```
        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]:          '/Users/xuchaoqi/Documents/GitHub/CSE253_PA2/
neuralnet.py'       ='/Users/xuchaoqi/Documents/GitHub/CSE253_PA2'
--- Config settings ---
layer_specs: [784, 100, 10]
activation: tanh
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, 100 outputs
Layer 2: 100 inputs, 10 outputs
--- Training set: x: (48001, 784) y: (48001, 10) ---
--- Validation set: x: (11999, 784) y: (11999, 10) ---
----Epoch 0 ---
        Train loss:0.820688, acc:0.727402
        Valid loss:0.583245, acc:0.805984
----Epoch 1 ---
        Train loss:0.548715, acc:0.812671
        Valid loss:0.504285, acc:0.828486
----Epoch 2 ---
        Train loss:0.492530, acc:0.829962
        Valid loss:0.469599, acc:0.838237
----Epoch 3 ---
        Train loss:0.461777, acc:0.839566
        Valid loss:0.447829, acc:0.845987
----Epoch 4 ---
        Train loss:0.440890, acc:0.846316
        Valid loss:0.430066, acc:0.850738
----Epoch 5 ---
        Train loss:0.424967, acc:0.852399
        Valid loss:0.418385, acc:0.855488
----Epoch 6 ---
        Train loss:0.412436, acc:0.857211
        Valid loss:0.410871, acc:0.857321
----Epoch 7 ---
        Train loss:0.402002, acc:0.859357
        Valid loss:0.399846, acc:0.859822
----Epoch 8 ---
        Train loss:0.392869, acc:0.862711
        Valid loss:0.393293, acc:0.861738
----Epoch 9 ---
        Train loss:0.385054, acc:0.864440
        Valid loss:0.388072, acc:0.862905
----Epoch 10 ---
        Train loss:0.377924, acc:0.867107
        Valid loss:0.382405, acc:0.863655
----Epoch 11 ---
        Train loss:0.371264, acc:0.869253
        Valid loss:0.377758, acc:0.866572
----Epoch 12 ---
        Train loss:0.365698, acc:0.871586
        Valid loss:0.374276, acc:0.865822
----Epoch 13 ---
        Train loss:0.360011, acc:0.873669
        Valid loss:0.371819, acc:0.867989
----Epoch 14 ---
```

```
        Train loss:0.355159, acc:0.874898
        Valid loss:0.367578, acc:0.868906
----Epoch 15 ---
        Train loss:0.349438, acc:0.877253
        Valid loss:0.365700, acc:0.869322
----Epoch 16 ---
        Train loss:0.345722, acc:0.878857
        Valid loss:0.361200, acc:0.869406
----Epoch 17 ---
        Train loss:0.341218, acc:0.880961
        Valid loss:0.358592, acc:0.871823
----Epoch 18 ---
        Train loss:0.337394, acc:0.881565
        Valid loss:0.354444, acc:0.872989
----Epoch 19 ---
        Train loss:0.333518, acc:0.883086
        Valid loss:0.351490, acc:0.873906
----Epoch 20 ---
        Train loss:0.329477, acc:0.885065
        Valid loss:0.352424, acc:0.873489
----Epoch 21 ---
        Train loss:0.326002, acc:0.886023
        Valid loss:0.349504, acc:0.876573
----Epoch 22 ---
        Train loss:0.322841, acc:0.887607
        Valid loss:0.345907, acc:0.877906
----Epoch 23 ---
        Train loss:0.319253, acc:0.887419
        Valid loss:0.344452, acc:0.876656
----Epoch 24 ---
        Train loss:0.315792, acc:0.889044
        Valid loss:0.343736, acc:0.876573
----Epoch 25 ---
        Train loss:0.312700, acc:0.890231
        Valid loss:0.343380, acc:0.878323
----Epoch 26 ---
        Train loss:0.309351, acc:0.891648
        Valid loss:0.340113, acc:0.877990
----Epoch 27 ---
        Train loss:0.307230, acc:0.892211
        Valid loss:0.337736, acc:0.879657
----Epoch 28 ---
        Train loss:0.304172, acc:0.893856
        Valid loss:0.335051, acc:0.880490
----Epoch 29 ---
        Train loss:0.301143, acc:0.894169
        Valid loss:0.335810, acc:0.879323
----Epoch 30 ---
```

```
        Train loss:0.299074, acc:0.895377
