

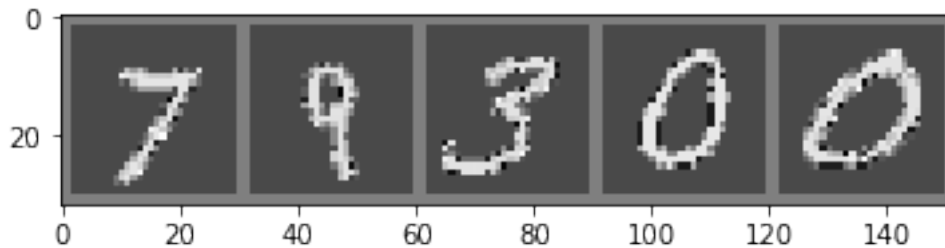
Report Assignment 2

Srinidhi Goud

Warmup

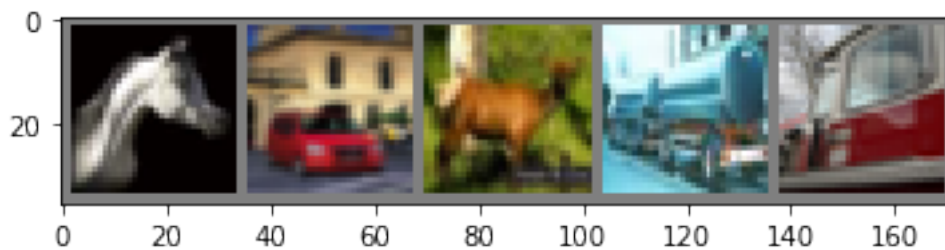
Result:

Displaying random 5 images for MNIST:



seven nine three zero zero

for CIFAR10:



horse car deer truck truck

Training a Single Layer Network on MNIST

Train Epoch: 1	[0/60000 (0%)]	Loss: 2.577542
Train Epoch: 1	[6400/60000 (11%)]	Loss: 0.544359
Train Epoch: 1	[12800/60000 (21%)]	Loss: 0.384140
Train Epoch: 1	[19200/60000 (32%)]	Loss: 0.408010
Train Epoch: 1	[25600/60000 (43%)]	Loss: 0.373087
Train Epoch: 1	[32000/60000 (53%)]	Loss: 0.392287
Train Epoch: 1	[38400/60000 (64%)]	Loss: 0.288197
Train Epoch: 1	[44800/60000 (75%)]	Loss: 0.288111
Train Epoch: 1	[51200/60000 (85%)]	Loss: 0.343845
Train Epoch: 1	[57600/60000 (96%)]	Loss: 0.274763

Test set: Average loss: 0.3300, Accuracy: 9067/10000 (91%)

Train Epoch: 2	[0/60000 (0%)]	Loss: 0.519524
Train Epoch: 2	[6400/60000 (11%)]	Loss: 0.337938
Train Epoch: 2	[12800/60000 (21%)]	Loss: 0.564827

Train Epoch: 2 [19200/60000 (32%)]	Loss: 0.486577
Train Epoch: 2 [25600/60000 (43%)]	Loss: 0.195081
Train Epoch: 2 [32000/60000 (53%)]	Loss: 0.215729
Train Epoch: 2 [38400/60000 (64%)]	Loss: 0.347026
Train Epoch: 2 [44800/60000 (75%)]	Loss: 0.244313
Train Epoch: 2 [51200/60000 (85%)]	Loss: 0.506253
Train Epoch: 2 [57600/60000 (96%)]	Loss: 0.349717

Test set: Average loss: 0.3032, Accuracy: 9157/10000 (92%)

Train Epoch: 3 [0/60000 (0%)]	Loss: 0.364705
Train Epoch: 3 [6400/60000 (11%)]	Loss: 0.272041
Train Epoch: 3 [12800/60000 (21%)]	Loss: 0.273759
Train Epoch: 3 [19200/60000 (32%)]	Loss: 0.230016
Train Epoch: 3 [25600/60000 (43%)]	Loss: 0.385725
Train Epoch: 3 [32000/60000 (53%)]	Loss: 0.435140
Train Epoch: 3 [38400/60000 (64%)]	Loss: 0.183245
Train Epoch: 3 [44800/60000 (75%)]	Loss: 0.258105
Train Epoch: 3 [51200/60000 (85%)]	Loss: 0.219853
Train Epoch: 3 [57600/60000 (96%)]	Loss: 0.323031

Test set: Average loss: 0.2973, Accuracy: 9161/10000 (92%)

Train Epoch: 4 [0/60000 (0%)]	Loss: 0.177594
Train Epoch: 4 [6400/60000 (11%)]	Loss: 0.331060
Train Epoch: 4 [12800/60000 (21%)]	Loss: 0.314664
Train Epoch: 4 [19200/60000 (32%)]	Loss: 0.315693
Train Epoch: 4 [25600/60000 (43%)]	Loss: 0.180037
Train Epoch: 4 [32000/60000 (53%)]	Loss: 0.217031
Train Epoch: 4 [38400/60000 (64%)]	Loss: 0.256635
Train Epoch: 4 [44800/60000 (75%)]	Loss: 0.297715
Train Epoch: 4 [51200/60000 (85%)]	Loss: 0.096892
Train Epoch: 4 [57600/60000 (96%)]	Loss: 0.208867

Test set: Average loss: 0.2884, Accuracy: 9188/10000 (92%)

Train Epoch: 5 [0/60000 (0%)]	Loss: 0.288036
Train Epoch: 5 [6400/60000 (11%)]	Loss: 0.139274
Train Epoch: 5 [12800/60000 (21%)]	Loss: 0.344348
Train Epoch: 5 [19200/60000 (32%)]	Loss: 0.471398
Train Epoch: 5 [25600/60000 (43%)]	Loss: 0.265125
Train Epoch: 5 [32000/60000 (53%)]	Loss: 0.234014
Train Epoch: 5 [38400/60000 (64%)]	Loss: 0.429277
Train Epoch: 5 [44800/60000 (75%)]	Loss: 0.363730
Train Epoch: 5 [51200/60000 (85%)]	Loss: 0.203440
Train Epoch: 5 [57600/60000 (96%)]	Loss: 0.380245

