

```

Valid loss:0.634514, acc:0.813984
----Epoch 22 ---
Train loss:0.639194, acc:0.814900
Valid loss:0.634192, acc:0.814318
----Epoch 23 ---
Train loss:0.638308, acc:0.815379
Valid loss:0.633960, acc:0.817151
----Epoch 24 ---
Train loss:0.638197, acc:0.815525
Valid loss:0.633363, acc:0.814151
----Epoch 25 ---
Train loss:0.638373, acc:0.814587
Valid loss:0.633809, acc:0.813984
----Epoch 26 ---
Train loss:0.637841, acc:0.815129
Valid loss:0.634362, acc:0.814735
----Epoch 27 ---
Train loss:0.638196, acc:0.816316
Valid loss:0.634492, acc:0.813901
Early Stopping
Test set: x:(10000, 784), y:(10000, 10)
Loss:    0.654763
Acc:     0.803500

```

```

In [3]:      '/Users/xuchaoqi/Documents/GitHub/CSE253_PA2/
neuralnet.py'      = '/Users/xuchaoqi/Documents/GitHub/CSE253_PA2'
--- Config settings ---
layer_specs: [784, 50, 10]
activation: tanh
learning_rate: 1e-05
batch_size: 128
epochs: 95
early_stop: False
early_stop_epoch: 4
L2_penalty: 0.001
momentum: True
momentum_gamma: 0.9
--- Model summary ---
Layer 0: 784 inputs, 50 outputs
Layer 2: 50 inputs, 10 outputs
--- Training set: x: (48001, 784) y: (48001, 10) ---
--- Validation set: x: (11999, 784) y: (11999, 10) ---
----Epoch 0 ---
Train loss:0.939902, acc:0.726172
Valid loss:0.727108, acc:0.790899
----Epoch 1 ---
Train loss:0.702658, acc:0.796963
Valid loss:0.676789, acc:0.803150

```

```

----Epoch 2 ---
    Train loss:0.673081, acc:0.807004
    Valid loss:0.663113, acc:0.810151
----Epoch 3 ---
    Train loss:0.661488, acc:0.810316
    Valid loss:0.653333, acc:0.811651
----Epoch 4 ---
    Train loss:0.654690, acc:0.812754
    Valid loss:0.647025, acc:0.811234
----Epoch 5 ---
    Train loss:0.650428, acc:0.811754
    Valid loss:0.643598, acc:0.814818
----Epoch 6 ---
    Train loss:0.647542, acc:0.813629
    Valid loss:0.642082, acc:0.815068
----Epoch 7 ---
    Train loss:0.646080, acc:0.813379
    Valid loss:0.640622, acc:0.814651
----Epoch 8 ---
    Train loss:0.644202, acc:0.814921
    Valid loss:0.639922, acc:0.815651
----Epoch 9 ---
    Train loss:0.643565, acc:0.814608
    Valid loss:0.637686, acc:0.815818
----Epoch 10 ---
    Train loss:0.642264, acc:0.814650
    Valid loss:0.637518, acc:0.817735
----Epoch 11 ---
    Train loss:0.642160, acc:0.814941
    Valid loss:0.636210, acc:0.815818
----Epoch 12 ---
    Train loss:0.640845, acc:0.815296
    Valid loss:0.639987, acc:0.815068
----Epoch 13 ---
    Train loss:0.641061, acc:0.814296
    Valid loss:0.637141, acc:0.815985
----Epoch 14 ---
    Train loss:0.640082, acc:0.815129
    Valid loss:0.636365, acc:0.816151
----Epoch 15 ---
    Train loss:0.640403, acc:0.815587
    Valid loss:0.635079, acc:0.813734
----Epoch 16 ---
    Train loss:0.640052, acc:0.815566
    Valid loss:0.635705, acc:0.815401
----Epoch 17 ---
    Train loss:0.639703, acc:0.815421
    Valid loss:0.634978, acc:0.814985

