Computer Vision CSCI-667
HW-4

Rupesh Kumar

USC ID: **4382673678**

Goal: Image classification on STL10 Dataset using LeNet-5 Architecture.

Framework: PyTorch, Python

Architecture

Architecture

Construct a LeNet-5 style CNN network, using PyTorch functions. LeNet-5 is shown in Figure 1. Note that the network is not exactly the same as described in the original, 1998, paper.

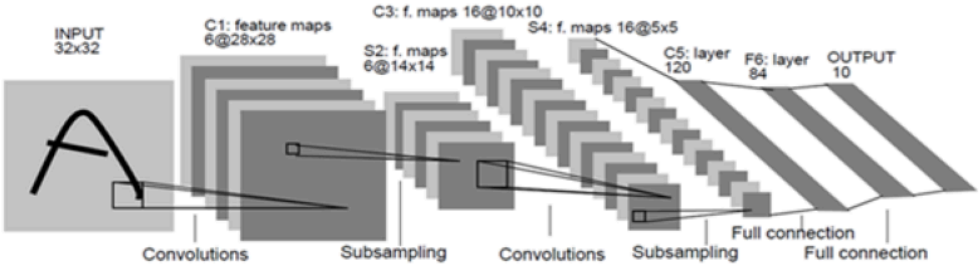


Figure 1: LeNet-5 Architecture

We ask you to experiment with varying the parameters of the network but use the following to start with (which is referred as 'main experiment' in the following sections):

1. The first layer has six 5×5 convolution filters, stride as 1, each followed by a max-pooling layer of 2×2 with stride as 2.
2. Second convolution layer has sixteen, 5×5 convolution filters, stride as 1, each followed by 2×2 max pooling with stride as 2.
3. Next is a fully connected layer of dimensions 120 followed by another fully connected layer of dimensions 84.
4. Next is a fully connected layer of dimensions 10 that gives unnormalized scores of the 10 classes.
5. All activation units should be ReLU.

Dataset: STL-10 dataset.

1. Result with given settings.

Loss Function: Cross entropy loss
Optimizer: ADAM
Learning rate: 0.001(initially and then decrementing by 50 % after every 20 epochs)
Number of epochs: 100
Batch-size: 128

CSCI_677_HW4$ python main.py

Model Summary

```
----------------------------------------------------------------
      Layer (type)               Output Shape         Param #
================================================================
          Conv2d-1            [-1, 6, 28, 28]             456
        MaxPool2d-2            [-1, 6, 14, 14]               0
          Conv2d-3           [-1, 16, 10, 10]           2,416
        MaxPool2d-4            [-1, 16, 5, 5]               0
          Linear-5                  [-1, 120]          48,120
          Linear-6                   [-1, 84]          10,164
          Linear-7                   [-1, 10]             850
================================================================
```

Total params: 62,006

```
Trainable params: 62,006
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.01
Forward/backward pass size (MB): 0.06
Params size (MB): 0.24
Estimated Total Size (MB): 0.31
----------------------------------------------------------------
None
############################################################
<generator object Module.parameters at 0x13319bb30>
############################################################
Total params:  62006
[1,     9] loss: 2.302
[1,    18] loss: 2.298
[1,    27] loss: 2.278
[1,    36] loss: 2.201
Testing the accuracy after epoch 0 on validation set for accuracy ...
**********************************************************
Accuracy of the network on the 3000 validation images: 19 %
F1 score on val dataset:  0.12315011315825201
Confusion Matrix on Validation dataset :  [[ 95   0   0   0 107   0   0   0  32  66]
 [  8   0   0   0 266   0   1   0   8  17]
 [  3   0   4   0 136   0   2   0  17 138]
 [  0   0   0   0 299   0   0   0   0   1]
 [  1   0   0   0 293   0   0   0   0   6]
 [  0   0   0   0 294   0   0   0   2   4]
 [  0   0   0   0 289   0   1   0   1   9]
 [  1   0   0   0 291   0   0   0   2   6]
 [ 68   0   0   0 103   0   0   0  43  86]]
```

```
 [  5   0   1   0 141   0   1   0  11 141]]
Epoch-0 lr: 0.001

[2,     9] loss: 2.094
[2,    18] loss: 1.969
[2,    27] loss: 1.959
[2,    36] loss: 1.889
Epoch-1 lr: 0.001
[3,     9] loss: 1.838
[3,    18] loss: 1.835
[3,    27] loss: 1.785
[3,    36] loss: 1.736
Epoch-2 lr: 0.001
[4,     9] loss: 1.916
[4,    18] loss: 1.827
[4,    27] loss: 1.771
[4,    36] loss: 1.720
Epoch-3 lr: 0.001
[5,     9] loss: 1.679
[5,    18] loss: 1.744
[5,    27] loss: 1.703
[5,    36] loss: 1.702
Epoch-4 lr: 0.001
[6,     9] loss: 1.812
[6,    18] loss: 1.747
[6,    27] loss: 1.721
[6,    36] loss: 1.619
Testing the accuracy after epoch 5 on validation set for accuracy ...
************************************************************
Accuracy of the network on the 3000 validation images: 35 %
F1 score on val dataset:  0.332008087625867
Confusion Matrix on Validation dataset :  [[149  34   2   1   3   0  19   2  53  37]
 [  9  77   1   4  69   0  78  43   7  12]
 [  6  15 130   0   4   0  20   7   8 110]
 [  0  69   3   8  91   0  82  40   1   6]
 [  1  42   2   4 140   0  79  29   1   2]
 [  0  35   2   6  92   0 126  35   0   4]
 [  0  26   0   1  37   0 200  22   1  13]
 [  1  39   1   4  91   0  86  68   0  10]
 [ 35  27  11   0   3   0  13   3 149  59]
 [  7  29  50   0   1   0  30   7  25 151]]
Epoch-5 lr: 0.001
[7,     9] loss: 1.720
[7,    18] loss: 1.669
[7,    27] loss: 1.673
[7,    36] loss: 1.606
Epoch-6 lr: 0.001
[8,     9] loss: 1.611
[8,    18] loss: 1.632
[8,    27] loss: 1.596
[8,    36] loss: 1.606
Epoch-7 lr: 0.001
[9,     9] loss: 1.615
```

