

Multi Modal Matching MNIST

Uthiralakshmi Sivaraman

label = 0



label = 4



label = 1



label = 1



label = 3



label = 1



label = 5



label = 3



label = 6



label = 2

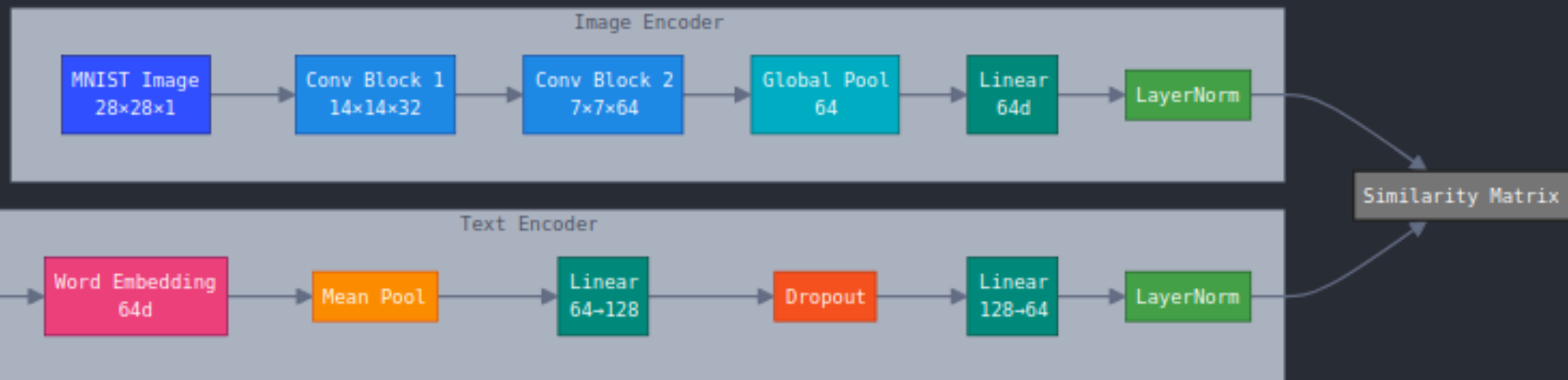


