative ease and can achieve/surpass human-level accuracy.
- With Transfer Learning, model can achieve ~60% accuracy and outperform humans by 30%.[1]

[1] Based on informal study conducted by author

# Can we capitalize on existing image data?

By the numbers…

- America's pet industry is estimated around $69.51 Billion, as of 2017, and estimated to grow by 3.7% to **$72.13 Billion** in 2018. [1]
- On social media, a sponsored post can command around $2,000. [2]
- Sponsored posts by high-profile pets can range from $10,000 to $15,000. [4]
- Boo the Pomeranian has 17 million followers on Facebook and earns $1 Million USD per year. [3]
- Grumpy Cat's Net Worth: ~$100 million. [3]

Sources: 1. Pet Industry Market Size & Ownership Statistics. http://www.americanpetproducts.org/press_industrytrends.asp
2. "The highly profitable, deeply adorable, and emotionally fraught world of Instagram's famous animals", Quartz. February 16, 2016. https://qz.com/612796/the-highly-profitable-deeply-adorable-and-emotionally-fraught-world-of-instagrams-famous-animals/
3. "20 Pets That Make Millions for Their Owners", Huffington Post. July 8, 2015. https://www.huffingtonpost.com/gobankingrates/20-pets-that-make-million_b_7755792.html
4. "Inside the glorious and lucrative world of Instagram's famous pups", Mashable. August 26. 2017. https://mashable.com/2017/08/26/instagram-famous-pets-money/#mv86uReL4mqw

# How can we capitalize?

- Social media platforms such as Facebook and Instagram can extract user data from their uploaded images and gather demographic data when previously not explicit (perhaps irrelevant or lack of caption or hashtags).
  - Serve better/more personalized ads and follow suggestions, increasing engagement
  - Help influencers expand their audience
  - Further expand platform’s ad audience
- Manually finding and labelling these images of dogs would be a massive undertaking…
- Can we automate and perform better than humans?

# Approach

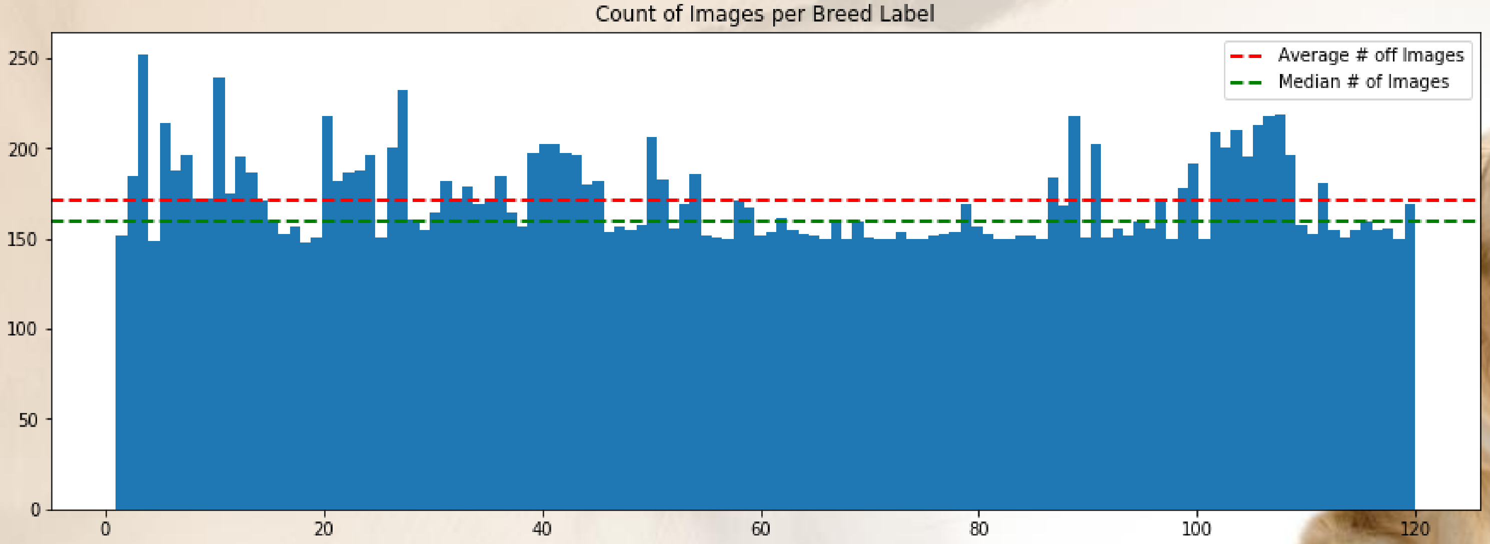
- Leverage use of Convolutional Neural Network and Transfer Learning
  - Significant advancements in field of computer vision
  - CNNs are explicitly designed for image data
  - Very similar to neural networks but uses a convolutional layer among the hidden layers
- Objective: Compare Neural Network performance on classifying dog breeds from images
  - Perform better than baseline and human

