

NAME: Sohini Gupta

Roll No: 20EE38032

Assignment-4: Convolutional Neural Network

Problem Statement: A social media company wants to design a method which differentiates images uploaded by different users in the social media into one of several classes. The purpose is to identify and block malicious images uploaded by any user in the social media before it spreads. The different classes are: (a) malicious, (b) acceptable, and (c) completely safe. While collecting and labelling the dataset, the engineers want to try their method on a toy dataset (e.g., MNIST data) to analyse different practical issues. Develop a CNN and study the effect of various aspects of the network to help the team.

Experiment 1: Comparing the performances based on percentage accuracy on the test set of CNN-Vanilla and CNN-Resnet.

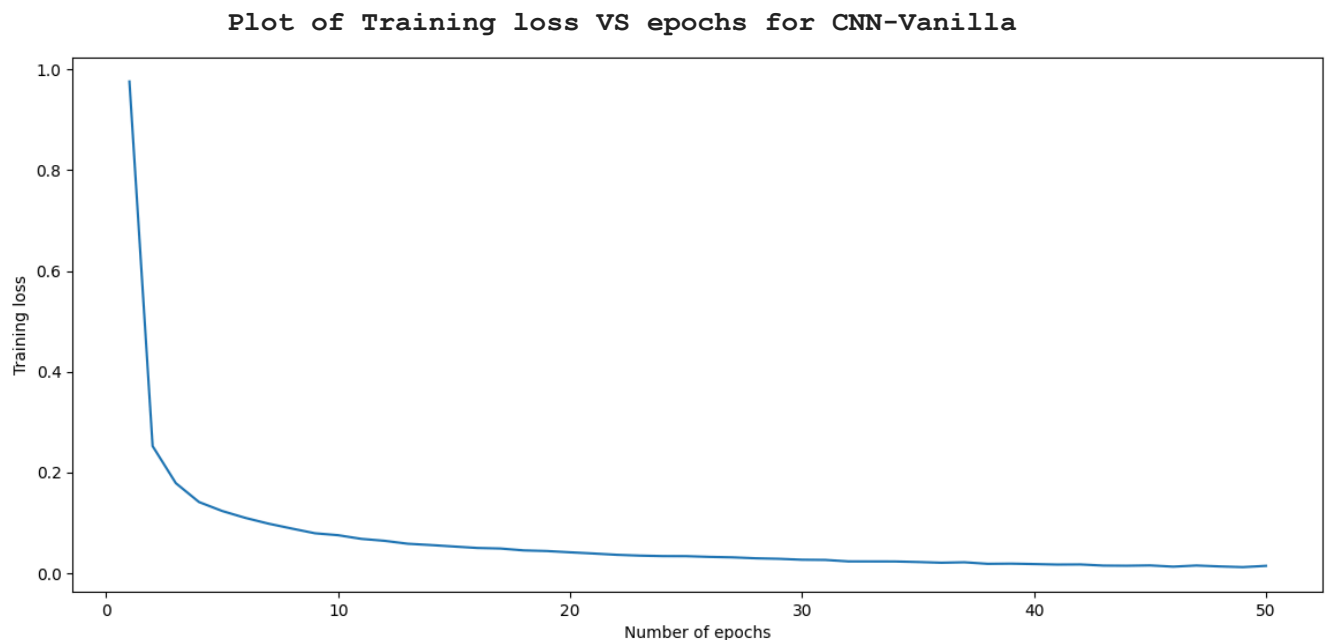
- Training CNN-Vanilla:

```
Epoch 1/50, Loss: 0.9761244615975847
Epoch 2/50, Loss: 0.2519268417612035
Epoch 3/50, Loss: 0.17838164482344973
Epoch 4/50, Loss: 0.1408456676817955
Epoch 5/50, Loss: 0.12304177063893765
Epoch 6/50, Loss: 0.10935748706155635
Epoch 7/50, Loss: 0.09785971357784373
Epoch 8/50, Loss: 0.08819040564939062
Epoch 9/50, Loss: 0.07878881459857555
Epoch 10/50, Loss: 0.07494902645653867
Epoch 11/50, Loss: 0.0676718099082404
Epoch 12/50, Loss: 0.06381981098588477
Epoch 13/50, Loss: 0.058021821152973684
Epoch 14/50, Loss: 0.05539839327176835
Epoch 15/50, Loss: 0.05247074815821141
Epoch 16/50, Loss: 0.04960295494170265
Epoch 17/50, Loss: 0.0485257507777436
Epoch 18/50, Loss: 0.04477119783454753
Epoch 19/50, Loss: 0.04353786984934135
Epoch 20/50, Loss: 0.04098673242322625
Epoch 21/50, Loss: 0.03863030393072899
Epoch 22/50, Loss: 0.03605512384959358
Epoch 23/50, Loss: 0.034411481738169776
Epoch 24/50, Loss: 0.033464404052876415
Epoch 25/50, Loss: 0.03334358222386304
Epoch 26/50, Loss: 0.03188398218852408
Epoch 27/50, Loss: 0.03098222258243155
```

Epoch 28/50, Loss: 0.029098123947079193
Epoch 29/50, Loss: 0.028097523671594706
Epoch 30/50, Loss: 0.026246623628515195
Epoch 31/50, Loss: 0.025861261435289017
Epoch 32/50, Loss: 0.02294220926160825
Epoch 33/50, Loss: 0.022868796773849332
Epoch 34/50, Loss: 0.02280892118504469
Epoch 35/50, Loss: 0.021703969182922167
Epoch 36/50, Loss: 0.02028699723230873
Epoch 37/50, Loss: 0.02117068012160784
Epoch 38/50, Loss: 0.018118816298095787
Epoch 39/50, Loss: 0.018490574401902393
Epoch 40/50, Loss: 0.017661896582811754
Epoch 41/50, Loss: 0.016654024491126233
Epoch 42/50, Loss: 0.016879550952890374
Epoch 43/50, Loss: 0.014584247405303919
Epoch 44/50, Loss: 0.01432416888097539
Epoch 45/50, Loss: 0.014960730051066963
Epoch 46/50, Loss: 0.012504421351378426
Epoch 47/50, Loss: 0.01476139828607321
Epoch 48/50, Loss: 0.012949271910041174
Epoch 49/50, Loss: 0.0115775622982294
Epoch 50/50, Loss: 0.01409410607475946

- Evaluation CNN-Vanilla:

Accuracy on test set: 98.43%



- Parameter and model summary for CNN-Vanilla:

Total number of parameters for CNN-Vanilla : 13050

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 8, 28, 28]	80
ReLU-2	[-1, 8, 28, 28]	0
MaxPool2d-3	[-1, 8, 14, 14]	0
Conv2d-4	[-1, 8, 14, 14]	584
ReLU-5	[-1, 8, 14, 14]	0
Conv2d-6	[-1, 8, 14, 14]	584
ReLU-7	[-1, 8, 14, 14]	0
MaxPool2d-8	[-1, 8, 7, 7]	0
Conv2d-9	[-1, 8, 7, 7]	584
ReLU-10	[-1, 8, 7, 7]	0
Conv2d-11	[-1, 8, 7, 7]	584
ReLU-12	[-1, 8, 7, 7]	0
MaxPool2d-13	[-1, 8, 3, 3]	0
Linear-14	[-1, 128]	9,344
ReLU-15	[-1, 128]	0
Linear-16	[-1, 10]	1,290

Total params: 13,050
 Trainable params: 13,050
 Non-trainable params: 0

Training CNN-ResNet:

