

## Data Sources

Stanford Dogs Dataset (image data only)

<http://vision.stanford.edu/aditya86/ImageNetDogs/>

## Load data

Dog-images.zip: contained five separate breed image folders

Five csv files prior to final merging:

- "n02085936-Maltese\_dog.csv.gz",
- "n02088364-beagle.csv.gz",
- "n02099601-golden\_retriever.csv.gz",
- "n02106662-German\_shepherd.csv.gz",
- "n02110958-pug.csv.gz"

Variables include rbg value, breed, breed\_name

Image data shape: (889, 49152)

Column	Variable Type	Description
0 - 49151	Float	Values representing the image's pixel intensities from RGB values normalized between 0 and 1
Breed	Integer	The numeric label for the breed, ranging from 0-4
Breed_name	String	The names of each breed

Figures provided below:

Figure 1: Number of images for each dog breed, original data.

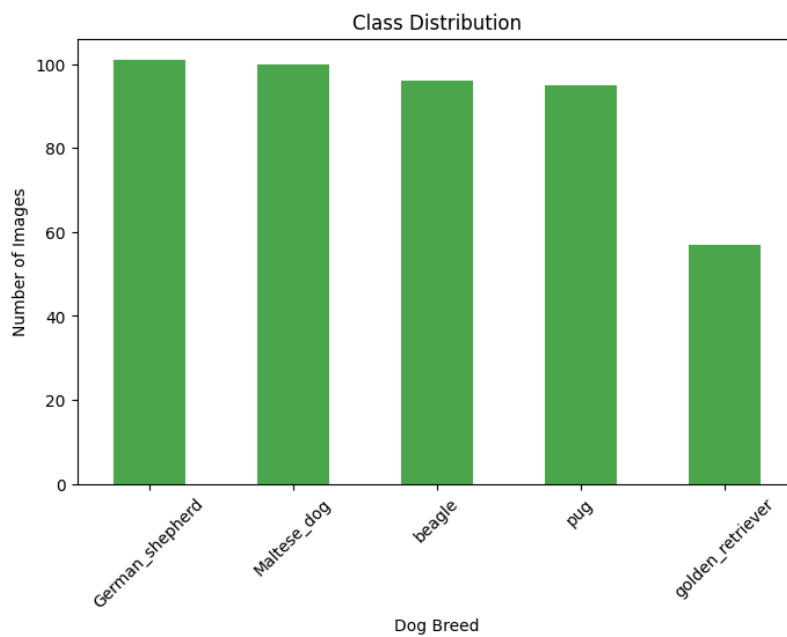


Figure 2: Number of images for each dog breed, updated data.

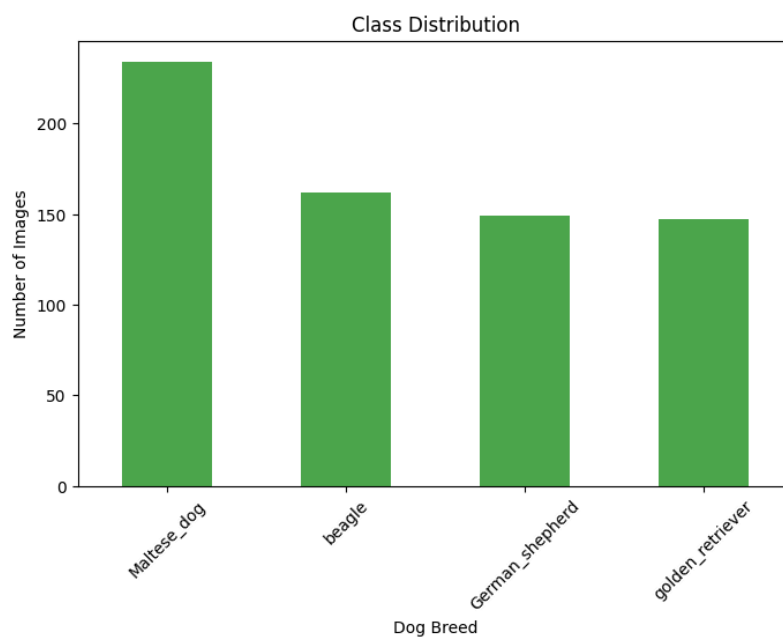


Figure 3: Breed class distribution bar graph, after reducing Maltese images to achieve equal distribution.

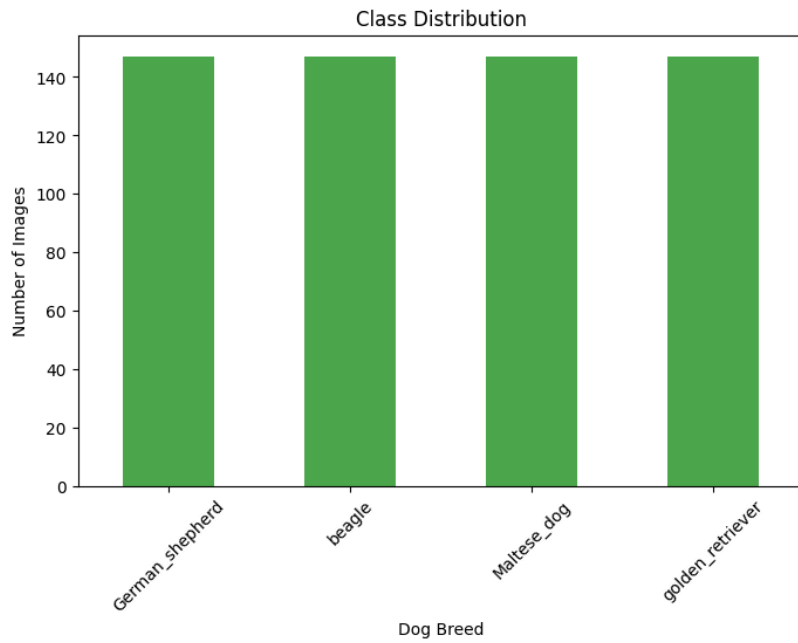


Figure 4: Mean and standard deviation of pixel values of each breed's images

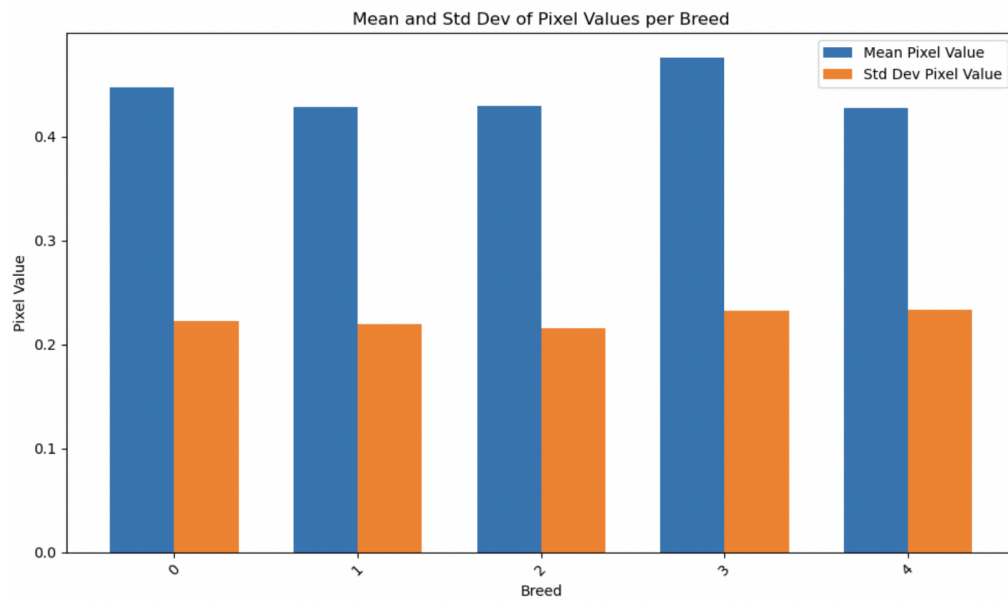


Figure 5: Confusion matrix of final model for training accuracy of 57% and test accuracy of 45%

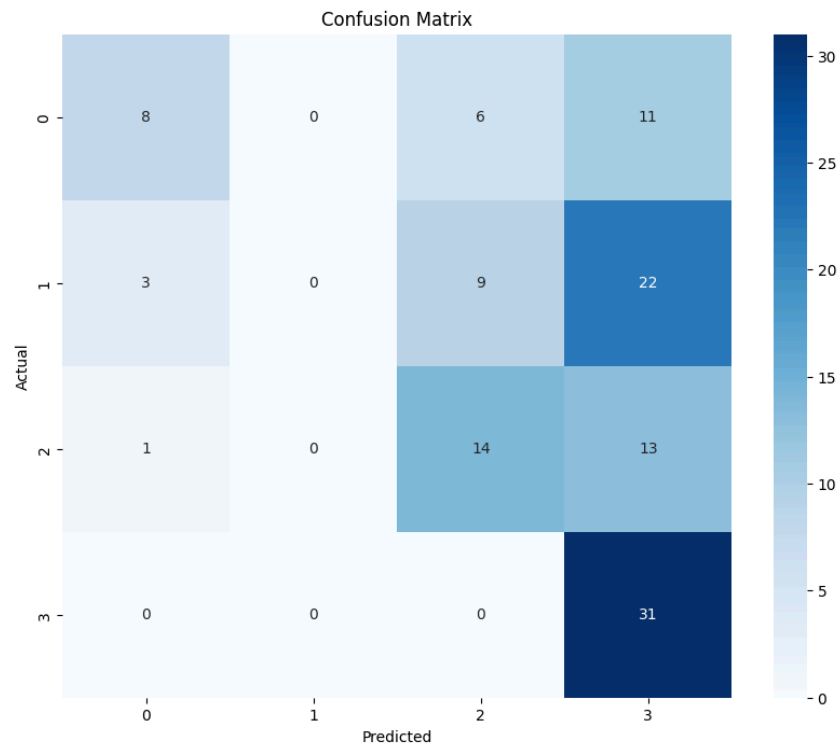


Figure 6: Accuracy per epoch for training and validation data

