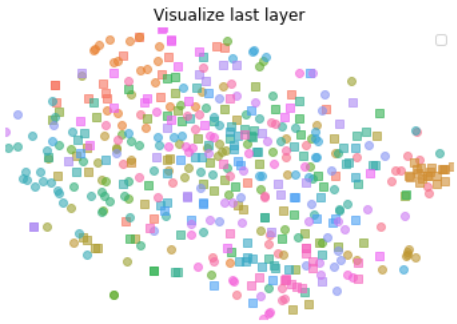
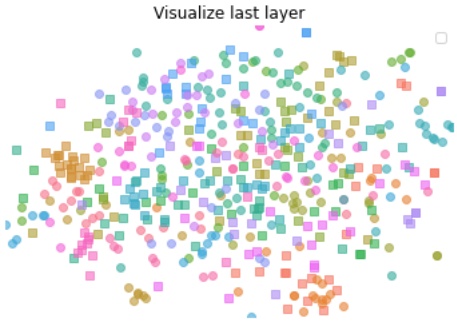
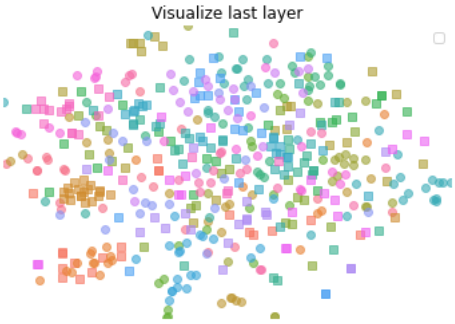
```
Epoch: 1
[ Train | 001/100 ] loss = 2.64559, acc = 0.22160
squeezing data
[ Test | 001/100 ] loss = 2.62661, acc = 0.23572
visualizing
```



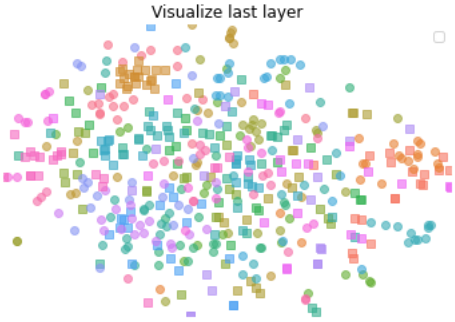

```
[Time cost | 001/100]: 883.8875s
Epoch: 2
[ Train | 002/100 ] loss = 2.19046, acc = 0.33082
squeezing data
[ Test | 002/100 ] loss = 2.27059, acc = 0.32146
visualizing
```



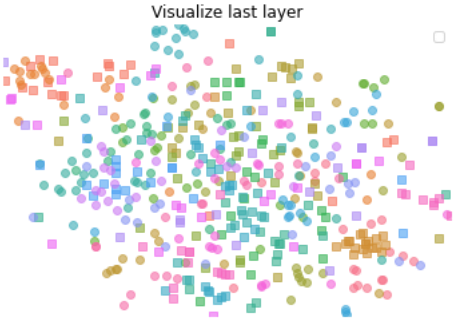

```
[Time cost | 002/100]: 940.9986s
Epoch: 3
[ Train | 003/100 ] loss = 1.97381, acc = 0.38828
squeezing data
[ Test | 003/100 ] loss = 2.95217, acc = 0.22932
[Time cost | 003/100]: 807.2090s
Epoch: 4
[ Train | 004/100 ] loss = 1.81311, acc = 0.43846
squeezing data
[ Test | 004/100 ] loss = 2.66428, acc = 0.29182
[Time cost | 004/100]: 782.2378s
Epoch: 5
[ Train | 005/100 ] loss = 1.68156, acc = 0.47628
squeezing data
[ Test | 005/100 ] loss = 2.38353, acc = 0.33762 visualizing
```