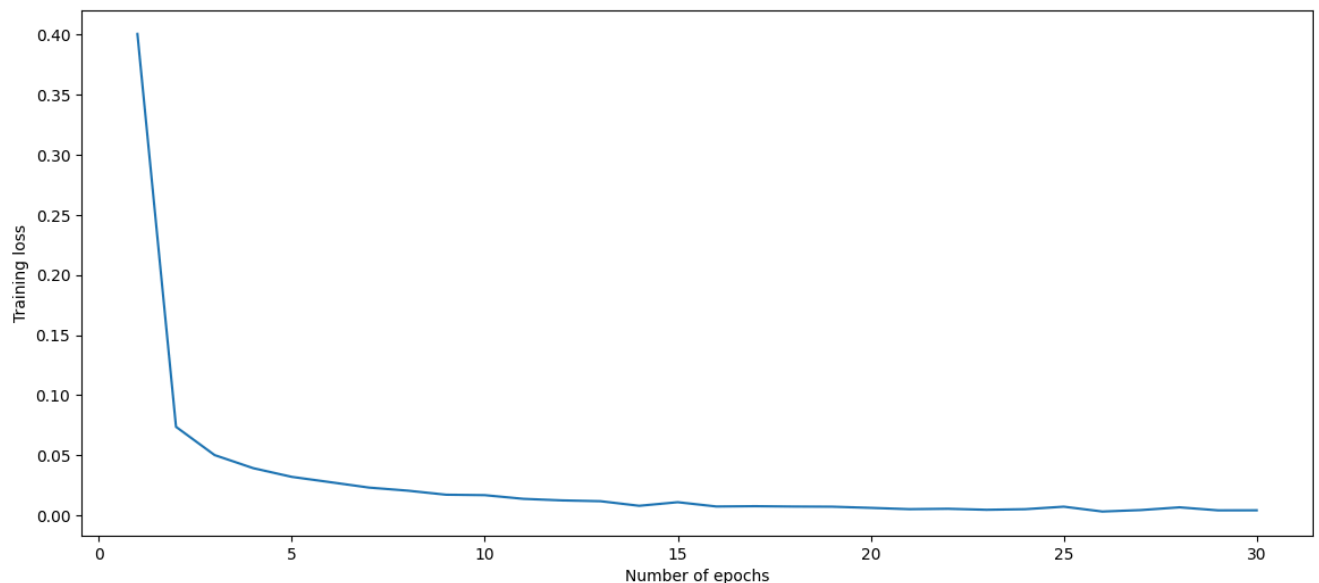
Epoch 1/50, Loss: 0.805587647760168
 Epoch 2/50, Loss: 0.20229874332534506
 Epoch 3/50, Loss: 0.13277557045221328
 Epoch 4/50, Loss: 0.10319299786648851
 Epoch 5/50, Loss: 0.08892580573546126
 Epoch 6/50, Loss: 0.07734136969802227
 Epoch 7/50, Loss: 0.07084020578322259
 Epoch 8/50, Loss: 0.06260268652534232
 Epoch 9/50, Loss: 0.059412377434683604
 Epoch 10/50, Loss: 0.056660340203249705
 Epoch 11/50, Loss: 0.05424518617305984
 Epoch 12/50, Loss: 0.049169216144512945
 Epoch 13/50, Loss: 0.04833057006385098
 Epoch 14/50, Loss: 0.04361765031167801
 Epoch 15/50, Loss: 0.042441547511422886
 Epoch 16/50, Loss: 0.0398213991578272
 Epoch 17/50, Loss: 0.03723645491842577
 Epoch 18/50, Loss: 0.03605773919162915
 Epoch 19/50, Loss: 0.034211552940665725
 Epoch 20/50, Loss: 0.03265385717312072
 Epoch 21/50, Loss: 0.03231986639030436
 Epoch 22/50, Loss: 0.03136911250730144
 Epoch 23/50, Loss: 0.029832398805252097
 Epoch 24/50, Loss: 0.02851523892914361

Epoch 25/50, Loss: 0.02579776517888333
Epoch 26/50, Loss: 0.02606196698356182
Epoch 27/50, Loss: 0.023760175871405196
Epoch 28/50, Loss: 0.024160647455682145
Epoch 29/50, Loss: 0.022599676300949872
Epoch 30/50, Loss: 0.022028549227546504
Epoch 31/50, Loss: 0.020787528320711027
Epoch 32/50, Loss: 0.019712826785720648
Epoch 33/50, Loss: 0.018905430856319343
Epoch 34/50, Loss: 0.019823413233252915
Epoch 35/50, Loss: 0.018076888078506638
Epoch 36/50, Loss: 0.019377835708709948
Epoch 37/50, Loss: 0.0162894758068897
Epoch 38/50, Loss: 0.015608794804404866
Epoch 39/50, Loss: 0.016385839686560582
Epoch 40/50, Loss: 0.015945075482050788
Epoch 41/50, Loss: 0.014046920696025754
Epoch 42/50, Loss: 0.013891659485562605
Epoch 43/50, Loss: 0.0132876122917762
Epoch 44/50, Loss: 0.01271031607744938
Epoch 45/50, Loss: 0.0119654926097397
Epoch 46/50, Loss: 0.011833688432965032
Epoch 47/50, Loss: 0.014424494779520134
Epoch 48/50, Loss: 0.012241905977196516
Epoch 49/50, Loss: 0.009968719807852237
Epoch 50/50, Loss: 0.00970513673927734

Evaluation CNN-ResNet:

Accuracy on test set: 98.66%

Plot of Training loss VS epochs for CNN-ResNet



- Parameter and model summary for CNN-ResNet:

```
-----
Total number of parameters for CNN-Resnet : 13050
-----
Layer (type)              Output Shape          Param #
-----
Conv2d-1                  [-1, 8, 28, 28]       80
MaxPool2d-2               [-1, 8, 14, 14]       0
ReLU-3                   [-1, 8, 14, 14]       0
Conv2d-4                  [-1, 8, 14, 14]       584
ReLU-5                   [-1, 8, 14, 14]       0
Conv2d-6                  [-1, 8, 14, 14]       584
ReLU-7                   [-1, 8, 14, 14]       0
ResidualBlock-8           [-1, 8, 14, 14]       0
MaxPool2d-9              [-1, 8, 7, 7]         0
Conv2d-10                 [-1, 8, 7, 7]         584
ReLU-11                  [-1, 8, 7, 7]         0
Conv2d-12                 [-1, 8, 7, 7]         584
ReLU-13                  [-1, 8, 7, 7]         0
ResidualBlock-14          [-1, 8, 7, 7]         0
MaxPool2d-15             [-1, 8, 3, 3]         0
Linear-16                 [-1, 128]             9,344
ReLU-17                  [-1, 128]             0
Linear-18                 [-1, 10]              1,290
=====
Total params: 13,050
Trainable params: 13,050
Non-trainable params: 0
```

Conclusion: We observe that CNN-ResNet has a better accuracy and has the same number of parameters as of CNN-Vanilla. So we will use CNN-ResNet for further experiments as our model.

Experiment 2: Study the Effect of Data Normalization

- Training CNN-ResNet with normalised dataset:

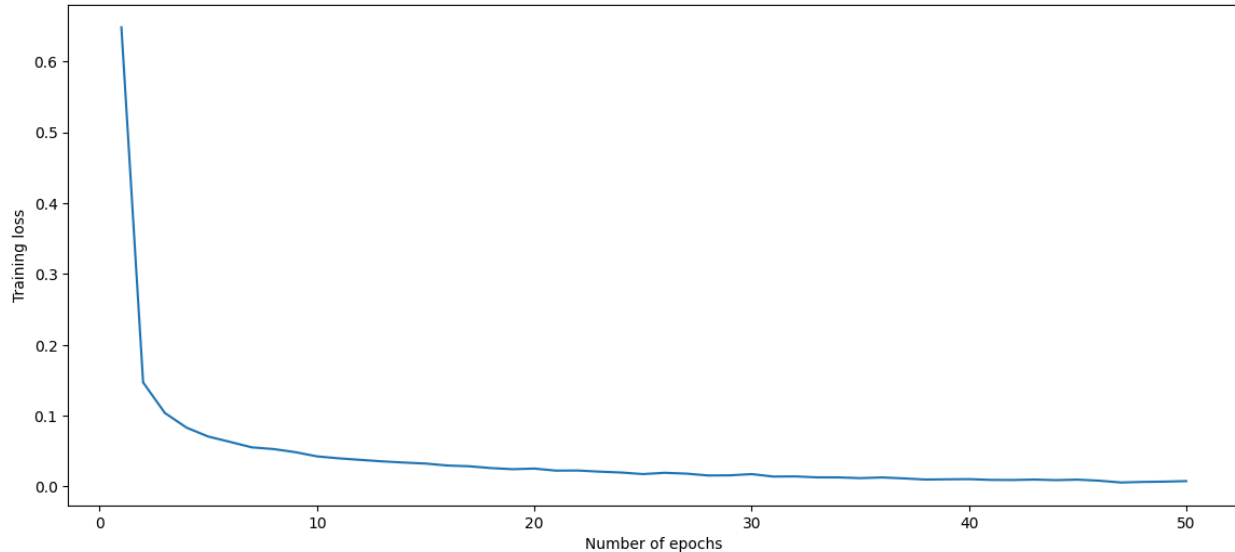
```
Epoch 1/50, Loss: 0.6484024709209483
Epoch 2/50, Loss: 0.14702002881689274
Epoch 3/50, Loss: 0.10393430898164181
Epoch 4/50, Loss: 0.08318741839141287
Epoch 5/50, Loss: 0.07060179858607181
Epoch 6/50, Loss: 0.06307626355518686
Epoch 7/50, Loss: 0.0553758335517759
Epoch 8/50, Loss: 0.05297935221423494
Epoch 9/50, Loss: 0.048595332025371966
Epoch 10/50, Loss: 0.042626540909739254
Epoch 11/50, Loss: 0.039872268928175275
```

Epoch 12/50, Loss: 0.03768440044703002
Epoch 13/50, Loss: 0.035619099608602676
Epoch 14/50, Loss: 0.03389414274629126
Epoch 15/50, Loss: 0.03248360952005742
Epoch 16/50, Loss: 0.02972621902863079
Epoch 17/50, Loss: 0.02867700754209085
Epoch 18/50, Loss: 0.02614593227057064
Epoch 19/50, Loss: 0.02453864065910432
Epoch 20/50, Loss: 0.025388863275540952
Epoch 21/50, Loss: 0.02251039818028345
Epoch 22/50, Loss: 0.022663360712257153
Epoch 23/50, Loss: 0.02107475284992301
Epoch 24/50, Loss: 0.01985214458768593
Epoch 25/50, Loss: 0.017650683222715048
Epoch 26/50, Loss: 0.01941754474958524
Epoch 27/50, Loss: 0.01829225951607557
Epoch 28/50, Loss: 0.01568454677892968
Epoch 29/50, Loss: 0.015878915534097145
Epoch 30/50, Loss: 0.017610655613045426
Epoch 31/50, Loss: 0.014121213843401996
Epoch 32/50, Loss: 0.014368912942727354
Epoch 33/50, Loss: 0.013020002887772514
Epoch 34/50, Loss: 0.012956364803323324
Epoch 35/50, Loss: 0.01190712738477506
Epoch 36/50, Loss: 0.012875318442758648
Epoch 37/50, Loss: 0.011557457250440146
Epoch 38/50, Loss: 0.009927480516118057
Epoch 39/50, Loss: 0.010253868848234019
Epoch 40/50, Loss: 0.010495138135268057
Epoch 41/50, Loss: 0.009463562389731684
Epoch 42/50, Loss: 0.009302818230130372
Epoch 43/50, Loss: 0.00992180288538951
Epoch 44/50, Loss: 0.00909292045695846
Epoch 45/50, Loss: 0.009808657078587628
Epoch 46/50, Loss: 0.00828739877968551
Epoch 47/50, Loss: 0.005760433790187272
Epoch 48/50, Loss: 0.006566614961434908
Epoch 49/50, Loss: 0.0069742055022791825
Epoch 50/50, Loss: 0.007735135946905815

- Evaluation CNN-ResNet with normalised dataset:

Accuracy on test set: 99.03%

Plot of Training loss VS epochs for CNN-ResNet with normalisation



*The plots for Resnet model without data normalisation is plotted in the previous part.

Conclusion: It is observed that by normalising the mnist dataset the performance/accuracy has increased. So we will use normalised dataset for further experiments.

Experiment 3: Comparing with different optimizers

- (a) Mini-batch gradient descent with no momentum,
- (b) Mini-batch gradient descent with momentum 0.9, and
- (c) ADAM optimizer.

a) Training and evaluating Mini-batch gradient descent (no momentum)

```
Epoch 1/50, Loss: 2.3051189077661394
Epoch 2/50, Loss: 2.251205005037024
Epoch 3/50, Loss: 2.2141021515460726
Epoch 4/50, Loss: 2.175894077788008
Epoch 5/50, Loss: 2.1302180919241396
Epoch 6/50, Loss: 2.072172954234671
Epoch 7/50, Loss: 1.996814801337871
Epoch 8/50, Loss: 1.8990167922161996
Epoch 9/50, Loss: 1.7755892418800516
Epoch 10/50, Loss: 1.6267614684206375
Epoch 11/50, Loss: 1.4573340324645347
Epoch 12/50, Loss: 1.2770408594861944
```

Epoch 13/50, Loss: 1.0961448732842791
Epoch 14/50, Loss: 0.9304797474374162
Epoch 15/50, Loss: 0.7927686384383669
Epoch 16/50, Loss: 0.6849908368384584
Epoch 17/50, Loss: 0.6033082286094097
Epoch 18/50, Loss: 0.5404425501823426
Epoch 19/50, Loss: 0.49029789503584514
Epoch 20/50, Loss: 0.44969758695744455
Epoch 21/50, Loss: 0.4159506599953834
Epoch 22/50, Loss: 0.3869010345732912
Epoch 23/50, Loss: 0.36220293666454073
Epoch 24/50, Loss: 0.341105761299742
Epoch 25/50, Loss: 0.3224016951119646
Epoch 26/50, Loss: 0.3063670238281818
Epoch 27/50, Loss: 0.2924469214170537
Epoch 28/50, Loss: 0.27945024802329693
Epoch 29/50, Loss: 0.2683769436592751
Epoch 30/50, Loss: 0.2581146326470882
Epoch 31/50, Loss: 0.24909031245302646
Epoch 32/50, Loss: 0.24035500605055626
Epoch 33/50, Loss: 0.233104330047648
Epoch 34/50, Loss: 0.226242004620268
Epoch 35/50, Loss: 0.21972827724319824
Epoch 36/50, Loss: 0.21417372068192098
Epoch 37/50, Loss: 0.20854432903071668
Epoch 38/50, Loss: 0.20372584617518363
Epoch 39/50, Loss: 0.19914626063818627
Epoch 40/50, Loss: 0.19486592947802645
Epoch 41/50, Loss: 0.1907511832866263
Epoch 42/50, Loss: 0.18681329095617255
Epoch 43/50, Loss: 0.18299022976388324
Epoch 44/50, Loss: 0.18016700231014415
Epoch 45/50, Loss: 0.17671565676623202
Epoch 46/50, Loss: 0.17366647713995995
Epoch 47/50, Loss: 0.17121942946885493
Epoch 48/50, Loss: 0.16799581225248092
Epoch 49/50, Loss: 0.1658839522841129
Epoch 50/50, Loss: 0.16300840653637622

Accuracy on test set: 95.35%

b) Training and evaluating Mini-batch gradient descent (momentum 0.9)

Epoch 1/50, Loss: 1.928415032650562
Epoch 2/50, Loss: 0.5726541136173492
Epoch 3/50, Loss: 0.2731964146203183
Epoch 4/50, Loss: 0.20458849541684415

Epoch 5/50, Loss: 0.1744706243593642
Epoch 6/50, Loss: 0.15218398591622392
Epoch 7/50, Loss: 0.13721647576448764
Epoch 8/50, Loss: 0.12442149668931961
Epoch 9/50, Loss: 0.11674724046854262
Epoch 10/50, Loss: 0.10794029096339611
Epoch 11/50, Loss: 0.102811378700302
Epoch 12/50, Loss: 0.09715269709679675
Epoch 13/50, Loss: 0.09242836939210587
Epoch 14/50, Loss: 0.0876563668647345
Epoch 15/50, Loss: 0.08434116302969608
Epoch 16/50, Loss: 0.08073519697373217
Epoch 17/50, Loss: 0.07926243925665287
Epoch 18/50, Loss: 0.07535338736277945
Epoch 19/50, Loss: 0.07230698136573142
Epoch 20/50, Loss: 0.07086031837190719
Epoch 21/50, Loss: 0.06939682037113828
Epoch 22/50, Loss: 0.06605454387658453
Epoch 23/50, Loss: 0.06491129166109764
Epoch 24/50, Loss: 0.06350759337557123
Epoch 25/50, Loss: 0.061981012212469225
Epoch 26/50, Loss: 0.061595966722419924
Epoch 27/50, Loss: 0.05884892733252429
Epoch 28/50, Loss: 0.059024996484847776
Epoch 29/50, Loss: 0.05686309192171122
Epoch 30/50, Loss: 0.05576980511083248
Epoch 31/50, Loss: 0.05502572642995956
Epoch 32/50, Loss: 0.05252141659247114
Epoch 33/50, Loss: 0.051426603141775794
Epoch 34/50, Loss: 0.052575820006747195
Epoch 35/50, Loss: 0.05123426447364878
Epoch 36/50, Loss: 0.04984635781258979
Epoch 37/50, Loss: 0.04935960080870923
Epoch 38/50, Loss: 0.04879905348841814
Epoch 39/50, Loss: 0.04699886300541619
Epoch 40/50, Loss: 0.047246283241250416
Epoch 41/50, Loss: 0.04578722709235042
Epoch 42/50, Loss: 0.04472317153072738
Epoch 43/50, Loss: 0.04462970772242927
Epoch 44/50, Loss: 0.04423435741361786
Epoch 45/50, Loss: 0.04346843706443906
Epoch 46/50, Loss: 0.043161289128375815
Epoch 47/50, Loss: 0.04223403904507769
Epoch 48/50, Loss: 0.042102372495734945
Epoch 49/50, Loss: 0.04189899903028569
Epoch 50/50, Loss: 0.040925678258088044

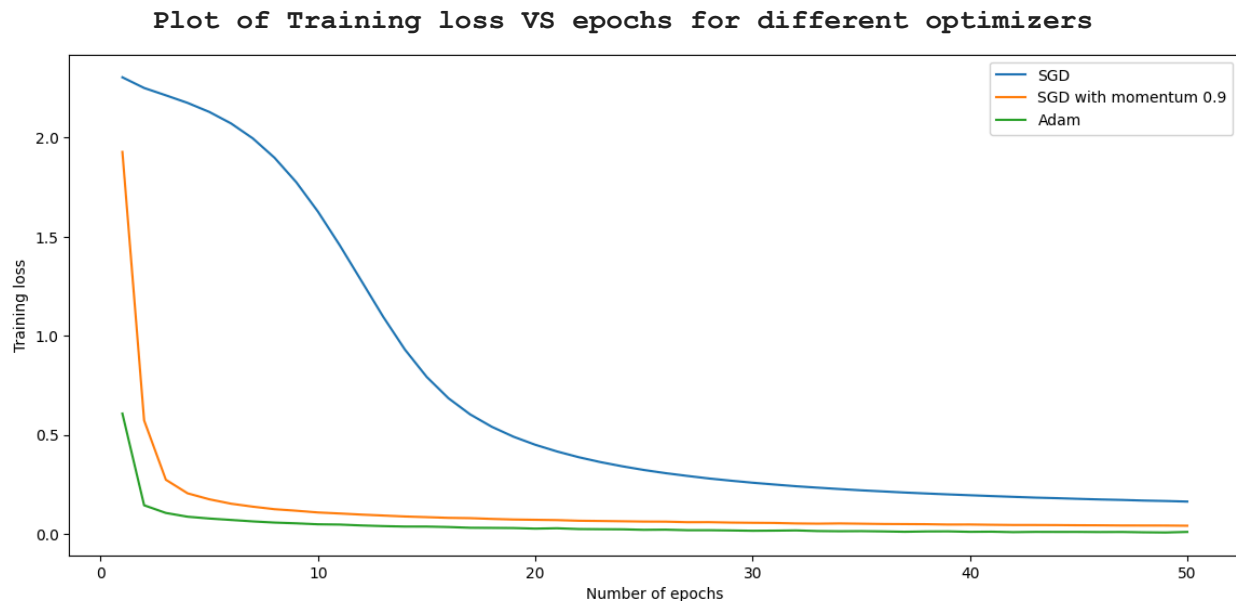
Accuracy on test set: 98.59%

c) Training and evaluating Adam optimizer

```
Epoch 1/50, Loss: 0.6073090026036222
Epoch 2/50, Loss: 0.144256937361144
Epoch 3/50, Loss: 0.10554063654643424
Epoch 4/50, Loss: 0.08645783526466248
Epoch 5/50, Loss: 0.07735604942637556
Epoch 6/50, Loss: 0.07012732576658116
Epoch 7/50, Loss: 0.06283265582107483
Epoch 8/50, Loss: 0.05691362143038435
Epoch 9/50, Loss: 0.05310243894445135
Epoch 10/50, Loss: 0.04829589892575081
Epoch 11/50, Loss: 0.04688991438201133
Epoch 12/50, Loss: 0.042404023427119915
Epoch 13/50, Loss: 0.03910866347636948
Epoch 14/50, Loss: 0.03684321497269767
Epoch 15/50, Loss: 0.03674800278163178
Epoch 16/50, Loss: 0.034453679349749015
Epoch 17/50, Loss: 0.03068767506331998
Epoch 18/50, Loss: 0.030051653518123514
Epoch 19/50, Loss: 0.029308441857946045
Epoch 20/50, Loss: 0.02637273715610834
Epoch 21/50, Loss: 0.028164375817125782
Epoch 22/50, Loss: 0.02422021493276066
Epoch 23/50, Loss: 0.02342410478344623
Epoch 24/50, Loss: 0.022886223826517768
Epoch 25/50, Loss: 0.02037169616966647
Epoch 26/50, Loss: 0.021086434668247053
Epoch 27/50, Loss: 0.018445922603751434
Epoch 28/50, Loss: 0.018366575280719614
Epoch 29/50, Loss: 0.017130665854610344
Epoch 30/50, Loss: 0.015515418743536352
Epoch 31/50, Loss: 0.016158311777411307
Epoch 32/50, Loss: 0.017277628651007692
Epoch 33/50, Loss: 0.014060793297261277
Epoch 34/50, Loss: 0.013159806763832556
Epoch 35/50, Loss: 0.013662382509519762
Epoch 36/50, Loss: 0.012353644325853662
Epoch 37/50, Loss: 0.010411273308861208
Epoch 38/50, Loss: 0.012165724742721687
Epoch 39/50, Loss: 0.012567574811846692
Epoch 40/50, Loss: 0.010142548875808914
Epoch 41/50, Loss: 0.010865845189162312
Epoch 42/50, Loss: 0.008611009197255121
Epoch 43/50, Loss: 0.009881512587456568
Epoch 44/50, Loss: 0.009673889451709754
Epoch 45/50, Loss: 0.009715690109086163
Epoch 46/50, Loss: 0.009052235031985935
```

Epoch 47/50, Loss: 0.00946286856203013
Epoch 48/50, Loss: 0.007728556769236209
Epoch 49/50, Loss: 0.006819160689477205
Epoch 50/50, Loss: 0.009751776445463815

Accuracy on test set: 98.97%



Conclusion: From the plot it is evident that the training loss converges much faster in case of **Adam Optimizer** and the absolute value of loss is much smaller as compared to the other two optimizers. Hence, we will use **Adam as our Optimizer**.