Test set: Average loss: 0.2816, Accuracy: 9200/10000 (92%)

Train Epoch: 6 [0/60000 (0%)]	Loss: 0.217563
Train Epoch: 6 [6400/60000 (11%)]	Loss: 0.254859

Train Epoch: 6 [12800/60000 (21%)]	Loss: 0.383338
Train Epoch: 6 [19200/60000 (32%)]	Loss: 0.346315
Train Epoch: 6 [25600/60000 (43%)]	Loss: 0.258452
Train Epoch: 6 [32000/60000 (53%)]	Loss: 0.253936
Train Epoch: 6 [38400/60000 (64%)]	Loss: 0.086758
Train Epoch: 6 [44800/60000 (75%)]	Loss: 0.217135
Train Epoch: 6 [51200/60000 (85%)]	Loss: 0.242411
Train Epoch: 6 [57600/60000 (96%)]	Loss: 0.442520

Test set: Average loss: 0.2816, Accuracy: 9199/10000 (92%)

Train Epoch: 7 [0/60000 (0%)]	Loss: 0.090015
Train Epoch: 7 [6400/60000 (11%)]	Loss: 0.332615
Train Epoch: 7 [12800/60000 (21%)]	Loss: 0.472020
Train Epoch: 7 [19200/60000 (32%)]	Loss: 0.137301
Train Epoch: 7 [25600/60000 (43%)]	Loss: 0.384776
Train Epoch: 7 [32000/60000 (53%)]	Loss: 0.333673
Train Epoch: 7 [38400/60000 (64%)]	Loss: 0.187276
Train Epoch: 7 [44800/60000 (75%)]	Loss: 0.226567
Train Epoch: 7 [51200/60000 (85%)]	Loss: 0.175371
Train Epoch: 7 [57600/60000 (96%)]	Loss: 0.259725

Test set: Average loss: 0.2783, Accuracy: 9191/10000 (92%)

Train Epoch: 8 [0/60000 (0%)]	Loss: 0.383326
Train Epoch: 8 [6400/60000 (11%)]	Loss: 0.222461
Train Epoch: 8 [12800/60000 (21%)]	Loss: 0.508418
Train Epoch: 8 [19200/60000 (32%)]	Loss: 0.210508
Train Epoch: 8 [25600/60000 (43%)]	Loss: 0.189479
Train Epoch: 8 [32000/60000 (53%)]	Loss: 0.195588
Train Epoch: 8 [38400/60000 (64%)]	Loss: 0.207123
Train Epoch: 8 [44800/60000 (75%)]	Loss: 0.260126
Train Epoch: 8 [51200/60000 (85%)]	Loss: 0.266143
Train Epoch: 8 [57600/60000 (96%)]	Loss: 0.462109

Test set: Average loss: 0.2783, Accuracy: 9203/10000 (92%)

Train Epoch: 9 [0/60000 (0%)]	Loss: 0.252528
Train Epoch: 9 [6400/60000 (11%)]	Loss: 0.215634
Train Epoch: 9 [12800/60000 (21%)]	Loss: 0.179456
Train Epoch: 9 [19200/60000 (32%)]	Loss: 0.304013
Train Epoch: 9 [25600/60000 (43%)]	Loss: 0.489238
Train Epoch: 9 [32000/60000 (53%)]	Loss: 0.451165
Train Epoch: 9 [38400/60000 (64%)]	Loss: 0.336278
Train Epoch: 9 [44800/60000 (75%)]	Loss: 0.188450
Train Epoch: 9 [51200/60000 (85%)]	Loss: 0.324191
Train Epoch: 9 [57600/60000 (96%)]	Loss: 0.299928

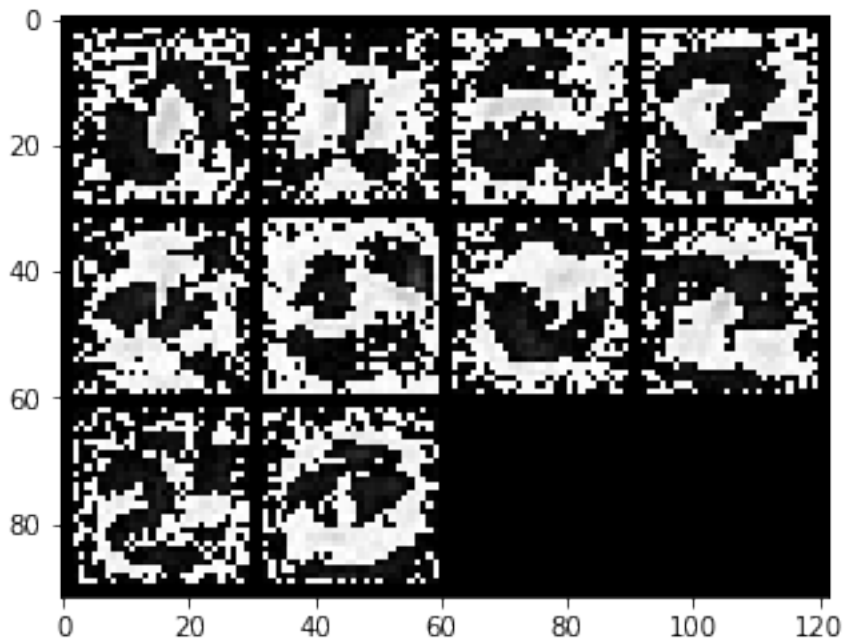
Test set: Average loss: 0.2739, Accuracy: 9219/10000 (92%)

Train Epoch: 10 [0/60000 (0%)] Loss: 0.215896

Train Epoch: 10 [6400/60000 (11%)]	Loss: 0.365033
Train Epoch: 10 [12800/60000 (21%)]	Loss: 0.270614
Train Epoch: 10 [19200/60000 (32%)]	Loss: 0.247973
Train Epoch: 10 [25600/60000 (43%)]	Loss: 0.335032
Train Epoch: 10 [32000/60000 (53%)]	Loss: 0.286341
Train Epoch: 10 [38400/60000 (64%)]	Loss: 0.165615
Train Epoch: 10 [44800/60000 (75%)]	Loss: 0.118591
Train Epoch: 10 [51200/60000 (85%)]	Loss: 0.392136
Train Epoch: 10 [57600/60000 (96%)]	Loss: 0.214605

