MLNS Deep Learning Assignment

Part 2

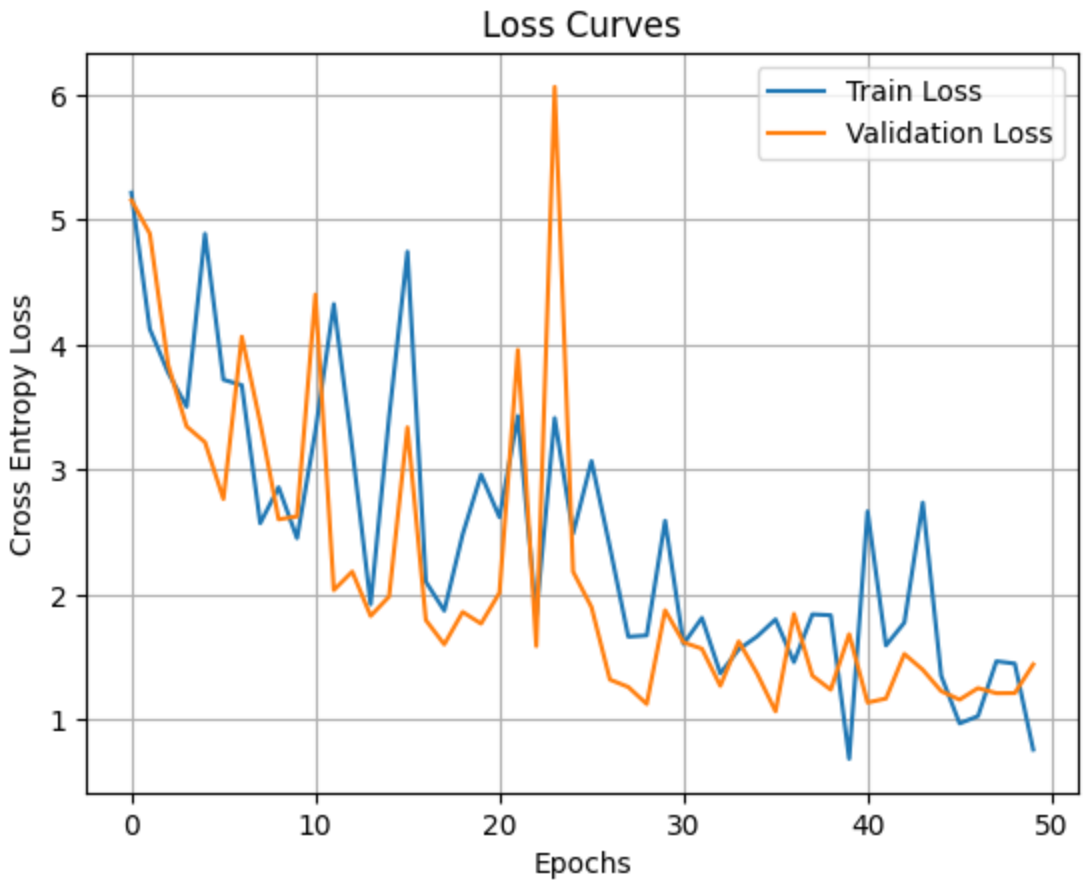
- **Task :** Predicting sum of digits based on image containing multiple digits
- **Dataset :** Images containing multiple digits and the sum of all digits as label
- **Improved Approach :** Now, we use a pretrained ResNet18 (or any other ResNet) and just change its final layer before training it over our own data
- **Model Used**

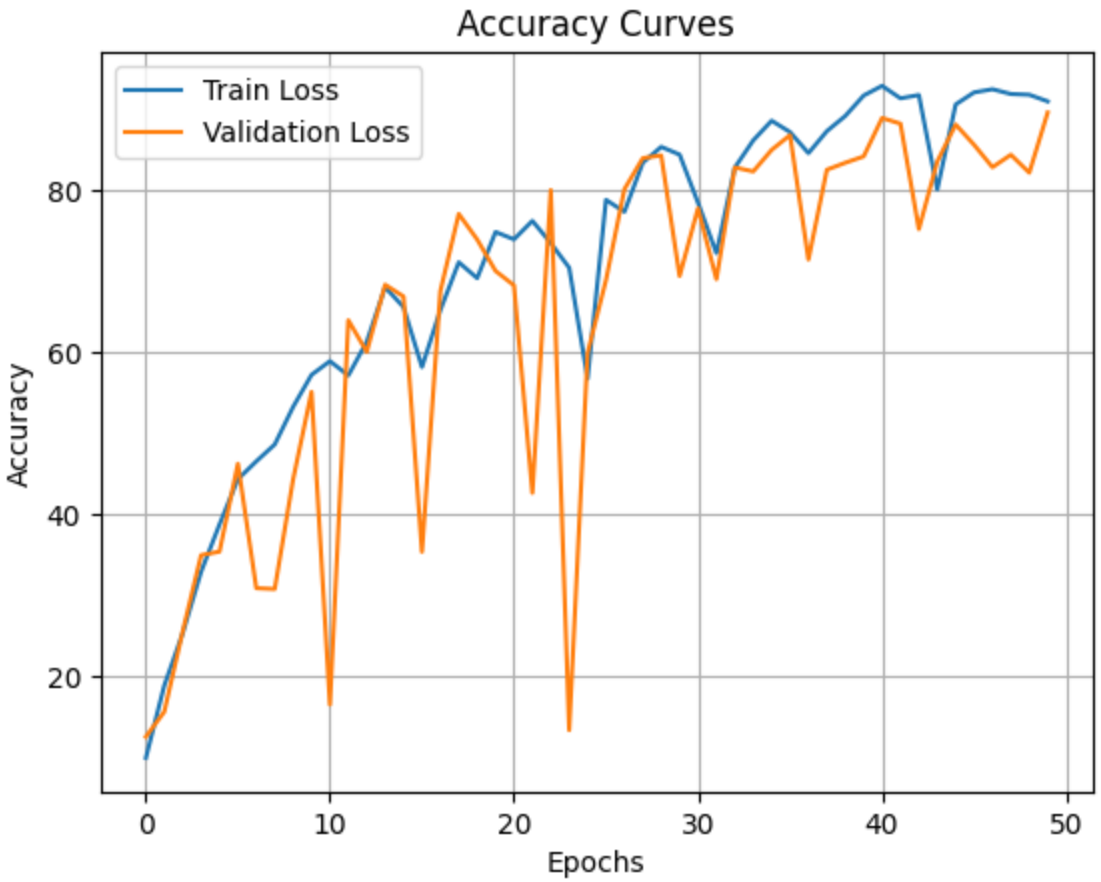
```
# Using pretrained resnet model
model = models.resnet18(weights=models.ResNet18_Weights.DEFAULT)
n_classes = 37 # all possible sums from 0 to 36
model.fc = nn.Sequential(
            nn.Linear(model.fc.in_features, 256),
            nn.ReLU(),
            nn.Dropout(0.2),
            nn.Linear(256, n_classes)
        )
```

- **File Structure**
    1. models/ : contains the saved models after training
    2. training.ipynb : loads datasets, defines models, trains, validates and saves model
    3. inference.ipynb : loads saved model and tests over testing instances
- **Other Approaches**
    1. Using OCR to identify the digits
    2. Using image processing techniques (opening/closing + dilation) to separate the digits and pass them all through a model trained on MNIST.

- **Results**

```
The accuracy over the test dataset is 89.69% and the average MAE is 0.16
```




Accuracy Curves
Train Loss
Validation Loss
Accuracy
Epochs
20
40
60
80
0
10
20
30
40
50