```

```
-----Epoch 18 ---
    Train loss:0.639407, acc:0.815379
    Valid loss:0.636115, acc:0.816318
-----Epoch 19 ---
    Train loss:0.639612, acc:0.815858
    Valid loss:0.634934, acc:0.815735
-----Epoch 20 ---
    Train loss:0.639100, acc:0.815087
    Valid loss:0.634795, acc:0.815318
-----Epoch 21 ---
    Train loss:0.638534, acc:0.815587
    Valid loss:0.634514, acc:0.813984
-----Epoch 22 ---
    Train loss:0.639194, acc:0.814900
    Valid loss:0.634192, acc:0.814318
-----Epoch 23 ---
    Train loss:0.638308, acc:0.815379
    Valid loss:0.633960, acc:0.817151
-----Epoch 24 ---
    Train loss:0.638197, acc:0.815525
    Valid loss:0.633363, acc:0.814151
-----Epoch 25 ---
    Train loss:0.638373, acc:0.814587
    Valid loss:0.633809, acc:0.813984
-----Epoch 26 ---
    Train loss:0.637841, acc:0.815129
    Valid loss:0.634362, acc:0.814735
-----Epoch 27 ---
    Train loss:0.638196, acc:0.816316
    Valid loss:0.634492, acc:0.813901
-----Epoch 28 ---
    Train loss:0.637891, acc:0.814796
    Valid loss:0.634624, acc:0.814568
-----Epoch 29 ---
    Train loss:0.638317, acc:0.814816
    Valid loss:0.636447, acc:0.815568
-----Epoch 30 ---
    Train loss:0.638308, acc:0.814566
    Valid loss:0.633741, acc:0.814818
-----Epoch 31 ---
    Train loss:0.637908, acc:0.815608
    Valid loss:0.634089, acc:0.815651
-----Epoch 32 ---
    Train loss:0.638496, acc:0.815212
    Valid loss:0.634334, acc:0.816485
-----Epoch 33 ---
    Train loss:0.638056, acc:0.814233
    Valid loss:0.634291, acc:0.815985
```

```
-----Epoch 34 ----
    Train loss:0.637339, acc:0.814837
    Valid loss:0.634180, acc:0.816485
-----Epoch 35 ----
    Train loss:0.638142, acc:0.814712
    Valid loss:0.634560, acc:0.816985
-----Epoch 36 ----
    Train loss:0.637756, acc:0.813921
    Valid loss:0.633332, acc:0.813818
-----Epoch 37 ----
    Train loss:0.637815, acc:0.815837
    Valid loss:0.633279, acc:0.813984
-----Epoch 38 ----
    Train loss:0.637380, acc:0.815046
    Valid loss:0.634718, acc:0.817068
-----Epoch 39 ----
    Train loss:0.637575, acc:0.815608
    Valid loss:0.634045, acc:0.814235
-----Epoch 40 ----
    Train loss:0.637601, acc:0.815066
    Valid loss:0.633435, acc:0.816818
-----Epoch 41 ----
    Train loss:0.637325, acc:0.816316
    Valid loss:0.633791, acc:0.815985
-----Epoch 42 ----
    Train loss:0.637917, acc:0.815212
    Valid loss:0.632968, acc:0.815318
-----Epoch 43 ----
    Train loss:0.637317, acc:0.815254
    Valid loss:0.632751, acc:0.814318
-----Epoch 44 ----
    Train loss:0.637505, acc:0.815754
    Valid loss:0.632148, acc:0.813901
-----Epoch 45 ----
    Train loss:0.637144, acc:0.814796
    Valid loss:0.633684, acc:0.815651
-----Epoch 46 ----
    Train loss:0.637569, acc:0.815316
    Valid loss:0.634301, acc:0.815735
-----Epoch 47 ----
    Train loss:0.636877, acc:0.815212
    Valid loss:0.633262, acc:0.814151
-----Epoch 48 ----
    Train loss:0.637659, acc:0.814921
    Valid loss:0.632406, acc:0.815818
-----Epoch 49 ----
    Train loss:0.637083, acc:0.815275
    Valid loss:0.633562, acc:0.813818
```

```

-----Epoch 50 ----
      Train loss:0.636947, acc:0.815046
      Valid loss:0.632873, acc:0.815401
-----Epoch 51 ----
      Train loss:0.637172, acc:0.815337
      Valid loss:0.632342, acc:0.814151
-----Epoch 52 ----
      Train loss:0.636892, acc:0.814983
      Valid loss:0.632553, acc:0.814651
-----Epoch 53 ----
      Train loss:0.637221, acc:0.815212
      Valid loss:0.632697, acc:0.815068
-----Epoch 54 ----
      Train loss:0.637185, acc:0.815650
      Valid loss:0.632807, acc:0.814985
-----Epoch 55 ----
      Train loss:0.637076, acc:0.814921
      Valid loss:0.632434, acc:0.817485
-----Epoch 56 ----
      Train loss:0.637115, acc:0.815358
      Valid loss:0.633343, acc:0.816401
-----Epoch 57 ----
      Train loss:0.637104, acc:0.815004
      Valid loss:0.632876, acc:0.815068
-----Epoch 58 ----
      Train loss:0.637035, acc:0.814900
      Valid loss:0.633087, acc:0.815818
-----Epoch 59 ----
      Train loss:0.637142, acc:0.815608
      Valid loss:0.633156, acc:0.814318
-----Epoch 60 ----
      Train loss:0.637294, acc:0.814983
      Valid loss:0.632131, acc:0.815818
-----Epoch 61 ----
      Train loss:0.637051, acc:0.814921
      Valid loss:0.634728, acc:0.815318
-----Epoch 62 ----
      Train loss:0.636801, acc:0.815108
      Valid loss:0.633670, acc:0.813901
-----Epoch 63 ----
      Train loss:0.636924, acc:0.815379
      Valid loss:0.633476, acc:0.817318
-----Epoch 64 ----
      Train loss:0.637457, acc:0.815566
      Valid loss:0.631957, acc:0.813818
-----Epoch 65 ----
      Train loss:0.636913, acc:0.815691
      Valid loss:0.632069, acc:0.816068