Test set: Average loss: 0.2737, Accuracy: 9224/10000 (92%)

a)Weights Image



b)

Train Epoch: 1 [0/60000 (0%)] Loss: 2.229349

Test set: Average loss: 0.7619, Accuracy: 8217/10000 (82%)

Train Epoch: 2 [0/60000 (0%)] Loss: 0.882909

Test set: Average loss: 0.5773, Accuracy: 8625/10000 (86%)

Train Epoch: 3 [0/60000 (0%)] Loss: 0.729428

Test set: Average loss: 0.4999, Accuracy: 8741/10000 (87%)

Train Epoch: 4 [0/60000 (0%)] Loss: 0.313826

Test set: Average loss: 0.4642, Accuracy: 8761/10000 (88%)

Train Epoch: 5 [0/60000 (0%)] Loss: 0.800690

Test set: Average loss: 0.4295, Accuracy: 8847/10000 (88%)

Train Epoch: 6 [0/60000 (0%)] Loss: 0.481635

Test set: Average loss: 0.4123, Accuracy: 8900/10000 (89%)

Train Epoch: 7 [0/60000 (0%)] Loss: 0.257481

Test set: Average loss: 0.3990, Accuracy: 8927/10000 (89%)

Train Epoch: 8 [0/60000 (0%)] Loss: 0.349227

Test set: Average loss: 0.3862, Accuracy: 8940/10000 (89%)

Train Epoch: 9 [0/60000 (0%)] Loss: 0.410699

Test set: Average loss: 0.3764, Accuracy: 8972/10000 (90%)

Train Epoch: 10 [0/60000 (0%)] Loss: 0.276375

Test set: Average loss: 0.3689, Accuracy: 8985/10000 (90%)

With only 50 training samples the, there is not constant reduction in the Loss function, it is not being minimized. However, it performed decently, avg Loss just fluctuating by a small value compared to the Loss value we got for all the training samples.

Training a Multi-Layer Network on MNIST

a)

Train Epoch: 1 [0/60000 (0%)] Loss: 2.359563

Train Epoch: 1 [6400/60000 (11%)] Loss: 0.746561

Train Epoch: 1 [12800/60000 (21%)] Loss: 0.584208

Train Epoch: 1 [19200/60000 (32%)] Loss: 0.525674

Train Epoch: 1 [25600/60000 (43%)] Loss: 0.373070

Train Epoch: 1 [32000/60000 (53%)] Loss: 0.415313

Train Epoch: 1 [38400/60000 (64%)] Loss: 0.319662

Train Epoch: 1 [44800/60000 (75%)] Loss: 0.497748

Train Epoch: 1 [51200/60000 (85%)] Loss: 0.429553

Train Epoch: 1 [57600/60000 (96%)] Loss: 0.351936

Test set: Average loss: 0.3237, Accuracy: 9111/10000 (91%)

Train Epoch: 2 [0/60000 (0%)] Loss: 0.318940

Train Epoch: 2 [6400/60000 (11%)] Loss: 0.440443

Train Epoch: 2 [12800/60000 (21%)] Loss: 0.281750

Train Epoch: 2 [19200/60000 (32%)]	Loss: 0.275298
Train Epoch: 2 [25600/60000 (43%)]	Loss: 0.364756
Train Epoch: 2 [32000/60000 (53%)]	Loss: 0.285660
Train Epoch: 2 [38400/60000 (64%)]	Loss: 0.224869
Train Epoch: 2 [44800/60000 (75%)]	Loss: 0.326156
Train Epoch: 2 [51200/60000 (85%)]	Loss: 0.198576
Train Epoch: 2 [57600/60000 (96%)]	Loss: 0.507915

Test set: Average loss: 0.2848, Accuracy: 9190/10000 (92%)

Train Epoch: 3 [0/60000 (0%)]	Loss: 0.316185
Train Epoch: 3 [6400/60000 (11%)]	Loss: 0.332845
Train Epoch: 3 [12800/60000 (21%)]	Loss: 0.355939
Train Epoch: 3 [19200/60000 (32%)]	Loss: 0.356583
Train Epoch: 3 [25600/60000 (43%)]	Loss: 0.183009
Train Epoch: 3 [32000/60000 (53%)]	Loss: 0.320532
Train Epoch: 3 [38400/60000 (64%)]	Loss: 0.360128
Train Epoch: 3 [44800/60000 (75%)]	Loss: 0.210640
Train Epoch: 3 [51200/60000 (85%)]	Loss: 0.291787
Train Epoch: 3 [57600/60000 (96%)]	Loss: 0.285156

Test set: Average loss: 0.2615, Accuracy: 9250/10000 (92%)

Train Epoch: 4 [0/60000 (0%)]	Loss: 0.157702
Train Epoch: 4 [6400/60000 (11%)]	Loss: 0.382375
Train Epoch: 4 [12800/60000 (21%)]	Loss: 0.349962
Train Epoch: 4 [19200/60000 (32%)]	Loss: 0.177547
Train Epoch: 4 [25600/60000 (43%)]	Loss: 0.237249
Train Epoch: 4 [32000/60000 (53%)]	Loss: 0.163370
Train Epoch: 4 [38400/60000 (64%)]	Loss: 0.190574
Train Epoch: 4 [44800/60000 (75%)]	Loss: 0.126067
Train Epoch: 4 [51200/60000 (85%)]	Loss: 0.086881
Train Epoch: 4 [57600/60000 (96%)]	Loss: 0.171461

Test set: Average loss: 0.2438, Accuracy: 9304/10000 (93%)

Train Epoch: 5 [0/60000 (0%)]	Loss: 0.465045
Train Epoch: 5 [6400/60000 (11%)]	Loss: 0.223884
Train Epoch: 5 [12800/60000 (21%)]	Loss: 0.472145
Train Epoch: 5 [19200/60000 (32%)]	Loss: 0.211174
Train Epoch: 5 [25600/60000 (43%)]	Loss: 0.252713
Train Epoch: 5 [32000/60000 (53%)]	Loss: 0.331133
Train Epoch: 5 [38400/60000 (64%)]	Loss: 0.142336
Train Epoch: 5 [44800/60000 (75%)]	Loss: 0.198892
Train Epoch: 5 [51200/60000 (85%)]	Loss: 0.155969
Train Epoch: 5 [57600/60000 (96%)]	Loss: 0.200978