[9, 18] loss: 1.598
[9, 27] loss: 1.568
[9, 36] loss: 1.643
Epoch-8 lr: 0.001
[10, 9] loss: 1.639
[10, 18] loss: 1.616
[10, 27] loss: 1.596
[10, 36] loss: 1.590
Epoch-9 lr: 0.001
[11, 9] loss: 1.590
[11, 18] loss: 1.569
[11, 27] loss: 1.553
[11, 36] loss: 1.582
Testing the accuracy after epoch 10 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 40 %
F1 score on val dataset: 0.38854511044712753
Confusion Matrix on Validation dataset : [[178 5 22 9 1 0 8 0 63 14]
 [17 28 7 83 34 18 36 50 17 10]
 [14 1 222 11 2 1 5 4 8 32]
 [2 14 8 105 33 31 34 52 4 17]
 [5 13 5 84 87 15 50 33 3 5]
 [6 8 9 56 33 39 79 58 1 11]
 [2 1 7 31 16 32 156 35 1 19]
 [1 4 7 70 34 26 41 102 5 10]
 [28 9 26 3 1 3 0 3 194 33]
 [9 1 115 10 1 1 6 7 46 104]]
Epoch-10 lr: 0.001
[12, 9] loss: 1.529
[12, 18] loss: 1.532
[12, 27] loss: 1.521
[12, 36] loss: 1.544
Epoch-11 lr: 0.001
[13, 9] loss: 1.512
[13, 18] loss: 1.550
[13, 27] loss: 1.497
[13, 36] loss: 1.507
Epoch-12 lr: 0.001
[14, 9] loss: 1.553
[14, 18] loss: 1.495
[14, 27] loss: 1.486
[14, 36] loss: 1.442
Epoch-13 lr: 0.001
[15, 9] loss: 1.456
[15, 18] loss: 1.509
[15, 27] loss: 1.527
[15, 36] loss: 1.465
Epoch-14 lr: 0.001
[16, 9] loss: 1.508
[16, 18] loss: 1.537
[16, 27] loss: 1.480
[16, 36] loss: 1.509
Testing the accuracy after epoch 15 on validation set for accuracy ...

```
************************************************************
Accuracy of the network on the 3000 validation images: 40 %
F1 score on val dataset:  0.37546418766121287
Confusion Matrix on Validation dataset :  [[236   6  27   4   2   2   2   0  15   6]
 [ 24  43  15  60  67  15  37  26   8   5]
 [ 19   5 249   3   3   3   1   2   6   9]
 [  3  24  15 103  69  22  35  22   2   5]
 [  9  11   6  39 157  11  55   9   0   3]
 [  6  13  14  47  77  30  81  28   0   4]
 [  4   1  17  23  34  25 173  14   1   8]
 [  3   9  10  57  75  29  50  59   3   5]
 [ 81  11  51   3   3   1   0   2 130  18]
 [ 27   8 178   6   2   1   7   2  24  45]]
Epoch-15 lr: 0.001
[17,     9] loss: 1.473
[17,    18] loss: 1.462
[17,    27] loss: 1.440
[17,    36] loss: 1.414
Epoch-16 lr: 0.001
[18,     9] loss: 1.499
[18,    18] loss: 1.487
[18,    27] loss: 1.439
[18,    36] loss: 1.398
Epoch-17 lr: 0.001
[19,     9] loss: 1.484
[19,    18] loss: 1.454
[19,    27] loss: 1.404
[19,    36] loss: 1.424
Epoch-18 lr: 0.001
[20,     9] loss: 1.453
[20,    18] loss: 1.432
[20,    27] loss: 1.363
[20,    36] loss: 1.392
Epoch-19 lr: 0.0005
[21,     9] loss: 1.392
[21,    18] loss: 1.349
[21,    27] loss: 1.376
[21,    36] loss: 1.412
Testing the accuracy after epoch 20 on validation set for accuracy ...
************************************************************
Accuracy of the network on the 3000 validation images: 42 %
F1 score on val dataset:  0.4080489732422352
Confusion Matrix on Validation dataset :  [[225   7   9   3   1   2   1   0  45   7]
 [ 22  79   8  64  75   3   6  18  17   8]
 [ 21   7 211   5   2   0   0   2  21  31]
 [  6  47  10  98  86  14   3  17  10   9]
 [  6  39   3  39 180   2  11  11   8   1]
 [  7  39   9  68  95  23  21  31   3   4]
 [  8  20  11  29  82  32  83  18   2  15]
 [  3  28   5  66  88  18  10  68   8   6]
 [ 54   8  12   0   2   1   0   1 206  16]
 [ 23  11  83   6   2   1   2   0  63 109]]
Epoch-20 lr: 0.0005
```

```
[22,     9] loss: 1.429
[22,    18] loss: 1.406
[22,    27] loss: 1.407
[22,    36] loss: 1.339
Epoch-21 lr: 0.0005
[23,     9] loss: 1.357
[23,    18] loss: 1.380
[23,    27] loss: 1.373
[23,    36] loss: 1.392
Epoch-22 lr: 0.0005
[24,     9] loss: 1.398
[24,    18] loss: 1.377
[24,    27] loss: 1.396
[24,    36] loss: 1.356
Epoch-23 lr: 0.0005
[25,     9] loss: 1.374
[25,    18] loss: 1.308
[25,    27] loss: 1.363
[25,    36] loss: 1.368
Epoch-24 lr: 0.0005
[26,     9] loss: 1.376
[26,    18] loss: 1.324
[26,    27] loss: 1.289
[26,    36] loss: 1.364
Testing the accuracy after epoch 25 on validation set for accuracy ...
************************************************************
Accuracy of the network on the 3000 validation images: 43 %
F1 score on val dataset:  0.43081568765864175
Confusion Matrix on Validation dataset :  [[203  14  17   7   3   2   1   0  37  16]
 [ 13  70   4  62  81   5  14  34   9   8]
 [ 17   8 205   4   5   3   1   5  12  40]
 [  2  43   4  98  84  16  15  30   1   7]
 [  1  28   3  30 186   6  21  23   1   1]
 [  2  28   7  58  91  31  31  48   0   4]
 [  1  11   5  22  70  26 118  31   1  15]
 [  0  21   3  57  90  12  14  97   2   4]
 [ 42  21  17   3   4   1   0   1 180  31]
 [ 14  17  84   8   3   1   5   2  36 130]]
Epoch-25 lr: 0.0005
[27,     9] loss: 1.398
[27,    18] loss: 1.347
[27,    27] loss: 1.356
[27,    36] loss: 1.309
Epoch-26 lr: 0.0005
[28,     9] loss: 1.322
[28,    18] loss: 1.345
[28,    27] loss: 1.354
[28,    36] loss: 1.368
Epoch-27 lr: 0.0005
[29,     9] loss: 1.356
[29,    18] loss: 1.316
[29,    27] loss: 1.303
```

[29, 36] loss: 1.338
Epoch-28 lr: 0.0005
[30, 9] loss: 1.346
[30, 18] loss: 1.349
[30, 27] loss: 1.292
[30, 36] loss: 1.335
Epoch-29 lr: 0.0005
[31, 9] loss: 1.328
[31, 18] loss: 1.322
[31, 27] loss: 1.275
[31, 36] loss: 1.288
Testing the accuracy after epoch 30 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 45 %
F1 score on val dataset: 0.4378999715342875
Confusion Matrix on Validation dataset : [[202 7 10 2 2 2 2 0 47 26]
 [15 84 4 32 68 14 30 28 10 15]
 [16 6 150 3 3 2 5 2 11 102]
 [3 52 4 59 75 17 34 37 5 14]
 [1 28 4 21 167 6 47 21 2 3]
 [4 32 3 25 67 42 67 45 2 13]
 [1 4 1 4 40 24 182 20 1 23]
 [2 21 1 24 72 30 48 91 3 8]
 [39 13 4 0 3 1 0 0 181 59]
 [13 11 34 3 3 1 5 3 33 194]]
Epoch-30 lr: 0.0005
[32, 9] loss: 1.338
[32, 18] loss: 1.271
[32, 27] loss: 1.334
[32, 36] loss: 1.353
Epoch-31 lr: 0.0005
[33, 9] loss: 1.352
[33, 18] loss: 1.328
[33, 27] loss: 1.326
[33, 36] loss: 1.304
Epoch-32 lr: 0.0005
[34, 9] loss: 1.331
[34, 18] loss: 1.303
[34, 27] loss: 1.313
[34, 36] loss: 1.303
Epoch-33 lr: 0.0005
[35, 9] loss: 1.304
[35, 18] loss: 1.228
[35, 27] loss: 1.309
[35, 36] loss: 1.308
Epoch-34 lr: 0.0005
[36, 9] loss: 1.276
[36, 18] loss: 1.285
[36, 27] loss: 1.322
[36, 36] loss: 1.263
Testing the accuracy after epoch 35 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 45 %

