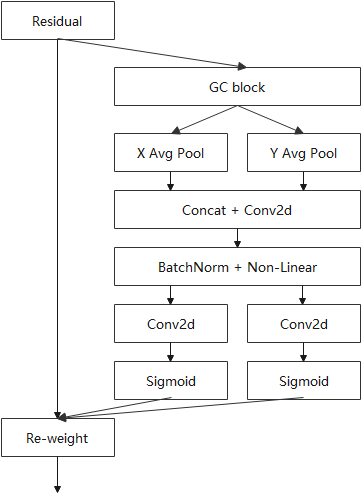
baseline参数量

FLOPs = 0.736403456G

Params = 14.143396M



CA

epochs =40

batch\_size = 8

lr=0.005 # ca 52%

momentum=0.9

[1, 2000] loss: 4.420,Train accuracy:4.18

[1, 4000] loss: 4.042,Train accuracy:7.39

[1, 6000] loss: 3.816,Train accuracy:10.86

Validation set: Average loss: 1.8199, Accuracy: 1361/10000 (14%)

[2, 2000] loss: 3.585,Train accuracy:14.06

[2, 4000] loss: 3.402,Train accuracy:17.50

[2, 6000] loss: 3.219,Train accuracy:20.54

Validation set: Average loss: 5.6406, Accuracy: 2096/10000 (21%)

[3, 2000] loss: 2.978,Train accuracy:24.51

[3, 4000] loss: 2.843,Train accuracy:27.44

[3, 6000] loss: 2.695,Train accuracy:30.46

Validation set: Average loss: 3.1619, Accuracy: 3259/10000 (33%)

[4, 2000] loss: 2.461,Train accuracy:34.86

[4, 4000] loss: 2.399,Train accuracy:36.47

[4, 6000] loss: 2.347,Train accuracy:37.86

Validation set: Average loss: 6.6301, Accuracy: 3653/10000 (37%)

[5, 2000] loss: 2.090,Train accuracy:43.56

[5, 4000] loss: 2.066,Train accuracy:44.02

[5, 6000] loss: 2.077,Train accuracy:43.84

Validation set: Average loss: 0.6050, Accuracy: 4333/10000 (43%)

[6, 2000] loss: 1.870,Train accuracy:48.99

[6, 4000] loss: 1.807,Train accuracy:50.24

[6, 6000] loss: 1.790,Train accuracy:50.51

Validation set: Average loss: 6.0028, Accuracy: 4367/10000 (44%)

[7, 2000] loss: 1.535,Train accuracy:56.27

[7, 4000] loss: 1.569,Train accuracy:55.50

[7, 6000] loss: 1.560,Train accuracy:56.03

Validation set: Average loss: 7.9500, Accuracy: 4691/10000 (47%)

[8, 2000] loss: 1.278,Train accuracy:62.54

[8, 4000] loss: 1.347,Train accuracy:61.45

[8, 6000] loss: 1.359,Train accuracy:60.42

Validation set: Average loss: 7.2303, Accuracy: 4584/10000 (46%)

[9, 2000] loss: 1.051,Train accuracy:68.86

[9, 4000] loss: 1.112,Train accuracy:67.11

[9, 6000] loss: 1.159,Train accuracy:66.09

Validation set: Average loss: 4.6323, Accuracy: 4876/10000 (49%)

[10, 2000] loss: 0.818,Train accuracy:74.74

[10, 4000] loss: 0.927,Train accuracy:71.65

[10, 6000] loss: 0.946,Train accuracy:71.56

Validation set: Average loss: 4.0290, Accuracy: 4917/10000 (49%)

[11, 2000] loss: 0.628,Train accuracy:80.04

[11, 4000] loss: 0.727,Train accuracy:77.41

[11, 6000] loss: 0.780,Train accuracy:76.29

Validation set: Average loss: 2.8330, Accuracy: 5153/10000 (52%)