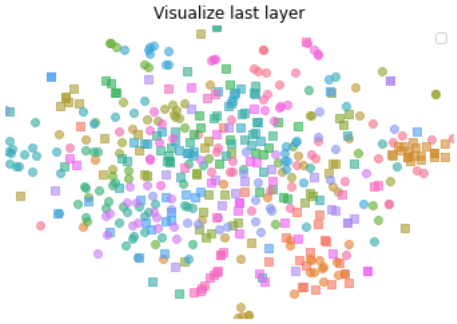
[Time cost | 005/100]: 810.2462s
Epoch: 6
[Train | 006/100] loss = 1.58544, acc = 0.50833
squeezing data
[Test | 006/100] loss = 2.14612, acc = 0.37668 visualizing



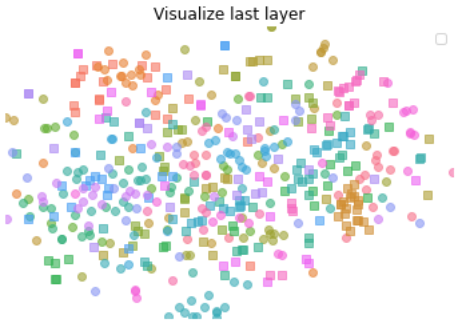
[Time cost | 006/100]: 805.1788s
Epoch: 7
[Train | 007/100] loss = 1.49446, acc = 0.52975
squeezing data
[Test | 007/100] loss = 2.15571, acc = 0.38564
visualizing



[Time cost | 007/100]: 817.7637s
Epoch: 8
[Train | 008/100] loss = 1.41291, acc = 0.55073
squeezing data
[Test | 008/100] loss = 1.95562, acc = 0.43629 visualizing

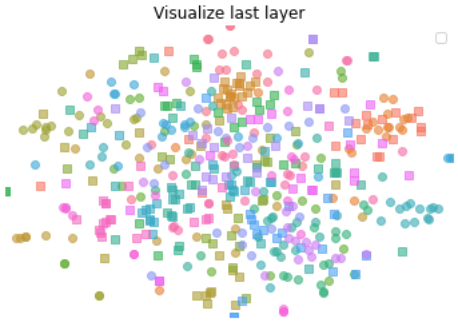


[Time cost | 008/100]: 849.5514s
Epoch: 9
[Train | 009/100] loss = 1.30748, acc = 0.58840
squeezing data
[Test | 009/100] loss = 1.89220, acc = 0.45057 visualizing

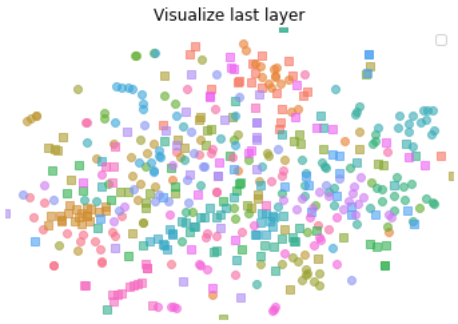


[Time cost | 009/100]: 805.8404s
Epoch: 10
[Train | 010/100] loss = 1.24923, acc = 0.60586
squeezing data
[Test | 010/100] loss = 1.97459, acc = 0.43939
[Time cost | 010/100]: 824.2583s
Epoch: 11
[Train | 011/100] loss = 1.17780, acc = 0.62317
squeezing data
[Test | 011/100] loss = 1.98006, acc = 0.42470
[Time cost | 011/100]: 808.2765s
Epoch: 12
[Train | 012/100] loss = 1.11014, acc = 0.64523
squeezing data
[Test | 012/100] loss = 2.14479, acc = 0.42652
[Time cost | 012/100]: 821.5544s
Epoch: 13
[Train | 013/100] loss = 1.03754, acc = 0.66961
squeezing data
[Test | 013/100] loss = 1.99140, acc = 0.43427
[Time cost | 013/100]: 778.6045s
Epoch: 14