Experiment 4: Study the Effect of Network Depth:

For the Resnet network (CNN-Resnet) and best choice of normalization and Adam Optimizer, we change the depth of the network (i.e., from two level Resnet blocks with two fully-connected layers) as follows: (a) Three level Resnet block with two fully-connected layers; (b) Two level Resnet blocks with four fully-connected layers

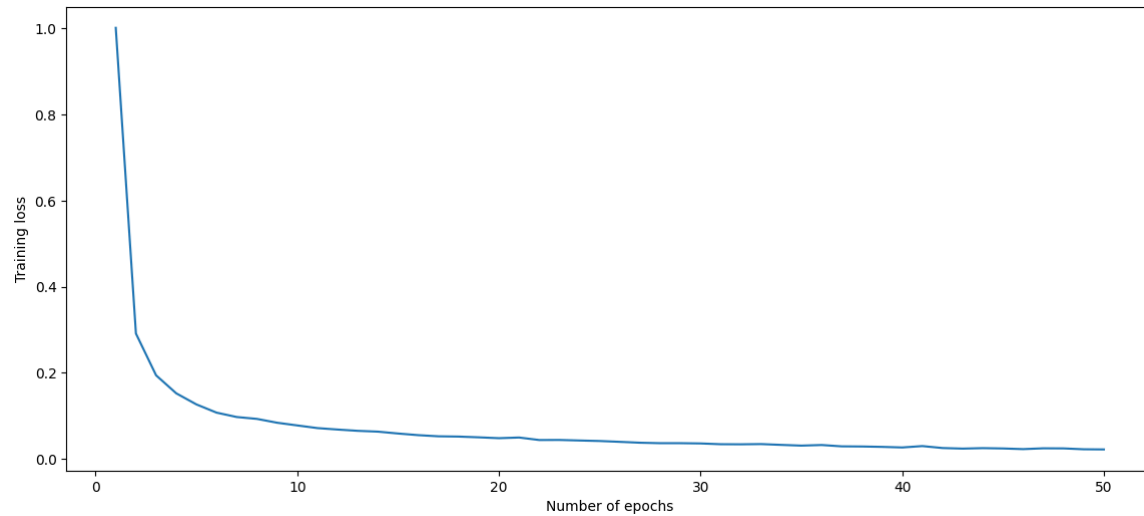
- **Training CNN-ResNet PART A:**

Epoch 1/50, Loss: 1.0014464392306957
Epoch 2/50, Loss: 0.2912398382704309
Epoch 3/50, Loss: 0.194161604408254
Epoch 4/50, Loss: 0.15233091127048148
Epoch 5/50, Loss: 0.1265899694029321
Epoch 6/50, Loss: 0.10749039187076244
Epoch 7/50, Loss: 0.09731506359703998
Epoch 8/50, Loss: 0.09297700491832926
Epoch 9/50, Loss: 0.08406925103924376
Epoch 10/50, Loss: 0.0777910570039394
Epoch 11/50, Loss: 0.07159972660084989

Epoch 12/50, Loss: 0.06824770229769514
Epoch 13/50, Loss: 0.06523650754639443
Epoch 14/50, Loss: 0.0633605286360104
Epoch 15/50, Loss: 0.059003287181258204
Epoch 16/50, Loss: 0.0552085664083666
Epoch 17/50, Loss: 0.05252900993491107
Epoch 18/50, Loss: 0.05189536445714692
Epoch 19/50, Loss: 0.05015117236353615
Epoch 20/50, Loss: 0.04815537022387094
Epoch 21/50, Loss: 0.04970740781343998
Epoch 22/50, Loss: 0.04389332658710315
Epoch 23/50, Loss: 0.044046681952920365
Epoch 24/50, Loss: 0.04278556890785694
Epoch 25/50, Loss: 0.041600161874072346
Epoch 26/50, Loss: 0.03965748340644418
Epoch 27/50, Loss: 0.03760983678016891
Epoch 28/50, Loss: 0.03653716751513012
Epoch 29/50, Loss: 0.03652938913474692
Epoch 30/50, Loss: 0.03594079629141599
Epoch 31/50, Loss: 0.03413633931220133
Epoch 32/50, Loss: 0.03392831566644476
Epoch 33/50, Loss: 0.0344687308898156
Epoch 34/50, Loss: 0.03259971872606176
Epoch 35/50, Loss: 0.030880318961522362
Epoch 36/50, Loss: 0.03231383206996512
Epoch 37/50, Loss: 0.0291806131621149
Epoch 38/50, Loss: 0.02885992286449417
Epoch 39/50, Loss: 0.02798536578787768
Epoch 40/50, Loss: 0.026598646728559695
Epoch 41/50, Loss: 0.029828345177179954
Epoch 42/50, Loss: 0.025285691358743513
Epoch 43/50, Loss: 0.023948007943901294
Epoch 44/50, Loss: 0.025042507955011853
Epoch 45/50, Loss: 0.024283145989311184
Epoch 46/50, Loss: 0.022775425144689197
Epoch 47/50, Loss: 0.02462695594579774
Epoch 48/50, Loss: 0.02443523506009753
Epoch 49/50, Loss: 0.022386134815144728
Epoch 50/50, Loss: 0.022006057386108217

- Evaluation CNN-ResNet PART A:
Accuracy on test set: 98.66%

Plot of Training loss VS epochs for CNN-ResNet part a



- Parameter and model summary for CNN-ResNet part a:

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 8, 28, 28]	80
MaxPool2d-2	[-1, 8, 14, 14]	0
ReLU-3	[-1, 8, 14, 14]	0
Conv2d-4	[-1, 8, 14, 14]	584
ReLU-5	[-1, 8, 14, 14]	0
Conv2d-6	[-1, 8, 14, 14]	584
ReLU-7	[-1, 8, 14, 14]	0
ResidualBlock-8	[-1, 8, 14, 14]	0
MaxPool2d-9	[-1, 8, 7, 7]	0
Conv2d-10	[-1, 8, 7, 7]	584
ReLU-11	[-1, 8, 7, 7]	0
Conv2d-12	[-1, 8, 7, 7]	584
ReLU-13	[-1, 8, 7, 7]	0
ResidualBlock-14	[-1, 8, 7, 7]	0
MaxPool2d-15	[-1, 8, 3, 3]	0
Conv2d-16	[-1, 8, 3, 3]	584
ReLU-17	[-1, 8, 3, 3]	0
Conv2d-18	[-1, 8, 3, 3]	584
ReLU-19	[-1, 8, 3, 3]	0
ResidualBlock-20	[-1, 8, 3, 3]	0
MaxPool2d-21	[-1, 8, 1, 1]	0
Linear-22	[-1, 128]	1,152
ReLU-23	[-1, 128]	0
Linear-24	[-1, 10]	1,290

=====
 Total params: 6,026
 Trainable params: 6,026
 Non-trainable params: 0

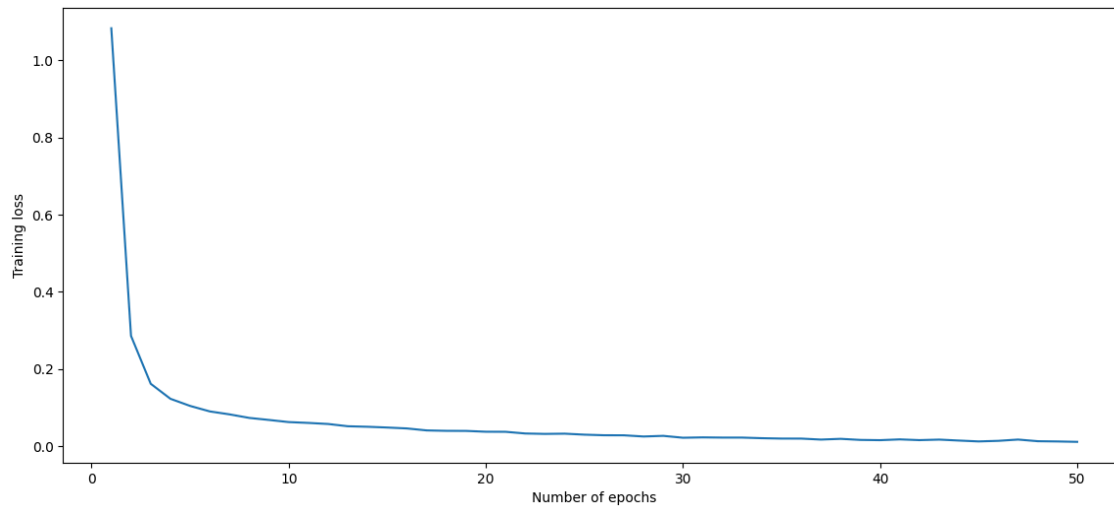
- Training CNN-ResNet PART B:

Epoch 1/50, Loss: 1.0827998600107558
 Epoch 2/50, Loss: 0.2856598934594621
 Epoch 3/50, Loss: 0.16145529842123071
 Epoch 4/50, Loss: 0.12255700937610992
 Epoch 5/50, Loss: 0.10404094104120072

Epoch 6/50, Loss: 0.08982029188662133
Epoch 7/50, Loss: 0.08216793337400924
Epoch 8/50, Loss: 0.0730063407979113
Epoch 9/50, Loss: 0.06772615613931036
Epoch 10/50, Loss: 0.06220176783885727
Epoch 11/50, Loss: 0.06003883671253286
Epoch 12/50, Loss: 0.057260573108462574
Epoch 13/50, Loss: 0.0513886413279366
Epoch 14/50, Loss: 0.05005850527197757
Epoch 15/50, Loss: 0.048080745291836716
Epoch 16/50, Loss: 0.045709623681738025
Epoch 17/50, Loss: 0.04058915986580418
Epoch 18/50, Loss: 0.03950802287324629
Epoch 19/50, Loss: 0.03932156174225693
Epoch 20/50, Loss: 0.03718476708027277
Epoch 21/50, Loss: 0.0370143132304099
Epoch 22/50, Loss: 0.03262847155887396
Epoch 23/50, Loss: 0.03156686909100477
Epoch 24/50, Loss: 0.032158281544144164
Epoch 25/50, Loss: 0.0296040799775894
Epoch 26/50, Loss: 0.02805362167431319
Epoch 27/50, Loss: 0.027930817059538465
Epoch 28/50, Loss: 0.024831074119565335
Epoch 29/50, Loss: 0.026488949949456816
Epoch 30/50, Loss: 0.02160987642970174
Epoch 31/50, Loss: 0.022538331211445495
Epoch 32/50, Loss: 0.02202694927976328
Epoch 33/50, Loss: 0.021998849872281734
Epoch 34/50, Loss: 0.020426115476367798
Epoch 35/50, Loss: 0.019468072078319545
Epoch 36/50, Loss: 0.019405853226681813
Epoch 37/50, Loss: 0.016862549520087766
Epoch 38/50, Loss: 0.018768443498621754
Epoch 39/50, Loss: 0.01598385597970535
Epoch 40/50, Loss: 0.01529988857466688
Epoch 41/50, Loss: 0.01736201328507803
Epoch 42/50, Loss: 0.015415851422585548
Epoch 43/50, Loss: 0.016807980256333472
Epoch 44/50, Loss: 0.014324657691944827
Epoch 45/50, Loss: 0.012066310278060747
Epoch 46/50, Loss: 0.013572310858054404
Epoch 47/50, Loss: 0.016923602164643718
Epoch 48/50, Loss: 0.012620393700025817
Epoch 49/50, Loss: 0.011988618237243846
Epoch 50/50, Loss: 0.010956342777415634

- **Evaluation CNN-ResNet PART B:**
Accuracy on test set: 98.66%

Plot of Training loss VS epochs for CNN-ResNet Part b



- Parameter and model summary for CNN-ResNet b:

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 8, 28, 28]	80
MaxPool2d-2	[-1, 8, 14, 14]	0
ReLU-3	[-1, 8, 14, 14]	0
Conv2d-4	[-1, 8, 14, 14]	584
ReLU-5	[-1, 8, 14, 14]	0
Conv2d-6	[-1, 8, 14, 14]	584
ReLU-7	[-1, 8, 14, 14]	0
ResidualBlock-8	[-1, 8, 14, 14]	0
MaxPool2d-9	[-1, 8, 7, 7]	0
Conv2d-10	[-1, 8, 7, 7]	584
ReLU-11	[-1, 8, 7, 7]	0
Conv2d-12	[-1, 8, 7, 7]	584
ReLU-13	[-1, 8, 7, 7]	0
ResidualBlock-14	[-1, 8, 7, 7]	0
MaxPool2d-15	[-1, 8, 3, 3]	0
Linear-16	[-1, 128]	9,344
ReLU-17	[-1, 128]	0
Linear-18	[-1, 64]	8,256
ReLU-19	[-1, 64]	0
Linear-20	[-1, 32]	2,080
ReLU-21	[-1, 32]	0
Linear-22	[-1, 10]	330
Total params: 22,426		
Trainable params: 22,426		
Non-trainable params: 0		

Conclusion: Accuracy on test sets for all the ResNet models are similar which is ~ 98.66%. Since this is a Mnist dataset we can't really conclude about which resnet is better. But the number of parameters shoots up in the case where we use four fully connected layers whereas the number of parameters lessens in the case where we add a convolution layer. So, I don't think there is any benefit of increasing either convolution layers or FC layers for the model because there is not much chance of accuracy.

