Dog Breed Classifier

My original project proposal focused on an NFL score total predictor, where there are tons of complex neural networks with a variety of datasets. After a couple of weeks of analyzing and messing with them, I was unable to improve on them. Since I was struggling so much, I decided to focus on a simpler neural network.

Overview: This project will use images from a Kaggle dataset of dogs of different breeds to train a model to predict which dog breed is in the image.

Background: Dogs are the most common pets in the world, and there are many different breeds based on their genetics. Dog breeding has been especially popular in the last century, and new hybrid breeds have been created. Knowing a dog's breed is essential, as it correlates to the dog's personality, behavior, and size. Making a model that could help classify these dogs would be important to buyers of dogs who are unsure of what breed they are looking at. Being able to know the breed of the dog could influence their decision to find a better-suited pet. Having a high-accuracy classifier would allow buyers to be sure of what breed their dog is or what breeds it could be mixed with.

Statement of Work: The Kaggle dataset I used has made many different models to determine the classification.
https://www.kaggle.com/code/yonatanrabinovich/dog-breed-classification-with-pytorch/notebook
https://www.kaggle.com/code/stpeteishii/dog-breeds-70-classify-fasiai
https://www.kaggle.com/code/megr25/your-dog-breed
These classifications are each unique and provided me insight into how I could add features from each to build a better model. For example, I could experiment with layers from the pre-trained model in the first link, add transformations to make the images easier to classify, and cut the training if the model is no longer improving. I am developing my neural network from scratch but I will use elements from the existing code from users on Kaggle and existing code from work I have done in this class. I plan to use the cross-entropy loss as the optimizer and the loss as the mean square errors to improve my model.

Dataset: The total number of samples is 9346. I am using a dataset with high-definition images, but I will resize them to make the model more convenient to run. The dataset consists of predetermined test, valid, and train images organized by their breed. I will use their predetermined organization to train and test my model. The goal of my model will be to accurately classify a test image into one of the seventy different breeds included in this dataset.

Evaluation: I hope to achieve an accuracy of around 90% in the classification of the test set to prove that I was able to make a model as good as the ones found on Kaggle. This accuracy means my model is extremely good for the 70 different classifications and the variety of images used.

Project Plan

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8
Setup Github and begin planning methods	X							
Finish processing Data Methods		X						

Build and train model			X					
Thanksgiving				X				
Project Progress Report				X				
Try to improve model, possibly add or delete variables				X	X	X		
Make user interface and Prepare for Presentation							X	X

https://www.kaggle.com/datasets/gpiosenka/70-dog-breedsimage-data-set/data?select=train