        Valid loss:0.334061, acc:0.879823
----Epoch 31 ---
        Train loss:0.296160, acc:0.895961
        Valid loss:0.331783, acc:0.880990
----Epoch 32 ---
        Train loss:0.293782, acc:0.897065
        Valid loss:0.333528, acc:0.880073
----Epoch 33 ---
        Train loss:0.291275, acc:0.897710
        Valid loss:0.330537, acc:0.881240
----Epoch 34 ---
        Train loss:0.288593, acc:0.898606
        Valid loss:0.330373, acc:0.882907
----Epoch 35 ---
        Train loss:0.286466, acc:0.899710
        Valid loss:0.330731, acc:0.880407
----Epoch 36 ---
        Train loss:0.283691, acc:0.900481
        Valid loss:0.326278, acc:0.883407
----Epoch 37 ---
        Train loss:0.281382, acc:0.902065
        Valid loss:0.325888, acc:0.883407
----Epoch 38 ---
        Train loss:0.279608, acc:0.902356
        Valid loss:0.324564, acc:0.882907
----Epoch 39 ---
        Train loss:0.277209, acc:0.904023
        Valid loss:0.327981, acc:0.882990
----Epoch 40 ---
        Train loss:0.275437, acc:0.903398
        Valid loss:0.322113, acc:0.884657
----Epoch 41 ---
        Train loss:0.273165, acc:0.904481
        Valid loss:0.322159, acc:0.883740
----Epoch 42 ---
        Train loss:0.270925, acc:0.905231
        Valid loss:0.321103, acc:0.885490
----Epoch 43 ---
        Train loss:0.269005, acc:0.905689
        Valid loss:0.320341, acc:0.884740
----Epoch 44 ---
        Train loss:0.267293, acc:0.906752
        Valid loss:0.319999, acc:0.885324
----Epoch 45 ---
        Train loss:0.265152, acc:0.907335
        Valid loss:0.320533, acc:0.885074
----Epoch 46 ---
```

```
        Train loss:0.263342, acc:0.908106
        Valid loss:0.318185, acc:0.886074
----Epoch 47 ---
        Train loss:0.261433, acc:0.908627
        Valid loss:0.320936, acc:0.884907
----Epoch 48 ---
        Train loss:0.259399, acc:0.909939
        Valid loss:0.316279, acc:0.887657
----Epoch 49 ---
        Train loss:0.257771, acc:0.910752
        Valid loss:0.315538, acc:0.886741
----Epoch 50 ---
        Train loss:0.255997, acc:0.910669
        Valid loss:0.317125, acc:0.886241
----Epoch 51 ---
        Train loss:0.254422, acc:0.910835
        Valid loss:0.314352, acc:0.887741
----Epoch 52 ---
        Train loss:0.252316, acc:0.912148
        Valid loss:0.315557, acc:0.887324
----Epoch 53 ---
        Train loss:0.250779, acc:0.912773
        Valid loss:0.316004, acc:0.885574
----Epoch 54 ---
        Train loss:0.248597, acc:0.913585
        Valid loss:0.314693, acc:0.888157
----Epoch 55 ---
        Train loss:0.247244, acc:0.913773
        Valid loss:0.312557, acc:0.887241
----Epoch 56 ---
        Train loss:0.245405, acc:0.915439
        Valid loss:0.313264, acc:0.889241
----Epoch 57 ---
        Train loss:0.243709, acc:0.915793
        Valid loss:0.311356, acc:0.889324
----Epoch 58 ---
        Train loss:0.242679, acc:0.915273
        Valid loss:0.312372, acc:0.887741
----Epoch 59 ---
        Train loss:0.240271, acc:0.916939
        Valid loss:0.311136, acc:0.887991
----Epoch 60 ---
        Train loss:0.239597, acc:0.917481
        Valid loss:0.310193, acc:0.890408
----Epoch 61 ---
        Train loss:0.237751, acc:0.917731
        Valid loss:0.311454, acc:0.889324
----Epoch 62 ---
```

```
        Train loss:0.235760, acc:0.919043
        Valid loss:0.310093, acc:0.887657
----Epoch 63 ---
        Train loss:0.234498, acc:0.918877
        Valid loss:0.310907, acc:0.889324
----Epoch 64 ---
        Train loss:0.232744, acc:0.919564
        Valid loss:0.308289, acc:0.890241
----Epoch 65 ---
        Train loss:0.231626, acc:0.920293
        Valid loss:0.308122, acc:0.889491
