Libraries

[ ] ↳ *13 cells hidden*

Visualization

↳ *2 cells hidden*

Evaluation

```
import numpy as np

# Place these lists outside the epoch loop to store metric values across all epochs
train_accuracies, val_accuracies, test_accuracies = [], [], []
train_precisions, val_precisions, test_precisions = [], [], []
train_recalls, val_recalls, test_recalls = [], [], []
train_f1s, val_f1s, test_f1s = [], [], []

for epoch in range(100):
    # Train
    train_loss, train_accuracy, train_precision, train_recall, train_f1 = train(train_loader, model, criterion, opt
    # Validation
    val_loss, val_accuracy, val_precision, val_recall, val_f1 = validate(valid_loader, model, criterion, device)
    # Test - to be run only after final training or less frequently
    test_accuracy, test_precision, test_recall, test_f1, test_cm = test(test_loader, model, device)

    # Append current epoch metrics to lists
    train_accuracies.append(train_accuracy)
    val_accuracies.append(val_accuracy)
    test_accuracies.append(test_accuracy)

    train_precisions.append(train_precision)
    val_precisions.append(val_precision)
    test_precisions.append(test_precision)

    train_recalls.append(train_recall)
    val_recalls.append(val_recall)
    test_recalls.append(test_recall)

    train_f1s.append(train_f1)
    val_f1s.append(val_f1)
    test_f1s.append(test_f1)

    print(f'Epoch {epoch+1}: Loss: {train_loss:.4f}, Train Acc: {train_accuracy:.4f}, Val Acc: {val_accuracy:.4f},

# After completing all epochs, calculate and print the average and standard deviation of metrics
def print_metrics(title, metrics):
    average = np.mean(metrics)
    std_dev = np.std(metrics)
    print(f'{title}: {average:.4f} ± {std_dev:.4f}')

print("\nFinal Performance Statistics:")
print_metrics("Average Train Accuracy", train_accuracies)
print_metrics("Average Validation Accuracy", val_accuracies)
print_metrics("Average Test Accuracy", test_accuracies)

print_metrics("Average Train Precision", train_precisions)
print_metrics("Average Validation Precision", val_precisions)
print_metrics("Average Test Precision", test_precisions)

print_metrics("Average Train Recall", train_recalls)
print_metrics("Average Validation Recall", val_recalls)
print_metrics("Average Test Recall", test_recalls)

print_metrics("Average Train F1 Score", train_f1s)
print_metrics("Average Validation F1 Score", val_f1s)
print_metrics("Average Test F1 Score", test_f1s)