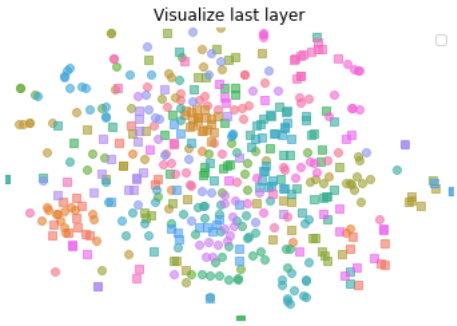
[Train | 014/100] loss = 0.98537, acc = 0.68427
squeezing data
[Test | 014/100] loss = 1.94744, acc = 0.40874
[Time cost | 014/100]: 772.4053s
Epoch: 15
[Train | 015/100] loss = 0.92206, acc = 0.70392
squeezing data
[Test | 015/100] loss = 1.87318, acc = 0.47124 visualizing



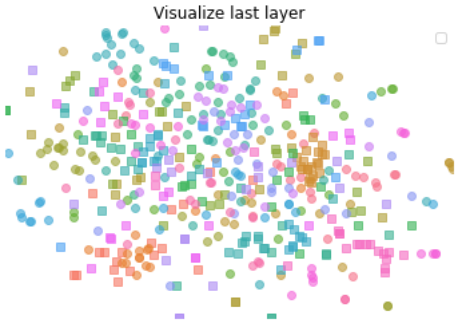
[Time cost | 015/100]: 805.2428s Epoch: 16
[Train | 016/100] loss = 0.86903, acc = 0.72445
squeezing data
[Test | 016/100] loss = 1.92525, acc = 0.47165
visualizing



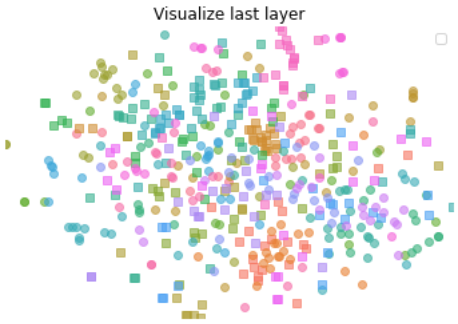
[Time cost | 016/100]: 795.8998s Epoch: 17
[Train | 017/100] loss = 0.79131, acc = 0.74983
squeezing data
[Test | 017/100] loss = 2.17223, acc = 0.44531
[Time cost | 017/100]: 759.0175s Epoch: 18
[Train | 018/100] loss = 0.75041, acc = 0.76309
squeezing data
[Test | 018/100] loss = 1.94982, acc = 0.50391 visualizing



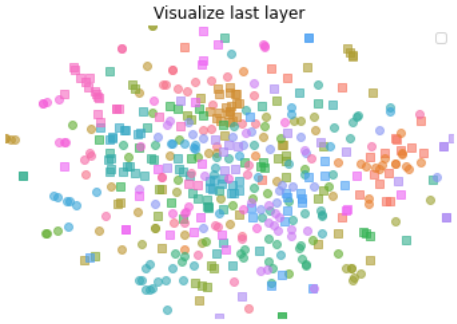

```
[Time cost | 018/100]: 768.6549s
Epoch: 19
[ Train | 019/100 ] loss = 0.68826, acc = 0.77764
squeezing data
[ Test | 019/100 ] loss = 2.03265, acc = 0.48182
[Time cost | 019/100]: 873.7185s
Epoch: 20
[ Train | 020/100 ] loss = 0.64796, acc = 0.78971
squeezing data
[ Test | 020/100 ] loss = 2.01073, acc = 0.49273
[Time cost | 020/100]: 876.3425s
Epoch: 21
[ Train | 021/100 ] loss = 0.59979, acc = 0.80760
squeezing data
[ Test | 021/100 ] loss = 2.27553, acc = 0.46969
[Time cost | 021/100]: 793.0388s
Epoch: 22
[ Train | 022/100 ] loss = 0.57457, acc = 0.81597
squeezing data
[ Test | 022/100 ] loss = 2.20797, acc = 0.46437
[Time cost | 022/100]: 780.8750s
Epoch: 23
[ Train | 023/100 ] loss = 0.57782, acc = 0.81713
squeezing data
[ Test | 023/100 ] loss = 2.05295, acc = 0.48296
[Time cost | 023/100]: 809.0315s
Epoch: 24
[ Train | 024/100 ] loss = 0.53646, acc = 0.82888
squeezing data
[ Test | 024/100 ] loss = 1.98735, acc = 0.48976
[Time cost | 024/100]: 809.4189s
Epoch: 25
[ Train | 025/100 ] loss = 0.45971, acc = 0.85625
squeezing data
[ Test | 025/100 ] loss = 2.08892, acc = 0.50546 visualizing
```