Test set: Average loss: 0.2262, Accuracy: 9355/10000 (94%)

Train Epoch: 6 [0/60000 (0%)]	Loss: 0.161606
Train Epoch: 6 [6400/60000 (11%)]	Loss: 0.067543

Train Epoch: 6 [12800/60000 (21%)]	Loss: 0.379206
Train Epoch: 6 [19200/60000 (32%)]	Loss: 0.183446
Train Epoch: 6 [25600/60000 (43%)]	Loss: 0.166491
Train Epoch: 6 [32000/60000 (53%)]	Loss: 0.182133
Train Epoch: 6 [38400/60000 (64%)]	Loss: 0.212980
Train Epoch: 6 [44800/60000 (75%)]	Loss: 0.237922
Train Epoch: 6 [51200/60000 (85%)]	Loss: 0.140050
Train Epoch: 6 [57600/60000 (96%)]	Loss: 0.203873

Test set: Average loss: 0.2110, Accuracy: 9422/10000 (94%)

Train Epoch: 7 [0/60000 (0%)]	Loss: 0.320934
Train Epoch: 7 [6400/60000 (11%)]	Loss: 0.166072
Train Epoch: 7 [12800/60000 (21%)]	Loss: 0.199063
Train Epoch: 7 [19200/60000 (32%)]	Loss: 0.143679
Train Epoch: 7 [25600/60000 (43%)]	Loss: 0.253772
Train Epoch: 7 [32000/60000 (53%)]	Loss: 0.174687
Train Epoch: 7 [38400/60000 (64%)]	Loss: 0.138602
Train Epoch: 7 [44800/60000 (75%)]	Loss: 0.102502
Train Epoch: 7 [51200/60000 (85%)]	Loss: 0.139515
Train Epoch: 7 [57600/60000 (96%)]	Loss: 0.212678

Test set: Average loss: 0.1956, Accuracy: 9460/10000 (95%)

Train Epoch: 8 [0/60000 (0%)]	Loss: 0.142527
Train Epoch: 8 [6400/60000 (11%)]	Loss: 0.097556
Train Epoch: 8 [12800/60000 (21%)]	Loss: 0.229534
Train Epoch: 8 [19200/60000 (32%)]	Loss: 0.098080
Train Epoch: 8 [25600/60000 (43%)]	Loss: 0.150019
Train Epoch: 8 [32000/60000 (53%)]	Loss: 0.182051
Train Epoch: 8 [38400/60000 (64%)]	Loss: 0.148059
Train Epoch: 8 [44800/60000 (75%)]	Loss: 0.133121
Train Epoch: 8 [51200/60000 (85%)]	Loss: 0.250748
Train Epoch: 8 [57600/60000 (96%)]	Loss: 0.155765

Test set: Average loss: 0.1831, Accuracy: 9487/10000 (95%)

Train Epoch: 9 [0/60000 (0%)]	Loss: 0.117860
Train Epoch: 9 [6400/60000 (11%)]	Loss: 0.260050
Train Epoch: 9 [12800/60000 (21%)]	Loss: 0.511214
Train Epoch: 9 [19200/60000 (32%)]	Loss: 0.203845
Train Epoch: 9 [25600/60000 (43%)]	Loss: 0.241523
Train Epoch: 9 [32000/60000 (53%)]	Loss: 0.133197
Train Epoch: 9 [38400/60000 (64%)]	Loss: 0.144057
Train Epoch: 9 [44800/60000 (75%)]	Loss: 0.123096
Train Epoch: 9 [51200/60000 (85%)]	Loss: 0.208286
Train Epoch: 9 [57600/60000 (96%)]	Loss: 0.263991

Test set: Average loss: 0.1711, Accuracy: 9513/10000 (95%)

Train Epoch: 10 [0/60000 (0%)]	Loss: 0.178349
--------------------------------	----------------

Train Epoch: 10 [6400/60000 (11%)]	Loss: 0.180430
Train Epoch: 10 [12800/60000 (21%)]	Loss: 0.133034
Train Epoch: 10 [19200/60000 (32%)]	Loss: 0.201327
Train Epoch: 10 [25600/60000 (43%)]	Loss: 0.168630
Train Epoch: 10 [32000/60000 (53%)]	Loss: 0.250597
Train Epoch: 10 [38400/60000 (64%)]	Loss: 0.114506
Train Epoch: 10 [44800/60000 (75%)]	Loss: 0.353448
Train Epoch: 10 [51200/60000 (85%)]	Loss: 0.137694
Train Epoch: 10 [57600/60000 (96%)]	Loss: 0.200292

Test set: Average loss: 0.1607, Accuracy: 9544/10000 (95%)

b)

Train Epoch: 1 [0/60000 (0%)]	Loss: 2.341428
Train Epoch: 1 [6400/60000 (11%)]	Loss: 643.909790
Train Epoch: 1 [12800/60000 (21%)]	Loss: 1229.932739
Train Epoch: 1 [19200/60000 (32%)]	Loss: 624.699219
Train Epoch: 1 [25600/60000 (43%)]	Loss: 405.776672
Train Epoch: 1 [32000/60000 (53%)]	Loss: 1078.665894
Train Epoch: 1 [38400/60000 (64%)]	Loss: 449.817017
Train Epoch: 1 [44800/60000 (75%)]	Loss: 537.274780
Train Epoch: 1 [51200/60000 (85%)]	Loss: 355.935974
Train Epoch: 1 [57600/60000 (96%)]	Loss: 487.709503

Test set: Average loss: 367.4333, Accuracy: 6858/10000 (69%)