# Introduction to Data

- Data for training, validation and testing was obtained from the Stanford Dogs Dataset (Aditya Khosla, Nityananda Jayadevaprakash, Bangpeng Yao, and Li Fei-Fei )
  - http://vision.stanford.edu/aditya86/ImageNetDogs/
- Data contains 20,580 images from ImageNet
- Across 120 dog breeds
  - Rough representation of the 167 breeds currently recognized by the American Kennel Club

# Exploratory Data Analysis

- The image data appears to be fairly balanced with about 160 images per breed.  The fewest samples of a breed is 150 images and the most is  about 250 images.

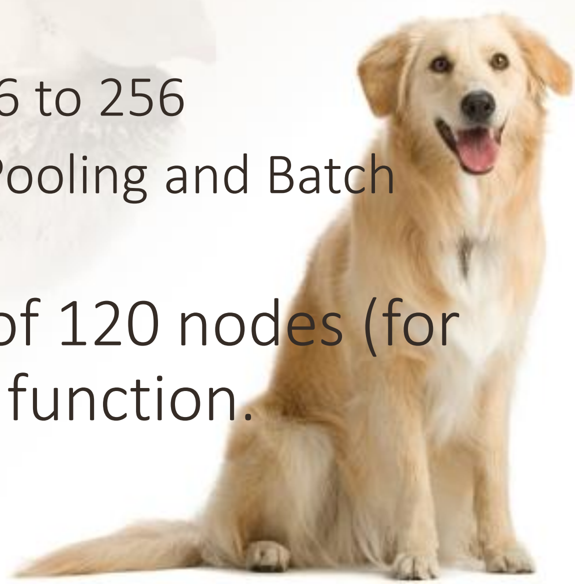


# Disclaimers

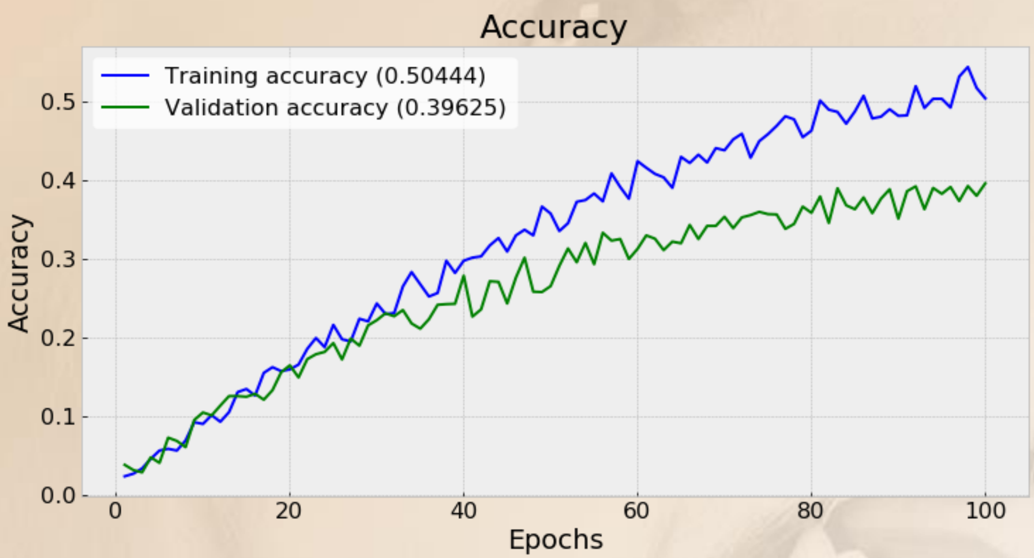
- Training a neural network requires a significant amount of computational power.
- Faster training and more accurate models could have been achieved by outsourcing the training tasks to a distributed computing system (such as an Amazon EC2 instance).