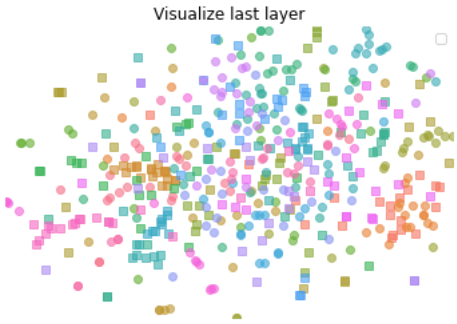
[Time cost | 025/100]: 801.1557s
Epoch: 26
[Train | 026/100] loss = 0.47010, acc = 0.85394
squeezing data
[Test | 026/100] loss = 2.17511, acc = 0.51583 visualizing



[Time cost | 026/100]: 798.2253s
Epoch: 27
[Train | 027/100] loss = 0.41104, acc = 0.87299
squeezing data
[Test | 027/100] loss = 2.20734, acc = 0.48532
[Time cost | 027/100]: 865.0295s
Epoch: 28
[Train | 028/100] loss = 0.39279, acc = 0.87789
squeezing data
[Test | 028/100] loss = 2.21434, acc = 0.47394
[Time cost | 028/100]: 894.7861s
Epoch: 29
[Train | 029/100] loss = 0.35688, acc = 0.89173
squeezing data
[Test | 029/100] loss = 2.37878, acc = 0.47777
[Time cost | 029/100]: 841.6865s
Epoch: 30
[Train | 030/100] loss = 0.38535, acc = 0.87608
squeezing data
[Test | 030/100] loss = 2.12970, acc = 0.52128 visualizing

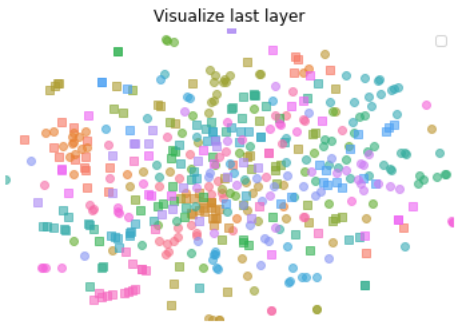


[Time cost | 030/100]: 831.7318s
Epoch: 31
[Train | 031/100] loss = 0.36248, acc = 0.88420
squeezing data
[Test | 031/100] loss = 2.12194, acc = 0.52324 visualizing