```

```
-----Epoch 66 ---
      Train loss:0.637127, acc:0.815066
      Valid loss:0.632169, acc:0.816651
-----Epoch 67 ---
      Train loss:0.636401, acc:0.814879
      Valid loss:0.632849, acc:0.816401
-----Epoch 68 ---
      Train loss:0.637103, acc:0.815587
      Valid loss:0.633229, acc:0.815151
-----Epoch 69 ---
      Train loss:0.636945, acc:0.815671
      Valid loss:0.633838, acc:0.818152
-----Epoch 70 ---
      Train loss:0.636908, acc:0.815316
      Valid loss:0.633223, acc:0.816485
-----Epoch 71 ---
      Train loss:0.637491, acc:0.814962
      Valid loss:0.632533, acc:0.817318
-----Epoch 72 ---
      Train loss:0.637055, acc:0.815525
      Valid loss:0.632730, acc:0.814568
-----Epoch 73 ---
      Train loss:0.636621, acc:0.815191
      Valid loss:0.633984, acc:0.815568
-----Epoch 74 ---
      Train loss:0.637094, acc:0.814733
      Valid loss:0.634195, acc:0.816151
-----Epoch 75 ---
      Train loss:0.637162, acc:0.815025
      Valid loss:0.632888, acc:0.815568
-----Epoch 76 ---
      Train loss:0.637135, acc:0.814921
      Valid loss:0.632642, acc:0.816901
-----Epoch 77 ---
      Train loss:0.636684, acc:0.815379
      Valid loss:0.633369, acc:0.817151
-----Epoch 78 ---
      Train loss:0.637142, acc:0.815129
      Valid loss:0.633788, acc:0.814985
-----Epoch 79 ---
      Train loss:0.636740, acc:0.815421
      Valid loss:0.632666, acc:0.815235
-----Epoch 80 ---
      Train loss:0.637245, acc:0.815504
      Valid loss:0.632112, acc:0.815485
-----Epoch 81 ---
      Train loss:0.636455, acc:0.815087
      Valid loss:0.633359, acc:0.815735
```

```

----Epoch 82 ---
    Train loss:0.637340, acc:0.815441
    Valid loss:0.632965, acc:0.814151
----Epoch 83 ---
    Train loss:0.637192, acc:0.815441
    Valid loss:0.634088, acc:0.816318
----Epoch 84 ---
    Train loss:0.636724, acc:0.815504
    Valid loss:0.633134, acc:0.814735
----Epoch 85 ---
    Train loss:0.637077, acc:0.815171
    Valid loss:0.634645, acc:0.815651
----Epoch 86 ---
    Train loss:0.637214, acc:0.814587
    Valid loss:0.632975, acc:0.816151
----Epoch 87 ---
    Train loss:0.636845, acc:0.815296
    Valid loss:0.632601, acc:0.815985
----Epoch 88 ---
    Train loss:0.636879, acc:0.814671
    Valid loss:0.633826, acc:0.815818
----Epoch 89 ---
    Train loss:0.637515, acc:0.815504
    Valid loss:0.633465, acc:0.816068
----Epoch 90 ---
    Train loss:0.637061, acc:0.814087
    Valid loss:0.633159, acc:0.816068
----Epoch 91 ---
    Train loss:0.636902, acc:0.814941
    Valid loss:0.633277, acc:0.816985
----Epoch 92 ---
    Train loss:0.637228, acc:0.814504
    Valid loss:0.632685, acc:0.814985
----Epoch 93 ---
    Train loss:0.637073, acc:0.815462
    Valid loss:0.634286, acc:0.814068
----Epoch 94 ---
    Train loss:0.636950, acc:0.814837
    Valid loss:0.632613, acc:0.816735
Test set: x:(10000, 784), y:(10000, 10)
    Loss:    0.652838
    Acc:     0.803400