Experiment 5: Study the Effect of Different Regularizers. Using the following regularizers: (a) Batch Normalization, (b) Dropout, (c) Both Batch Normalization and Dropout.

a) Training and evaluating with Batch Normalization

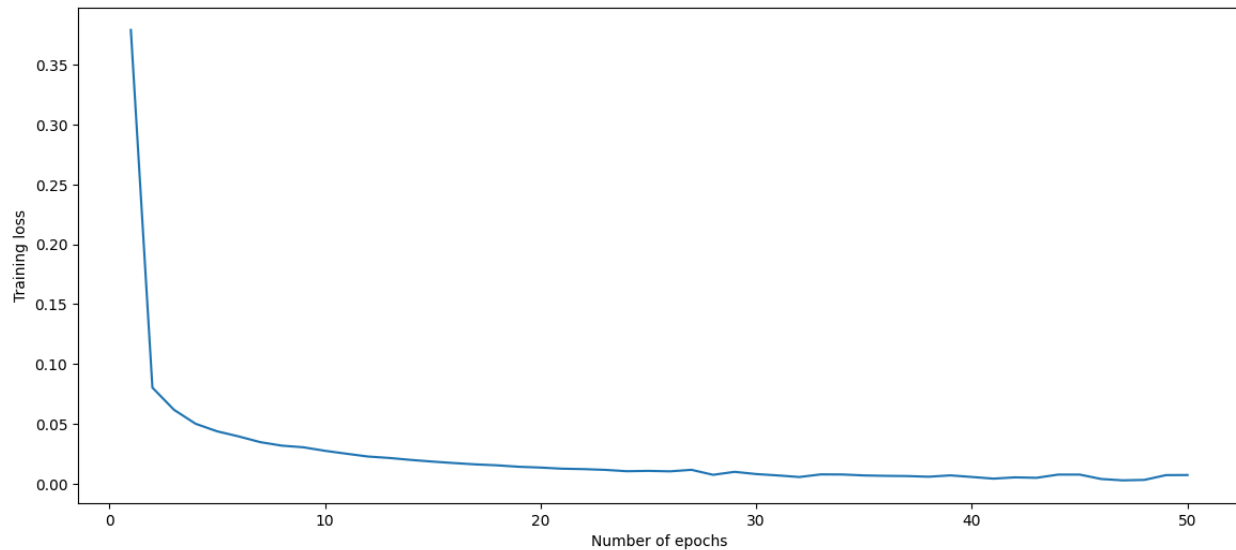
Epoch 1/50, Loss: 0.3790447359706493
Epoch 2/50, Loss: 0.0801699254899583
Epoch 3/50, Loss: 0.06165989764985886
Epoch 4/50, Loss: 0.05000122404122289
Epoch 5/50, Loss: 0.04366818060900303
Epoch 6/50, Loss: 0.0393985482289436
Epoch 7/50, Loss: 0.034625770547922624
Epoch 8/50, Loss: 0.03169367125733419
Epoch 9/50, Loss: 0.0303114901752548
Epoch 10/50, Loss: 0.02739085231293389
Epoch 11/50, Loss: 0.024937564611157523
Epoch 12/50, Loss: 0.02258580492046523
Epoch 13/50, Loss: 0.021375893422977087
Epoch 14/50, Loss: 0.019733421064596225
Epoch 15/50, Loss: 0.018348625581216145
Epoch 16/50, Loss: 0.017115485034585792
Epoch 17/50, Loss: 0.016020117607463073
Epoch 18/50, Loss: 0.015232437736097169
Epoch 19/50, Loss: 0.01403815910923592
Epoch 20/50, Loss: 0.013373457318905027
Epoch 21/50, Loss: 0.012424735801462876
Epoch 22/50, Loss: 0.012081899362160487
Epoch 23/50, Loss: 0.011435333285678891
Epoch 24/50, Loss: 0.01031886404355433
Epoch 25/50, Loss: 0.01060208028796664
Epoch 26/50, Loss: 0.010213280591291396
Epoch 27/50, Loss: 0.011483071613412827
Epoch 28/50, Loss: 0.0073120091440517735
Epoch 29/50, Loss: 0.009809721435762031
Epoch 30/50, Loss: 0.007946461723792922
Epoch 31/50, Loss: 0.006798638230774235
Epoch 32/50, Loss: 0.005473207853154894
Epoch 33/50, Loss: 0.007673872009036131
Epoch 34/50, Loss: 0.007585199975205506
Epoch 35/50, Loss: 0.006794606884687574
Epoch 36/50, Loss: 0.006453535089113234
Epoch 37/50, Loss: 0.006261281170473592
Epoch 38/50, Loss: 0.005721465532452443
Epoch 39/50, Loss: 0.006835106357794622
Epoch 40/50, Loss: 0.00551118205586
Epoch 41/50, Loss: 0.004120670009201313
Epoch 42/50, Loss: 0.005182541535171541
Epoch 43/50, Loss: 0.004806286540839751

Epoch 44/50, Loss: 0.007494402285312313
Epoch 45/50, Loss: 0.007506210905960464
Epoch 46/50, Loss: 0.003833154436684476
Epoch 47/50, Loss: 0.002694297703667271
Epoch 48/50, Loss: 0.0030738522818400597
Epoch 49/50, Loss: 0.007043031587183693
Epoch 50/50, Loss: 0.007128943502314715

Evaluation:

Accuracy on test set: 98.89%

Plot of Training loss VS epochs for CNN-ResNet with batch normalisation



b) Training and evaluating with dropouts

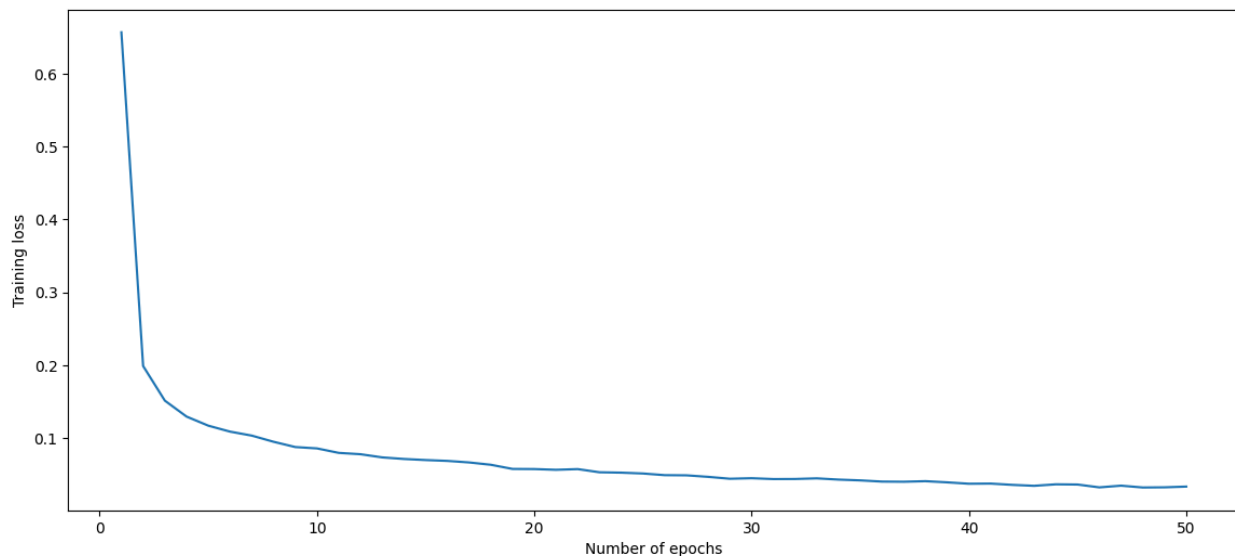
Epoch 1/50, Loss: 0.6571810764835236
Epoch 2/50, Loss: 0.1986889033558521
Epoch 3/50, Loss: 0.15108177918702997
Epoch 4/50, Loss: 0.12917554470452858
Epoch 5/50, Loss: 0.11670822929194632
Epoch 6/50, Loss: 0.1085387073615764
Epoch 7/50, Loss: 0.10305495303361974
Epoch 8/50, Loss: 0.09461521049763294
Epoch 9/50, Loss: 0.08734844685234922
Epoch 10/50, Loss: 0.08548441245517832
Epoch 11/50, Loss: 0.07934926099124107
Epoch 12/50, Loss: 0.07749325940583615
Epoch 13/50, Loss: 0.07313774403263913
Epoch 14/50, Loss: 0.07096484458826958
Epoch 15/50, Loss: 0.06947712611724087
Epoch 16/50, Loss: 0.0683098186203774
Epoch 17/50, Loss: 0.06618495254599034
Epoch 18/50, Loss: 0.06300129495719646
Epoch 19/50, Loss: 0.05727721250992506
Epoch 20/50, Loss: 0.05710693451952427
Epoch 21/50, Loss: 0.05609601483224554

