Assignment 1 - Part 2

Name - Arihant Jain

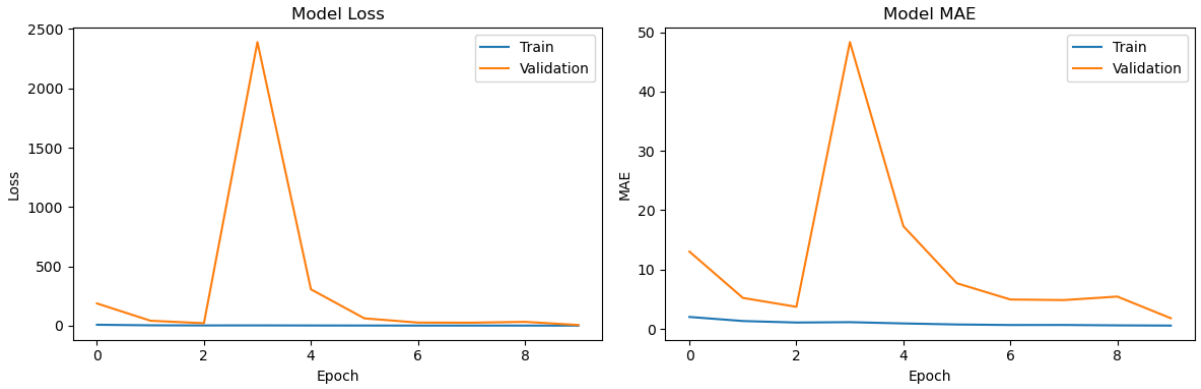
Roll Number - 2021112018

Problem statement

Given an image (and its label value) predict the sum of the digits in the image.

Training File

- TensorFlow threading is configured to optimize performance by utilizing multiple CPUs
- I combined all three data files to create the full training set.
- I normalized the pixel values to prevent the loss function from blowing up.
- A pre-trained MobileNetV2 model is used as the backbone.
- Custom layers are added for fine-tuning the model to the task (e.g., a dense layer for classification or regression).
- The model directly predicts the final sum.
- I am using Mean Squared Error as the loss function because this is a regression problem.
- I also track Mean Absolute Error as a metric to measure the difference between predicted and actual values.
- The model trains for 10 epochs with a batch size of 32.



Inference File

- I used the data2 file for testing.
- I calculated Mean Absolute Error, Root Mean Square Error, and Accuracy (predictions within ±0.5, ±1, ±2 of the true value).
- Accuracy within ±0.5: 19.83%
- Accuracy within ±1: 38.36%
- Accuracy within ±2: 66.94%
- Finally, I showed the results for 5 random samples.