Train Epoch: 2 [0/60000 (0%)]	Loss: 320.170135
Train Epoch: 2 [6400/60000 (11%)]	Loss: 541.442078
Train Epoch: 2 [12800/60000 (21%)]	Loss: 266.904297
Train Epoch: 2 [19200/60000 (32%)]	Loss: 385.677338
Train Epoch: 2 [25600/60000 (43%)]	Loss: 966.936096
Train Epoch: 2 [32000/60000 (53%)]	Loss: 328.864777
Train Epoch: 2 [38400/60000 (64%)]	Loss: 119.439735
Train Epoch: 2 [44800/60000 (75%)]	Loss: 262.330048
Train Epoch: 2 [51200/60000 (85%)]	Loss: 257.858337
Train Epoch: 2 [57600/60000 (96%)]	Loss: 294.054596

Test set: Average loss: 184.1136, Accuracy: 6468/10000 (65%)

Train Epoch: 3 [0/60000 (0%)]	Loss: 221.696945
Train Epoch: 3 [6400/60000 (11%)]	Loss: 273.126404
Train Epoch: 3 [12800/60000 (21%)]	Loss: 426.852997
Train Epoch: 3 [19200/60000 (32%)]	Loss: 112.051590
Train Epoch: 3 [25600/60000 (43%)]	Loss: 147.900146
Train Epoch: 3 [32000/60000 (53%)]	Loss: 54.386509
Train Epoch: 3 [38400/60000 (64%)]	Loss: 321.812103
Train Epoch: 3 [44800/60000 (75%)]	Loss: 62.969486
Train Epoch: 3 [51200/60000 (85%)]	Loss: 127.888474
Train Epoch: 3 [57600/60000 (96%)]	Loss: 308.651154

Test set: Average loss: 181.1004, Accuracy: 6823/10000 (68%)

Train Epoch: 4 [0/60000 (0%)]	Loss: 105.925713
Train Epoch: 4 [6400/60000 (11%)]	Loss: 116.411247
Train Epoch: 4 [12800/60000 (21%)]	Loss: 180.314819
Train Epoch: 4 [19200/60000 (32%)]	Loss: 495.344299
Train Epoch: 4 [25600/60000 (43%)]	Loss: 167.326675
Train Epoch: 4 [32000/60000 (53%)]	Loss: 295.745361
Train Epoch: 4 [38400/60000 (64%)]	Loss: 239.620850
Train Epoch: 4 [44800/60000 (75%)]	Loss: 425.028931
Train Epoch: 4 [51200/60000 (85%)]	Loss: 123.585579
Train Epoch: 4 [57600/60000 (96%)]	Loss: 386.244019

Test set: Average loss: 281.2910, Accuracy: 5964/10000 (60%)

Train Epoch: 5 [0/60000 (0%)]	Loss: 221.374542
Train Epoch: 5 [6400/60000 (11%)]	Loss: 138.870758
Train Epoch: 5 [12800/60000 (21%)]	Loss: 428.118622
Train Epoch: 5 [19200/60000 (32%)]	Loss: 192.079712
Train Epoch: 5 [25600/60000 (43%)]	Loss: 160.129395
Train Epoch: 5 [32000/60000 (53%)]	Loss: 313.292480
Train Epoch: 5 [38400/60000 (64%)]	Loss: 120.291512
Train Epoch: 5 [44800/60000 (75%)]	Loss: 145.360733
Train Epoch: 5 [51200/60000 (85%)]	Loss: 267.305603
Train Epoch: 5 [57600/60000 (96%)]	Loss: 200.828110

Test set: Average loss: 263.3243, Accuracy: 5649/10000 (56%)

Train Epoch: 6 [0/60000 (0%)]	Loss: 223.111343
Train Epoch: 6 [6400/60000 (11%)]	Loss: 190.686691
Train Epoch: 6 [12800/60000 (21%)]	Loss: 152.523026
Train Epoch: 6 [19200/60000 (32%)]	Loss: 241.166916
Train Epoch: 6 [25600/60000 (43%)]	Loss: 334.007965
Train Epoch: 6 [32000/60000 (53%)]	Loss: 83.850349
Train Epoch: 6 [38400/60000 (64%)]	Loss: 94.571564
Train Epoch: 6 [44800/60000 (75%)]	Loss: 125.385765
Train Epoch: 6 [51200/60000 (85%)]	Loss: 197.824158
Train Epoch: 6 [57600/60000 (96%)]	Loss: 356.980133

Test set: Average loss: 151.1076, Accuracy: 7312/10000 (73%)

Train Epoch: 7 [0/60000 (0%)]	Loss: 160.406494
Train Epoch: 7 [6400/60000 (11%)]	Loss: 473.676697
Train Epoch: 7 [12800/60000 (21%)]	Loss: 133.810333
Train Epoch: 7 [19200/60000 (32%)]	Loss: 121.952042
Train Epoch: 7 [25600/60000 (43%)]	Loss: 353.745544
Train Epoch: 7 [32000/60000 (53%)]	Loss: 112.768364
Train Epoch: 7 [38400/60000 (64%)]	Loss: 280.423737
Train Epoch: 7 [44800/60000 (75%)]	Loss: 357.795380
Train Epoch: 7 [51200/60000 (85%)]	Loss: 313.408051
Train Epoch: 7 [57600/60000 (96%)]	Loss: 156.856155

Test set: Average loss: 126.1487, Accuracy: 7679/10000 (77%)

Train Epoch: 8 [0/60000 (0%)]	Loss: 150.887405
Train Epoch: 8 [6400/60000 (11%)]	Loss: 134.070374
Train Epoch: 8 [12800/60000 (21%)]	Loss: 383.127716
Train Epoch: 8 [19200/60000 (32%)]	Loss: 146.818542
Train Epoch: 8 [25600/60000 (43%)]	Loss: 52.439949
Train Epoch: 8 [32000/60000 (53%)]	Loss: 63.128204
Train Epoch: 8 [38400/60000 (64%)]	Loss: 154.694946
Train Epoch: 8 [44800/60000 (75%)]	Loss: 114.792641
Train Epoch: 8 [51200/60000 (85%)]	Loss: 207.356979
Train Epoch: 8 [57600/60000 (96%)]	Loss: 78.472572

Test set: Average loss: 125.9353, Accuracy: 7261/10000 (73%)