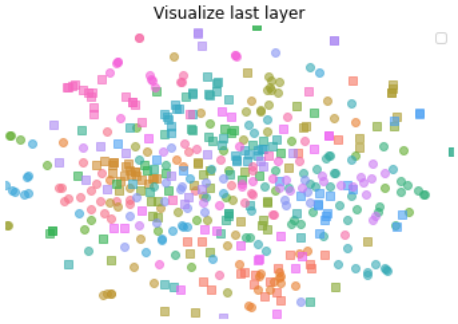


[Time cost | 031/100]: 796.9128s
Epoch: 32
[Train | 032/100] loss = 0.32950, acc = 0.89932
squeezing data
[Test | 032/100] loss = 2.43699, acc = 0.49333
[Time cost | 032/100]: 789.0220s
Epoch: 33
[Train | 033/100] loss = 0.33745, acc = 0.89179
squeezing data
[Test | 033/100] loss = 2.26090, acc = 0.51758
[Time cost | 033/100]: 798.1073s
Epoch: 34
[Train | 034/100] loss = 0.30015, acc = 0.90238
squeezing data
[Test | 034/100] loss = 2.18097, acc = 0.50977
[Time cost | 034/100]: 785.6654s
Epoch: 35
[Train | 035/100] loss = 0.29126, acc = 0.91051
squeezing data
[Test | 035/100] loss = 2.40740, acc = 0.47178
[Time cost | 035/100]: 785.5645s

```
Epoch: 36
[ Train | 036/100 ] loss = 0.29318, acc = 0.90687
squeezing data
[ Test | 036/100 ] loss = 2.26557, acc = 0.48633
[Time cost | 036/100]: 790.9252s
Epoch: 37
[ Train | 037/100 ] loss = 0.25266, acc = 0.92430
squeezing data
[ Test | 037/100 ] loss = 2.20232, acc = 0.52270
[Time cost | 037/100]: 816.6476s
Epoch: 38
[ Train | 038/100 ] loss = 0.26478, acc = 0.91731
squeezing data
[ Test | 038/100 ] loss = 2.39077, acc = 0.50101
[Time cost | 038/100]: 784.5520s
Epoch: 39
[ Train | 039/100 ] loss = 0.28234, acc = 0.90976
squeezing data
[ Test | 039/100 ] loss = 2.15166, acc = 0.52559 visualizing
```



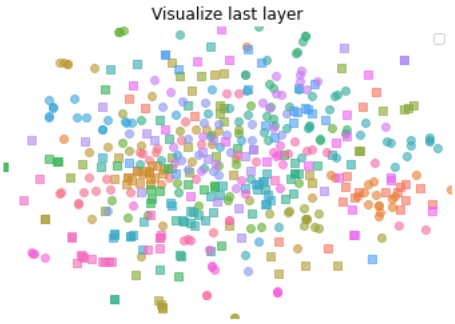

```
[Time cost | 039/100]: 786.3632s
Epoch: 40
[ Train | 040/100 ] loss = 0.22007, acc = 0.92975
squeezing data
[ Test | 040/100 ] loss = 2.24422, acc = 0.52054
[Time cost | 040/100]: 779.1450s
Epoch: 41
[ Train | 041/100 ] loss = 0.23667, acc = 0.92420
squeezing data
[ Test | 041/100 ] loss = 2.20548, acc = 0.52856 visualizing
```



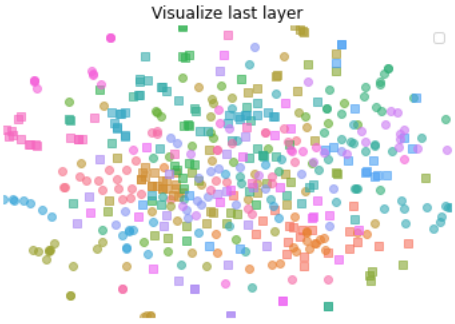

```
[Time cost | 041/100]: 775.6000s
Epoch: 42
[ Train | 042/100 ] loss = 0.25630, acc = 0.92116
squeezing data
[ Test | 042/100 ] loss = 2.43915, acc = 0.51852
[Time cost | 042/100]: 793.4128s
Epoch: 43
[ Train | 043/100 ] loss = 0.21064, acc = 0.93606
squeezing data
[ Test | 043/100 ] loss = 2.29078, acc = 0.51542
[Time cost | 043/100]: 839.3237s
Epoch: 44
[ Train | 044/100 ] loss = 0.19388, acc = 0.93998
squeezing data
[ Test | 044/100 ] loss = 2.61395, acc = 0.49529
[Time cost | 044/100]: 779.8460s
Epoch: 45
[ Train | 045/100 ] loss = 0.19108, acc = 0.94277
squeezing data
[ Test | 045/100 ] loss = 2.55500, acc = 0.48478
[Time cost | 045/100]: 845.0452s
Epoch: 46
[ Train | 046/100 ] loss = 0.25216, acc = 0.92002
squeezing data
[ Test | 046/100 ] loss = 2.46336, acc = 0.50741
[Time cost | 046/100]: 845.1104s
Epoch: 47
[ Train | 047/100 ] loss = 0.17794, acc = 0.94562
squeezing data
[ Test | 047/100 ] loss = 2.36581, acc = 0.52128
[Time cost | 047/100]: 813.6679s
Epoch: 48
[ Train | 048/100 ] loss = 0.16849, acc = 0.94727
squeezing data
[ Test | 048/100 ] loss = 2.43736, acc = 0.50795
```

```
[Time cost | 048/100]: 806.9272s
Epoch: 49
[ Train | 049/100 ] loss = 0.17345, acc = 0.94678
squeezing data
[ Test | 049/100 ] loss = 2.58558, acc = 0.52849
[Time cost | 049/100]: 798.2962s
Epoch: 50
[ Train | 050/100 ] loss = 0.21942, acc = 0.93064
squeezing data
[ Test | 050/100 ] loss = 2.42392, acc = 0.52189
[Time cost | 050/100]: 783.2388s
Epoch: 51
[ Train | 051/100 ] loss = 0.17615, acc = 0.94523
squeezing data
[ Test | 051/100 ] loss = 2.57533, acc = 0.50411
[Time cost | 051/100]: 795.5900s
Epoch: 52
[ Train | 052/100 ] loss = 0.17005, acc = 0.94697
squeezing data
[ Test | 052/100 ] loss = 2.80436, acc = 0.51111
[Time cost | 052/100]: 834.9593s
Epoch: 53
[ Train | 053/100 ] loss = 0.20099, acc = 0.93928
squeezing data
[ Test | 053/100 ] loss = 2.50626, acc = 0.50216
[Time cost | 053/100]: 821.3839s
Epoch: 54
[ Train | 054/100 ] loss = 0.19883, acc = 0.93922
squeezing data
[ Test | 054/100 ] loss = 2.49773, acc = 0.48963
[Time cost | 054/100]: 856.8529s
Epoch: 55
[ Train | 055/100 ] loss = 0.22781, acc = 0.93239
squeezing data
[ Test | 055/100 ] loss = 2.71554, acc = 0.50330
[Time cost | 055/100]: 848.5419s
Epoch: 56
[ Train | 056/100 ] loss = 0.18909, acc = 0.94356
squeezing data
```

[Test | 056/100] loss = 2.38596, acc = 0.53361 visualizing

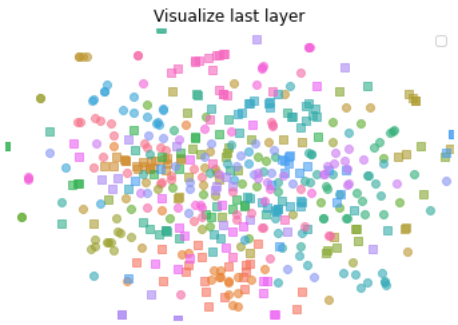


[Time cost | 056/100]: 833.8460s
Epoch: 57
[Train | 057/100] loss = 0.17142, acc = 0.94855
squeezing data
[Test | 057/100] loss = 2.43785, acc = 0.51717
[Time cost | 057/100]: 799.1309s
Epoch: 58
[Train | 058/100] loss = 0.18231, acc = 0.94361
squeezing data
[Test | 058/100] loss = 2.67089, acc = 0.50350
[Time cost | 058/100]: 794.3386s
Epoch: 59
[Train | 059/100] loss = 0.17987, acc = 0.94380
squeezing data
[Test | 059/100] loss = 2.52309, acc = 0.51933
[Time cost | 059/100]: 795.2907s
Epoch: 60
[Train | 060/100] loss = 0.16192, acc = 0.95241
squeezing data
[Test | 060/100] loss = 2.37549, acc = 0.54708 visualizing



[Time cost | 060/100]: 924.4600s
Epoch: 61
[Train | 061/100] loss = 0.16056, acc = 0.95232
squeezing data
[Test | 061/100] loss = 2.66718, acc = 0.50855