----Epoch 66 ---
        Train loss:0.230096, acc:0.920772
        Valid loss:0.309898, acc:0.889324
----Epoch 67 ---
        Train loss:0.228423, acc:0.921252
        Valid loss:0.313259, acc:0.886741
----Epoch 68 ---
        Train loss:0.227284, acc:0.921252
        Valid loss:0.307068, acc:0.890241
----Epoch 69 ---
        Train loss:0.225466, acc:0.922189
        Valid loss:0.307193, acc:0.890491
----Epoch 70 ---
        Train loss:0.223966, acc:0.922939
        Valid loss:0.308303, acc:0.889907
----Epoch 71 ---
        Train loss:0.222868, acc:0.922877
        Valid loss:0.307422, acc:0.889824
----Epoch 72 ---
        Train loss:0.221322, acc:0.923293
        Valid loss:0.307597, acc:0.888491
----Epoch 73 ---
        Train loss:0.219758, acc:0.924106
        Valid loss:0.305742, acc:0.890408
----Epoch 74 ---
        Train loss:0.218485, acc:0.924647
        Valid loss:0.306385, acc:0.889324
----Epoch 75 ---
        Train loss:0.217555, acc:0.924606
        Valid loss:0.305635, acc:0.890824
----Epoch 76 ---
        Train loss:0.215904, acc:0.926043
        Valid loss:0.306014, acc:0.890158
----Epoch 77 ---
        Train loss:0.214413, acc:0.925689
        Valid loss:0.305163, acc:0.891241
----Epoch 78 ---
```

```
        Train loss:0.212875, acc:0.926189
        Valid loss:0.306666, acc:0.888491
----Epoch 79 ---
        Train loss:0.211965, acc:0.927752
        Valid loss:0.310936, acc:0.885324
----Epoch 80 ---
        Train loss:0.210983, acc:0.926981
        Valid loss:0.304690, acc:0.891658
----Epoch 81 ---
        Train loss:0.209375, acc:0.928856
        Valid loss:0.305801, acc:0.887574
----Epoch 82 ---
        Train loss:0.208180, acc:0.928856
        Valid loss:0.305936, acc:0.889907
----Epoch 83 ---
        Train loss:0.206999, acc:0.928939
        Valid loss:0.305551, acc:0.891074
----Epoch 84 ---
        Train loss:0.205408, acc:0.928918
        Valid loss:0.306003, acc:0.889824
----Epoch 85 ---
        Train loss:0.203957, acc:0.929939
        Valid loss:0.303061, acc:0.891658
----Epoch 86 ---
        Train loss:0.203016, acc:0.930689
        Valid loss:0.304651, acc:0.891408
----Epoch 87 ---
        Train loss:0.201972, acc:0.930668
        Valid loss:0.305649, acc:0.890408
----Epoch 88 ---
        Train loss:0.200493, acc:0.931626
        Valid loss:0.305113, acc:0.890324
----Epoch 89 ---
        Train loss:0.199415, acc:0.932981
        Valid loss:0.304419, acc:0.891074
----Epoch 90 ---
        Train loss:0.198140, acc:0.932710
        Valid loss:0.305065, acc:0.891074
----Epoch 91 ---
        Train loss:0.196724, acc:0.933522
        Valid loss:0.309486, acc:0.891074
----Epoch 92 ---
        Train loss:0.195904, acc:0.933439
        Valid loss:0.304682, acc:0.890241
----Epoch 93 ---
        Train loss:0.194395, acc:0.934126
        Valid loss:0.304266, acc:0.891074
----Epoch 94 ---
```

```
        Train loss:0.193845, acc:0.934126
        Valid loss:0.309529, acc:0.885407
----Epoch 95 ---
        Train loss:0.192164, acc:0.935126
        Valid loss:0.303419, acc:0.893574
----Epoch 96 ---
        Train loss:0.191244, acc:0.935918
        Valid loss:0.303822, acc:0.891324
----Epoch 97 ---
        Train loss:0.189799, acc:0.935481
        Valid loss:0.303271, acc:0.891158
----Epoch 98 ---
        Train loss:0.188325, acc:0.936085
        Valid loss:0.303601, acc:0.891324
----Epoch 99 ---
        Train loss:0.187391, acc:0.936876
        Valid loss:0.305178, acc:0.891908
Test set: x:(10000, 784), y:(10000, 10)
        Loss:    0.336503
        Acc:     0.880900
```

In [**6**]: