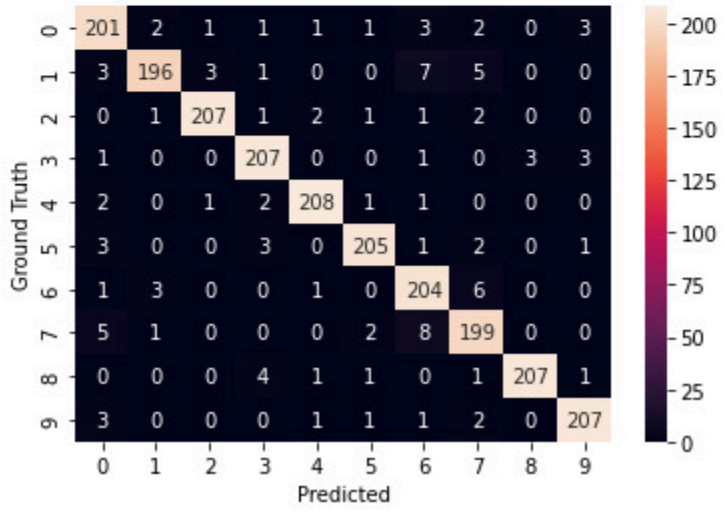
We experimented with a 10 class animal classification dataset.

We used resnet pretrained model and fine tuned using 5000 images (500 images each) to make predictions. The Test and Validation dataset each has 215 images.

When trained model was obtained after running it for 15 epochs we got following results on test data:

	precision	recall	f1-score	support
0	0.92	0.93	0.93	215
1	0.97	0.91	0.94	215
2	0.98	0.96	0.97	215
3	0.95	0.96	0.95	215
4	0.97	0.97	0.97	215
5	0.97	0.95	0.96	215
6	0.90	0.95	0.92	215
7	0.91	0.93	0.92	215
8	0.99	0.96	0.97	215
9	0.96	0.96	0.96	215
accuracy			0.95	2150
macro avg	0.95	0.95	0.95	2150
weighted avg	0.95	0.95	0.95	2150

Text(0.5, 15.0, 'Predicted')



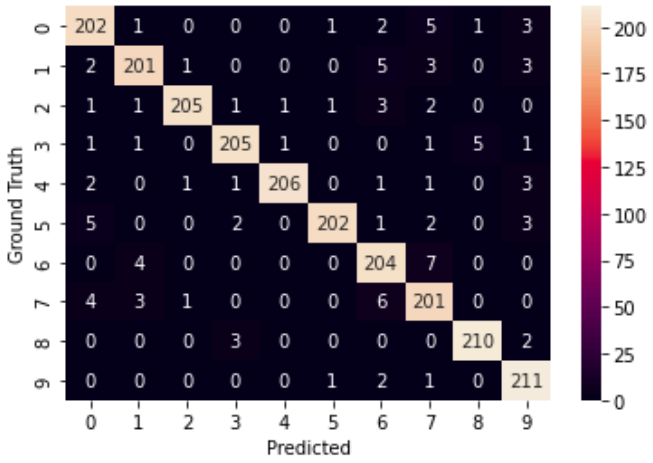
Here there were 109 images misclassified out of 2150 images.

I tried to see the predictions of the training set images when passed to the trained DL model. I found that 468/5000 points were predicted wrong by the model and only 6 points were having confidence probability less than 0.80 and lowest was 0.68. Also 5 out of 6 points were wrongly predicted. The training images seem fine.

Following are the results I got on test data after using the model trained using 25 epochs:

	precision	recall	f1-score	support
0	0.93	0.94	0.94	215
1	0.95	0.93	0.94	215
2	0.99	0.95	0.97	215
3	0.97	0.95	0.96	215
4	0.99	0.96	0.97	215
5	0.99	0.94	0.96	215
6	0.91	0.95	0.93	215
7	0.90	0.93	0.92	215
8	0.97	0.98	0.97	215
9	0.93	0.98	0.96	215
accuracy			0.95	2150
macro avg	0.95	0.95	0.95	2150
weighted avg	0.95	0.95	0.95	2150

```
Text(0.5, 15.0, 'Predicted')
```



Here we saw there were 103 images misclassified during test time. Thus there is a very slight improvement in predictions.

I further performed the augmentations and resizing of images using Albumentation. Following is my data transform block for all train, test and validation data:

```
data_transforms =  A.Compose([
        A.HorizontalFlip(p=0.5),
        A.ShiftScaleRotate(shift_limit=0.0625, scale_limit=0.50, rotate_limit=45, p=.75),
        A.RandomBrightnessContrast(brightness_limit=0.2, contrast_limit=0.2,
brightness_by_max=True, always_apply=False, p=0.5),
        A.Resize(height= 250, width = 320),
        A.Normalize(mean=(0.485, 0.456, 0.406), std=(0.229, 0.224, 0.225)),
        ToTensorV2()
        ])
```

I got the following results after training it for 25 epochs:

	precision	recall	f1-score	support
0.0	0.88	0.91	0.89	215
1.0	0.94	0.93	0.94	215
2.0	0.94	0.97	0.95	215
3.0	0.98	0.95	0.97	215
4.0	0.99	0.96	0.97	215
5.0	0.95	0.89	0.92	215
6.0	0.91	0.91	0.91	215
7.0	0.89	0.91	0.90	215
8.0	0.97	0.97	0.97	215
9.0	0.95	0.97	0.96	215
accuracy			0.94	2150
macro avg	0.94	0.94	0.94	2150
weighted avg	0.94	0.94	0.94	2150