F1 score on val dataset: 0.4459591096283828
Confusion Matrix on Validation dataset : [[204 11 11   8   3   2   6   0 35 20]
 [14 79   3 34 71 21 41 23   8   6]
 [15   7 183   5   4   2   5   4 12 63]
 [3 41   3 66 76 19 48 35   2   7]
 [1 24   2 21 172   7 54 16   1   2]
 [3 22   4 28 65 42 89 41   0   6]
 [1   5   2   5 36 22 201 18   1   9]
 [1 17   1 25 69 36 57 88   3   3]
 [44 17   9   3   5   1   1   1 181 38]
 [10 14 57   4   3   1 14   2 37 158]]
Epoch-35 lr: 0.0005
[37,   9] loss: 1.284
[37,   18] loss: 1.259
[37,   27] loss: 1.260
[37,   36] loss: 1.273
Epoch-36 lr: 0.0005
[38,   9] loss: 1.310
[38,   18] loss: 1.255
[38,   27] loss: 1.299
[38,   36] loss: 1.240
Epoch-37 lr: 0.0005
[39,   9] loss: 1.292
[39,   18] loss: 1.267
[39,   27] loss: 1.282
[39,   36] loss: 1.257
Epoch-38 lr: 0.0005
[40,   9] loss: 1.254
[40,   18] loss: 1.255
[40,   27] loss: 1.259
[40,   36] loss: 1.245
Epoch-39 lr: 0.00025
[41,   9] loss: 1.273
[41,   18] loss: 1.244
[41,   27] loss: 1.188
[41,   36] loss: 1.238
Testing the accuracy after epoch 40 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 48 %
F1 score on val dataset: 0.47261129170427074
Confusion Matrix on Validation dataset : [[220 14 12   2   1   2   2   0 34 13]
 [16 103   6 33 52 22 19 34   7   8]
 [15   6 192   4   3   3   2   5 15 55]
 [3 50   6 72 52 24 29 51   6   7]
 [3 29   4 22 161 14 36 27   3   1]
 [4 38   5 22 51 59 53 61   2   5]
 [1   9   2   5 37 41 161 28   1 15]
 [2 19   2 25 50 38 33 124   4   3]
 [43 14   9   0   4   1   0   1 199 29]
 [16 14 61   4   2   1   5   3 44 150]]
Epoch-40 lr: 0.00025

[42,   9] loss: 1.221

```
[42,    18] loss: 1.237
[42,    27] loss: 1.204
[42,    36] loss: 1.262
Epoch-41 lr: 0.00025
[43,     9] loss: 1.210
[43,    18] loss: 1.234
[43,    27] loss: 1.237
[43,    36] loss: 1.221
Epoch-42 lr: 0.00025
[44,     9] loss: 1.212
[44,    18] loss: 1.209
[44,    27] loss: 1.216
[44,    36] loss: 1.247
Epoch-43 lr: 0.00025
[45,     9] loss: 1.247
[45,    18] loss: 1.198
[45,    27] loss: 1.224
[45,    36] loss: 1.226
Epoch-44 lr: 0.00025
[46,     9] loss: 1.196
[46,    18] loss: 1.193
[46,    27] loss: 1.247
[46,    36] loss: 1.263
Testing the accuracy after epoch 45 on validation set for accuracy ...
************************************************************
Accuracy of the network on the 3000 validation images: 47 %
F1 score on val dataset:  0.46247543054273227
Confusion Matrix on Validation dataset :  [[224  11  13   4   1   2   2   0  32  11]
 [ 19  95   6  41  54  29  19  22   7   8]
 [ 18   7 199   4   3   3   3   3  15  45]
 [  3  49   7  83  55  28  32  31   5   7]
 [  3  28   4  27 165  17  38  16   1   1]
 [  3  36   5  35  54  61  60  38   2   6]
 [  1   8   4  10  34  44 169  16   1  13]
 [  2  20   3  40  61  40  40  87   2   5]
 [ 52  13  12   1   4   1   0   0 187  30]
 [ 18  12  70   7   2   2   4   1  41 143]]
Epoch-45 lr: 0.00025
[47,     9] loss: 1.209
[47,    18] loss: 1.187
[47,    27] loss: 1.230
[47,    36] loss: 1.236
Epoch-46 lr: 0.00025
[48,     9] loss: 1.179
[48,    18] loss: 1.189
[48,    27] loss: 1.213
[48,    36] loss: 1.258
Epoch-47 lr: 0.00025
[49,     9] loss: 1.250
[49,    18] loss: 1.195
[49,    27] loss: 1.213
[49,    36] loss: 1.202
Epoch-48 lr: 0.00025
```

[50, 9] loss: 1.208
[50, 18] loss: 1.213
[50, 27] loss: 1.192
[50, 36] loss: 1.196
Epoch-49 lr: 0.00025
[51, 9] loss: 1.196
[51, 18] loss: 1.201
[51, 27] loss: 1.204
[51, 36] loss: 1.197
Testing the accuracy after epoch 50 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 47 %
F1 score on val dataset: 0.47763891831113553
Confusion Matrix on Validation dataset : [[213 12 9 5 0 2 2 0 38 19]
 [15 114 6 39 36 32 13 29 8 8]
 [16 8 173 4 2 3 3 4 15 72]
 [3 52 5 90 26 37 24 48 6 9]
 [2 37 4 36 132 24 34 26 3 2]
 [4 39 4 30 35 68 51 59 2 8]
 [1 11 2 11 23 57 157 22 1 15]
 [1 21 1 35 27 44 35 127 4 5]
 [41 19 5 0 3 0 0 0 197 35]
 [15 12 50 5 2 2 4 2 40 168]]
Epoch-50 lr: 0.00025
[52, 9] loss: 1.222
[52, 18] loss: 1.204
[52, 27] loss: 1.201
[52, 36] loss: 1.165
Epoch-51 lr: 0.00025
[53, 9] loss: 1.176
[53, 18] loss: 1.179
[53, 27] loss: 1.184
[53, 36] loss: 1.223
Epoch-52 lr: 0.00025
[54, 9] loss: 1.172
[54, 18] loss: 1.232
[54, 27] loss: 1.168
[54, 36] loss: 1.194
Epoch-53 lr: 0.00025
[55, 9] loss: 1.173
[55, 18] loss: 1.155
[55, 27] loss: 1.190
[55, 36] loss: 1.221
Epoch-54 lr: 0.00025
[56, 9] loss: 1.179
[56, 18] loss: 1.179
[56, 27] loss: 1.176
[56, 36] loss: 1.188
Testing the accuracy after epoch 55 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 47 %
F1 score on val dataset: 0.469453180463928
Confusion Matrix on Validation dataset : [[190 19 12 2 0 2 2 0 59 14]

[14 129 2 23 33 25 14 44 9 7]
[14 9 184 3 2 3 3 6 19 57]
[3 60 4 60 26 25 29 81 6 6]
[1 46 3 22 128 20 33 43 3 1]
[2 43 4 21 34 49 57 82 3 5]
[1 12 4 6 19 38 166 44 1 9]
[1 21 2 15 24 32 31 167 4 3]
[33 18 7 0 3 0 0 1 208 30]
[11 17 55 3 2 1 5 5 50 151]]
Epoch-55 lr: 0.00025
[57, 9] loss: 1.137
[57, 18] loss: 1.192
[57, 27] loss: 1.178
[57, 36] loss: 1.217
Epoch-56 lr: 0.00025
[58, 9] loss: 1.213
[58, 18] loss: 1.130
[58, 27] loss: 1.214
[58, 36] loss: 1.173
Epoch-57 lr: 0.00025
[59, 9] loss: 1.146
[59, 18] loss: 1.246
[59, 27] loss: 1.129
[59, 36] loss: 1.186
Epoch-58 lr: 0.00025
[60, 9] loss: 1.189
[60, 18] loss: 1.158
[60, 27] loss: 1.180
[60, 36] loss: 1.196
Epoch-59 lr: 0.000125
[61, 9] loss: 1.173
[61, 18] loss: 1.163
[61, 27] loss: 1.149
[61, 36] loss: 1.160
Testing the accuracy after epoch 60 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 48 %
F1 score on val dataset: 0.4753150183148185
Confusion Matrix on Validation dataset : [[209 14 11 5 2 2 2 0 42 13]
[15 100 2 44 52 25 16 34 8 4]
[13 7 194 6 3 2 3 6 16 50]
[3 40 2 84 54 22 32 54 4 5]
[2 28 2 28 169 12 31 25 2 1]
[2 33 4 34 50 42 61 68 1 5]
[1 4 4 12 36 33 169 32 1 8]
[1 14 3 32 48 31 33 132 3 3]
[38 17 11 6 4 0 1 0 198 25]
[12 12 64 8 2 0 6 5 39 152]]
Epoch-60 lr: 0.000125