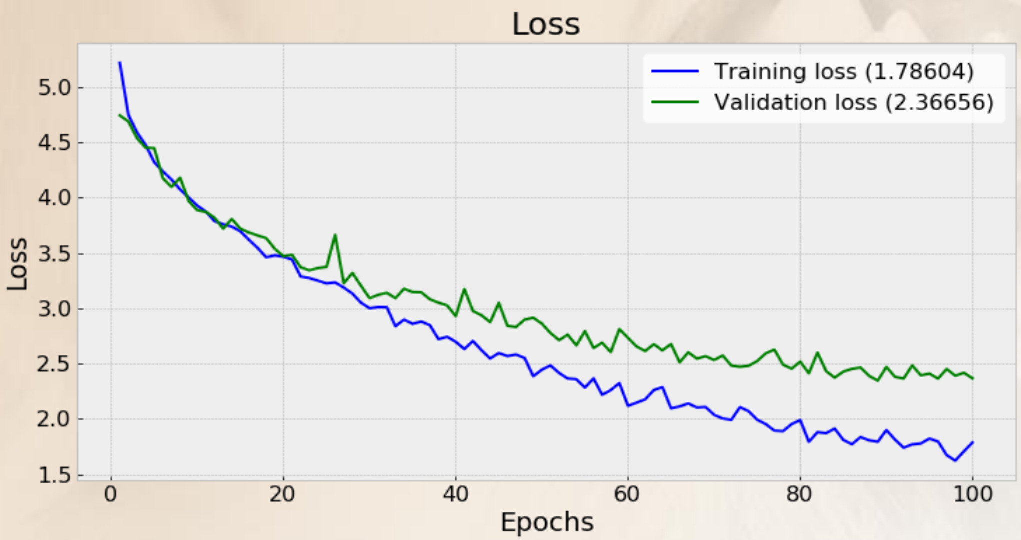
# Method #1: Convolution Neural Network from Scratch

- To establish a baseline before leveraging pre-trained models, let's evaluate a CNN built from scratch without the use of pre-trained weights.
- Model architecture:
  - – Input Layer: Batch Normalization
  - – 5 Convolutional Hidden Layers
    - Increasing magnitude of kernels from 16 to 256
    - Each paired with Dropout of 20%, MaxPooling and Batch Normalization to prevent overfitting
  - – Output layer: Fully Connected layer of 120 nodes (for each breed) with SoftMax activation function.



# CNN from scratch – Results

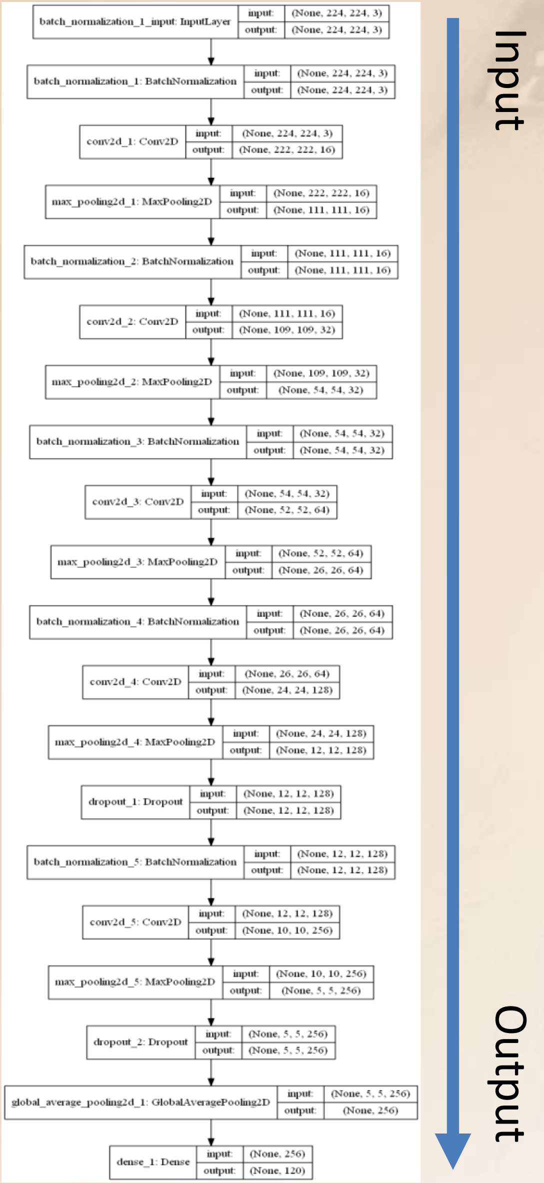



- After 100 epochs of training, our CNN model achieved an accuracy on validation images of 39.6% and a categorical cross-entropy loss of 2.36.