[12, 2000] loss: 0.478,Train accuracy:85.04

[12, 4000] loss: 0.537,Train accuracy:82.81

[12, 6000] loss: 0.597,Train accuracy:80.91

Validation set: Average loss: 8.9122, Accuracy: 4807/10000 (48%)

[13, 2000] loss: 0.369,Train accuracy:88.20

[13, 4000] loss: 0.416,Train accuracy:86.67

[13, 6000] loss: 0.453,Train accuracy:85.47

Validation set: Average loss: 7.5284, Accuracy: 4986/10000 (50%)

[14, 2000] loss: 0.265,Train accuracy:91.61

[14, 4000] loss: 0.298,Train accuracy:90.31

[14, 6000] loss: 0.334,Train accuracy:89.50

Validation set: Average loss: 30.4952, Accuracy: 4804/10000 (48%)

[15, 2000] loss: 0.194,Train accuracy:93.87

[15, 4000] loss: 0.235,Train accuracy:92.45

[15, 6000] loss: 0.255,Train accuracy:92.03

[16, 2000] loss: 0.154,Train accuracy:95.33

[16, 4000] loss: 0.174,Train accuracy:94.51

[16, 6000] loss: 0.198,Train accuracy:93.84

Validation set: Average loss: 8.1980, Accuracy: 5098/10000 (51%)

[17, 2000] loss: 0.118,Train accuracy:96.48

[17, 4000] loss: 0.130,Train accuracy:96.03

[17, 6000] loss: 0.155,Train accuracy:95.15

Validation set: Average loss: 32.8530, Accuracy: 5020/10000 (50%)

[18, 2000] loss: 0.091,Train accuracy:97.38

[18, 4000] loss: 0.102,Train accuracy:96.77

[18, 6000] loss: 0.113,Train accuracy:96.43

Validation set: Average loss: 14.9202, Accuracy: 4942/10000 (49%)

[19, 2000] loss: 0.074,Train accuracy:97.90

[19, 4000] loss: 0.089,Train accuracy:97.24

[19, 6000] loss: 0.096,Train accuracy:97.04

Validation set: Average loss: 19.3704, Accuracy: 5079/10000 (51%)

[20, 2000] loss: 0.052,Train accuracy:98.46

[20, 4000] loss: 0.066,Train accuracy:98.06

[20, 6000] loss: 0.073,Train accuracy:97.96

Validation set: Average loss: 8.5065, Accuracy: 5295/10000 (53%)

[21, 2000] loss: 0.049,Train accuracy:98.67

[21, 4000] loss: 0.056,Train accuracy:98.43

[21, 6000] loss: 0.059,Train accuracy:98.35

Validation set: Average loss: 61.0517, Accuracy: 4897/10000 (49%)

[22, 2000] loss: 0.039,Train accuracy:98.93

[22, 4000] loss: 0.049,Train accuracy:98.60

[22, 6000] loss: 0.052,Train accuracy:98.55

Validation set: Average loss: 20.2754, Accuracy: 5199/10000 (52%)

[23, 2000] loss: 0.042,Train accuracy:98.78

[23, 4000] loss: 0.039,Train accuracy:98.97

[23, 6000] loss: 0.040,Train accuracy:98.93

Validation set: Average loss: 17.0414, Accuracy: 5195/10000 (52%)

[24, 2000] loss: 0.028,Train accuracy:99.33

[24, 4000] loss: 0.031,Train accuracy:99.28

[24, 6000] loss: 0.032,Train accuracy:99.07

Validation set: Average loss: 4.2757, Accuracy: 5323/10000 (53%)

[25, 2000] loss: 0.029,Train accuracy:99.28

[25, 4000] loss: 0.028,Train accuracy:99.33

[25, 6000] loss: 0.031,Train accuracy:99.15

Validation set: Average loss: 8.4393, Accuracy: 5337/10000 (53%)

[26, 2000] loss: 0.026,Train accuracy:99.28

[26, 4000] loss: 0.020,Train accuracy:99.51

[26, 6000] loss: 0.025,Train accuracy:99.42

Validation set: Average loss: 10.4686, Accuracy: 5268/10000 (53%)

[27, 2000] loss: 0.019,Train accuracy:99.51

[27, 4000] loss: 0.026,Train accuracy:99.42

[27, 6000] loss: 0.028,Train accuracy:99.30

Validation set: Average loss: 14.8152, Accuracy: 5179/10000 (52%)

[28, 2000] loss: 0.019,Train accuracy:99.59

[28, 4000] loss: 0.018,Train accuracy:99.48

[28, 6000] loss: 0.016,Train accuracy:99.60

Validation set: Average loss: 15.7367, Accuracy: 5242/10000 (52%)

[29, 2000] loss: 0.018,Train accuracy:99.61

[29, 4000] loss: 0.018,Train accuracy:99.59

[29, 6000] loss: 0.015,Train accuracy:99.71

Validation set: Average loss: 10.7421, Accuracy: 5269/10000 (53%)

[30, 2000] loss: 0.012,Train accuracy:99.76

[30, 4000] loss: 0.012,Train accuracy:99.73

[30, 6000] loss: 0.014,Train accuracy:99.65

Validation set: Average loss: 14.2660, Accuracy: 5314/10000 (53%)

[31, 2000] loss: 0.013,Train accuracy:99.74

[31, 4000] loss: 0.014,Train accuracy:99.63

[31, 6000] loss: 0.013,Train accuracy:99.72

Validation set: Average loss: 12.0891, Accuracy: 5176/10000 (52%)

[32, 2000] loss: 0.013,Train accuracy:99.71

[32, 4000] loss: 0.012,Train accuracy:99.71

[32, 6000] loss: 0.013,Train accuracy:99.71

Validation set: Average loss: 13.9503, Accuracy: 5319/10000 (53%)

[33, 2000] loss: 0.011,Train accuracy:99.73

[33, 4000] loss: 0.012,Train accuracy:99.74

[33, 6000] loss: 0.013,Train accuracy:99.66

Validation set: Average loss: 35.2238, Accuracy: 5059/10000 (51%)

[34, 2000] loss: 0.012,Train accuracy:99.71

[34, 4000] loss: 0.012,Train accuracy:99.77

[34, 6000] loss: 0.010,Train accuracy:99.74

Validation set: Average loss: 25.7356, Accuracy: 5021/10000 (50%)

[35, 2000] loss: 0.007,Train accuracy:99.84

[35, 4000] loss: 0.009,Train accuracy:99.78

[35, 6000] loss: 0.009,Train accuracy:99.82

Validation set: Average loss: 11.2099, Accuracy: 5250/10000 (52%)

[36, 2000] loss: 0.005,Train accuracy:99.89

[36, 4000] loss: 0.007,Train accuracy:99.83

[36, 6000] loss: 0.009,Train accuracy:99.78

Validation set: Average loss: 45.9676, Accuracy: 4855/10000 (49%)

[37, 2000] loss: 0.007,Train accuracy:99.85

[37, 4000] loss: 0.010,Train accuracy:99.76

[37, 6000] loss: 0.007,Train accuracy:99.84

Validation set: Average loss: 33.2636, Accuracy: 5178/10000 (52%)

[38, 2000] loss: 0.008,Train accuracy:99.79

[38, 4000] loss: 0.008,Train accuracy:99.83

[38, 6000] loss: 0.009,Train accuracy:99.80

Validation set: Average loss: 10.0997, Accuracy: 5255/10000 (53%)