[Time cost | 061/100]: 986.0419s
Epoch: 62
[Train | 062/100] loss = 0.14208, acc = 0.95552
squeezing data
[Test | 062/100] loss = 2.40867, acc = 0.53926
[Time cost | 062/100]: 830.0875s
Epoch: 63
[Train | 063/100] loss = 0.16387, acc = 0.94825
squeezing data
[Test | 063/100] loss = 2.49446, acc = 0.53805
[Time cost | 063/100]: 842.8397s
Epoch: 64
[Train | 064/100] loss = 0.17401, acc = 0.94737
squeezing data
[Test | 064/100] loss = 2.48649, acc = 0.55199 visualizing

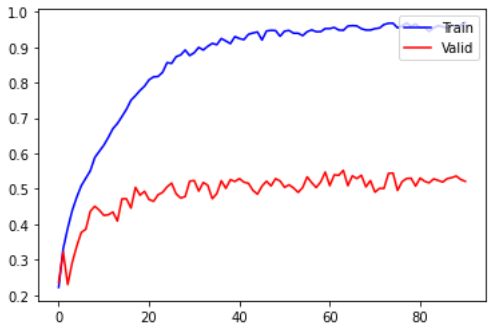


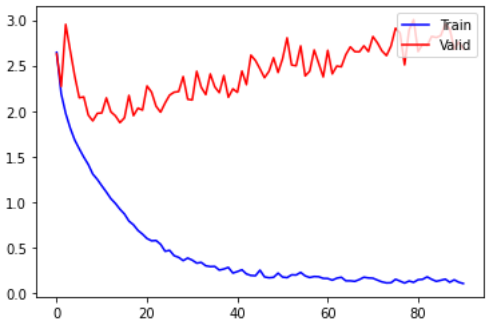
[Time cost | 064/100]: 822.8652s
Epoch: 65
[Train | 065/100] loss = 0.13521, acc = 0.95931
squeezing data
[Test | 065/100] loss = 2.62025, acc = 0.50862
[Time cost | 065/100]: 865.1630s
Epoch: 66
[Train | 066/100] loss = 0.13486, acc = 0.96075
squeezing data
[Test | 066/100] loss = 2.70519, acc = 0.53637
[Time cost | 066/100]: 866.9494s
Epoch: 67
[Train | 067/100] loss = 0.12925, acc = 0.95982
squeezing data
[Test | 067/100] loss = 2.65266, acc = 0.52869
[Time cost | 067/100]: 932.3244s
Epoch: 68
[Train | 068/100] loss = 0.14911, acc = 0.95195
squeezing data
[Test | 068/100] loss = 2.65114, acc = 0.53805

```
[Time cost | 068/100]: 936.9432s
Epoch: 69
[ Train | 069/100 ] loss = 0.17450, acc = 0.94822
squeezing data
[ Test | 069/100 ] loss = 2.71656, acc = 0.50512
[Time cost | 069/100]: 890.2309s
Epoch: 70
[ Train | 070/100 ] loss = 0.16660, acc = 0.94810
squeezing data
[ Test | 070/100 ] loss = 2.65397, acc = 0.52324
[Time cost | 070/100]: 870.9369s
Epoch: 71
[ Train | 071/100 ] loss = 0.16362, acc = 0.95222
squeezing data
[ Test | 071/100 ] loss = 2.81831, acc = 0.49044
[Time cost | 071/100]: 887.3069s
Epoch: 72
[ Train | 072/100 ] loss = 0.14245, acc = 0.95384
squeezing data
[ Test | 072/100 ] loss = 2.74857, acc = 0.50094
[Time cost | 072/100]: 850.7632s
Epoch: 73
[ Train | 073/100 ] loss = 0.12329, acc = 0.96371
squeezing data
[ Test | 073/100 ] loss = 2.66011, acc = 0.50094
[Time cost | 073/100]: 887.6339s
Epoch: 74
[ Train | 074/100 ] loss = 0.11223, acc = 0.96731
squeezing data
[ Test | 074/100 ] loss = 2.60896, acc = 0.54337
[Time cost | 074/100]: 964.3235s
Epoch: 75
[ Train | 075/100 ] loss = 0.11397, acc = 0.96755
squeezing data
[ Test | 075/100 ] loss = 2.71475, acc = 0.54432
[Time cost | 075/100]: 818.9128s
Epoch: 76
[ Train | 076/100 ] loss = 0.15157, acc = 0.95448
squeezing data
[ Test | 076/100 ] loss = 2.91036, acc = 0.49508
[Time cost | 076/100]: 812.3410s