```

```

In [4]:      '/Users/xuchaoqi/Documents/GitHub/CSE253_PA2/
neuralnet.py'      = '/Users/xuchaoqi/Documents/GitHub/CSE253_PA2'
--- Config settings ---
layer_specs: [784, 50, 10]
activation: ReLU

```

```

learning_rate: 1e-05
batch_size: 128
epochs: 100
early_stop: True
early_stop_epoch: 4
L2_penalty: 0.0
momentum: True
momentum_gamma: 0.9
--- Model summary ---
Layer 0: 784 inputs, 50 outputs
Layer 2: 50 inputs, 10 outputs
--- Training set: x: (48001, 784) y: (48001, 10) ---
--- Validation set: x: (11999, 784) y: (11999, 10) ---
----Epoch 0 ---
      Train loss:0.778632, acc:0.734881
      Valid loss:0.553666, acc:0.807734
----Epoch 1 ---
      Train loss:0.521468, acc:0.816295
      Valid loss:0.487706, acc:0.828152
----Epoch 2 ---
      Train loss:0.475123, acc:0.832295
      Valid loss:0.458231, acc:0.838987
----Epoch 3 ---
      Train loss:0.450969, acc:0.839545
      Valid loss:0.437851, acc:0.847154
----Epoch 4 ---
      Train loss:0.431733, acc:0.848607
      Valid loss:0.423763, acc:0.852904
----Epoch 5 ---
      Train loss:0.418024, acc:0.851691
      Valid loss:0.414391, acc:0.854488
----Epoch 6 ---
      Train loss:0.406977, acc:0.856336
      Valid loss:0.405921, acc:0.857155
----Epoch 7 ---
      Train loss:0.397357, acc:0.859982
      Valid loss:0.401125, acc:0.858572
----Epoch 8 ---
      Train loss:0.389032, acc:0.862336
      Valid loss:0.393478, acc:0.859905
----Epoch 9 ---
      Train loss:0.381613, acc:0.865690
      Valid loss:0.386753, acc:0.863239
----Epoch 10 ---
      Train loss:0.374959, acc:0.867253
      Valid loss:0.384020, acc:0.865489
----Epoch 11 ---
      Train loss:0.369199, acc:0.870107