Train Epoch: 9 [0/60000 (0%)]	Loss: 164.281876
Train Epoch: 9 [6400/60000 (11%)]	Loss: 317.200012
Train Epoch: 9 [12800/60000 (21%)]	Loss: 161.288177
Train Epoch: 9 [19200/60000 (32%)]	Loss: 147.372543
Train Epoch: 9 [25600/60000 (43%)]	Loss: 78.335640
Train Epoch: 9 [32000/60000 (53%)]	Loss: 158.172577
Train Epoch: 9 [38400/60000 (64%)]	Loss: 293.732941
Train Epoch: 9 [44800/60000 (75%)]	Loss: 198.052551
Train Epoch: 9 [51200/60000 (85%)]	Loss: 165.585175
Train Epoch: 9 [57600/60000 (96%)]	Loss: 123.850441

Test set: Average loss: 177.1925, Accuracy: 6431/10000 (64%)

Train Epoch: 10 [0/60000 (0%)]	Loss: 177.098984
Train Epoch: 10 [6400/60000 (11%)]	Loss: 205.369400
Train Epoch: 10 [12800/60000 (21%)]	Loss: 143.926056
Train Epoch: 10 [19200/60000 (32%)]	Loss: 211.533707
Train Epoch: 10 [25600/60000 (43%)]	Loss: 93.161285
Train Epoch: 10 [32000/60000 (53%)]	Loss: 65.667114
Train Epoch: 10 [38400/60000 (64%)]	Loss: 158.963715
Train Epoch: 10 [44800/60000 (75%)]	Loss: 107.197067
Train Epoch: 10 [51200/60000 (85%)]	Loss: 51.244072
Train Epoch: 10 [57600/60000 (96%)]	Loss: 264.380341

Test set: Average loss: 414.7021, Accuracy: 5790/10000 (58%)

As gradient descent is not slow, it jumps over the minima, and never actually getting minimized.
The jumps are very large.

Training a Convolutional Network on CIFAR

Train Epoch: 1 [0/50000 (0%)]	Loss: 2.322038
-------------------------------	----------------

Train Epoch: 1 [6400/50000 (13%)]	Loss: 2.183053
Train Epoch: 1 [12800/50000 (26%)]	Loss: 2.148056
Train Epoch: 1 [19200/50000 (38%)]	Loss: 2.011059
Train Epoch: 1 [25600/50000 (51%)]	Loss: 1.844144
Train Epoch: 1 [32000/50000 (64%)]	Loss: 1.968429
Train Epoch: 1 [38400/50000 (77%)]	Loss: 1.835537
Train Epoch: 1 [44800/50000 (90%)]	Loss: 1.853805

Test set: Average loss: 1.8378, Accuracy: 17830/50000 (36%)

Train Epoch: 2 [0/50000 (0%)]	Loss: 1.792717
Train Epoch: 2 [6400/50000 (13%)]	Loss: 1.825227
Train Epoch: 2 [12800/50000 (26%)]	Loss: 1.699355
Train Epoch: 2 [19200/50000 (38%)]	Loss: 1.668300
Train Epoch: 2 [25600/50000 (51%)]	Loss: 1.876743
Train Epoch: 2 [32000/50000 (64%)]	Loss: 1.678932
Train Epoch: 2 [38400/50000 (77%)]	Loss: 1.718377
Train Epoch: 2 [44800/50000 (90%)]	Loss: 1.808040

Test set: Average loss: 1.6248, Accuracy: 21307/50000 (43%)

Train Epoch: 3 [0/50000 (0%)]	Loss: 1.637457
Train Epoch: 3 [6400/50000 (13%)]	Loss: 1.695967
Train Epoch: 3 [12800/50000 (26%)]	Loss: 1.433261
Train Epoch: 3 [19200/50000 (38%)]	Loss: 1.553680
Train Epoch: 3 [25600/50000 (51%)]	Loss: 1.642020
Train Epoch: 3 [32000/50000 (64%)]	Loss: 1.531573
Train Epoch: 3 [38400/50000 (77%)]	Loss: 1.457323
Train Epoch: 3 [44800/50000 (90%)]	Loss: 1.734312

Test set: Average loss: 1.4576, Accuracy: 24135/50000 (48%)

Train Epoch: 4 [0/50000 (0%)]	Loss: 1.498209
Train Epoch: 4 [6400/50000 (13%)]	Loss: 1.470921
Train Epoch: 4 [12800/50000 (26%)]	Loss: 1.384445
Train Epoch: 4 [19200/50000 (38%)]	Loss: 1.443344
Train Epoch: 4 [25600/50000 (51%)]	Loss: 1.451047
Train Epoch: 4 [32000/50000 (64%)]	Loss: 1.417928
Train Epoch: 4 [38400/50000 (77%)]	Loss: 1.306242
Train Epoch: 4 [44800/50000 (90%)]	Loss: 1.245950

Test set: Average loss: 1.4263, Accuracy: 24341/50000 (49%)

Train Epoch: 5 [0/50000 (0%)]	Loss: 1.413039
-------------------------------	----------------

Train Epoch: 5 [6400/50000 (13%)]	Loss: 1.504374
Train Epoch: 5 [12800/50000 (26%)]	Loss: 1.697311
Train Epoch: 5 [19200/50000 (38%)]	Loss: 1.134848
Train Epoch: 5 [25600/50000 (51%)]	Loss: 1.264056
Train Epoch: 5 [32000/50000 (64%)]	Loss: 1.277785
Train Epoch: 5 [38400/50000 (77%)]	Loss: 1.259776
Train Epoch: 5 [44800/50000 (90%)]	Loss: 1.305818

Test set: Average loss: 1.3106, Accuracy: 26470/50000 (53%)

Train Epoch: 6 [0/50000 (0%)]	Loss: 1.276134
Train Epoch: 6 [6400/50000 (13%)]	Loss: 1.337098
Train Epoch: 6 [12800/50000 (26%)]	Loss: 1.319898
Train Epoch: 6 [19200/50000 (38%)]	Loss: 1.190057
Train Epoch: 6 [25600/50000 (51%)]	Loss: 1.303985
Train Epoch: 6 [32000/50000 (64%)]	Loss: 1.157348
Train Epoch: 6 [38400/50000 (77%)]	Loss: 1.384315
Train Epoch: 6 [44800/50000 (90%)]	Loss: 1.276033

Test set: Average loss: 1.2984, Accuracy: 27179/50000 (54%)

