CS510 Assignment 4
Gauthami, Sanket, Shreya

The wildlife dataset given has in total 1643 images of different classes namely: Bighorn (31), Bobcat (200), Coyote (200), Gray Fox (210), Javelina (200), Mule Deer (200), Raptor (402), White-tailed Deer (200). Some sample images from each class are:

Fig. 1: Sample images from the dataset

Analysis of the data set:

The White-Tailed Deer dataset has different dimensions of images: (256, 256, 3) (1024, 1024, 3) (512, 512, 3) (1014, 1024, 3) (1024, 1025, 3) (1017, 1024, 3) (512, 513, 3) (1024, 975, 3) (1024, 1014, 3) (1005, 1024, 3). The rest of the dataset has dimensions (256, 256, 3). We also observed that all the images of Gray Fox have been captured at night while the rest like Mule Deer, Javelina, etc., have a combination of daylight and night images as shown in Fig 1. Moreover, initially, the images were not well organized in classes. We noticed that the image names contain the class name. Hence, we wrote a small tool that can segregate images into appropriate classes.

To discuss the issues, we found that the resolution of some images is not that good, and some images were blurry. This has happened mainly because the animal is in motion. Also, the face of the animal is occluded in some cases (by other object or by its own body) and the animal's body also camouflages with the background as well. Particularly birds can be difficult to recognize, wings being the only visible portion of the body. All these issues can be a good explanation if CNN (Convolutional Neural Network) does not give satisfactory results. Furthermore, we found class imbalance to be a major issue. There are only 31 samples of Bighorn Sheep but 402 for Raptor. This can lead to CNN not performing well on the ones which have a small number of images.

Experiments:

Despite the issues discussed above, we started with experiments. First, we had to normalize every image to the same dimension before we started the training process. Hence, we resized all of them to a size of (256, 256). This size was chosen as it is the least one an image has in the dataset. We also did data augmentation to increase the performance of our network. Images are horizontally flipped and rotated at an angle of 20 degrees. Further, we built a vanilla CNN with 3 convolutional layers followed by a linear layer for our classification task. We decided to use 5-fold cross validation for this task to make use of our dataset more effectively. With this in mind, we trained our model for two cases: 1] Resized images are converted to grayscale and then trained and 2] CNN is trained on the resized original images; and got the following results:

	Accuracy
Training set	100%
Test set	71.4%

Images are not grayscale

	Accuracy
Training set	100%
Test set	85.71%

Images are grayscale

Table 1: Results

As we can see, we are getting satisfactory results when images are grayscale, but, **for part 2 of the assignment**, we are planning to improve by using transfer learning and training some well-known models like ResNet or VGG for our dataset. We are confident that the accuracy of our classification task may increase if these well-known models are harnessed for our purpose. Also, if possible, we will try to get more samples for our Bighorn Sheep class as the number of images we have is too small to get a satisfactory accuracy for the same.