- On unseen test images, the model achieved an accuracy of 40.3% and a loss of 2.29.

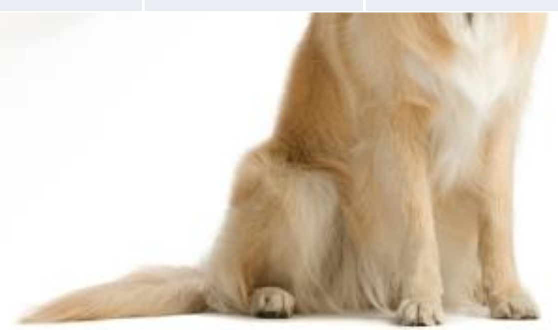
# More on CNN from Scratch…




This CNN model from scratch was result of 2 major prior iterations….

| CNN from Scratch #1 | After 20 Epochs | …After 29 | …After 39 | After 100 |
|---|---|---|---|---|
| Validation Loss | 3.13 | 3.07 (-0.06) | 2.96 (-0.11) | (n/a) |
| Validation Accuracy | 21.7% | 23.9% (+2.2%) | 25.4% (+1.5%) | (n/a) |
| | | | | |
| CNN Scratch #2 | | | | |
| Validation Loss | 3.5* | 3.25* (-0.25) | 3.0* (-0.25) | 2.48 (-0.52) |
| Validation Accuracy | 15%* | 20%* (+5%) | 24%* (+4%) | 36.0% (+12%) |
| | | | | |
| CNN Scratch #3 | | | | |
| Validation Loss | 3.5* | 3.3* (-0.2) | 3.0* (-0.3) | 2.36 (-0.64) |
| Validation Accuracy | 16%* | 21%* (+5%) | 25%* (+4%) | 39.6% (14.6%) |