[39, 2000] loss: 0.008,Train accuracy:99.79

[39, 4000] loss: 0.007,Train accuracy:99.87

[39, 6000] loss: 0.006,Train accuracy:99.86

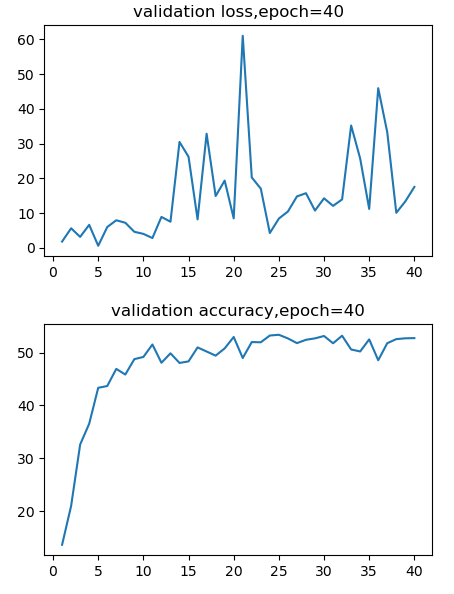
Validation set: Average loss: 13.3809, Accuracy: 5270/10000 (53%)

[40, 2000] loss: 0.004,Train accuracy:99.91

[40, 4000] loss: 0.007,Train accuracy:99.81

[40, 6000] loss: 0.009,Train accuracy:99.78

Validation set: Average loss: 17.5461, Accuracy: 5274/10000 (53%)



CA

epochs =40

batch\_size = 8

lr=0.0008 # ca 52%

momentum=0.9

[1, 2000] loss: 4.286,Train accuracy:6.36

[1, 4000] loss: 3.735,Train accuracy:13.34

[1, 6000] loss: 3.431,Train accuracy:17.75

Validation set: Average loss: 0.3868, Accuracy: 2335/10000 (23%)

[2, 2000] loss: 3.077,Train accuracy:23.54

[2, 4000] loss: 2.853,Train accuracy:28.39

[2, 6000] loss: 2.721,Train accuracy:30.37

Validation set: Average loss: 0.3025, Accuracy: 3736/10000 (37%)

[3, 2000] loss: 2.385,Train accuracy:37.33

[3, 4000] loss: 2.333,Train accuracy:38.85

[3, 6000] loss: 2.244,Train accuracy:41.02

Validation set: Average loss: 0.2616, Accuracy: 4464/10000 (45%)

[4, 2000] loss: 1.933,Train accuracy:47.14

[4, 4000] loss: 1.963,Train accuracy:46.78

[4, 6000] loss: 1.936,Train accuracy:47.21

Validation set: Average loss: 0.2468, Accuracy: 4875/10000 (49%)

[5, 2000] loss: 1.568,Train accuracy:55.97

[5, 4000] loss: 1.640,Train accuracy:54.34

[5, 6000] loss: 1.663,Train accuracy:53.90

Validation set: Average loss: 0.2299, Accuracy: 5147/10000 (51%)

[6, 2000] loss: 1.242,Train accuracy:64.42

[6, 4000] loss: 1.339,Train accuracy:61.13

[6, 6000] loss: 1.399,Train accuracy:60.06

Validation set: Average loss: 0.2437, Accuracy: 5088/10000 (51%)

[7, 2000] loss: 0.954,Train accuracy:71.30

[7, 4000] loss: 1.070,Train accuracy:68.46

[7, 6000] loss: 1.165,Train accuracy:65.90

Validation set: Average loss: 0.2491, Accuracy: 5082/10000 (51%)

[8, 2000] loss: 0.717,Train accuracy:78.43

[8, 4000] loss: 0.821,Train accuracy:74.96

[8, 6000] loss: 0.934,Train accuracy:71.71

Validation set: Average loss: 0.2462, Accuracy: 5314/10000 (53%)

[9, 2000] loss: 0.506,Train accuracy:84.79

[9, 4000] loss: 0.582,Train accuracy:81.92

[9, 6000] loss: 0.689,Train accuracy:79.12

Validation set: Average loss: 0.2613, Accuracy: 5255/10000 (53%)

[10, 2000] loss: 0.358,Train accuracy:89.33

[10, 4000] loss: 0.410,Train accuracy:87.12

[10, 6000] loss: 0.500,Train accuracy:84.49

Validation set: Average loss: 0.2792, Accuracy: 5215/10000 (52%)