```



```

Valid loss:0.377787, acc:0.866906
----Epoch 12 ---
Train loss:0.363266, acc:0.872253
Valid loss:0.377666, acc:0.865739
----Epoch 13 ---
Train loss:0.358959, acc:0.874023
Valid loss:0.369894, acc:0.869656
----Epoch 14 ---
Train loss:0.353080, acc:0.876044
Valid loss:0.378503, acc:0.866322
----Epoch 15 ---
Train loss:0.348920, acc:0.877482
Valid loss:0.363428, acc:0.871406
----Epoch 16 ---
Train loss:0.344375, acc:0.878398
Valid loss:0.363516, acc:0.871989
----Epoch 17 ---
Train loss:0.340560, acc:0.880377
Valid loss:0.360333, acc:0.872656
----Epoch 18 ---
Train loss:0.337202, acc:0.881148
Valid loss:0.358279, acc:0.873739
----Epoch 19 ---
Train loss:0.333770, acc:0.882482
Valid loss:0.352684, acc:0.875323
----Epoch 20 ---
Train loss:0.330098, acc:0.883752
Valid loss:0.350487, acc:0.874490
----Epoch 21 ---
Train loss:0.325596, acc:0.885294
Valid loss:0.353073, acc:0.873906
----Epoch 22 ---
Train loss:0.322811, acc:0.886586
Valid loss:0.354675, acc:0.876073
----Epoch 23 ---
Train loss:0.319857, acc:0.887315
Valid loss:0.347872, acc:0.876073
----Epoch 24 ---
Train loss:0.317297, acc:0.888127
Valid loss:0.352073, acc:0.875073
----Epoch 25 ---
Train loss:0.314279, acc:0.889211
Valid loss:0.344973, acc:0.875156
----Epoch 26 ---
Train loss:0.311509, acc:0.889981
Valid loss:0.346984, acc:0.877240
----Epoch 27 ---
Train loss:0.307955, acc:0.891856

```

```

Valid loss:0.341546, acc:0.877406
----Epoch 28 ---
Train loss:0.306064, acc:0.891669
Valid loss:0.344761, acc:0.875906
----Epoch 29 ---
Train loss:0.303749, acc:0.892336
Valid loss:0.339360, acc:0.879990
----Epoch 30 ---
Train loss:0.301397, acc:0.893169
Valid loss:0.339523, acc:0.879990
----Epoch 31 ---
Train loss:0.298789, acc:0.895002
Valid loss:0.338511, acc:0.876740
----Epoch 32 ---
Train loss:0.296002, acc:0.895961
Valid loss:0.336192, acc:0.879407
----Epoch 33 ---
Train loss:0.293219, acc:0.896231
Valid loss:0.334972, acc:0.880490
----Epoch 34 ---
Train loss:0.291466, acc:0.897419
Valid loss:0.337047, acc:0.878073
----Epoch 35 ---
Train loss:0.290522, acc:0.897002
Valid loss:0.332576, acc:0.881573
----Epoch 36 ---
Train loss:0.288030, acc:0.897940
Valid loss:0.331978, acc:0.881407
----Epoch 37 ---
Train loss:0.285472, acc:0.898773
Valid loss:0.333842, acc:0.880657
----Epoch 38 ---
Train loss:0.283750, acc:0.899690
Valid loss:0.329664, acc:0.882157
----Epoch 39 ---
Train loss:0.281624, acc:0.900648
Valid loss:0.328137, acc:0.883990
----Epoch 40 ---
Train loss:0.279699, acc:0.902398
Valid loss:0.335394, acc:0.879657
----Epoch 41 ---
Train loss:0.278311, acc:0.901606
Valid loss:0.331381, acc:0.883240
----Epoch 42 ---
Train loss:0.275989, acc:0.902773
Valid loss:0.332015, acc:0.882657
----Epoch 43 ---
Train loss:0.275015, acc:0.903315

```

```
Valid loss:0.325924, acc:0.884740
----Epoch 44 ---
Train loss:0.272838, acc:0.902815
Valid loss:0.328861, acc:0.882907
----Epoch 45 ---
Train loss:0.270546, acc:0.904814
Valid loss:0.325043, acc:0.884407
----Epoch 46 ---
Train loss:0.269235, acc:0.904419
Valid loss:0.325086, acc:0.884574
----Epoch 47 ---
Train loss:0.267551, acc:0.904814
Valid loss:0.325298, acc:0.884490
----Epoch 48 ---
Train loss:0.265955, acc:0.905294
Valid loss:0.327435, acc:0.883907
Early Stopping
Test set: x:(10000, 784), y:(10000, 10)
Loss:    0.363365
Acc:     0.873900
```

In [5]: