DS 6050 Project Report

Project Title: **Sit, Stay; Training the Dog (Model)**

Group 6 Project Members:

Lauren Bassett, jme9rt

Cullan Bedwell, qja3fh

Maggie Houck, meh7za

Frank Vasquez, fav3ba

**Abstract**

Pet owners in the United States and all around the world are interested in knowing their pet’s genetic makeup, so much so that the animal genetics industry is valued at billions of dollars. There are existing ways to use DNA to understand your pet’s genetic makeup, though it can be prohibitively expensive for the average household. Outside of DNA sequencing research, there is some existing machine learning work that allows you to predict the breed of dogs based on image; however, that work focuses specifically on training and testing images of purebred dogs. Since the majority of dogs are not purebred, this project focuses on identifying the breed makeup of mixed breed dog images. In order to do so, an existing widely used dataset of purebred dog images was used, then CutMixing data augmentation was performed and trained on. The resulting model actually had lower accuracy than the model that was not trained using CutMixing, however, it did not overfit nearly as much. Additionally, it allowed the model to predict more than one dog breed, which is necessary for mixed breed dog identification. Overall, the resulting accuracy was not as high as hoped, but these methods provided promising results and could be seen as a stepping stone for future similar work.

**Motivation**

Animal Genetics Research is a rapidly growing industry. According to a report released by Global Market Insights, the industry is projected to be valued at 6.5 Billion USD by 2027 (Ugalmugle & Swain, 2021). Dog DNA testing services determine the genetic makeup of your dog, specifically, what type of mixed breed, in order to determine potential health risks, explain the dog’s behavior, and provide other insights. However, these kits are not cheap. Dog DNA tests can cost upwards of $70 per test.

Our project hoped to expand on the pre-existing data modeling set out in the Stanford Dog Dataset to determine whether or not we could train a neural net to identify the potential breeds in a mixed breed dog based on images. This architecture could potentially allow dog owners to better understand their dogs’ mix of breeds without having to go through the hassle of paying for a genetic testing kit. We knew that this service most likely would not be as accurate as a full DNA test, but we hoped to see if there is a potential for a lower-cost alternative to determine the possible ancestors of mixed breed dogs based on photos alone. If successful, this type of lower-cost or even free alternative could be implemented via a mobile application and made accessible to people everywhere.

**Methods**

*Cut Mixing for Data Augmentation*

CutMixing is an approach to data augmentation that was implemented in order to increase the generalization and classification accuracy for mixed breed dogs. CutMixing works by overlaying parts of an image on top of other images, effectively mixing their features. This is in contrast to more basic data augmentation methods that slightly change the values of the individual pixels. By putting parts of images on top of other images and keeping the associated labels associated with said pixels, we essentially created a new augmented dataset of synthetic mixed breed dog images with features from both images. This method was first proposed in *CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features* (Yun et al., 2019). While the mixed data created by the CutMix algorithm is obviously altered to the human eye, this type of data augmentation causes a neural network to learn a mix of latent features with mixed labels for each image, ultimately lending itself well to the goal of mixed breed dog classification.

****

*MobileNetV2 Architecture*

MobileNetV2 is the convolutional neural network (CNN) architecture that was used at the backbone of the models. This architecture is based on an inverted residual structure. Within this architecture, the CutMixing and a normal version were both implemented.

In order to make the layers of the MobileNet trainable, we did not use the imagenet parameters. We also stacked on a pooling layer, a dropout layer, and a dense layer before the softmax output layer.

**Experiments**

*The Pitbull Problem*

We collected the mixed breed images and DNA results from employees at Capital One and the MITRE Corporation, and were able to identify 27 individual dogs, with 41 total photos. These 23 dogs consisted of 42 specific breeds. 27 of these breeds can be found in the Stanford Dataset, but the remaining 14 breeds are not. However, each mixed breed dog we are analyzing has at least one main breed that is included in the Stanford dataset. The most critical missing breed is the pitbull terrier. The Stanford Dogs Dataset does not have any training data on pitbull terriers.

**Results on Stanford Dogs Dataset**

Before moving on to examining the performance on mixed breed dogs, we examined how both the baseline and CutMix models performed on the Stanford Dogs dataset. Overall, we saw that the baseline model had a higher accuracy than the CutMix model; however, it had a significant degree of overfitting and was likely to predict a single class with high probability in spite of uncertainty. This can be illustrated in the drop in accuracy from 99.98% in training to 76.63% in validation for the baseline model. Meanwhile, the CutMix model had a lower accuracy, 76.40% in training and 67.97% in validation. However, it isn’t as overfit and is much more likely to give multiple classes a chance in the face of uncertainty.

Additionally, both models increased in accuracy and got more similar as more predictions were introduced. The top 3 and 5 accuracies for the baseline model were 92.71% and 95.99%, and they were 87.09% and 92.22% in the CutMix model.

Given our main application on predicting on the average, mixed breed dog, these results give hope to good results on our completely new data. These drop in accuracies were on purebred images with minimal obstructions and good lighting and framing. With CutMix being better at showing uncertainty in its predictions and the more similar top k accuracies, we have hope in it performing better in our next analysis.

**Coworker’s Results**

To bring in mixed breed dogs to test, we gathered DNA results from 27 different dogs. These dogs had 41 different breeds in their genetic makeup. Out of these breeds, 14 are not included in the Stanford Dogs training data, which influenced the results. Looking at the results of the predictions, we are able to identify some weaknesses in the model.

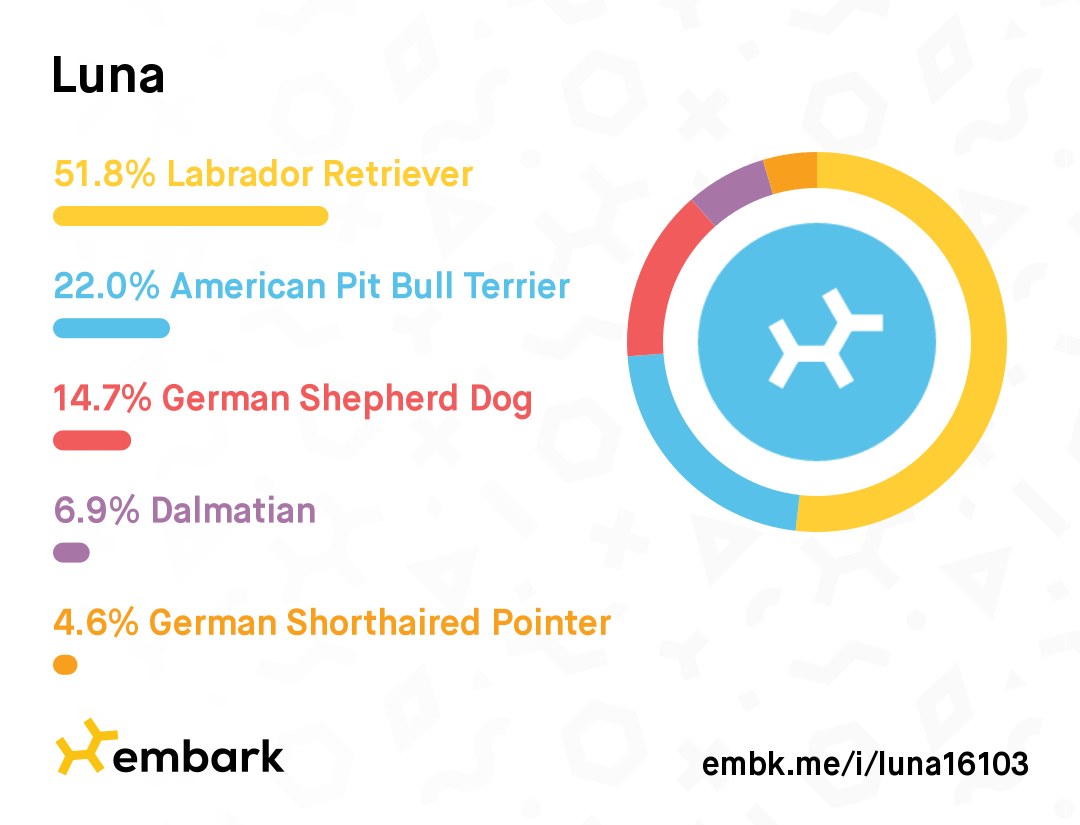
*Scale:* Photos of small dogs, such as chihuahua mixes, without any objects to reference the size of the animal were often mislabeled.

Uncommon Predictions: Many mixed breeds have similar characteristics to pure-bred dogs. The Dingo, a wild dog indigenous to Australia, was commonly predicted. In reality, the likelihood of a mixed breed dog having Dingo is very small.

Covering the body: Around 20% of the mixed breed images included dogs wearing some sort of clothing, harness, or costume which changed the color and texture of their coats. This could possibly influence the results.

Overall, the model was usually able to predict at least one of the breeds present in the images provided by the coworkers, but the results were not as accurate as we would have hoped. However, there are clear improvements that can be made to the images provided to the model that help improve the prediction accuracy.

**Luna’s Results**



We ordered a DNA test kit through Embark. The results show Luna is primarily Labrador Retriever, followed by American Pit Bull Terrier, German Shepherd, Dalmatian, and German Shorthaired Pointer.

American Pitbull Terrier and Dalmatian are not present in the Stanford Dogs dataset. To see how different features in images affected the results, we looked at several photos of Luna in different situations. Experimenting with the lighting, angles, features visible, and environmental factors had an affect on how the model predicted her breed makeup.

**Photos with Accurate Results:**

|  |  |  |  |
| --- | --- | --- | --- |

The pictures of Luna that were the most accurate either showed her full body, or clearly showed the depth of field in her face and snout. These photos were correctly identified mostly as labrador retriever, some terrier breed, and German Short-Haired Pointer.

**Photos with inaccurate Results:**

|  |  |  |  |
| --- | --- | --- | --- |

These pictures were not identified correctly. In the first image, her face is not visible at all. The second picture is backlit, and there is not much depth-of-field in her snout. In the third image, her body is covered in a t-shirt, which changes the color and texture of her coat. In the last image, the snow disrupts her fur pattern. The final image was actually identified as a husky mix, which *may* be impacted by the snow as well.

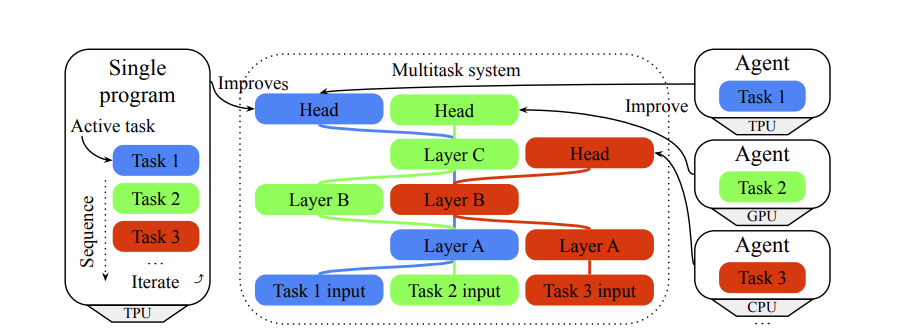
The algorithm never detected German Shepherd, which makes up 14.7% of her genetic makeup, but Luna does not have any physical characteristics of that breed. This is a limitation of this algorithm. The physical recessive traits of mixed breeds will not be identified by this algorithm. Anecdotally, some of the predicted breeds are breeds that Luna gets compared to often. Specifically, black and tan coonhound, greyhound, and doberman. Even though Luna does not have any of these breeds DNA, it is interesting that the model predicted breeds she is frequently compared to.

**Future Improvements - Predictions**

1. *Add More Common Breeds to the Dataset* - Missing common breeds such as the pitbull in the training data makes the model run at a huge disadvantage. Including more common breeds to the training data would help improve the model.
2. *Weight the common breeds more heavily -* Uncommon breeds such as the dingo, a wild Australian dog, are less likely to be present in a dog than something like a beagle or lab. Weighting the breeds based on how common they are might improve the predictions when breeds have physical similarities
3. *Require multiple images of a single dog before making a prediction* - A single image may not accurately represent the features of the dog. By requiring multiple images in different contexts, the prediction may be more accurate.

**Future Improvements - ViT Agent**

When investigating different Deep Learning frameworks that we could implement in our project, we came across the paper *A Multi-Agent Framework for the Asynchronous and Collaborative Extension of Multitask ML Systems* published by Andrea Gesmundo in September of 2022.

[[1]](#footnote-0)

The framework runs multiple agents in parallel, with each agent working asynchronously. The paper is part of an effort to demonstrate the “creation of dynamic large-scale multi-modal multitask intelligent systems that can be indefinitely and collaboratively extended” (Gesmundo, 2022). The ViT Agent framework aims to generate efficient deep learning algorithms by automating knowledge transfer of one subset of tasks to another.

The final model represents the findings from the stanford dataset alone. In future iterations, we would hope to implement this framework, and pull in multiple relevant datasets, or ‘tasks’ to finely tune the model to the images provided. Datasets like the tensorflow pet finder dataset, or Oxford IIIT dataset could be used in this parallel training framework, both of which record relevant breed information.

The novel paper and methodology were released too recently to apply this framework to our project in time. However, future iterations of this project would benefit from the cross-task knowledge transfer described to create an even more sophisticated model.

**Conclusion**

The model is able to generally predict what breeds might be present in a mixed breed dog. However, due to the variability in features of mixed breeds, this model cannot replace a DNA test kit. The use of a quality photo can greatly increase the accuracy of the prediction. Mixed breeds with distinct characteristics were more likely to be accurately predicted than ‘super-mutt’ breeds with ambiguous characteristics. Increasing the number of breeds in the training data, and analyzing multiple photos of the same dog will greatly improve the accuracy of the prediction.

If this model were to be deployed in a real-world environment like a mobile application, a few improvements and recommendations would need to be implemented. First, the future improvements mentioned previously would be executed to see if any of them could improve the accuracy of the models. In addition to model improvements, the app itself could offer instructions to the user to help them provide the app and model with better testing images. For example, as mentioned previously, some results were likely impacted by photo quality. So the app could instruct the user to make sure that the dog is the only thing in the photo, its face and body are visible, it’s not wearing anything on its body, etc.

Overall, the project and models were successful. While the CutMix model was slightly less accurate at predicting purebred dog breeds, it was able to predict multiple dog breeds for potential mixed breed dogs. With more training data and previously mentioned future improvements, this could potentially be an accurate way to estimate dog breeds of mixed breed dogs, ideally reducing the need for DNA test kits.

**Contributions**

Lauren Bassett

* Technical Research on the ViT Agent Model based on Frank’s literature review, including phone call with primary researchers for the methodology.
* Analysis of Mixed Breed Results
* Powerpoint slide deck

Frank Vasquez

* Implementation of CutMix Method
* State of the Art literature review that led to ViT Agent Model
* Implementation of lightweight baseline model

Maggie Houck

* Attempt at SuperMix implementation (ended up not being included)
* Accuracy code implementation
* Writing report

Cullan Bedwell

* Implemented CutMix Model from CutMix Method on a batch of images into a custom Model Training Algorithm
* Cleaned up Coding Repository
* Trained + Saved final models for Prediction Analysis

**References**

Comtex. (2022, August 24). *Dog DNA Test Service Market Growth 2022 Size, Share and Recent Developments, Industry Segments, Regions, Leading Players, and Exciting Opportunities till 2028*. MarketWatch. Retrieved October 9, 2022, from <https://www.marketwatch.com/press-release/dog-dna-test-service-market-growth-2022-size-share-and-recent-developments-industry-segments-regions-leading-players-and-exciting-opportunities-till-2028-2022-08-24>

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

Kyathanahally, S., Hardeman, T., Reyes, M., Merz, E., Bulas, T., Brun, P., Baity-Jesi, M. (2022, September 22). Ensembles of Vision Transformers as a New Paradigm for Automated Classification in Ecology. doi:<https://doi.org/10.48550/arXiv.2203.01726>

Morrill, K., Hekman, J., Li, X., McClure, J., Logan, B., Goodman, L., Gao, M., Dong, Y., Alonso, M., Carmichael, E., Snyder-Mackler, N., Alonso, J., Noh, H. J., Johnson, J., Koltookian, M., Lieu, C., Megquier, K., Swofford, R., Turner-Maier, J., & White, M. E. (2022). Ancestry-inclusive dog genomics challenges popular breed stereotypes. *Science*, *376*(6592). <https://doi.org/10.1126/science.abk0639>

Turcsán, B., Miklósi, Á., & Kubinyi, E. (2017). Owner perceived differences between mixed-breed and purebred dogs. *PloS one*, *12*(2), e0172720. <https://doi.org/10.1371/journal.pone.0172720>

S. Yun, D. Han, S. Chun, S. J. Oh, Y. Yoo and J. Choe, "CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), 2019, pp. 6022-6031, doi: 10.1109/ICCV.2019.00612.

1. Gesmundo (2022) provides a graphical representation of the sequential execution on the left, and the multi-agent parallel execution on the right. This graphic demonstrates how the agents run asynchronously and in parallel to improve the model. [↑](#footnote-ref-0)