Train Epoch: 7 [0/50000 (0%)]	Loss: 1.402281
Train Epoch: 7 [6400/50000 (13%)]	Loss: 1.339730
Train Epoch: 7 [12800/50000 (26%)]	Loss: 1.242891
Train Epoch: 7 [19200/50000 (38%)]	Loss: 1.535021
Train Epoch: 7 [25600/50000 (51%)]	Loss: 1.136907
Train Epoch: 7 [32000/50000 (64%)]	Loss: 1.036229
Train Epoch: 7 [38400/50000 (77%)]	Loss: 1.078800
Train Epoch: 7 [44800/50000 (90%)]	Loss: 1.206226

Test set: Average loss: 1.1892, Accuracy: 28885/50000 (58%)

Train Epoch: 8 [0/50000 (0%)]	Loss: 1.323325
Train Epoch: 8 [6400/50000 (13%)]	Loss: 1.310774
Train Epoch: 8 [12800/50000 (26%)]	Loss: 1.167133
Train Epoch: 8 [19200/50000 (38%)]	Loss: 1.390817
Train Epoch: 8 [25600/50000 (51%)]	Loss: 1.139728
Train Epoch: 8 [32000/50000 (64%)]	Loss: 1.013611
Train Epoch: 8 [38400/50000 (77%)]	Loss: 1.040779
Train Epoch: 8 [44800/50000 (90%)]	Loss: 1.114083

Test set: Average loss: 1.1902, Accuracy: 28835/50000 (58%)

Train Epoch: 9 [0/50000 (0%)]	Loss: 1.010815
-------------------------------	----------------

Train Epoch: 9 [6400/50000 (13%)]	Loss: 1.236927
Train Epoch: 9 [12800/50000 (26%)]	Loss: 1.088752
Train Epoch: 9 [19200/50000 (38%)]	Loss: 1.177638
Train Epoch: 9 [25600/50000 (51%)]	Loss: 0.860759
Train Epoch: 9 [32000/50000 (64%)]	Loss: 1.035243
Train Epoch: 9 [38400/50000 (77%)]	Loss: 1.251297
Train Epoch: 9 [44800/50000 (90%)]	Loss: 1.197291

Test set: Average loss: 1.1182, Accuracy: 30366/50000 (61%)

Train Epoch: 10 [0/50000 (0%)]	Loss: 1.203449
Train Epoch: 10 [6400/50000 (13%)]	Loss: 1.150770
Train Epoch: 10 [12800/50000 (26%)]	Loss: 1.051200
Train Epoch: 10 [19200/50000 (38%)]	Loss: 1.140940
Train Epoch: 10 [25600/50000 (51%)]	Loss: 1.369101
Train Epoch: 10 [32000/50000 (64%)]	Loss: 1.110441
Train Epoch: 10 [38400/50000 (77%)]	Loss: 1.201700
Train Epoch: 10 [44800/50000 (90%)]	Loss: 1.135780

Test set: Average loss: 1.0914, Accuracy: 30820/50000 (62%)

Train Epoch: 11 [0/50000 (0%)]	Loss: 1.155916
Train Epoch: 11 [6400/50000 (13%)]	Loss: 1.045684
Train Epoch: 11 [12800/50000 (26%)]	Loss: 1.135876
Train Epoch: 11 [19200/50000 (38%)]	Loss: 0.862280
Train Epoch: 11 [25600/50000 (51%)]	Loss: 1.007160
Train Epoch: 11 [32000/50000 (64%)]	Loss: 1.224550
Train Epoch: 11 [38400/50000 (77%)]	Loss: 0.906788
Train Epoch: 11 [44800/50000 (90%)]	Loss: 1.103378

Test set: Average loss: 1.0135, Accuracy: 32492/50000 (65%)

Train Epoch: 12 [0/50000 (0%)]	Loss: 1.229802
Train Epoch: 12 [6400/50000 (13%)]	Loss: 1.198027
Train Epoch: 12 [12800/50000 (26%)]	Loss: 1.090834
Train Epoch: 12 [19200/50000 (38%)]	Loss: 1.023619
Train Epoch: 12 [25600/50000 (51%)]	Loss: 1.106481
Train Epoch: 12 [32000/50000 (64%)]	Loss: 1.025354
Train Epoch: 12 [38400/50000 (77%)]	Loss: 0.967274
Train Epoch: 12 [44800/50000 (90%)]	Loss: 0.922827

Test set: Average loss: 1.0207, Accuracy: 32003/50000 (64%)

Train Epoch: 13 [0/50000 (0%)]	Loss: 1.117485
--------------------------------	----------------

Train Epoch: 13 [6400/50000 (13%)]	Loss: 1.056878
Train Epoch: 13 [12800/50000 (26%)]	Loss: 0.919123
Train Epoch: 13 [19200/50000 (38%)]	Loss: 0.941757
Train Epoch: 13 [25600/50000 (51%)]	Loss: 0.686299
Train Epoch: 13 [32000/50000 (64%)]	Loss: 1.021570
Train Epoch: 13 [38400/50000 (77%)]	Loss: 0.950681
Train Epoch: 13 [44800/50000 (90%)]	Loss: 1.178713

Test set: Average loss: 0.9643, Accuracy: 33129/50000 (66%)

Train Epoch: 14 [0/50000 (0%)]	Loss: 0.961063
Train Epoch: 14 [6400/50000 (13%)]	Loss: 1.211912
Train Epoch: 14 [12800/50000 (26%)]	Loss: 0.801523
Train Epoch: 14 [19200/50000 (38%)]	Loss: 0.836658
Train Epoch: 14 [25600/50000 (51%)]	Loss: 0.761864
Train Epoch: 14 [32000/50000 (64%)]	Loss: 0.882734
Train Epoch: 14 [38400/50000 (77%)]	Loss: 0.955647
Train Epoch: 14 [44800/50000 (90%)]	Loss: 0.868048

Test set: Average loss: 1.0008, Accuracy: 32264/50000 (65%)

