CS 203: Software Tools & Techniques for AI
IIT Gandhinagar
Sem-II - 2024-25

LAB 05

Total marks: 10
Submission deadline:
Submission guidelines:
1. Code should be added to a GitHub repository, and the repository details should be shared.
2. Late submissions will be penalized 20% per day.
3. Google form submission link:

Note: By submitting this assignment solution you confirm to follow the IITGN's honor code. We shall strictly penalize the submissions containing plagiarized text/code.
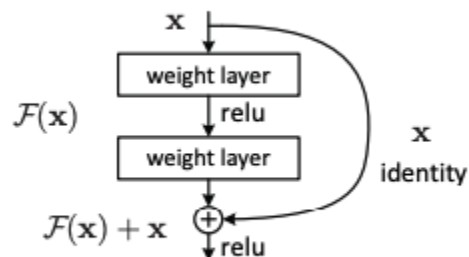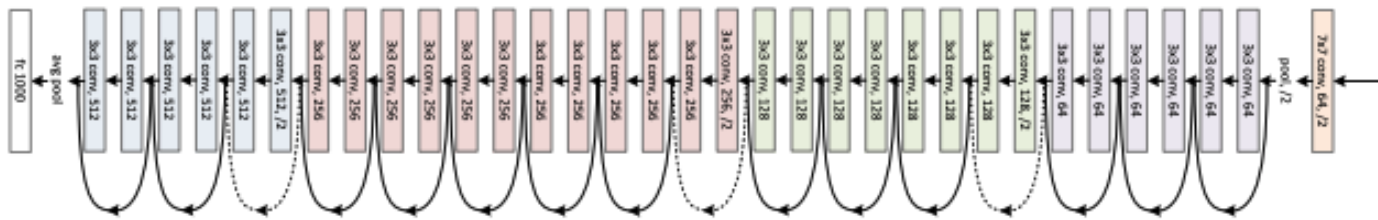

Model Architecture

Figure 2. Residual learning: a building block.

Source: Deep Residual Learning for Image Recognition, Microsoft Research

Resnet 34 Network (50 is a larger version of this)