label = 8

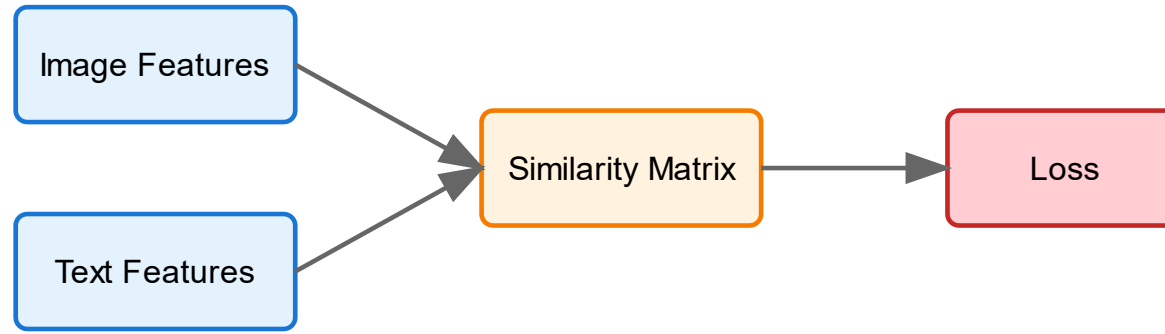


label = 6



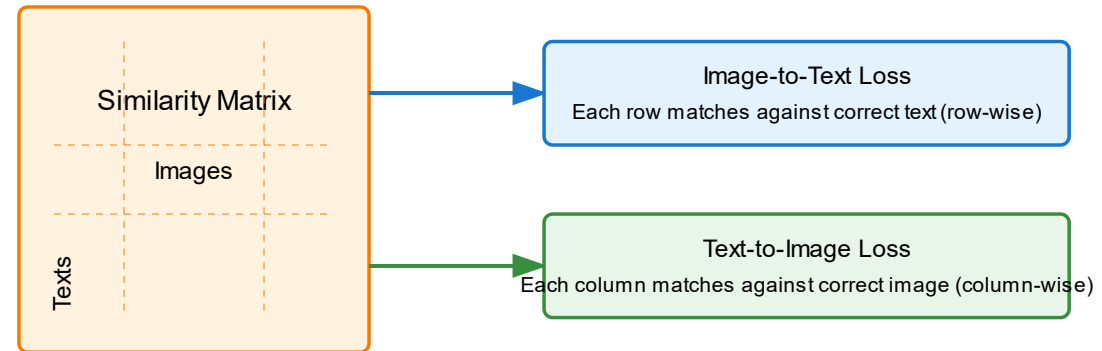


Contrastive Loss



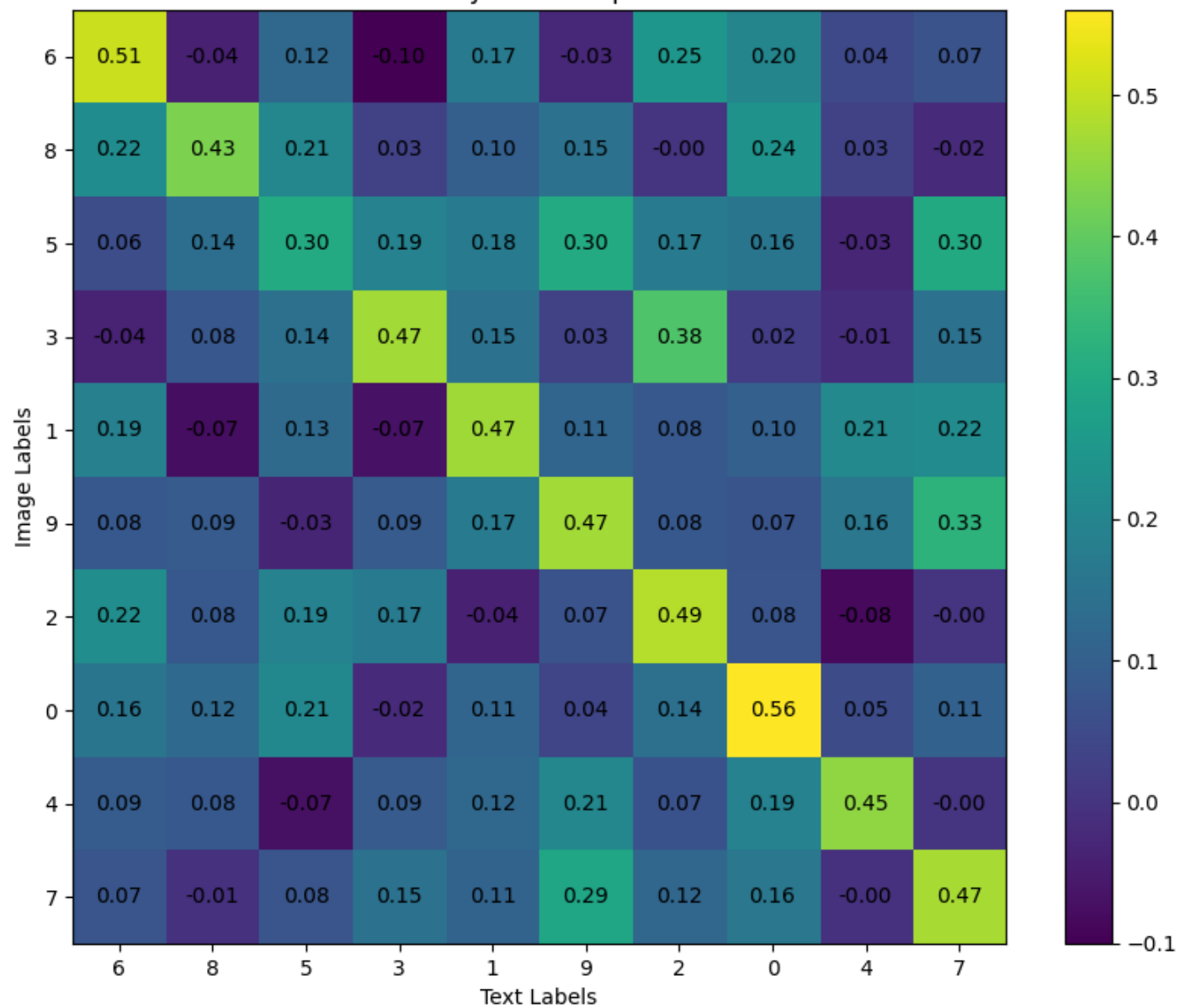
$$\text{loss} = (\text{I2T_loss} + \text{T2I_loss}) / 2$$

Image-to-Text vs Text-to-Image Loss

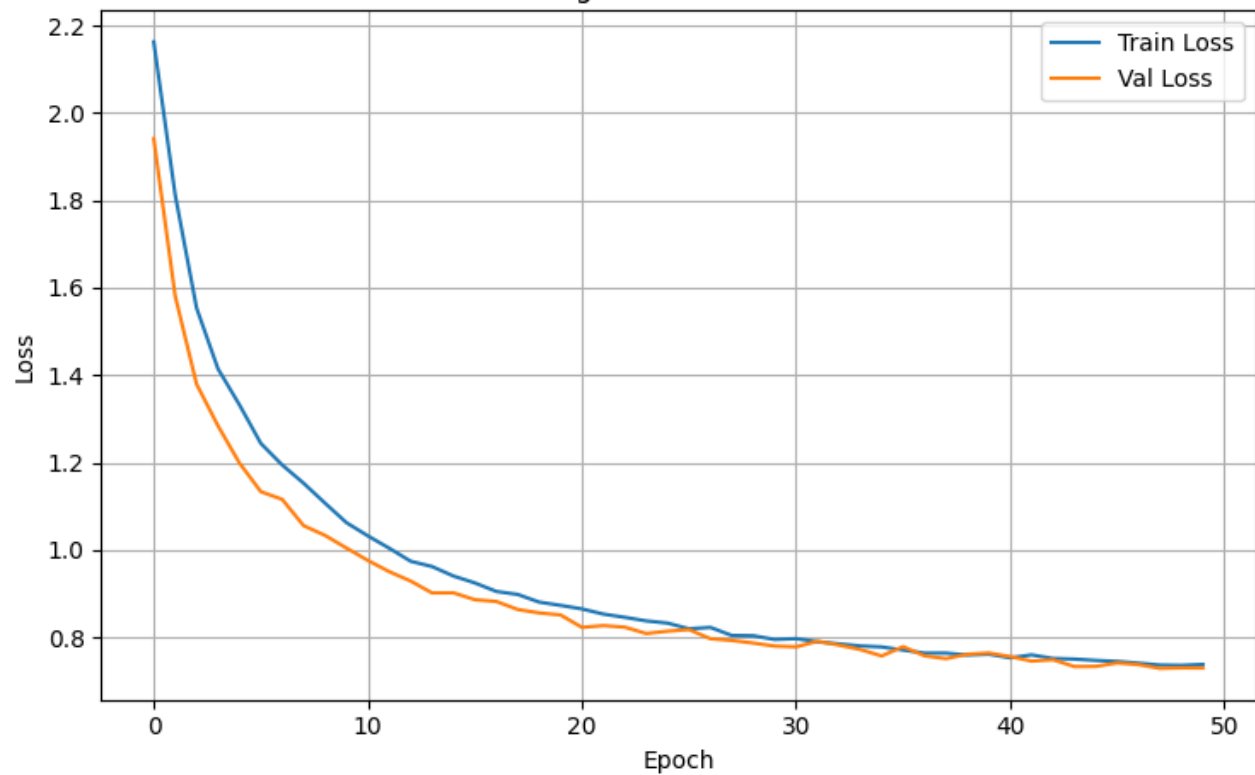


$$\text{Similarity}[i,j] = \cos(\text{image_i}, \text{text_j})$$

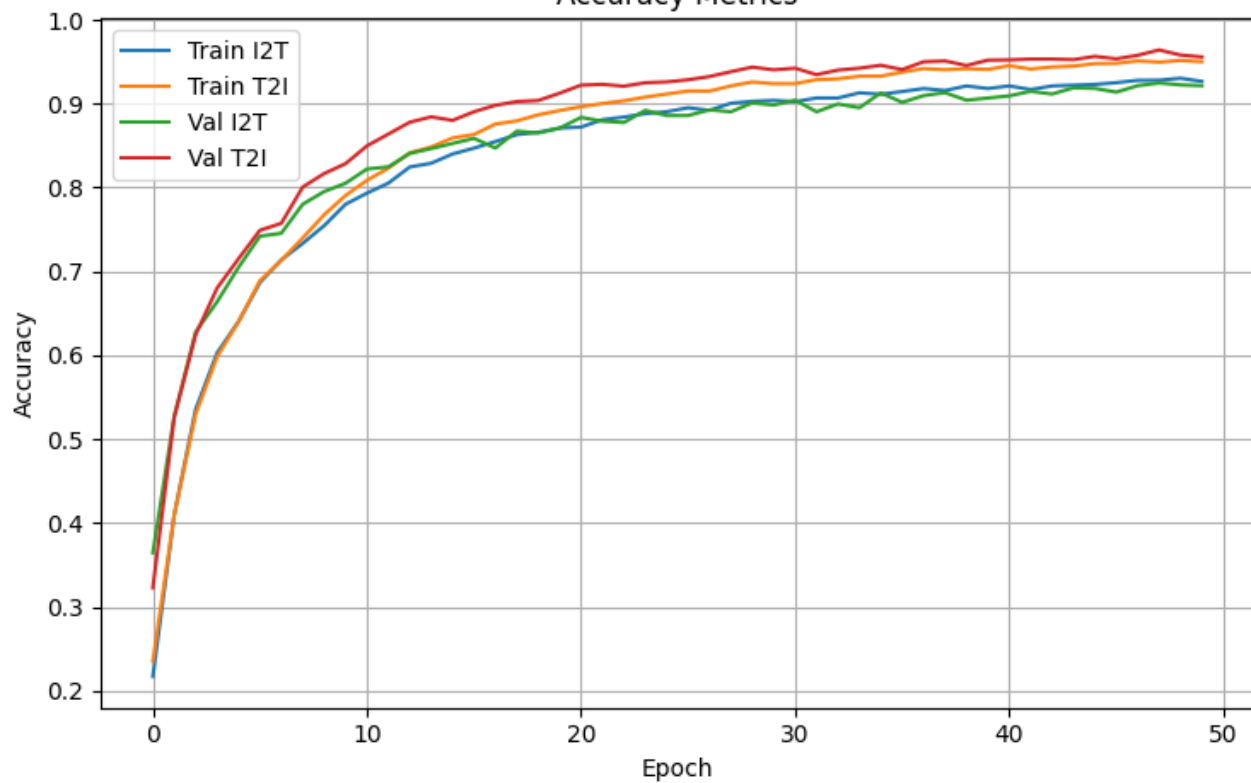
Similarity Matrix - Epoch 50



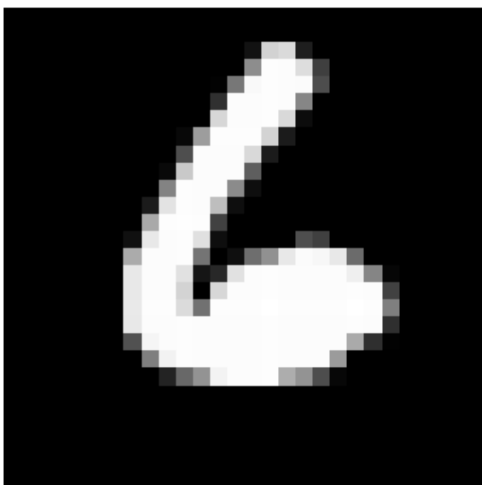