Train Epoch: 15 [0/50000 (0%)]	Loss: 1.043757
Train Epoch: 15 [6400/50000 (13%)]	Loss: 0.786045
Train Epoch: 15 [12800/50000 (26%)]	Loss: 0.973085
Train Epoch: 15 [19200/50000 (38%)]	Loss: 1.267984
Train Epoch: 15 [25600/50000 (51%)]	Loss: 0.823255
Train Epoch: 15 [32000/50000 (64%)]	Loss: 0.796976
Train Epoch: 15 [38400/50000 (77%)]	Loss: 0.955850
Train Epoch: 15 [44800/50000 (90%)]	Loss: 0.745473

Test set: Average loss: 0.9408, Accuracy: 33443/50000 (67%)

Train Epoch: 16 [0/50000 (0%)]	Loss: 0.864455
Train Epoch: 16 [6400/50000 (13%)]	Loss: 1.029907
Train Epoch: 16 [12800/50000 (26%)]	Loss: 1.026490
Train Epoch: 16 [19200/50000 (38%)]	Loss: 0.945166
Train Epoch: 16 [25600/50000 (51%)]	Loss: 0.854361
Train Epoch: 16 [32000/50000 (64%)]	Loss: 0.937528
Train Epoch: 16 [38400/50000 (77%)]	Loss: 0.960435
Train Epoch: 16 [44800/50000 (90%)]	Loss: 0.942537

Test set: Average loss: 0.8823, Accuracy: 34671/50000 (69%)

Train Epoch: 17 [0/50000 (0%)]	Loss: 0.674867
--------------------------------	----------------

Train Epoch: 17 [6400/50000 (13%)]	Loss: 0.783870
Train Epoch: 17 [12800/50000 (26%)]	Loss: 0.853094
Train Epoch: 17 [19200/50000 (38%)]	Loss: 1.020453
Train Epoch: 17 [25600/50000 (51%)]	Loss: 0.905458
Train Epoch: 17 [32000/50000 (64%)]	Loss: 0.846606
Train Epoch: 17 [38400/50000 (77%)]	Loss: 0.776157
Train Epoch: 17 [44800/50000 (90%)]	Loss: 0.772932

Test set: Average loss: 1.0236, Accuracy: 32319/50000 (65%)

Train Epoch: 18 [0/50000 (0%)]	Loss: 1.177680
Train Epoch: 18 [6400/50000 (13%)]	Loss: 0.820801
Train Epoch: 18 [12800/50000 (26%)]	Loss: 0.799447
Train Epoch: 18 [19200/50000 (38%)]	Loss: 1.019677
Train Epoch: 18 [25600/50000 (51%)]	Loss: 0.965439
Train Epoch: 18 [32000/50000 (64%)]	Loss: 0.856691
Train Epoch: 18 [38400/50000 (77%)]	Loss: 0.857407
Train Epoch: 18 [44800/50000 (90%)]	Loss: 0.910402

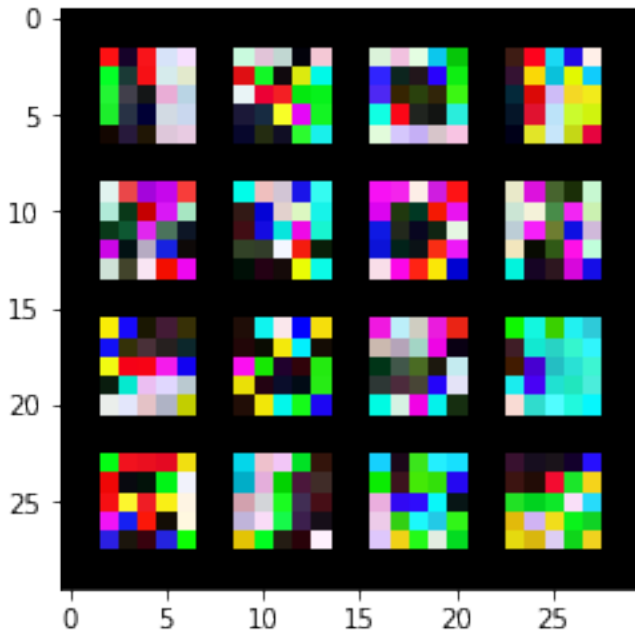
Test set: Average loss: 0.9046, Accuracy: 34295/50000 (69%)

Train Epoch: 19 [0/50000 (0%)]	Loss: 0.768344
Train Epoch: 19 [6400/50000 (13%)]	Loss: 1.172663
Train Epoch: 19 [12800/50000 (26%)]	Loss: 0.951083
Train Epoch: 19 [19200/50000 (38%)]	Loss: 0.825030
Train Epoch: 19 [25600/50000 (51%)]	Loss: 0.737163
Train Epoch: 19 [32000/50000 (64%)]	Loss: 0.984638
Train Epoch: 19 [38400/50000 (77%)]	Loss: 0.993044
Train Epoch: 19 [44800/50000 (90%)]	Loss: 0.780604

Test set: Average loss: 0.8341, Accuracy: 35466/50000 (71%)

Train Epoch: 20 [0/50000 (0%)]	Loss: 0.833307
Train Epoch: 20 [6400/50000 (13%)]	Loss: 1.105089
Train Epoch: 20 [12800/50000 (26%)]	Loss: 0.818836
Train Epoch: 20 [19200/50000 (38%)]	Loss: 0.785232
Train Epoch: 20 [25600/50000 (51%)]	Loss: 0.954037
Train Epoch: 20 [32000/50000 (64%)]	Loss: 0.724312
Train Epoch: 20 [38400/50000 (77%)]	Loss: 0.738534
Train Epoch: 20 [44800/50000 (90%)]	Loss: 0.771691

Test set: Average loss: 0.8010, Accuracy: 36100/50000 (72%)



Weights vector for layer 1:conv1

Break-Down of parameters:Img_size=32x32

Layer 1: conv1: **16x3x5x5**(3 layer input, 16 channels, 5x5 convolution filter)Img_size=28x28

Pooling: img_size=14x14

Layer 2: conv2: **128x16x5x5** (16 layer input, 128 channels, 5x5 convolution filter)img_size=10x10

Pooling: img_size=5x5

Layer 3: linear1: **128x5x5x64** (128 layer inputx 5x5 input img, 64 hidden units)

Layer 4: linear2: **64x10**(64 hidden units, 10 output units)

Total Number of parameters: **257840**