ResNet50 is a deep convolutional neural network with **50 layers**, designed for image recognition. It uses **residual connections (skip connections)** to prevent vanishing gradients and enable training of very deep networks. Pre-trained on **ImageNet**, it achieves high accuracy on various vision tasks. It's widely used in **transfer learning** for object detection, classification, and feature extraction.
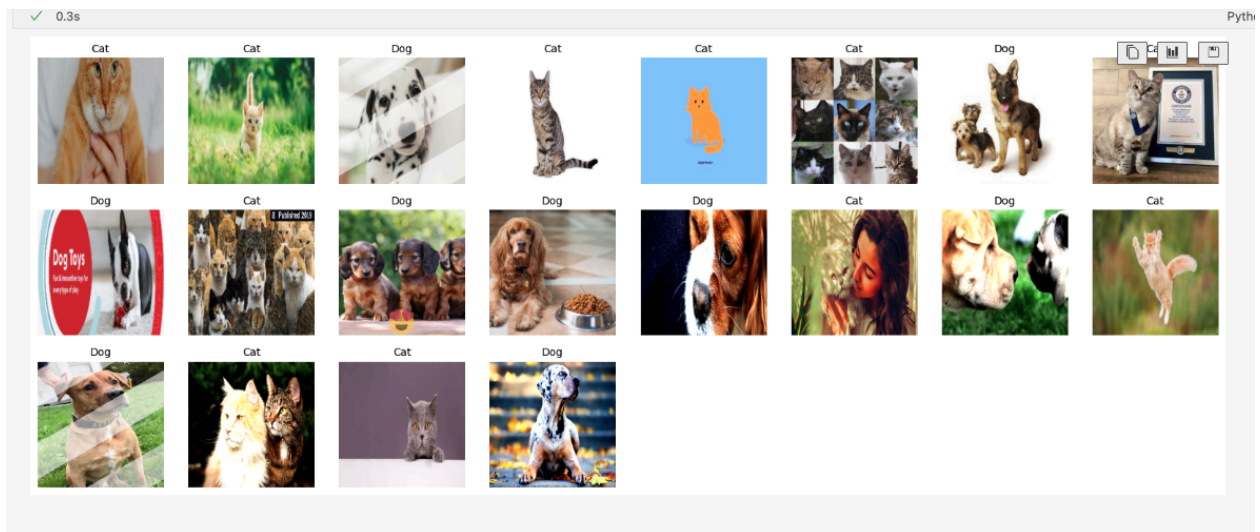
Additionally, our setup uses

```
Training Parameters:
- Model: ResNet50
- Optimizer: Adam
- Learning Rate: 0.001
- Epochs: 20
- Loss Function: CrossEntropyLoss
```
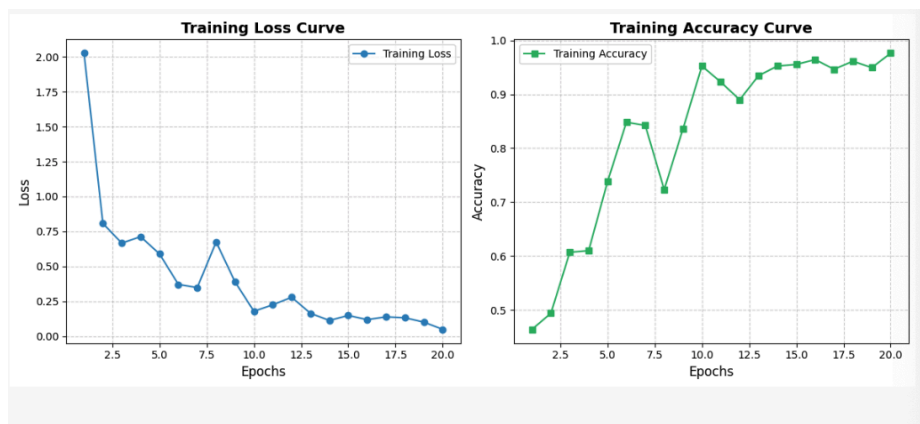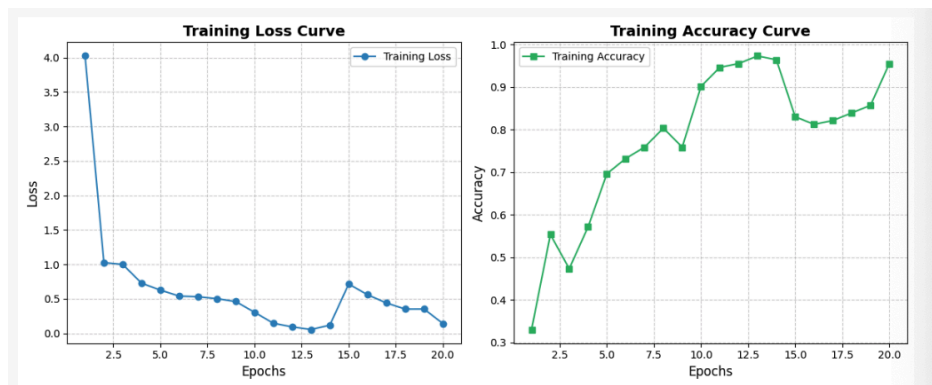
Graphs
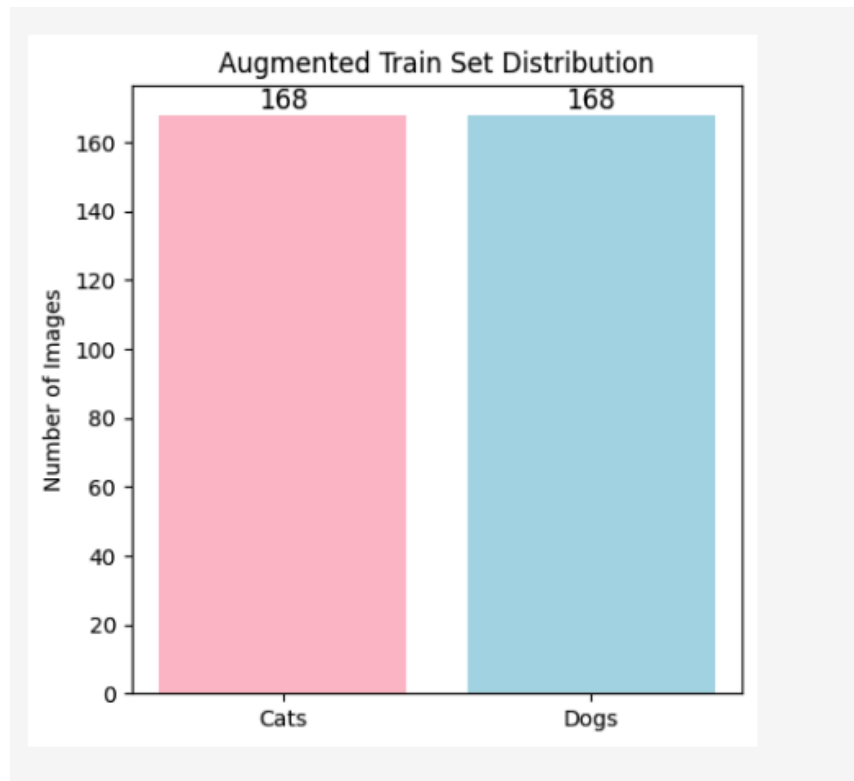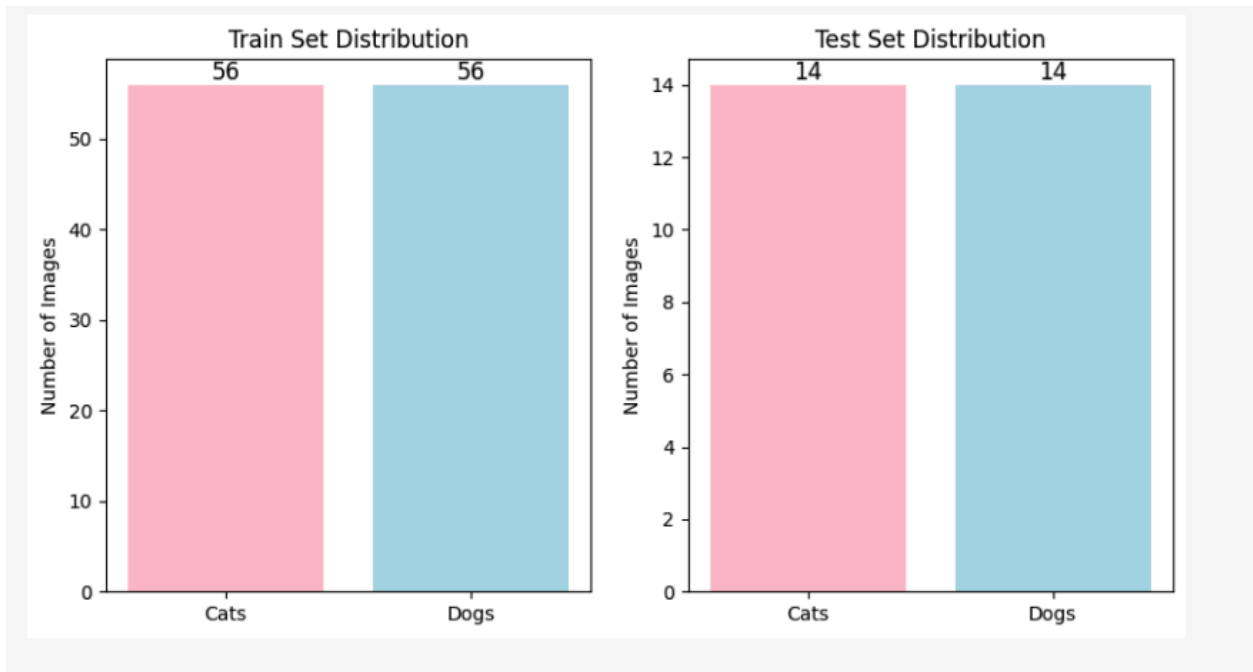
Augmented images (few)