```

```
Epoch 57: Loss: 0.0802, Train Acc: 0.9713, Val Acc: 0.8986, Test Acc: 0.9362
Epoch 58: Loss: 0.0612, Train Acc: 0.9799, Val Acc: 0.9565, Test Acc: 0.9362
Epoch 59: Loss: 0.0317, Train Acc: 0.9971, Val Acc: 0.9565, Test Acc: 0.8936
Epoch 60: Loss: 0.0209, Train Acc: 0.9943, Val Acc: 0.9275, Test Acc: 0.9149
Epoch 61: Loss: 0.0122, Train Acc: 1.0000, Val Acc: 0.9130, Test Acc: 0.8936
Epoch 62: Loss: 0.0606, Train Acc: 0.9914, Val Acc: 0.9130, Test Acc: 0.8936
Epoch 63: Loss: 0.0176, Train Acc: 0.9971, Val Acc: 0.9275, Test Acc: 0.9149
Epoch 64: Loss: 0.0149, Train Acc: 1.0000, Val Acc: 0.9420, Test Acc: 0.9149
Epoch 65: Loss: 0.0108, Train Acc: 0.9971, Val Acc: 0.9420, Test Acc: 0.9149
Epoch 66: Loss: 0.0324, Train Acc: 0.9943, Val Acc: 0.9275, Test Acc: 0.9149
Epoch 67: Loss: 0.0101, Train Acc: 1.0000, Val Acc: 0.9565, Test Acc: 0.9362
Epoch 68: Loss: 0.0076, Train Acc: 1.0000, Val Acc: 0.9565, Test Acc: 0.9362
Epoch 69: Loss: 0.0671, Train Acc: 0.9828, Val Acc: 0.9420, Test Acc: 0.8723
Epoch 70: Loss: 0.1681, Train Acc: 0.9655, Val Acc: 0.9130, Test Acc: 0.8723
Epoch 71: Loss: 0.1693, Train Acc: 0.9397, Val Acc: 0.9130, Test Acc: 0.8723
Epoch 72: Loss: 0.0447, Train Acc: 0.9885, Val Acc: 0.9565, Test Acc: 0.9362
Epoch 73: Loss: 0.0698, Train Acc: 0.9741, Val Acc: 0.9710, Test Acc: 0.9149
Epoch 74: Loss: 0.0806, Train Acc: 0.9684, Val Acc: 0.9420, Test Acc: 0.9362
Epoch 75: Loss: 0.0938, Train Acc: 0.9885, Val Acc: 0.9420, Test Acc: 0.8723
Epoch 76: Loss: 0.0235, Train Acc: 0.9943, Val Acc: 0.9275, Test Acc: 0.9149
Epoch 77: Loss: 0.0734, Train Acc: 0.9828, Val Acc: 0.9275, Test Acc: 0.9149
Epoch 78: Loss: 0.0404, Train Acc: 0.9856, Val Acc: 0.9420, Test Acc: 0.8936
Epoch 79: Loss: 0.0177, Train Acc: 1.0000, Val Acc: 0.9565, Test Acc: 0.8936
Epoch 80: Loss: 0.0749, Train Acc: 0.9856, Val Acc: 0.9710, Test Acc: 0.9362
Epoch 81: Loss: 0.0731, Train Acc: 0.9856, Val Acc: 0.9565, Test Acc: 0.9149
Epoch 82: Loss: 0.1097, Train Acc: 0.9626, Val Acc: 0.9275, Test Acc: 0.9149
Epoch 83: Loss: 0.0253, Train Acc: 0.9943, Val Acc: 0.9275, Test Acc: 0.8936
Epoch 84: Loss: 0.0156, Train Acc: 0.9943, Val Acc: 0.9130, Test Acc: 0.9149
Epoch 85: Loss: 0.0732, Train Acc: 0.9684, Val Acc: 0.9130, Test Acc: 0.8936
Epoch 86: Loss: 0.0224, Train Acc: 0.9914, Val Acc: 0.9275, Test Acc: 0.9362
Epoch 87: Loss: 0.0153, Train Acc: 0.9971, Val Acc: 0.8986, Test Acc: 0.9362
Epoch 88: Loss: 0.0825, Train Acc: 0.9799, Val Acc: 0.9420, Test Acc: 0.8936
Epoch 89: Loss: 0.0736, Train Acc: 0.9684, Val Acc: 0.9130, Test Acc: 0.8298
Epoch 90: Loss: 0.0185, Train Acc: 1.0000, Val Acc: 0.9420, Test Acc: 0.9362
Epoch 91: Loss: 0.0152, Train Acc: 0.9971, Val Acc: 0.9565, Test Acc: 0.9362
Epoch 92: Loss: 0.0585, Train Acc: 0.9828, Val Acc: 0.9565, Test Acc: 0.9149
Epoch 93: Loss: 0.1749, Train Acc: 0.9626, Val Acc: 0.9420, Test Acc: 0.8936
Epoch 94: Loss: 0.1172, Train Acc: 0.9655, Val Acc: 0.9130, Test Acc: 0.9149
Epoch 95: Loss: 0.0495, Train Acc: 0.9914, Val Acc: 0.9420, Test Acc: 0.9149
Epoch 96: Loss: 0.0149, Train Acc: 1.0000, Val Acc: 0.9275, Test Acc: 0.9149
Epoch 97: Loss: 0.0256, Train Acc: 0.9914, Val Acc: 0.9565, Test Acc: 0.9149
Epoch 98: Loss: 0.0755, Train Acc: 0.9799, Val Acc: 0.9130, Test Acc: 0.8936
Epoch 99: Loss: 0.1065, Train Acc: 0.9741, Val Acc: 0.9565, Test Acc: 0.9362
Epoch 100: Loss: 0.1022, Train Acc: 0.9684, Val Acc: 0.9565, Test Acc: 0.9149

Final Performance Statistics:
Average Train Accuracy: 0.9722 ± 0.0349
Average Validation Accuracy: 0.9330 ± 0.0265
Average Test Accuracy: 0.8989 ± 0.0309
Average Train Precision: 0.9723 ± 0.0348
Average Validation Precision: 0.9356 ± 0.0249
Average Test Precision: 0.9025 ± 0.0289
Average Train Recall: 0.9722 ± 0.0351
Average Validation Recall: 0.9327 ± 0.0268
Average Test Recall: 0.8993 ± 0.0305
Average Train F1 Score: 0.9722 ± 0.0351
Average Validation F1 Score: 0.9329 ± 0.0268
Average Test F1 Score: 0.8987 ± 0.0312
```