[11, 2000] loss: 0.260,Train accuracy:92.43

[11, 4000] loss: 0.293,Train accuracy:91.18

[11, 6000] loss: 0.345,Train accuracy:89.40

Validation set: Average loss: 0.2846, Accuracy: 5293/10000 (53%)

[12, 2000] loss: 0.201,Train accuracy:94.08

[12, 4000] loss: 0.227,Train accuracy:93.19

[12, 6000] loss: 0.257,Train accuracy:91.97

Validation set: Average loss: 0.2957, Accuracy: 5282/10000 (53%)

[13, 2000] loss: 0.153,Train accuracy:95.78

[13, 4000] loss: 0.142,Train accuracy:95.85

[13, 6000] loss: 0.167,Train accuracy:95.16

Validation set: Average loss: 0.2838, Accuracy: 5332/10000 (53%)

[14, 2000] loss: 0.109,Train accuracy:96.91

[14, 4000] loss: 0.116,Train accuracy:96.65

[14, 6000] loss: 0.143,Train accuracy:95.83

Validation set: Average loss: 0.3031, Accuracy: 5330/10000 (53%)

[15, 2000] loss: 0.092,Train accuracy:97.59

[15, 4000] loss: 0.084,Train accuracy:97.88

[15, 6000] loss: 0.102,Train accuracy:97.09

Validation set: Average loss: 0.2986, Accuracy: 5453/10000 (55%)

[16, 2000] loss: 0.064,Train accuracy:98.36

[16, 4000] loss: 0.070,Train accuracy:98.16

[16, 6000] loss: 0.082,Train accuracy:97.93

Validation set: Average loss: 0.3082, Accuracy: 5476/10000 (55%)

[17, 2000] loss: 0.057,Train accuracy:98.57

[17, 4000] loss: 0.051,Train accuracy:98.78

[17, 6000] loss: 0.068,Train accuracy:98.34

Validation set: Average loss: 0.3079, Accuracy: 5499/10000 (55%)

[18, 2000] loss: 0.047,Train accuracy:98.89

[18, 4000] loss: 0.046,Train accuracy:98.90

[18, 6000] loss: 0.053,Train accuracy:98.56

Validation set: Average loss: 0.3028, Accuracy: 5529/10000 (55%)

[19, 2000] loss: 0.032,Train accuracy:99.30

[19, 4000] loss: 0.033,Train accuracy:99.28

[19, 6000] loss: 0.038,Train accuracy:99.20

Validation set: Average loss: 0.3040, Accuracy: 5577/10000 (56%)

[20, 2000] loss: 0.029,Train accuracy:99.42

[20, 4000] loss: 0.033,Train accuracy:99.34

[20, 6000] loss: 0.030,Train accuracy:99.33

Validation set: Average loss: 0.3061, Accuracy: 5542/10000 (55%)

[21, 2000] loss: 0.022,Train accuracy:99.67

[21, 4000] loss: 0.024,Train accuracy:99.48

[21, 6000] loss: 0.025,Train accuracy:99.43

Validation set: Average loss: 0.3078, Accuracy: 5534/10000 (55%)

[22, 2000] loss: 0.021,Train accuracy:99.61

[22, 4000] loss: 0.025,Train accuracy:99.46

[22, 6000] loss: 0.026,Train accuracy:99.41

Validation set: Average loss: 0.3206, Accuracy: 5514/10000 (55%)

[23, 2000] loss: 0.020,Train accuracy:99.56

[23, 4000] loss: 0.022,Train accuracy:99.49

[23, 6000] loss: 0.024,Train accuracy:99.53

Validation set: Average loss: 0.3131, Accuracy: 5481/10000 (55%)

[24, 2000] loss: 0.017,Train accuracy:99.68

[24, 4000] loss: 0.020,Train accuracy:99.68

[24, 6000] loss: 0.016,Train accuracy:99.64

Validation set: Average loss: 0.3156, Accuracy: 5560/10000 (56%)