[62, 9] loss: 1.126
[62, 18] loss: 1.142
[62, 27] loss: 1.175

[62,   36] loss: 1.147
Epoch-61 lr: 0.000125
[63,   9] loss: 1.151
[63,   18] loss: 1.130
[63,   27] loss: 1.178
[63,   36] loss: 1.131
Epoch-62 lr: 0.000125
[64,   9] loss: 1.117
[64,   18] loss: 1.183
[64,   27] loss: 1.109
[64,   36] loss: 1.160
Epoch-63 lr: 0.000125
[65,   9] loss: 1.173
[65,   18] loss: 1.124
[65,   27] loss: 1.174
[65,   36] loss: 1.144
Epoch-64 lr: 0.000125
[66,   9] loss: 1.161
[66,   18] loss: 1.180
[66,   27] loss: 1.138
[66,   36] loss: 1.160
Testing the accuracy after epoch 65 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 48 %
F1 score on val dataset: 0.47181463928082573
Confusion Matrix on Validation dataset : [[208 13 15   5   3   2   2   0 39 13]
 [15 105   5 34 54 28 17 29   8   5]
 [13   8 202   5   4   2   3   6 13 44]
 [3 41   5 72 61 25 30 53   4   6]
 [2 29   4 20 174 13 30 25   2   1]
 [2 33   4 31 52 44 62 66   1   5]
 [1   6   4   9 37 38 169 27   1   8]
 [1 14   3 34 48 32 32 129   3   4]
 [39 18 16   5   4   0   0   0 193 25]
 [12 12 77   7   2   1   5   3 35 146]]
Epoch-65 lr: 0.000125
[67,   9] loss: 1.118
[67,   18] loss: 1.117
[67,   27] loss: 1.161
[67,   36] loss: 1.179
Epoch-66 lr: 0.000125
[68,   9] loss: 1.116
[68,   18] loss: 1.169
[68,   27] loss: 1.141
[68,   36] loss: 1.160
Epoch-67 lr: 0.000125
[69,   9] loss: 1.154
[69,   18] loss: 1.143
[69,   27] loss: 1.146
[69,   36] loss: 1.141
Epoch-68 lr: 0.000125
[70,   9] loss: 1.099
[70,   18] loss: 1.164

```
[70,    27] loss: 1.136
[70,    36] loss: 1.156
Epoch-69 lr: 0.000125
[71,     9] loss: 1.159
[71,    18] loss: 1.157
[71,    27] loss: 1.097
[71,    36] loss: 1.139
Testing the accuracy after epoch 70 on validation set for accuracy ...
************************************************************
Accuracy of the network on the 3000 validation images: 49 %
F1 score on val dataset:  0.48345992110180613
Confusion Matrix on Validation dataset :  [[224   9  11   3   0   2   2   0  38  11]
 [ 16 106   5  43  47  28  20  20   8   7]
 [ 19   5 189   4   2   3   4   4  16  54]
 [  3  44   4  90  47  21  38  40   6   7]
 [  3  26   4  30 163  12  37  21   2   2]
 [  3  33   4  37  46  49  65  55   3   5]
 [  1   5   4  12  33  34 182  17   1  11]
 [  1  17   2  38  44  33  41 116   4   4]
 [ 45  11  10   2   3   0   0   0 205  24]
 [ 17   9  62   6   2   1   4   2  42 155]]
Epoch-70 lr: 0.000125
[72,     9] loss: 1.136
[72,    18] loss: 1.133
[72,    27] loss: 1.172
[72,    36] loss: 1.143
Epoch-71 lr: 0.000125
[73,     9] loss: 1.112
[73,    18] loss: 1.126
[73,    27] loss: 1.176
[73,    36] loss: 1.124
Epoch-72 lr: 0.000125
[74,     9] loss: 1.160
[74,    18] loss: 1.121
[74,    27] loss: 1.100
[74,    36] loss: 1.147
Epoch-73 lr: 0.000125
[75,     9] loss: 1.123
[75,    18] loss: 1.109
[75,    27] loss: 1.176
[75,    36] loss: 1.138
Epoch-74 lr: 0.000125
[76,     9] loss: 1.126
[76,    18] loss: 1.151
[76,    27] loss: 1.100
[76,    36] loss: 1.111
Testing the accuracy after epoch 75 on validation set for accuracy ...
************************************************************
Accuracy of the network on the 3000 validation images: 48 %
F1 score on val dataset:  0.47358040630284265
Confusion Matrix on Validation dataset :  [[203  14  16   5   2   2   3   0  40  15]
 [ 15 106   5  33  51  23  19  33   8   7]
 [ 13   7 194   3   4   3   3   6  12  55]
```

```
 [  3  44   4  71  52  18  36  61   4   7]
 [  2  30   3  22 167  10  34  29   2   1]
 [  2  33   4  27  48  39  64  77   1   5]
 [  1   6   4   9  32  29 178  32   1   8]
 [  1  14   3  29  45  26  35 140   3   4]
 [ 40  15  15   5   3   0   0   0 194  28]
 [ 11  12  70   7   2   0   6   3  29 160]]
Epoch-75 lr: 0.000125
[77,     9] loss: 1.144
[77,    18] loss: 1.120
[77,    27] loss: 1.101
[77,    36] loss: 1.117
Epoch-76 lr: 0.000125
[78,     9] loss: 1.144
[78,    18] loss: 1.116
[78,    27] loss: 1.095
[78,    36] loss: 1.136
Epoch-77 lr: 0.000125
[79,     9] loss: 1.135
[79,    18] loss: 1.076
[79,    27] loss: 1.120
[79,    36] loss: 1.128
Epoch-78 lr: 0.000125
[80,     9] loss: 1.170
[80,    18] loss: 1.111
[80,    27] loss: 1.098
[80,    36] loss: 1.102
Epoch-79 lr: 6.25e-05
[81,     9] loss: 1.114
[81,    18] loss: 1.129
[81,    27] loss: 1.089
[81,    36] loss: 1.124
Testing the accuracy after epoch 80 on validation set for accuracy ...
************************************************************
Accuracy of the network on the 3000 validation images: 48 %
F1 score on val dataset:  0.4808664663379533
Confusion Matrix on Validation dataset :  [[212  11  12   4   1   2   2   0  43  13]
 [ 16 103   5  42  50  29  15  25   8   7]
 [ 16   5 185   5   2   2   3   6  17  59]
 [  3  39   4  95  49  27  29  42   6   6]
 [  2  28   4  31 165  13  29  24   2   2]
 [  2  34   4  37  45  54  54  59   4   7]
 [  1   7   4  12  36  45 160  23   1  11]
 [  1  15   2  39  45  33  32 125   4   4]
 [ 42  11   9   3   3   0   0   0 206  26]
 [ 16  10  59   6   2   1   4   2  43 157]]
Epoch-80 lr: 6.25e-05

[82,     9] loss: 1.102
[82,    18] loss: 1.090
[82,    27] loss: 1.139
[82,    36] loss: 1.128
Epoch-81 lr: 6.25e-05
```

[83, 9] loss: 1.108
[83, 18] loss: 1.096
[83, 27] loss: 1.127
[83, 36] loss: 1.102
Epoch-82 lr: 6.25e-05
[84, 9] loss: 1.100
[84, 18] loss: 1.107
[84, 27] loss: 1.108
[84, 36] loss: 1.121
Epoch-83 lr: 6.25e-05
[85, 9] loss: 1.127
[85, 18] loss: 1.126
[85, 27] loss: 1.102
[85, 36] loss: 1.105
Epoch-84 lr: 6.25e-05
[86, 9] loss: 1.095
[86, 18] loss: 1.159
[86, 27] loss: 1.119
[86, 36] loss: 1.090
Testing the accuracy after epoch 85 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 48 %
F1 score on val dataset: 0.4777141998231604
Confusion Matrix on Validation dataset : [[208 14 13 4 2 2 2 0 42 13]
 [15 108 4 37 51 28 15 29 8 5]
 [14 7 186 4 3 3 3 6 16 58]
 [3 42 2 83 53 26 27 52 5 7]
 [2 30 3 27 169 12 27 27 2 1]
 [2 36 4 33 48 47 56 67 2 5]
 [1 8 4 12 37 45 158 25 1 9]
 [1 15 2 34 45 31 31 133 4 4]
 [39 16 10 4 3 0 0 0 203 25]
 [15 13 60 7 2 0 4 3 38 158]]
Epoch-85 lr: 6.25e-05
[87, 9] loss: 1.133
[87, 18] loss: 1.075
[87, 27] loss: 1.121
[87, 36] loss: 1.090
Epoch-86 lr: 6.25e-05
[88, 9] loss: 1.093
[88, 18] loss: 1.141
[88, 27] loss: 1.089
[88, 36] loss: 1.079
Epoch-87 lr: 6.25e-05
[89, 9] loss: 1.118
[89, 18] loss: 1.063
[89, 27] loss: 1.138
[89, 36] loss: 1.125
Epoch-88 lr: 6.25e-05
[90, 9] loss: 1.100
[90, 18] loss: 1.092
[90, 27] loss: 1.123
[90, 36] loss: 1.114

Epoch-89 lr: 6.25e-05
[91, 9] loss: 1.098
[91, 18] loss: 1.090
[91, 27] loss: 1.110
[91, 36] loss: 1.096
Testing the accuracy after epoch 90 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 48 %
F1 score on val dataset: 0.4784640637425183
Confusion Matrix on Validation dataset : [[210 14 14 3 0 2 2 0 44 11]
 [16 118 5 36 44 26 15 26 8 6]
 [17 5 196 5 2 2 3 6 16 48]
 [3 48 4 89 45 25 28 46 6 6]
 [2 33 4 32 157 13 26 29 2 2]
 [2 38 4 37 40 47 53 68 4 7]
 [1 9 5 12 34 46 157 27 1 8]
 [1 18 3 36 43 31 31 130 3 4]
 [40 12 12 3 3 0 0 0 207 23]
 [17 11 69 5 2 1 4 2 43 146]]
Epoch-90 lr: 6.25e-05
[92, 9] loss: 1.105
[92, 18] loss: 1.102
[92, 27] loss: 1.125
[92, 36] loss: 1.099
Epoch-91 lr: 6.25e-05
[93, 9] loss: 1.049
[93, 18] loss: 1.126
[93, 27] loss: 1.103
[93, 36] loss: 1.138
Epoch-92 lr: 6.25e-05
[94, 9] loss: 1.051
[94, 18] loss: 1.087
[94, 27] loss: 1.142
[94, 36] loss: 1.122
Epoch-93 lr: 6.25e-05
[95, 9] loss: 1.138
[95, 18] loss: 1.125
[95, 27] loss: 1.093
[95, 36] loss: 1.065
Epoch-94 lr: 6.25e-05
[96, 9] loss: 1.065
[96, 18] loss: 1.122
[96, 27] loss: 1.076
[96, 36] loss: 1.118
Testing the accuracy after epoch 95 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 49 %
F1 score on val dataset: 0.482477408158922
Confusion Matrix on Validation dataset : [[211 13 15 3 1 2 2 0 39 14]
 [15 119 5 33 42 27 19 25 8 7]
 [15 5 193 3 2 3 3 6 13 57]
 [3 51 4 83 43 25 33 45 5 8]
 [2 32 4 26 156 14 37 25 2 2]

[2 34 4 35 44 47 62 60 2 10]
[1 7 5 7 31 35 177 22 1 14]
[1 18 3 35 41 32 38 125 3 4]
[41 13 14 4 3 0 0 0 196 29]
[15 10 69 5 2 1 4 2 27 165]]
Epoch-95 lr: 6.25e-05
[97, 9] loss: 1.090
[97, 18] loss: 1.121
[97, 27] loss: 1.082
[97, 36] loss: 1.112
Epoch-96 lr: 6.25e-05
[98, 9] loss: 1.070
[98, 18] loss: 1.101
[98, 27] loss: 1.133
[98, 36] loss: 1.082
Epoch-97 lr: 6.25e-05
[99, 9] loss: 1.067
[99, 18] loss: 1.106
[99, 27] loss: 1.114
[99, 36] loss: 1.093
Epoch-98 lr: 6.25e-05
[100, 9] loss: 1.094
[100, 18] loss: 1.102
[100, 27] loss: 1.071
[100, 36] loss: 1.092
Epoch-99 lr: 3.125e-05
Finished Training

Final Result on Test Dataset

Accuracy of the network on the 5000 test images: 48 %
F1 score on test dataset: 0.48287278669948247
Confusion Matrix on Test dataset : [
[335 36 20 10 9 3 7 1 63 16]
[24 181 8 56 79 41 16 80 9 6]
[18 20 325 8 1 6 7 3 28 84]
[2 43 5 134 123 48 26 105 8 6]
[7 38 5 58 262 28 48 47 3 4]
[3 37 2 51 85 108 80 124 4 6]
[3 22 4 20 46 53 274 67 2 9]
[2 37 4 57 72 40 42 240 2 4]
[49 14 19 13 5 3 4 3 346 44]
[32 22 107 11 4 4 11 9 72 228]
]

Graphs

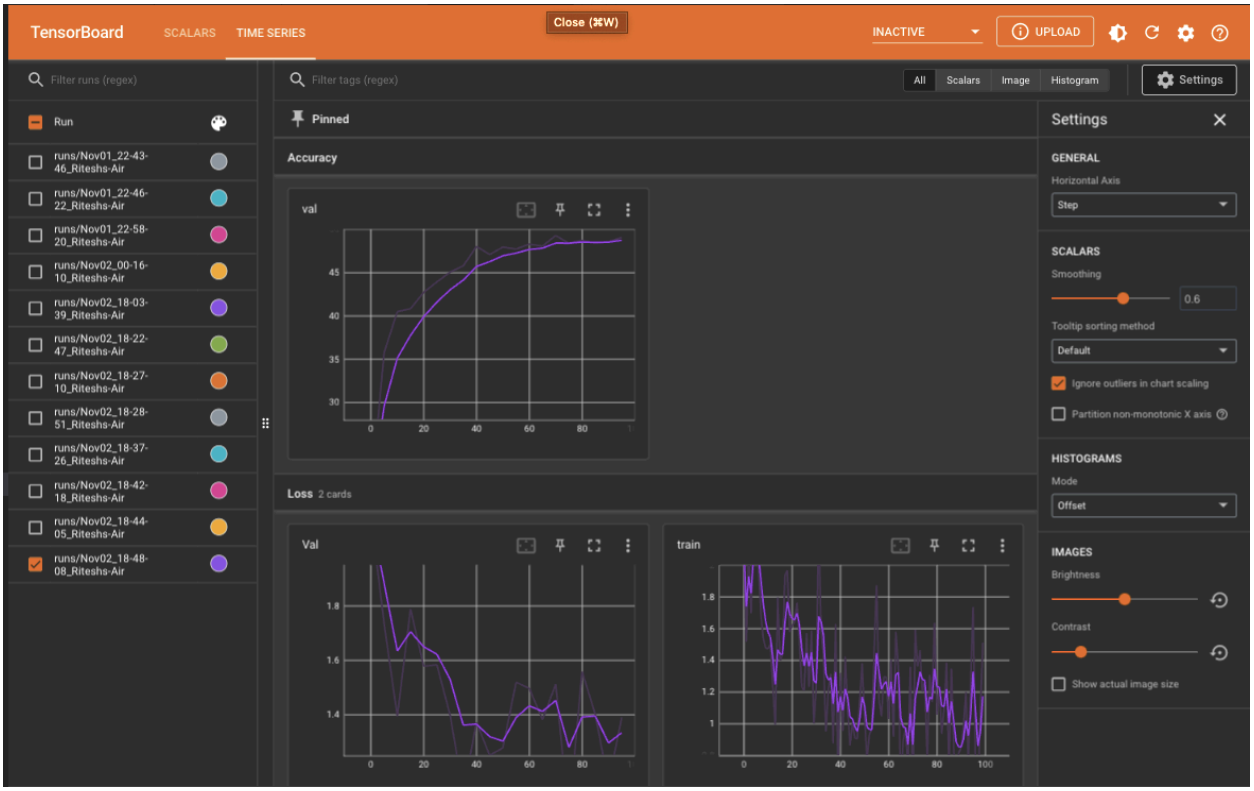


Fig 1. Showing Accuracy graph over epochs for validation dataset (on top), validation loss (on left-below), and train loss (on right-below)

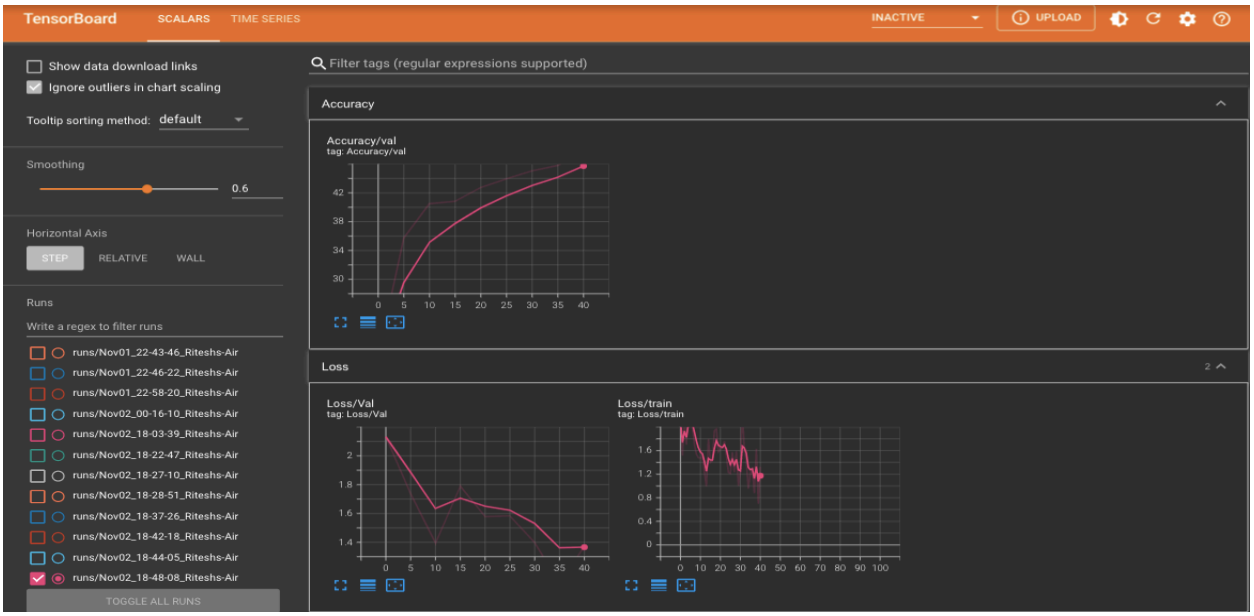


Fig 2. Time-series representation of above three graphs.

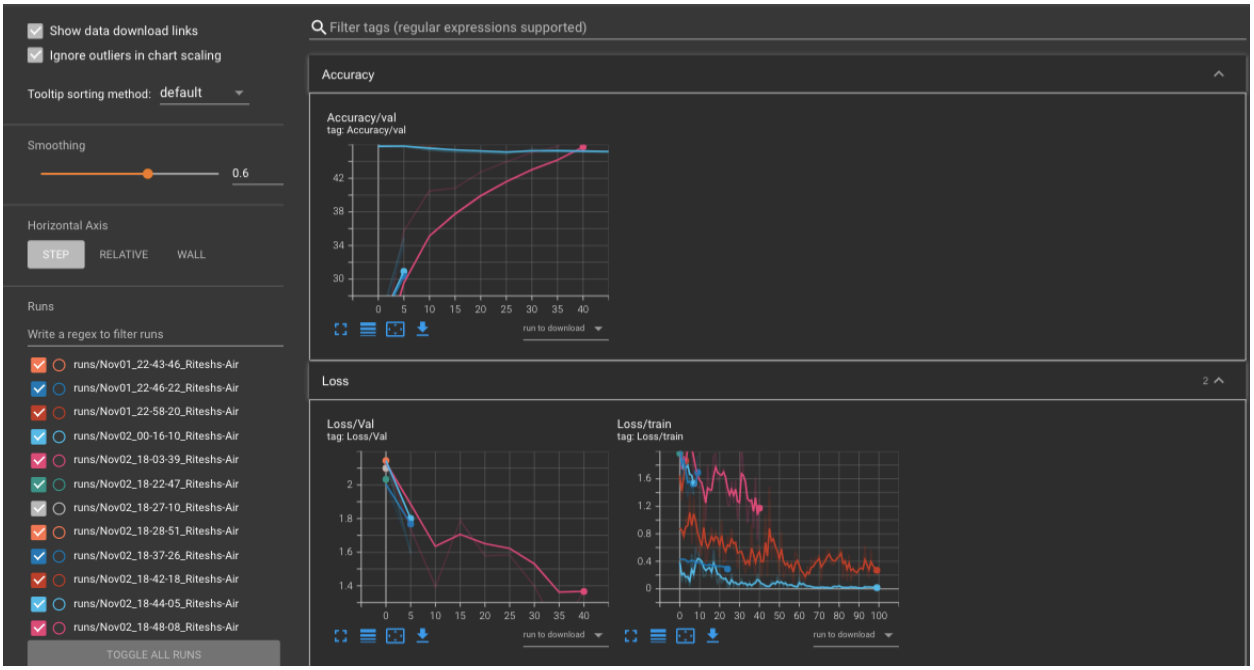


Fig 3: Training with different hyper-parameters values.

Accuracy in the above figure shows one curve which starts from 47 % and remains horizontal. This training was done with the help of loading the previous best model during the start of the training (sort of using pre-trained model).

2. Result with below setting.

Loss Function: Cross entropy loss
Optimizer: SGD
Learning rate: 0.00001(initially and then decrementing by 20 % after every 10 epochs)
Number of epochs: 100
Batch-size: 128

```
CSCI_677_HW4$ python main.py
Model Summary
----------------------------------------------------------------
      Layer (type)           Output Shape        Param #
================================================================
         Conv2d-1          [-1, 6, 28, 28]           456
       MaxPool2d-2          [-1, 6, 14, 14]             0
         Conv2d-3         [-1, 16, 10, 10]         2,416
       MaxPool2d-4          [-1, 16, 5, 5]             0
         Linear-5              [-1, 120]        48,120
         Linear-6               [-1, 84]        10,164
         Linear-7               [-1, 10]           850
================================================================
Total params: 62,006
Trainable params: 62,006
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.01
Forward/backward pass size (MB): 0.06
Params size (MB): 0.24
Estimated Total Size (MB): 0.31
----------------------------------------------------------------
None
##############################################################
<generator object Module.parameters at 0x13466eb30>
##############################################################
Total params:  62006
[1,     9] loss: 2.303
[1,    18] loss: 2.306
[1,    27] loss: 2.305
[1,    36] loss: 2.305
Testing the accuracy after epoch 0 on validation set for accuracy ...
**********************************************************
Accuracy of the network on the 3000 validation images: 10 %
F1 score on val dataset:  0.01818181818181818
Confusion Matrix on Validation dataset :  [[  0   0   0   0   0 300   0   0   0   0]
```

[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]]
Epoch-0 lr: 0.0001

[2, 9] loss: 2.305
[2, 18] loss: 2.308
[2, 27] loss: 2.305
[2, 36] loss: 2.302
Epoch-1 lr: 0.0001
[3, 9] loss: 2.306
[3, 18] loss: 2.303
[3, 27] loss: 2.306
[3, 36] loss: 2.303
Epoch-2 lr: 0.0001
[4, 9] loss: 2.307
[4, 18] loss: 2.304
[4, 27] loss: 2.303
[4, 36] loss: 2.307
Epoch-3 lr: 0.0001
[5, 9] loss: 2.306
[5, 18] loss: 2.304
[5, 27] loss: 2.302
[5, 36] loss: 2.306
Epoch-4 lr: 0.0001
[6, 9] loss: 2.306
[6, 18] loss: 2.303
[6, 27] loss: 2.306
[6, 36] loss: 2.305
Testing the accuracy after epoch 5 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 10 %
F1 score on val dataset: 0.01818181818181818
Confusion Matrix on Validation dataset : [[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]