*Visual estimation based on plotted values

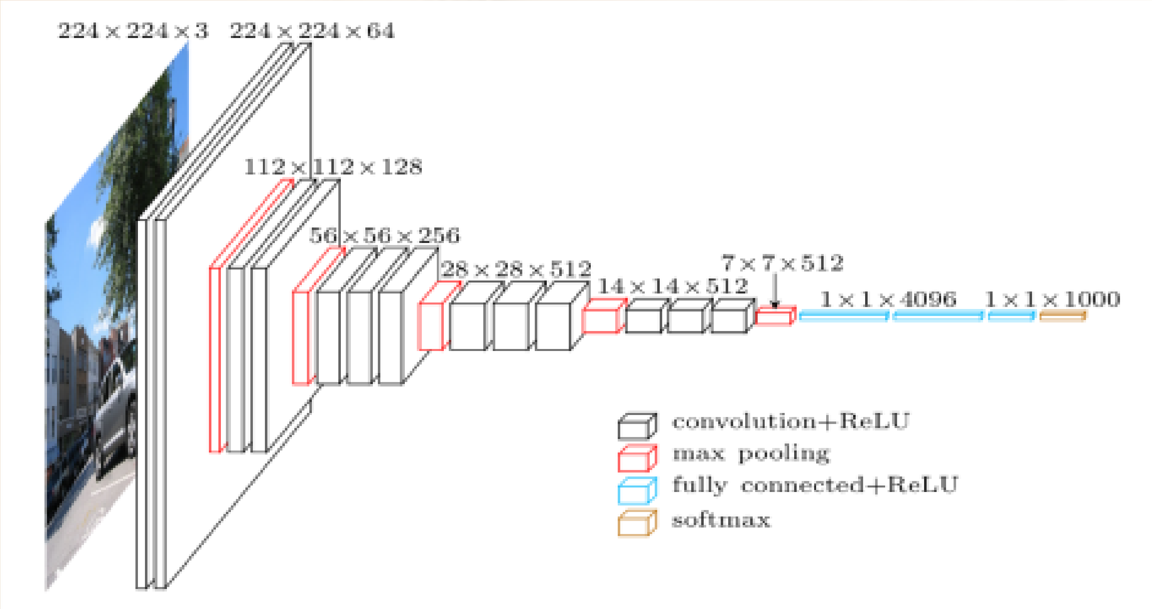


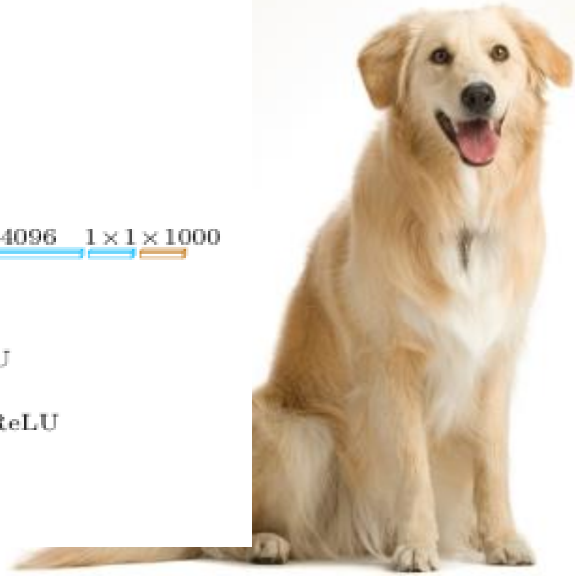
# Leveraging Transfer Learning

- What is Transfer Learning?
  - Using an existing CNN architecture with trained model weights.
  - Model weights derived from ImageNet, which is a database of 1.2 million images and trained over several weeks across multiple GPUs.
- Benefits of Transfer Learning:
  - Model can be repurposed for one's own application.
  - High degree of customization.
  - Effective with small datasets.
    - This is the case here with only ~100 images per class available for training.