Epoch: 77
[ Train | 077/100 ] loss = 0.13196, acc = 0.95981
squeezing data
```

```
[ Test | 077/100 ] loss = 2.85684, acc = 0.51994
[Time cost | 077/100]: 818.4602s
Epoch: 78
[ Train | 078/100 ] loss = 0.11045, acc = 0.96816
squeezing data
[ Test | 078/100 ] loss = 2.50808, acc = 0.52876
[Time cost | 078/100]: 785.8637s
Epoch: 79
[ Train | 079/100 ] loss = 0.13304, acc = 0.95878
squeezing data
[ Test | 079/100 ] loss = 2.89048, acc = 0.52990
[Time cost | 079/100]: 831.5814s
Epoch: 80
[ Train | 080/100 ] loss = 0.11842, acc = 0.96518
squeezing data
[ Test | 080/100 ] loss = 3.00404, acc = 0.50741
[Time cost | 080/100]: 820.1841s
Epoch: 81
[ Train | 081/100 ] loss = 0.14700, acc = 0.95417
squeezing data
[ Test | 081/100 ] loss = 2.65390, acc = 0.53024
[Time cost | 081/100]: 872.4435s
Epoch: 82
[ Train | 082/100 ] loss = 0.15169, acc = 0.95603
squeezing data
[ Test | 082/100 ] loss = 2.73492, acc = 0.52108
[Time cost | 082/100]: 837.5750s
Epoch: 83
[ Train | 083/100 ] loss = 0.17872, acc = 0.94452
squeezing data
[ Test | 083/100 ] loss = 2.72081, acc = 0.51657
[Time cost | 083/100]: 843.8426s
Epoch: 84
[ Train | 084/100 ] loss = 0.15116, acc = 0.95458
squeezing data
[ Test | 084/100 ] loss = 2.82052, acc = 0.52788
[Time cost | 084/100]: 921.7098s
Epoch: 85
[ Train | 085/100 ] loss = 0.12951, acc = 0.96056
squeezing data
[ Test | 085/100 ] loss = 2.80830, acc = 0.52344
[Time cost | 085/100]: 946.8614s
Epoch: 86
[ Train | 086/100 ] loss = 0.14031, acc = 0.95777
```

```
squeezing data
[ Test | 086/100 ] loss = 2.82822, acc = 0.51893
[Time cost | 086/100]: 888.4441s
Epoch: 87
[ Train | 087/100 ] loss = 0.15339, acc = 0.95396
squeezing data
[ Test | 087/100 ] loss = 2.95468, acc = 0.52849
[Time cost | 087/100]: 848.5547s
Epoch: 88
[ Train | 088/100 ] loss = 0.11940, acc = 0.96170
squeezing data
[ Test | 088/100 ] loss = 2.90282, acc = 0.53145
[Time cost | 088/100]: 829.8262s
Epoch: 89
[ Train | 089/100 ] loss = 0.14588, acc = 0.95766
squeezing data
[ Test | 089/100 ] loss = 2.67927, acc = 0.53590
[Time cost | 089/100]: 965.2562s
Epoch: 90
[ Train | 090/100 ] loss = 0.11938, acc = 0.96197
squeezing data
[ Test | 090/100 ] loss = 2.75908, acc = 0.52613
[Time cost | 090/100]: 980.1152s
Epoch: 91
[ Train | 091/100 ] loss = 0.10550, acc = 0.96969
squeezing data
[ Test | 091/100 ] loss = 2.68416, acc = 0.52088
[Time cost | 091/100]: 1012.6739s
```




Train
Valid
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
20
40
60
80