[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]]
Epoch-5 lr: 0.0001
[7, 9] loss: 2.307
[7, 18] loss: 2.302
[7, 27] loss: 2.304
[7, 36] loss: 2.306
Epoch-6 lr: 0.0001
[8, 9] loss: 2.303
[8, 18] loss: 2.305
[8, 27] loss: 2.304
[8, 36] loss: 2.307
Epoch-7 lr: 0.0001
[9, 9] loss: 2.304
[9, 18] loss: 2.305
[9, 27] loss: 2.305
[9, 36] loss: 2.306
Epoch-8 lr: 0.0001
[10, 9] loss: 2.307
[10, 18] loss: 2.303
[10, 27] loss: 2.305
[10, 36] loss: 2.304
Epoch-9 lr: 2e-05
[11, 9] loss: 2.305
[11, 18] loss: 2.306
[11, 27] loss: 2.302
[11, 36] loss: 2.305
Testing the accuracy after epoch 10 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 10 %
F1 score on val dataset: 0.01818181818181818
Confusion Matrix on Validation dataset : [[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]
[0 0 0 0 0 300 0 0 0 0]]

Epoch-10 lr: 2e-05
[12, 9] loss: 2.303
[12, 18] loss: 2.306
[12, 27] loss: 2.306
[12, 36] loss: 2.305
Epoch-11 lr: 2e-05
[13, 9] loss: 2.302
[13, 18] loss: 2.304
[13, 27] loss: 2.307
[13, 36] loss: 2.306
Epoch-12 lr: 2e-05
[14, 9] loss: 2.307
[14, 18] loss: 2.302
[14, 27] loss: 2.304
[14, 36] loss: 2.306
Epoch-13 lr: 2e-05
[15, 9] loss: 2.304
[15, 18] loss: 2.306
[15, 27] loss: 2.306
[15, 36] loss: 2.303
Epoch-14 lr: 2e-05
[16, 9] loss: 2.304
[16, 18] loss: 2.303
[16, 27] loss: 2.305
[16, 36] loss: 2.307
Testing the accuracy after epoch 15 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 10 %
F1 score on val dataset: 0.01818181818181818
Confusion Matrix on Validation dataset : [[0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]]
Epoch-15 lr: 2e-05
[17, 9] loss: 2.306
[17, 18] loss: 2.303
[17, 27] loss: 2.302
[17, 36] loss: 2.307

Epoch-16 lr: 2e-05
[18, 9] loss: 2.304
[18, 18] loss: 2.307
[18, 27] loss: 2.303
[18, 36] loss: 2.305
Epoch-17 lr: 2e-05
[19, 9] loss: 2.302
[19, 18] loss: 2.308
[19, 27] loss: 2.307
[19, 36] loss: 2.302
Epoch-18 lr: 2e-05
[20, 9] loss: 2.304
[20, 18] loss: 2.307
[20, 27] loss: 2.302
[20, 36] loss: 2.306
Epoch-19 lr: 4.000000000000001e-06
[21, 9] loss: 2.303
[21, 18] loss: 2.306
[21, 27] loss: 2.304
[21, 36] loss: 2.307
Testing the accuracy after epoch 20 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 10 %
F1 score on val dataset: 0.01818181818181818
Confusion Matrix on Validation dataset : [[0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]]
Epoch-20 lr: 4.000000000000001e-06

[22, 9] loss: 2.305
[22, 18] loss: 2.304
[22, 27] loss: 2.303
[22, 36] loss: 2.308
Epoch-21 lr: 4.000000000000001e-06
[23, 9] loss: 2.309
[23, 18] loss: 2.306
[23, 27] loss: 2.303

[23, 36] loss: 2.301
Epoch-22 lr: 4.000000000000001e-06
[24,   9] loss: 2.303
[24, 18] loss: 2.305
[24, 27] loss: 2.304
[24, 36] loss: 2.305
Epoch-23 lr: 4.000000000000001e-06
[25,   9] loss: 2.304
[25, 18] loss: 2.306
[25, 27] loss: 2.304
[25, 36] loss: 2.303
Epoch-24 lr: 4.000000000000001e-06
[26,   9] loss: 2.305
[26, 18] loss: 2.304
[26, 27] loss: 2.307
[26, 36] loss: 2.305
Testing the accuracy after epoch 25 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 10 %
F1 score on val dataset: 0.01818181818181818
Confusion Matrix on Validation dataset : [[0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]]
Epoch-25 lr: 4.000000000000001e-06
[27,   9] loss: 2.304
[27, 18] loss: 2.306
[27, 27] loss: 2.305
[27, 36] loss: 2.304
Epoch-26 lr: 4.000000000000001e-06
[28,   9] loss: 2.303
[28, 18] loss: 2.306
[28, 27] loss: 2.304
[28, 36] loss: 2.305
Epoch-27 lr: 4.000000000000001e-06
[29,   9] loss: 2.304
[29, 18] loss: 2.303
[29, 27] loss: 2.303

[29, 36] loss: 2.307
Epoch-28 lr: 4.000000000000001e-06
[30,   9] loss: 2.307
[30, 18] loss: 2.305
[30, 27] loss: 2.304
[30, 36] loss: 2.304
Epoch-29 lr: 8.000000000000002e-07
[31,   9] loss: 2.304
[31, 18] loss: 2.306
[31, 27] loss: 2.303
[31, 36] loss: 2.306
Testing the accuracy after epoch 30 on validation set for accuracy ...
****
Accuracy of the network on the 3000 validation images: 10 %
F1 score on val dataset: 0.01818181818181818
Confusion Matrix on Validation dataset : [[0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]
 [0   0   0   0   0 300   0   0   0   0]]
Epoch-30 lr: 8.000000000000002e-07
[32,   9] loss: 2.305
[32, 18] loss: 2.305
[32, 27] loss: 2.306
[32, 36] loss: 2.302
Epoch-31 lr: 8.000000000000002e-07
[33,   9] loss: 2.304
[33, 18] loss: 2.301