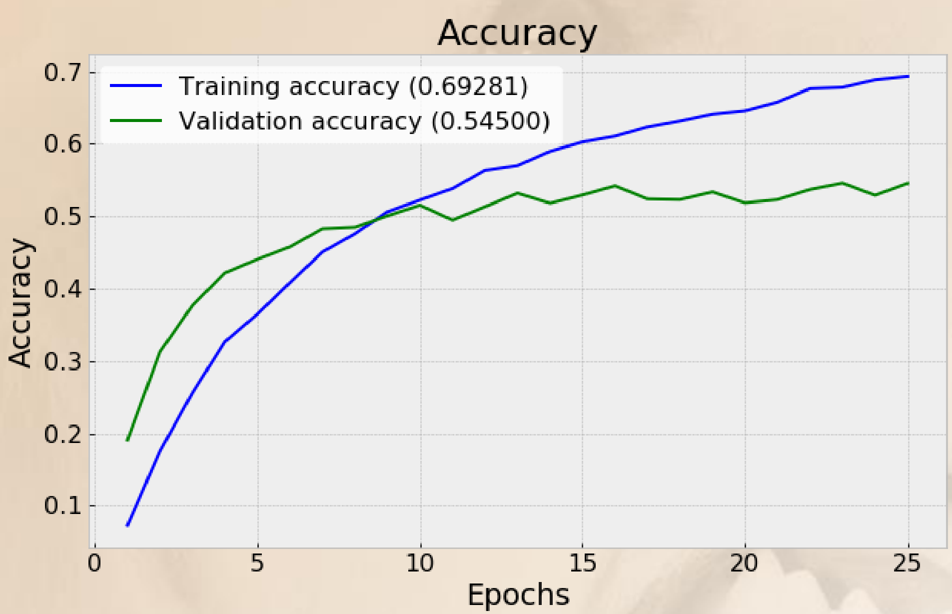
# Transfer Learning with VGG16

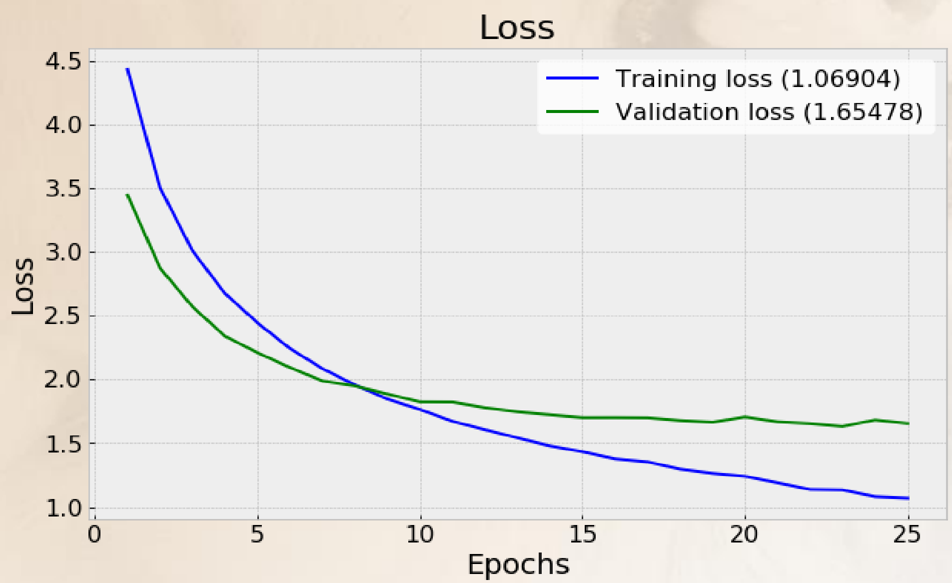
- Utilizing the Visual Geometry Group's VGG16 (with Batch Normalization) model with weights trained from ImageNet.
  - Popular choice for an array of image classification applications
- Repurposed default output FC layers (of 4096 and 1000 nodes) with own FC layers of 256 nodes and 120 nodes (for each respective breed).



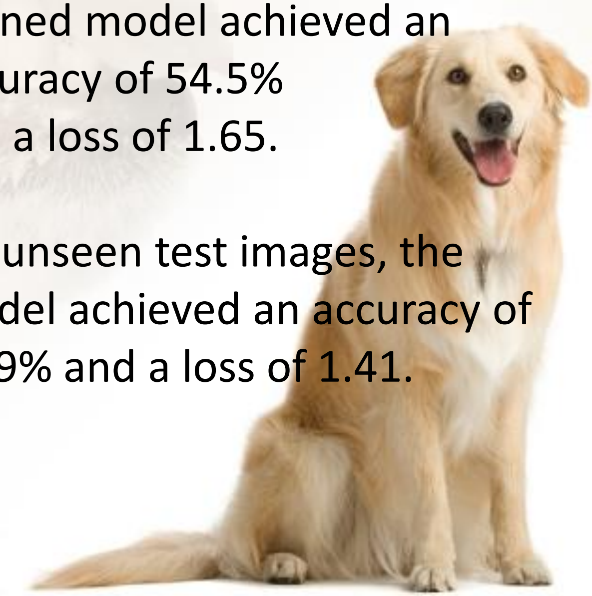



# VGG16 Model – Results




- Only 25 epochs of training were required to reach this accuracy level (as compared to 100 epochs for 39.6% accuracy).
- On validation data, the pre-trained model achieved an accuracy of 54.5% and a loss of 1.65.
- On unseen test images, the model achieved an accuracy of 58.9% and a loss of 1.41.



