

CSCI 677 – Advanced Computer VisionAssignment 4

This assignment uses LeNet-5 and ResNet-9 to classify images in CINIC 10 dataset.

Code available at: <https://colab.research.google.com/drive/1yM3pmW43feKOYZOpkuY8-O4ehtL6JLlv?usp=sharing>

The CINIC dataset has 10 classes:

1. automobile
2. bird
3. cat
4. deer
5. dog
6. frog
7. horse
8. ship
9. truck
10. airplane

It has 3 folders, each containing 9000 images

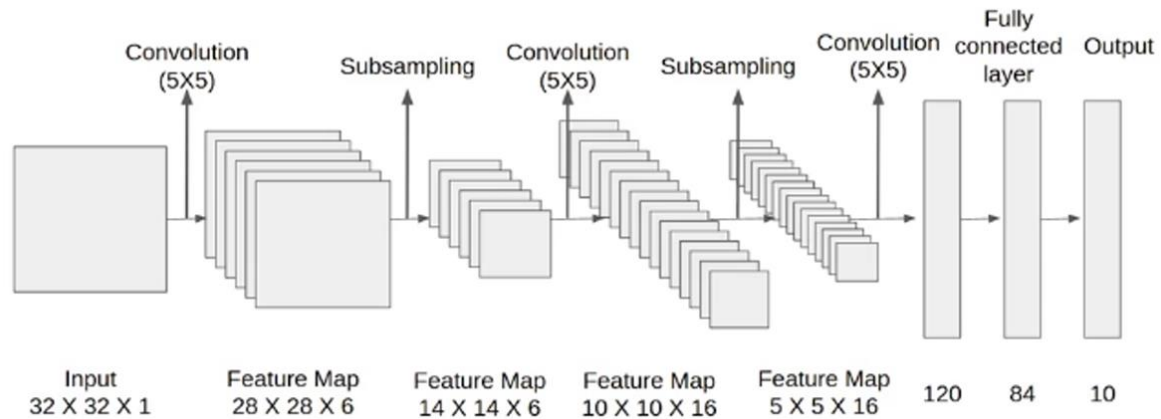
- train
- test
- valid

MAIN EXPERIMENT

General Parameters:

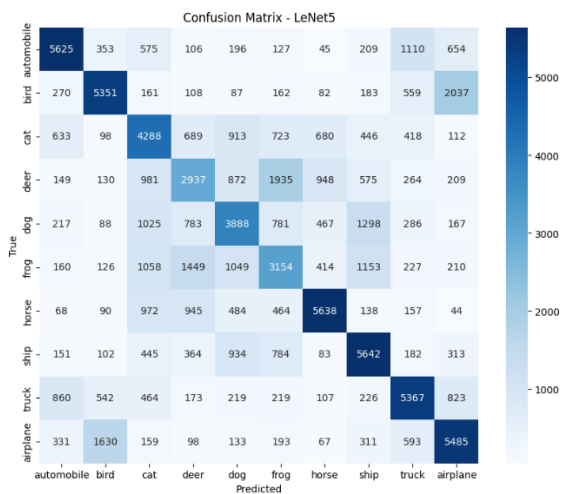
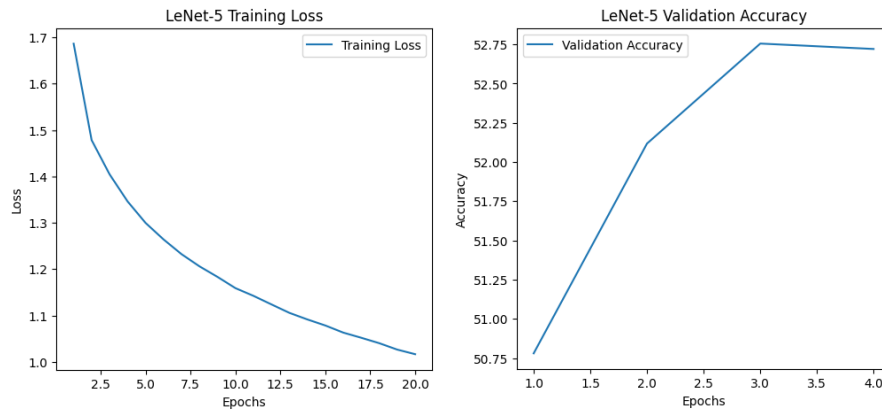
- Epochs = 20
- Batch size = 128
- Optimizer = Adam
- Learning rate = 0.001
- Decay rate = 0.9, 0.999
- Loss function = Cross entropy

Train on LeNet 5 architecture:



### Results:

Epoch [1/20] - LeNet-5: Training Loss = 1.6861  
 Epoch [2/20] - LeNet-5: Training Loss = 1.4782  
 Epoch [3/20] - LeNet-5: Training Loss = 1.4044  
 Epoch [4/20] - LeNet-5: Training Loss = 1.3461  
 Epoch [5/20] - LeNet-5: Training Loss = 1.2998, Validation Accuracy = 50.78%  
 Epoch [6/20] - LeNet-5: Training Loss = 1.2641  
 Epoch [7/20] - LeNet-5: Training Loss = 1.2324  
 Epoch [8/20] - LeNet-5: Training Loss = 1.2061  
 Epoch [9/20] - LeNet-5: Training Loss = 1.1834  
 Epoch [10/20] - LeNet-5: Training Loss = 1.1591, Validation Accuracy = 52.12%  
 Epoch [11/20] - LeNet-5: Training Loss = 1.1425  
 Epoch [12/20] - LeNet-5: Training Loss = 1.1241  
 Epoch [13/20] - LeNet-5: Training Loss = 1.1059  
 Epoch [14/20] - LeNet-5: Training Loss = 1.0916  
 Epoch [15/20] - LeNet-5: Training Loss = 1.0786, Validation Accuracy = 52.76%  
 Epoch [16/20] - LeNet-5: Training Loss = 1.0632  
 Epoch [17/20] - LeNet-5: Training Loss = 1.0521  
 Epoch [18/20] - LeNet-5: Training Loss = 1.0404  
 Epoch [19/20] - LeNet-5: Training Loss = 1.0266  
 Epoch [20/20] - LeNet-5: Training Loss = 1.0167, Validation Accuracy = 52.72%



Time Taken: 18 mins 53 secs

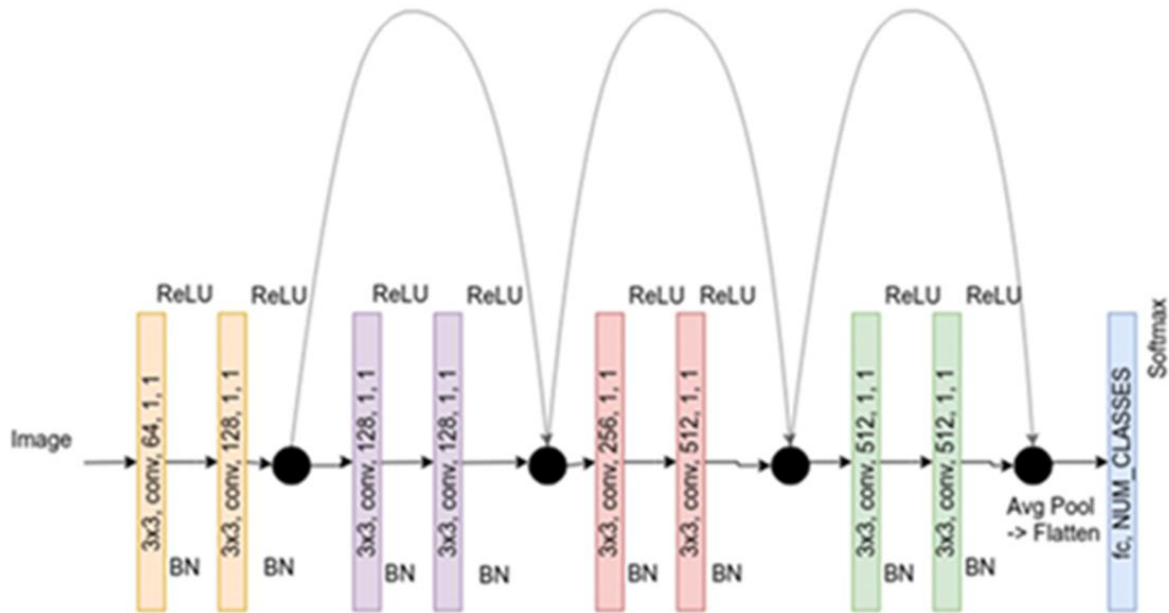
Overall Test Accuracy: 52.64%

Some misclassified images:



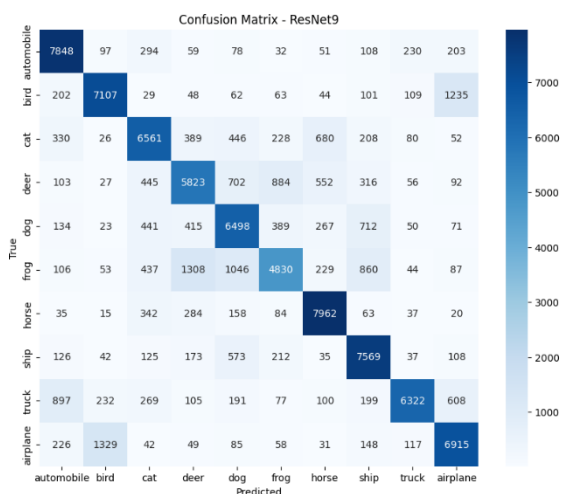
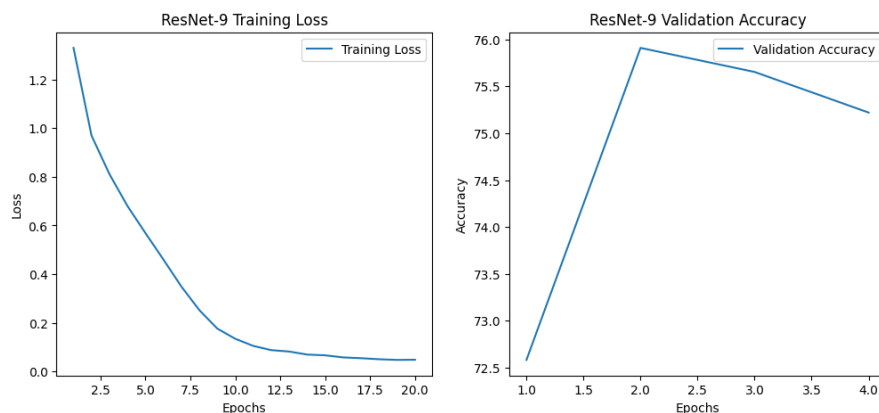
There is misclassification because the images are super blurry and from even human perspective, the shape of plane and colour looks like a cat or a dog on the ground. The angle also makes the model misclassify the images of airplane into bird or cat.

Train on ResNet 9 architecture:



### Results:

Epoch [1/20] - Resnet-9: Training Loss = 1.3300  
 Epoch [2/20] - Resnet-9: Training Loss = 0.9698  
 Epoch [3/20] - Resnet-9: Training Loss = 0.8110  
 Epoch [4/20] - Resnet-9: Training Loss = 0.6803  
 Epoch [5/20] - ResNet-9: Training Loss = 0.5690, Validation Accuracy = 72.58%  
 Epoch [6/20] - Resnet-9: Training Loss = 0.4603  
 Epoch [7/20] - Resnet-9: Training Loss = 0.3492  
 Epoch [8/20] - Resnet-9: Training Loss = 0.2522  
 Epoch [9/20] - Resnet-9: Training Loss = 0.1764  
 Epoch [10/20] - ResNet-9: Training Loss = 0.1345, Validation Accuracy = 75.91%  
 Epoch [11/20] - Resnet-9: Training Loss = 0.1053  
 Epoch [12/20] - Resnet-9: Training Loss = 0.0880  
 Epoch [13/20] - Resnet-9: Training Loss = 0.0821  
 Epoch [14/20] - Resnet-9: Training Loss = 0.0696  
 Epoch [15/20] - ResNet-9: Training Loss = 0.0666, Validation Accuracy = 75.65%  
 Epoch [16/20] - Resnet-9: Training Loss = 0.0580  
 Epoch [17/20] - Resnet-9: Training Loss = 0.0548  
 Epoch [18/20] - Resnet-9: Training Loss = 0.0506  
 Epoch [19/20] - Resnet-9: Training Loss = 0.0478  
 Epoch [20/20] - ResNet-9: Training Loss = 0.0484, Validation Accuracy = 75.22%



Time Taken: 24 m 05 s

Overall Test Accuracy: 74.93%

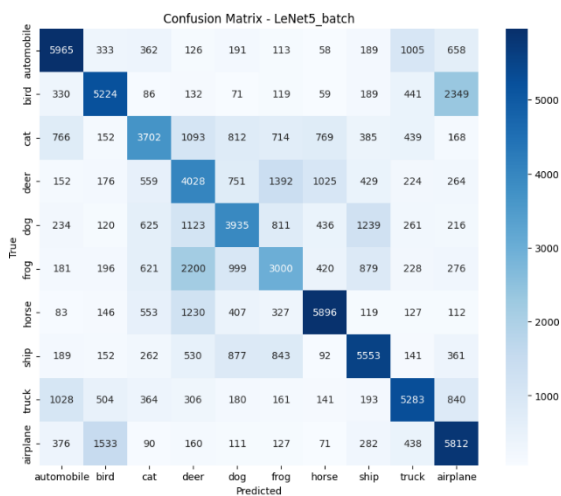
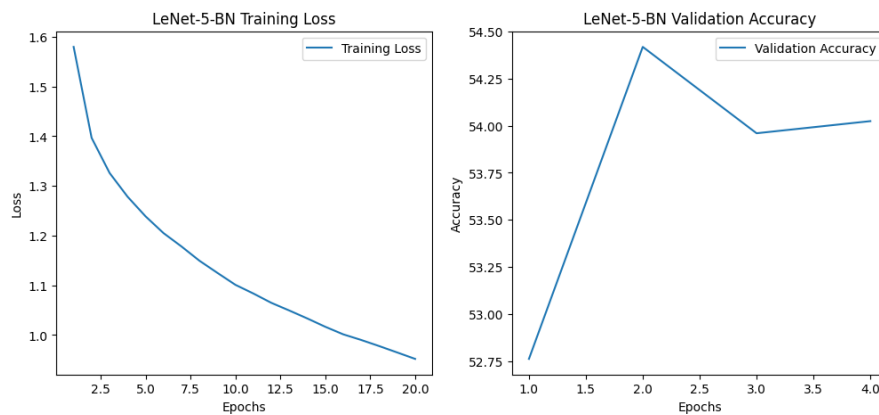
Some misclassified images:



There is close relationship between automobile and trucks as both of them are vehicles and since the images are blurry, the model isn't able to extract significant information to distinguish between the two classes. However, it is able to distinguish clearly between birds, cats and airplanes which the above Lenet model was unable to due to its ability to learn both complex and simpler mappings, avoiding overfitting in some cases.

### VARIATIONS

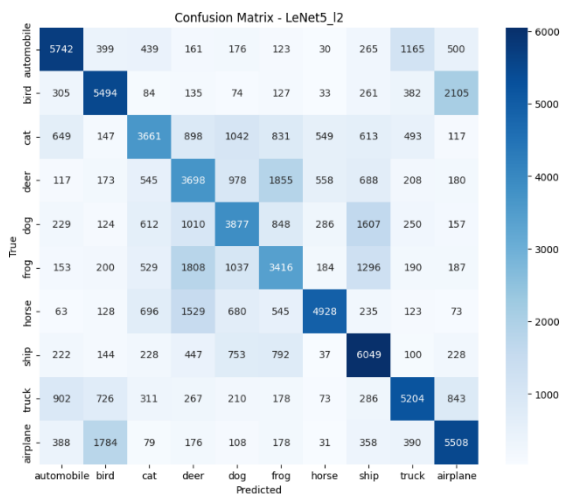
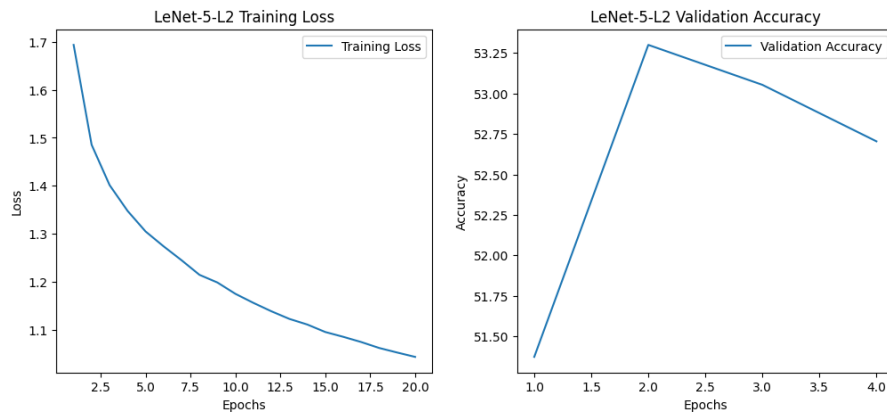
### Training original LeNet-5 with batch normalization



Time taken: 20 mins 10 sec

LeNet-5 with batch normalization Test Accuracy: 53.78%

Training original LetNet-5 with L2 regularization (0.000001)



Time Taken: 20 mins 7 sec

Overall Test Accuracy: 52.86%

### ADDITIONAL TESTING

	Optimizer	Learning Rate	Test accuracy	Time Taken
Original LeNet	Adam	0.001	52.64%	18 m 53 s
Original ResNet	Adam	0.001	74.93%	24 m 05 s
LeNet	SGD	0.01	50.59%	18 m 45 s
LeNet	SGD	0.001	53.12%	18 m 38 s
ResNet	SGD	0.01	74.71%	23 m 02 s
ResNet	SGD	0.001	72.64%	23 m 48 s

A lower learning rate helps in reaching the global minimum faster, or to converge faster. Hence, models with learning rate = 0.001 have a better test accuracy.

For the same learning rate, SGD optimizes Lenet better while Adam optimizes Resnet better. This is because SGD works well due to simplicity and shallower architecture and its stability and regularization help.

Adam with ResNet is Better for deeper, more complex networks. Adaptive learning rate and momentum aid training stability and efficiency in challenging architectures.