Epoch 22/50, Loss: 0.05703967824340501
Epoch 23/50, Loss: 0.05268480611687645
Epoch 24/50, Loss: 0.05215653414897462
Epoch 25/50, Loss: 0.05099965068016281
Epoch 26/50, Loss: 0.04871924656661267
Epoch 27/50, Loss: 0.04848115696631213
Epoch 28/50, Loss: 0.04639456943073805
Epoch 29/50, Loss: 0.043798005121185424
Epoch 30/50, Loss: 0.04452785054974733
Epoch 31/50, Loss: 0.043378385691408144
Epoch 32/50, Loss: 0.04348265818974122
Epoch 33/50, Loss: 0.04433513787040051
Epoch 34/50, Loss: 0.04264955583642772
Epoch 35/50, Loss: 0.041547670706789545
Epoch 36/50, Loss: 0.03989603674792229
Epoch 37/50, Loss: 0.03968271494426943
Epoch 38/50, Loss: 0.04046291494742036
Epoch 39/50, Loss: 0.03884489775853271
Epoch 40/50, Loss: 0.036935967179213436
Epoch 41/50, Loss: 0.03715171945935234
Epoch 42/50, Loss: 0.035383388017641106
Epoch 43/50, Loss: 0.034076352334244456
Epoch 44/50, Loss: 0.036129971292424706
Epoch 45/50, Loss: 0.035843256940232944
Epoch 46/50, Loss: 0.031934212594035456
Epoch 47/50, Loss: 0.03428865784303622
Epoch 48/50, Loss: 0.031772714336105484
Epoch 49/50, Loss: 0.03200833206064999
Epoch 50/50, Loss: 0.03300814757401005

Evaluation :

Accuracy on test set: 98.00%

Plot of Training loss VS epochs for CNN-ResNet with dropout

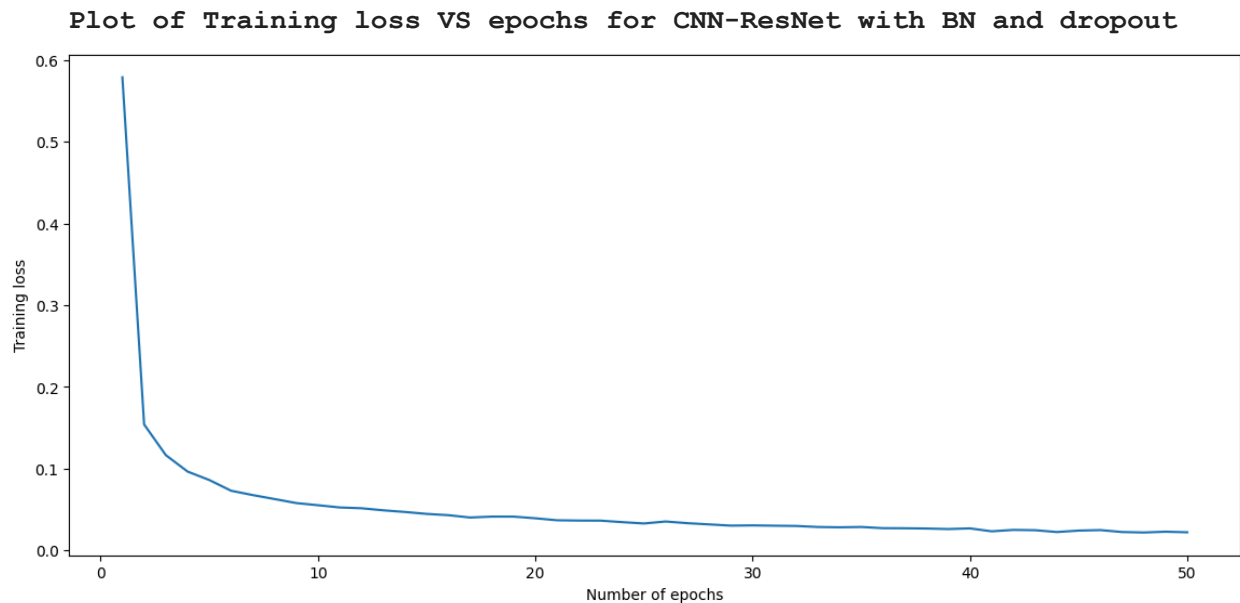


c) Training and evaluating with Batch Normalization and dropouts

Epoch 1/50, Loss: 0.5789327381456152
Epoch 2/50, Loss: 0.1540459338971909
Epoch 3/50, Loss: 0.11647601986819126
Epoch 4/50, Loss: 0.09639386600477898
Epoch 5/50, Loss: 0.08602759355401739
Epoch 6/50, Loss: 0.07294359640713702
Epoch 7/50, Loss: 0.0675905145110285
Epoch 8/50, Loss: 0.06277678172956122
Epoch 9/50, Loss: 0.05782856972135128
Epoch 10/50, Loss: 0.05519077599524184
Epoch 11/50, Loss: 0.05244276235474551
Epoch 12/50, Loss: 0.05148923553209356
Epoch 13/50, Loss: 0.04896592815267913
Epoch 14/50, Loss: 0.04695157576352358
Epoch 15/50, Loss: 0.04457515769816459
Epoch 16/50, Loss: 0.04300814203600934
Epoch 17/50, Loss: 0.040151738302108456
Epoch 18/50, Loss: 0.041285887283292855
Epoch 19/50, Loss: 0.04125227261176135
Epoch 20/50, Loss: 0.039225761998603316
Epoch 21/50, Loss: 0.036762547342384114
Epoch 22/50, Loss: 0.03640552310649543
Epoch 23/50, Loss: 0.03631485831190614
Epoch 24/50, Loss: 0.034497185603299355
Epoch 25/50, Loss: 0.032868215041433245
Epoch 26/50, Loss: 0.03527219260428497
Epoch 27/50, Loss: 0.033218971574480866
Epoch 28/50, Loss: 0.031733723886390314
Epoch 29/50, Loss: 0.030238033841027225
Epoch 30/50, Loss: 0.030516310673246675
Epoch 31/50, Loss: 0.03012844078995763
Epoch 32/50, Loss: 0.02981952812146158
Epoch 33/50, Loss: 0.028553571197343
Epoch 34/50, Loss: 0.028173160120012596
Epoch 35/50, Loss: 0.028569667072015557
Epoch 36/50, Loss: 0.027112166378448935
Epoch 37/50, Loss: 0.02698830751147359
Epoch 38/50, Loss: 0.02668825531418019
Epoch 39/50, Loss: 0.026028482954790617
Epoch 40/50, Loss: 0.026841685281907942
Epoch 41/50, Loss: 0.023254996979054303
Epoch 42/50, Loss: 0.025006243599420216
Epoch 43/50, Loss: 0.024620083554528972
Epoch 44/50, Loss: 0.022385686077177523
Epoch 45/50, Loss: 0.02413795780132901
Epoch 46/50, Loss: 0.02477022296293302

Epoch 47/50, Loss: 0.02239133589306569
Epoch 48/50, Loss: 0.021815316582177866
Epoch 49/50, Loss: 0.022794892105888177
Epoch 50/50, Loss: 0.02206510845076372

Accuracy on test set: 99.06%



Batch Normalisation and dropout gives the best accuracy of 99.06%

Conclusion: Overall we conclude that a CNN-ResNet using Batch normalisation and dropouts model with normalised dataset , optimizer as Adam has the best accuracy.