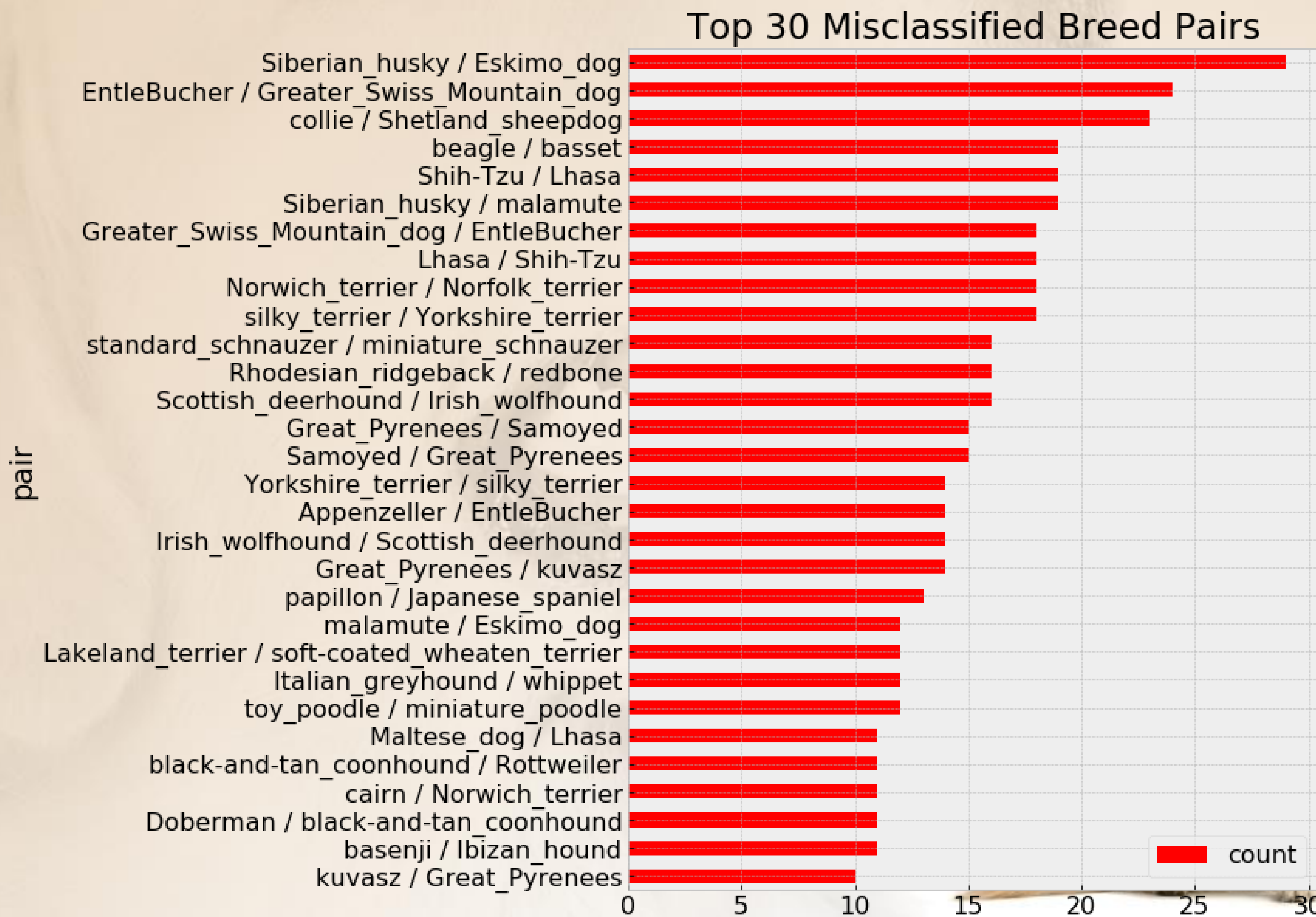
# Results Comparison

| | CNN from Scratch | Pre-Trained VGG16 | (Difference) |
|---|---|---|---|
| Validation Accuracy | 39.6% | 54.5% | +14.9% |
| Validation Loss | 2.36 | 1.65 | -0.71 |
| Test Accuracy | 40.3% | 58.9% | +18.6% |
| Test Loss | 2.29 | 1.41 | -0.88 |

Note: A lower loss value is desired.

# Where does our CNN fall short?

Let’s examine the Top 30 Misclassified breed pairs…



# Top misclassified breed pairs

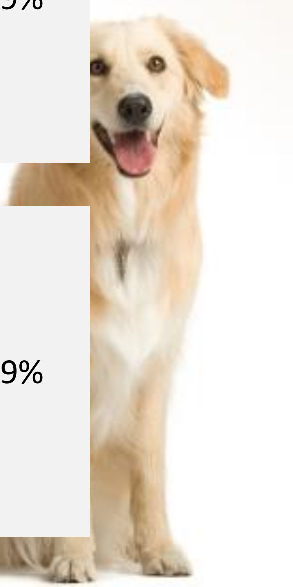
Husky (above) vs.
Eskimo Dog (below)

EntleBucher (above) vs.
Greater Swiss Mountain Dog
(below)

Collie (above) vs.
Shetland Sheepdog
(below)

# Model Precision, Recall and F-1 Score

| Classification Metric | Lowest | Highest | Average |
|---|---|---|---|
| **Precision *(TP/TP+FP)*** |  Briard – 31% |  Komondor – 89% | 60% |
| **Recall *(TP/FN + TP)*** |  Collie – 21% |  Norwegian Elkhound – 91% | 59% |
| **F1-Score *(2) / (1/Precision + 1/Recall)*** |  Collie – 28% |  Komondor – 88% | 59% |