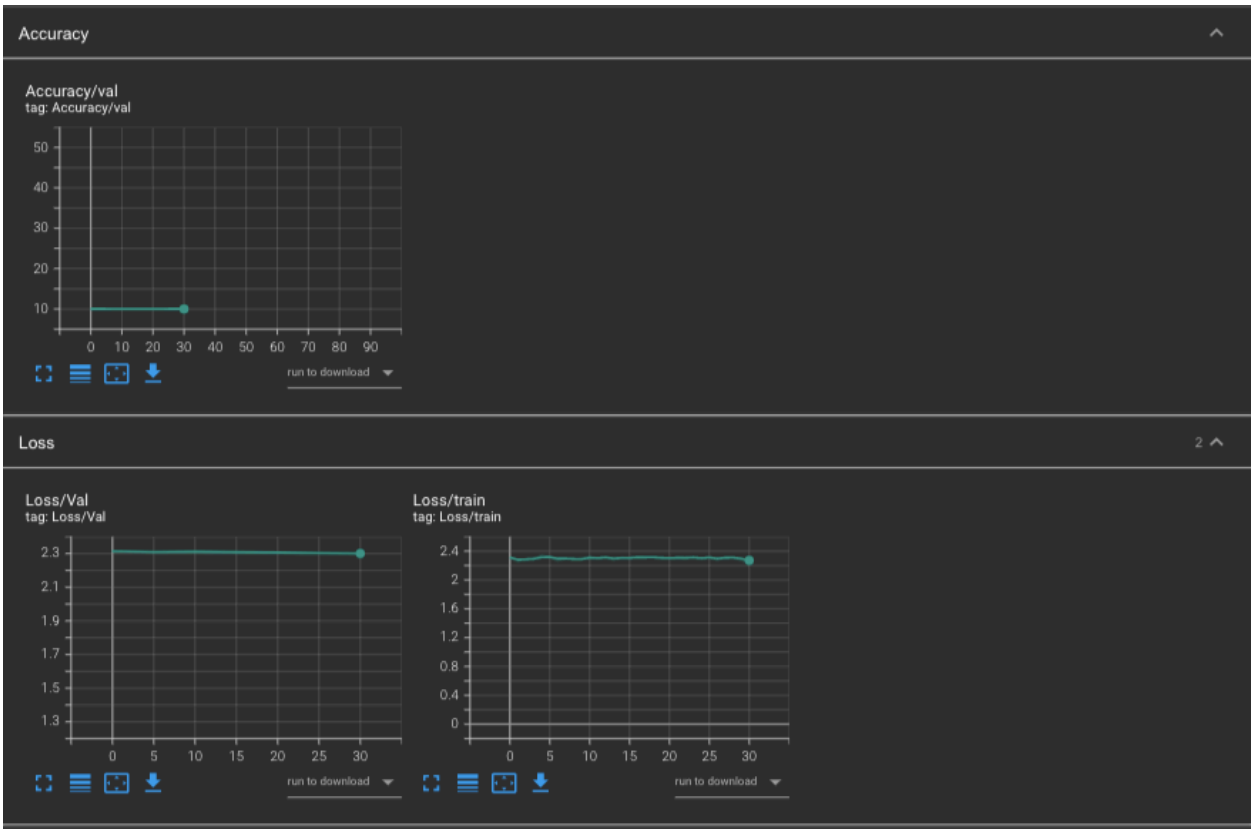


Fig4: Stopped the training after few iterations as it was not showing good progress in terms of loss or accuracy.

This experiment shows that decreasing learning rate to too small value may take longer time to converge. Stopped this experiment as it was taking longer time to finish to a decent value of accuracy.

3. Result with below setting.

Loss Function: Cross entropy loss
Optimizer: ADAM
Learning rate: 0.001(initially and then decrementing by 50 % after every 20 epochs)
Number of epochs: 100
Batch-size: 12

Observation: Seeing overfitting as training loss = 0.

****

Accuracy of the network on the 3000 validation images: 46 %
F1 score on val dataset: 0.4628030493145562

Finished Training

****

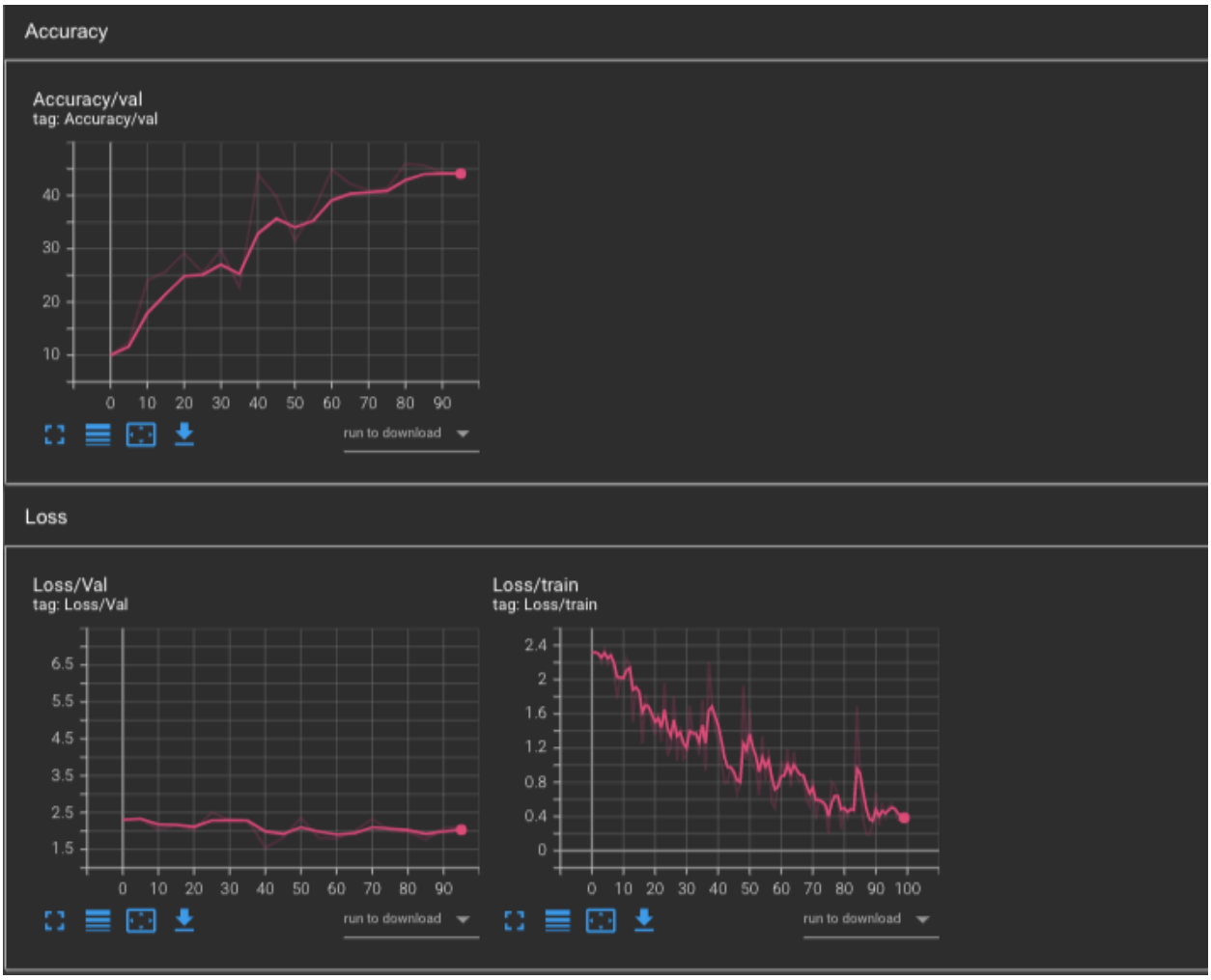
****

Accuracy of the network on the 5000 test images: 46 %
F1 score on test dataset: 0.4654998140801104
Confusion Matrix on Test dataset : [[341 27 18 7 17 4 4 3 45 34]
[20 173 12 49 51 76 30 69 8 12]
[9 12 317 10 3 7 3 6 28 105]
[5 39 7 140 92 80 34 74 15 14]
[14 40 4 49 231 51 53 42 8 8]
[5 60 4 57 61 128 92 81 5 7]
[12 33 9 34 59 56 228 51 1 17]
[1 49 4 81 51 91 41 173 2 7]
[51 11 11 17 2 0 1 6 340 61]
[40 4 89 19 4 8 6 6 64 260]]

4. Result with below setting.

Loss Function: Cross entropy loss
Optimizer: SGD
Learning rate: 0.1(initially and then decrementing by 50 % after every 20 epochs)
Number of epochs: 100
Batch-size: 128

Accuracy of the network on the 5000 test images: 44 %
F1 score on test dataset: 0.43376431770071266
Confusion Matrix on Test dataset : [[309 23 22 8 25 6 6 9 73 19]
[19 159 12 56 71 17 21 123 17 5]
[20 7 314 5 10 7 3 18 60 56]
[4 35 10 130 105 24 20 150 16 6]
[15 23 6 46 235 22 27 108 14 4]
[4 47 8 61 76 58 48 185 9 4]
[9 25 12 23 73 21 191 132 4 10]
[7 36 9 55 62 17 32 272 6 4]
[39 5 28 21 11 1 4 10 359 22]
[32 7 107 13 12 4 11 40 94 180]]



Analysis

- The final accuracy on test data is 47-48% approx.
- The final f1-score on test data is 0.4812.
- ADAM is working best for this best with 0.001 lr and then decaying lr.
- SGD is also giving the similar accuracy but with learning rate = 0.01. Also, SGD seems to be overfitting after 45 epochs also early stopping is needed to avoid the overfitting.
- The accuracy/f1-score should get better if we increase the depth of the hidden layers. As our network has only 2 CONV layers so accuracy seems to be coming around 46-48%.
- As per the confusion matrix, it shows that class labels with car(label=1) and truck(label=9) are high False positive. It seems that the network is confused with some of the common features like wheels or something and making a wrong prediction.