

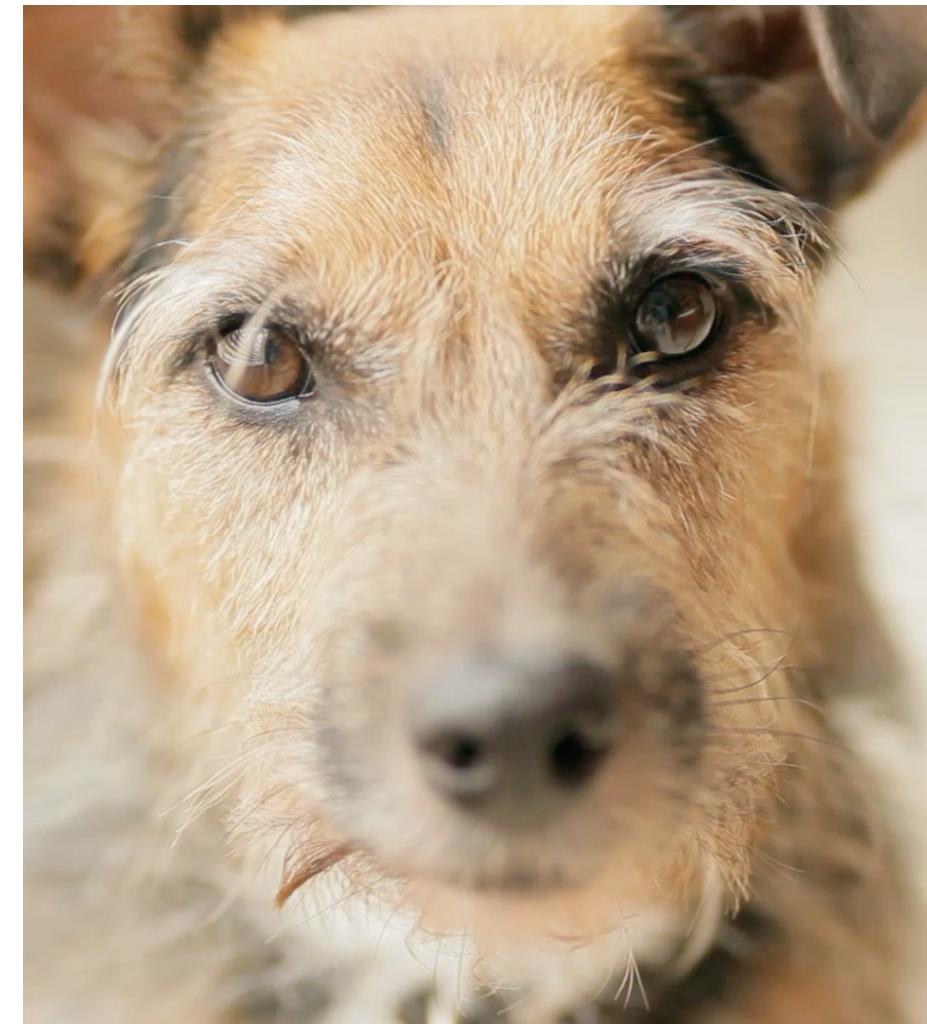


Using CNN to classify dog breeds

Helping animal shelters estimate dog breeds with more confidence

Problem Statement

- Approximately 2 million dogs are adopted each in the US each year.
- Between 75% - 80% of dogs in shelters are mixed breed
- And as many as 25% of dogs in shelters are purebred dogs
- A study found that when trying to identify the primary breed in a mixed dog, shelter staff got it right 56% of the time. But when they tried to identify the secondary breed, their success rate dropped to 10%.
- Correctly identifying breeds can be a matter of life and death at kill shelters. Dogs lacking any Pit Bull DNA were mislabeled as such roughly 50%.



Can CNN Improve classification rates?

Resources at dog shelters are scarce, providing each dog with a DNA test is not feasible. However, I set out to train a model that can improve the classification rates of shelter dogs.



Training and Evaluating the Dog Classifying model

Data

The Stanford Dog Data –
Includes 120 dog breeds
from around the world

The model – InceptionV3

Transfer learning was used
to structure the model.

I fine tuned the model to the
dog data and use image
augmentation, other
regularization techniques

Analysis of single breed predictions

Globally – the model
achieved an accuracy
scores of 86%

Analysis of mixed breed predictions

When tested with dogs of
two breeds, the model, in its
top 10 predictions, the
model made at least one
match 88% of the time

Recommendation for next steps

There is so much room for
improvement here, I have a
recommendations for next
steps

The Data

About 21,500 images in total, relatively balanced. Each breed had at least 150 images and on average, we had about 180 per class.



Cairn



French Bulldog



Chihuahua



Cocker Spaniel

Transfer Learning with InceptionV3

InceptionV3

- Benefit over other pretrained models is that it's design to lower the parameter.
- Trains relatively quickly
- Train in over 21000 images in less than two hours

General Structure

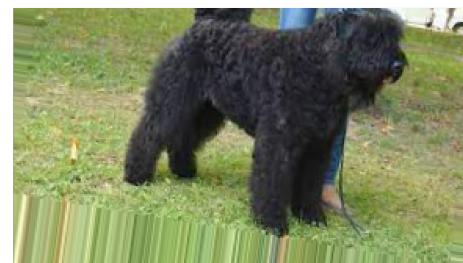
- InceptionV3
- Last layer of softmax
- Early Stopping
- ReduceLROnPlateau which is used to reduce the learning rate once learning stagnates



Image Augmentation

Image augmentation was made easy with TF ImageDataGenerator

- Flip
- Shifts
- Shear range
- Zoom range



Model Performance – Single breed classification

Globally

- Model enjoyed 86% accuracy score

Breed by breed

- Over 60% of breeds had recall scores of 0.90 or above
- Only 9 breeds had recall scores of less than 0.50

Misclassifications – Too similar

Brittany Spaniel



Welsh Springer Spaniel



Misclassification – Same breed group

Toy Poodle



Miniature Poodle



Misclassification – Botched data

Eskimo Dog



Eskimo Dog????!!!



Samoyed



Model Performance – Mixed breed classification

Any Matches

- Out of the top ten predictions, the model correctly matches at least one of the two breeds 88% of the time.

All Matches

- Out of the top then predictions, the model matches both breeds at a rate of 30%.

Conclusions

- When used as a single breed classifier, the model performs well with an 85% accuracy score, and a global precision score of 86%
- When classifying pit bulls, shelters tend to have a very low precision score, approximately 50%. Whereas my model was able to identify bullterriers with precisions scores of 71% or more and can provide more confidence when estimating a breed.
- Shelters are not great at classifying mixed breed dogs, and with my model matching at least one breed (in the top ten predictions), 86%, and matching both breeds with approximately 30% rate, I think it can be a useful tool to give shelters a little more confidence when making guesses.



Recommendations for next steps

Clean the data

Do a thorough clean-up of the data. Remove other breeds and make sure only the right breed is included.

See if model improves

Cleaner data should produce better predictions, right?

Get more mixed breed data

I want to properly collect data for mixed breeds that is properly confirmed

Build a proper multilabel model

Mixed breed dogs deserve a real try at predicting their breeds

Thank you