# Prime Examples of Class Label as determined by model

## Popular US Dog Breeds


Pug


Pembroke Welsh Corgi


Labrador Retriever

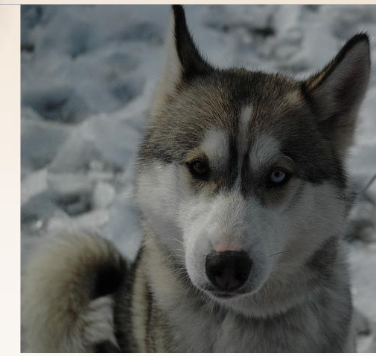

German Shepherd


Golden Retriever

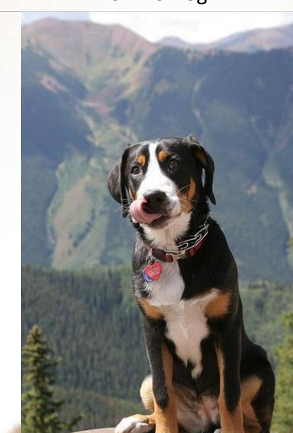

French Bulldog

## Top Misclassified


Siberian Husky


Eskimo Dog**


EntleBucher


Greater Swiss Mountain Dog

# Model Performance vs. Humans

- Conducted informal study in quiz format.
  - Images consisted of both popular and less popular breeds, contained equal mix of breeds that model excelled at and struggled with.
  - Contained only images from test data set (and no training data)
  - Responder presented with 10 images, 5 multiple choice answers.
    - Multiple choice answers consisted of 1 correct breed, 2 breeds which the model most often confuses, and 2 breed based of random number generator.
  - Sample Size: 51 responders

- Tested the same images on model

Link to quiz: https://goo.gl/forms/pxUGJwv3IGZ9xNNj1

# Results

| Correct Label | Human Prediction* | Model Prediction** |
|---|---|---|
| Papillion | Papillion (70.0%) ✓ | Papillion (99.50%) ✓ |
| Eskimo Dog | Husky (90.0%) ✗ | Saluki (18.06%) ✗ |
| EntleBucher | Greater Swiss Mountain Dog (40.0%) ✗ | EntleBucher (86.99%) ✓ |
| Cairn | Norwich Terrier (76.0%) ✗ | West Highland Terrier (54.76%) ✗ |
| Malamute | Malamute (46.9%) ✓ | Malamute (51.96%) ✓ |
| Gordon Setter | Brittany Spaniel (62.0%) ✗ | Gordon Setter (97.43%) ✓ |
| Border Collie | Shetland Sheepdog (50.0%) ✗ | Border Collie (45.00%) ✓ |
| Shih-Tzu | Shih Tzu (84.0%) ✓ | Shih Tzu (15.16%) ✓ |
| French Bulldog | French Bulldog (80.0%) ✓ | French Bulldog (97.15%) ✓ |
| Tibetan Mastiff | Tibetan Mastiff (70.0%) ✓ | Tibetan Mastiff (85.87%) ✓ |
| **(Overall Score)** | **Collective: 50% (Average: 44.9%)** | **80.0%** |

* Majority Response (% of all responses)

** Model Prediction with Confidence

# Conclusion

- It is possible to build a model that outperforms better than humans at identifying a dog’s breed from an image.
- Transfer Learning and Pre-Trained weights are the crux to achieving this level of accuracy.
  - Applicable to real world where data sets are small.
  - Can be repurposed to a wide array of applications.

# Further Research and Work

- Leveraging High-Performance Computing platforms (like Amazon EC2 instance) for accelerating training and model fine-tuning.
- Obtain more labeled training data
  - Original dataset includes 25,850 images equates to just ~80 images for training, ~20 for validation and ~60 for testing for each breed.
- Other pre-trained models to experiment with
  - Such as VGG19, Xception, ResNet50, InceptionV3 and many more.

# Additional References and Resources

- Github Repo: https://github.com/thegarrickchu/Springboard-Dog_Breed_Classifier

- Detailed Report: https://github.com/thegarrickchu/Springboard-Dog_Breed_Classifier/blob/master/Springboard%20Capstone%20Project%202%20-%20Final%20Report.pdf

- Email: chu.garrick@gmail.com

# Dedicated to Pepsi

April 22, 2007 – April 7, 2018