Training and Validation Loss



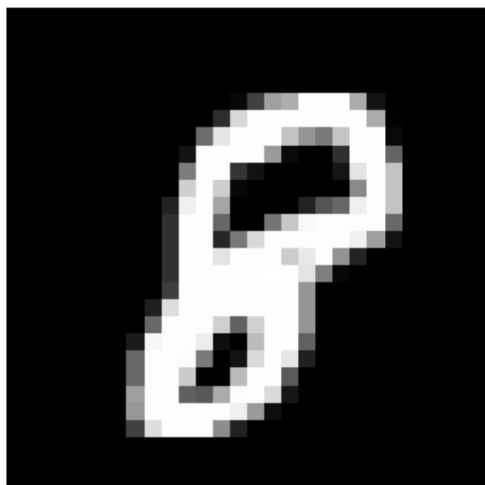
Accuracy Metrics



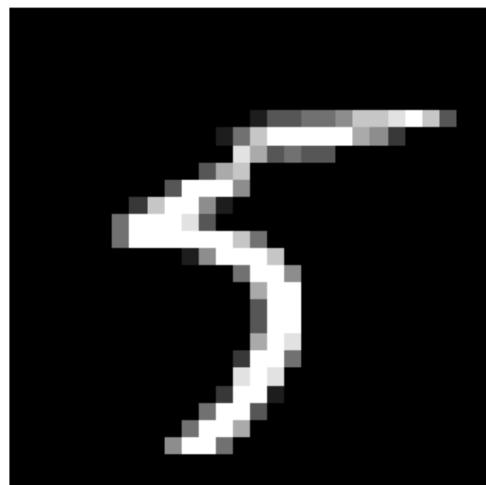
Label: 6
Pred: 6



Label: 8
Pred: 8



Label: 5
Pred: 5



Label: 3
Pred: 3



Label: 1
Pred: 1

