Project Proposal

Luke Trinity and Nat Shenton
CS 254

October 29, 2019

1 Introduction

Online sources of high resolution images including Google, Image-net [?], and Flickr are providing new opportunities to explore applications of image categorization using computer vision. One interesting application in the field of fine-grained image classification is identification of dog breed. Dogs faces are considered to be highly differentiable for each breed, which makes classification an approachable problem. However, the work is challenging due to the varying combinations of facial characteristics and color patterns between species, as well as intra-class variation within a single species. Another aspect of the problem that makes it more complicated is the presence of humans and man-made backgrounds, which are not as common in other animal datasets. In addition, many attempts to classify dog breeds fail to corroborate the results with actual DNA evidence. Our approach will be to utilize an existing dog breed image classification dataset [?], in combination with a novel dataset that includes paired DNA test results and photos [?], to classify non-pure bred dog breeds.

2 Problem Definition and Algorithm

2.1 Task Definition

The task is to take dog images where each image is labeled with a breed and classify what breed the dog could possibly be. This will is to normalize these images as the inputs. Each input image is defined as a matrix of $250 \times 250 \times 3$. The output will the probability that the dog is a certain breed. This is an interesting problem as it could provide dog owners an easy way to classify a dog breed. With innovations such as Googles Opensource API which allows for object detection. Allowing dog owners the ability to classify their pet in terms of the percentages of possible breeds would be a useful tool for dog owners. The final input for this task would allow for a dog owner to upload a picture of their dog to a website or app and then the machine learning algorithm will classify their dog as a percentage of each breed. The final output of be the probability of the most likely breeds.

2.2 Dataset

The Stanford Dogs Dataset has 20,580 images each classified into one of 120 classes of dog breed. [?] Each image also has an associated bounding box identifying the location of the dog within the

photo. The mixed-breed DNA identified dog photo dataset from Voith et al. [**?**] has 20 images, each paired with DNA percentage breakdown of dog breed. There are no associated bounding boxes for the Voith dataset [**?**]. We will need to outfit this second dataset with bounding boxes to improve the accuracy of our classification algorithm, when the second phase of the project is reached. Later in the semester we hope to have other mixed breed training data, our backup plan is to scrape dog show images hosted online to generate our own dataset.

2.3 Algorithm Definition

First we will convert each of the images in the Stanford Dogs dataset [**?**] into a single feature vector based on the trimmed height and width of the images and the (r,g,b) values in each cell. We will then pair the images with their associated class labels, and utilize a neural network as a black box to train our model to accurately predict the class of our test images. Later in the semester we hope to unbox the neural network with more advanced knowledge to achieve better accuracy. One method we will work to implement, which was utilized by both Kohsla et al. [**?**] and Liu et al. [**?**], is scale-invariant feature transform (SIFT). This algorithm detects and describes local features in images by finding candidate matching features based on Euclidean distance of feature vectors. [**?**] After achieving a high level of accuracy classifying pure-bred dogs, we will introduce our other dataset [**?**] that carries interesting practical concerns in our implementation of SIFT. Here we will attempt to classify mixed breed dogs using images annotated with precise DNA testing of breed composition. Hopefully we can augment this dataset with more samples by reaching out to different labs. This is an interesting computer vision problem because dogs often share features, whether it be color patterns or facial structure, that lead to breed misrepresentation. Training an algorithm in this capacity could lead to novel conclusions related to classifying a dog as purebred or a mixed breed.

3 Experimental Evaluation

3.1 Methodology

3.2 Results

3.3 Discussion

4 Related Work

In the initial article utilizing the novel dataset by Khosla et al. [**?**], the mean accuracy reached 22% when 100 training images were utilized within the SIFT methodology. Liu et al. [**?**] reached 69% accuracy, with greater success than Khosla et al. [**?**] due to their partially localized approach. We hope to leverage the bounding boxes and augment the work of Liu et al. [**?**], possibly improving on their methodology to achieve higher accuracy. Several bounding box approaches are currently being reviewed, although all are endlessly modifiable from simple image segmentation. One is a slight adaptation to a method often employed for training algorithms to images of different shade and dimension called DeepCut. [**?**] There are several standard variations on the basic DeepCut methodology, which hones in on the nature and dimension of the bounding in a cookie-cutter, traced-out way as opposed to any kind of geometric shape. A very similar methodology is known as GrabCut which does not need to even be a single, continuous cookie-cutter-style line bounded
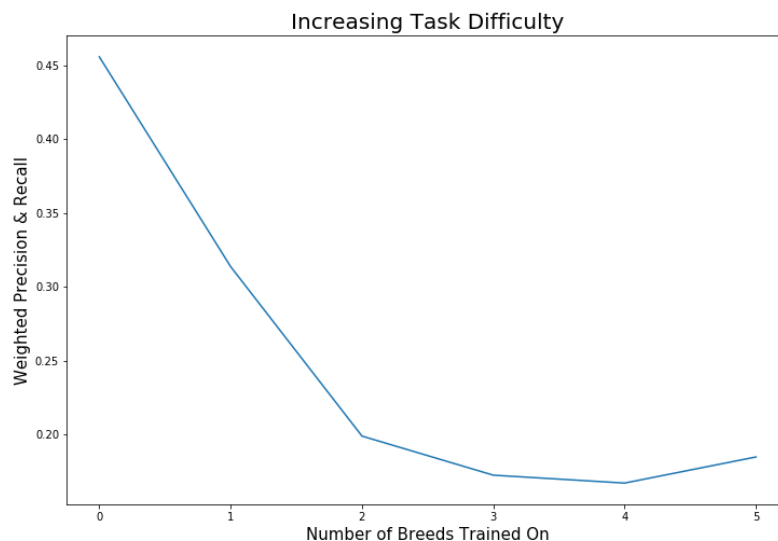
Figure 1: The wieght precision and recall the classification from Keras model

around a desired segment of an image. GrabCut is standardly designed to grab multiple disparate segments of an image at once for a single, desired analysis. In addition to one of these approaches or some modification thereof, we plan to incorporate some aspect of KL Loss to account for localization uncertainty which could greatly assist with interference from background objects such as people, buildings, trees, etc. [?] There are, of course, a number of other modifiable 'standard' or 'familiar' combinations of approaches to deal with bounding and uncertainty to also be explored. Even so, most will lead to an approximately similar result if implemented correctly. Deep/GrabCut is simple and can narrow the selected image segment in a convenient, efficient manner. KL Loss will be further researched and incorporated to deal with expected shading difficulties, image blurring and background/foreground image 'noise'.

5 Next Steps

These initial results show the next steps, one is to improve the learning algorithm. The initial Neural network trained using Keras shows the classification of dog breeds get worse and worse as the more breeds get introduced. From these results there are two main directions this project will have to go. The first is it move this project onto the VACC as there when running any more than a 6 breed classification there was a memory error on the local machine. The next step will be to utilize the VACC, once the current model is working on the VACC, the model will be scaled from the 6 breed classification to the 120 breed classification. The model will return the prediction probability from the most possible breeds. This will be Luke's main task.

The second task will be to improve the model. Thru cross validation, the pre weighted neural networks from the Keras libraries will be used to improve the model. Keras has an extensive library of pre weighted networks specifically designed from image classifications. This will include the tuning of hyper parameters to improve the prediction probability thru cross validation. This will be Nat's main task.

Once the algorithm has been optimized, the final step of the project will be a user interface

where a user will upload a image of their dog, and the algrorithm will return the prediction probability. This will either be down with a simple HTML page or thru Unity.

6   Code and Dataset

The code can be found in our github repo: https://github.com/nshenton/CS_251_DogProject. Please see the Readme in the repo for more information. The Stanford Dogs Dataset can be downloaded from this page: http://vision.stanford.edu/aditya86/ImageNetDogs/.

7   Conclusion