Augmented train set training

Normal train set training

Scores

```
...
      📊 ResNet50 (Normal Training) Results:
         Accuracy:  0.5357
         Precision: 0.5200
         Recall:    0.9286
         F1 Score:  0.6667

      📊 ResNet50 (Augmented Training) Results:
         Accuracy:  0.6071
         Precision: 0.6000
         Recall:    0.6429
         F1 Score:  0.6207
```

Interpretation

Traintime
- We observe that the accuracy for augmented dataset rises quicker on average as compared to normal.
- This is because our augmented data offers more diversity and improves learning.

Testtime
- Here we observe, normal gives an accuracy of around 50%. This indicates that there is near to no learning that helps us, as this is binary classification and you can be right ½ times.
- However, for the augmented dataset, our accuracy is better. And it is reasonable.
- We also feel that we could get a much higher accuracy if we used other strategies like more augmentation, validation, k-folds etc.
- For augmented dataset training, **the model generalizes better.**

Extra
- Recall drops after augmentation (from 92.86% → 64.29%), which means the first model was aggressively predicting "dog" even when unsure. Because dog is positive class.
- F1 score is more balanced (from 66.67% → 62.07%), suggesting the second model is more reliable overall rather than just favoring recall.
- Precision improves (from 52% → 60%), reducing false positives.