Report

Model

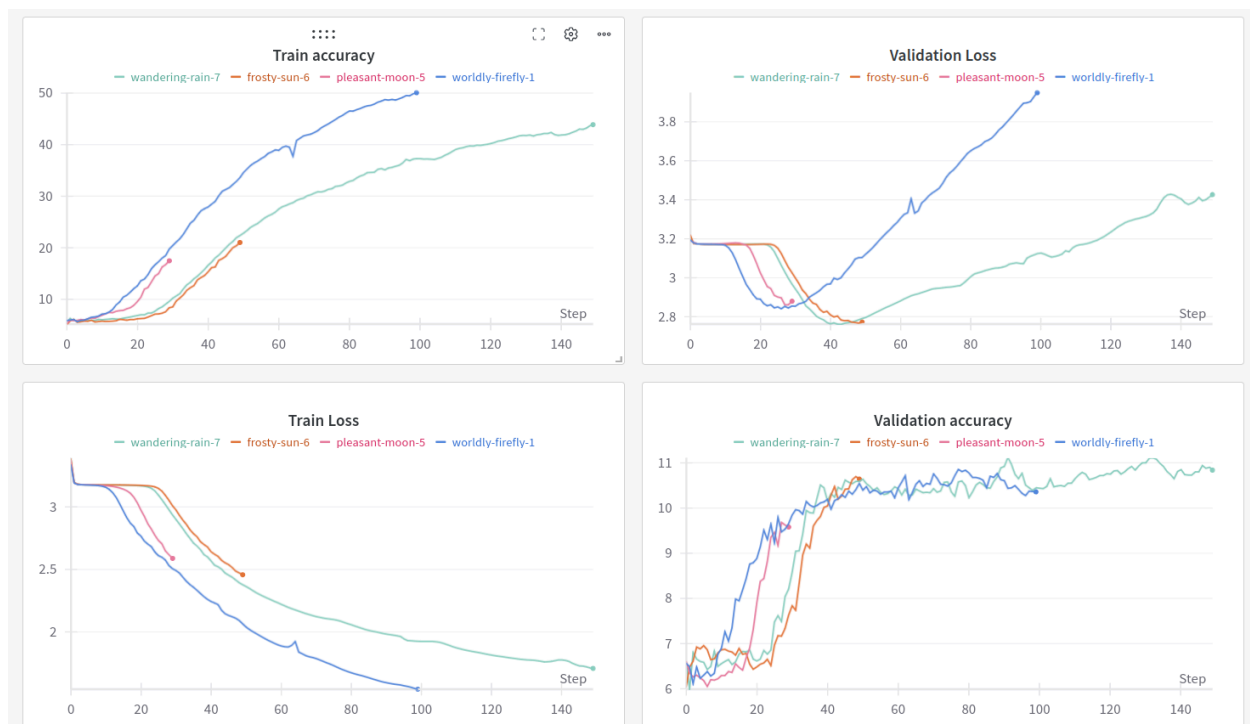- A 4 layer CNN followed by MLP layer.

Training strategy

- Model directly predicts the sum once image is seen.

Hyperparameters

- Lr = 0.001
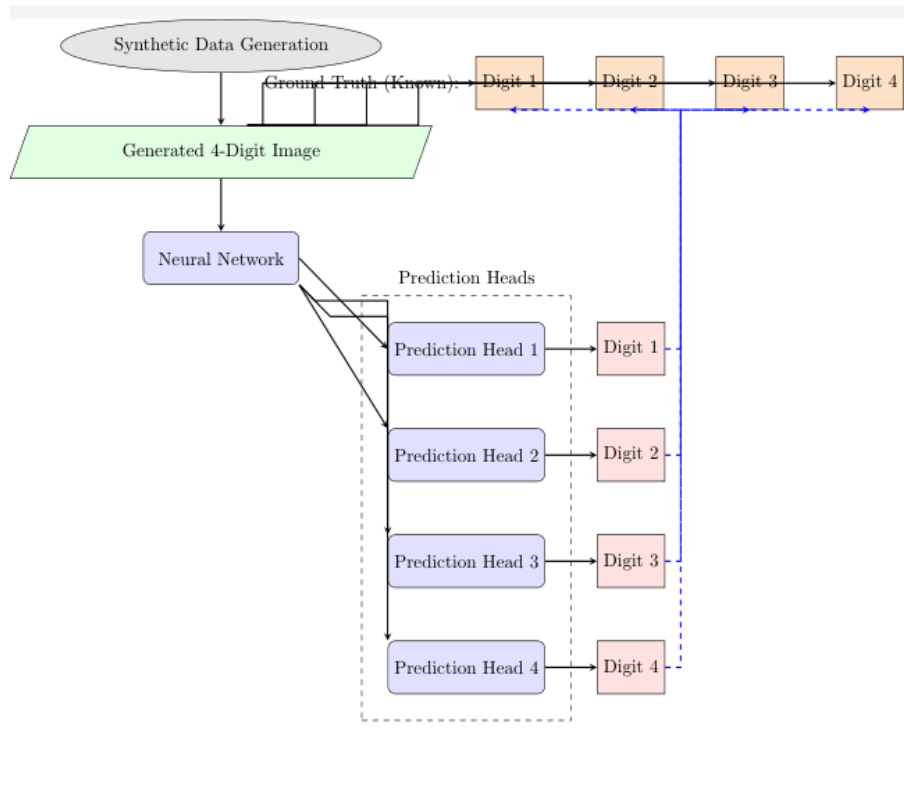- Weight decay in regularization helps in reducing overfitting

Best accuracy reached is 10-11%

- The plots show that training process is not stable, and highly depends in initialization. Overfitting occurs soon.
- Interestingly, the model does not train at all, for first 10 epochs, where loss plateaus, and then starts decreasing again

Future approach

- The data for individual digits is not given. We can create a synthetically generated dataset, where we know all 4 digits and teach network to predict those

4 digits using 4 different prediction heads.

- Other approach is pretrain on normal MNIST. And then apply it on our dataset. We can select top4 choices in a network pretrained on multi-label classification task.