[25, 4000] loss: 0.016,Train accuracy:99.66

[25, 6000] loss: 0.018,Train accuracy:99.53

Validation set: Average loss: 0.3167, Accuracy: 5542/10000 (55%)

[26, 2000] loss: 0.014,Train accuracy:99.72

[26, 4000] loss: 0.014,Train accuracy:99.73

[26, 6000] loss: 0.017,Train accuracy:99.63

Validation set: Average loss: 0.3142, Accuracy: 5564/10000 (56%)

[27, 2000] loss: 0.012,Train accuracy:99.74

[27, 4000] loss: 0.015,Train accuracy:99.72

[27, 6000] loss: 0.014,Train accuracy:99.79

Validation set: Average loss: 0.3135, Accuracy: 5571/10000 (56%)

[28, 2000] loss: 0.011,Train accuracy:99.84

[28, 4000] loss: 0.013,Train accuracy:99.74

[28, 6000] loss: 0.014,Train accuracy:99.72

Validation set: Average loss: 0.3140, Accuracy: 5593/10000 (56%)

[29, 2000] loss: 0.012,Train accuracy:99.76

[29, 4000] loss: 0.013,Train accuracy:99.72

[29, 6000] loss: 0.013,Train accuracy:99.76

Validation set: Average loss: 0.3156, Accuracy: 5588/10000 (56%)

[30, 2000] loss: 0.008,Train accuracy:99.87

[30, 4000] loss: 0.011,Train accuracy:99.79

[30, 6000] loss: 0.012,Train accuracy:99.79

Validation set: Average loss: 0.3141, Accuracy: 5585/10000 (56%)

[31, 2000] loss: 0.010,Train accuracy:99.80

[31, 4000] loss: 0.010,Train accuracy:99.82

[31, 6000] loss: 0.013,Train accuracy:99.74

Validation set: Average loss: 0.3176, Accuracy: 5577/10000 (56%)

[32, 2000] loss: 0.010,Train accuracy:99.84

[32, 4000] loss: 0.007,Train accuracy:99.89

[32, 6000] loss: 0.012,Train accuracy:99.73

Validation set: Average loss: 0.3194, Accuracy: 5643/10000 (56%)

[33, 2000] loss: 0.008,Train accuracy:99.90

[33, 4000] loss: 0.007,Train accuracy:99.86

[33, 6000] loss: 0.009,Train accuracy:99.76

[34, 2000] loss: 0.009,Train accuracy:99.86

[34, 4000] loss: 0.011,Train accuracy:99.78

[34, 6000] loss: 0.009,Train accuracy:99.90

Validation set: Average loss: 0.3218, Accuracy: 5594/10000 (56%)

[35, 2000] loss: 0.006,Train accuracy:99.91

[35, 4000] loss: 0.008,Train accuracy:99.84

[35, 6000] loss: 0.008,Train accuracy:99.83

Validation set: Average loss: 0.3205, Accuracy: 5580/10000 (56%)

[36, 2000] loss: 0.007,Train accuracy:99.87

[36, 4000] loss: 0.008,Train accuracy:99.82

[36, 6000] loss: 0.009,Train accuracy:99.86

Validation set: Average loss: 0.3237, Accuracy: 5578/10000 (56%)

[37, 2000] loss: 0.006,Train accuracy:99.91

[37, 4000] loss: 0.006,Train accuracy:99.93

[37, 6000] loss: 0.006,Train accuracy:99.91

Validation set: Average loss: 0.3211, Accuracy: 5630/10000 (56%)

[38, 2000] loss: 0.006,Train accuracy:99.91

[38, 4000] loss: 0.006,Train accuracy:99.93

[38, 6000] loss: 0.008,Train accuracy:99.86

Validation set: Average loss: 0.3211, Accuracy: 5624/10000 (56%)

[39, 2000] loss: 0.005,Train accuracy:99.93

[39, 4000] loss: 0.005,Train accuracy:99.91

[39, 6000] loss: 0.005,Train accuracy:99.92

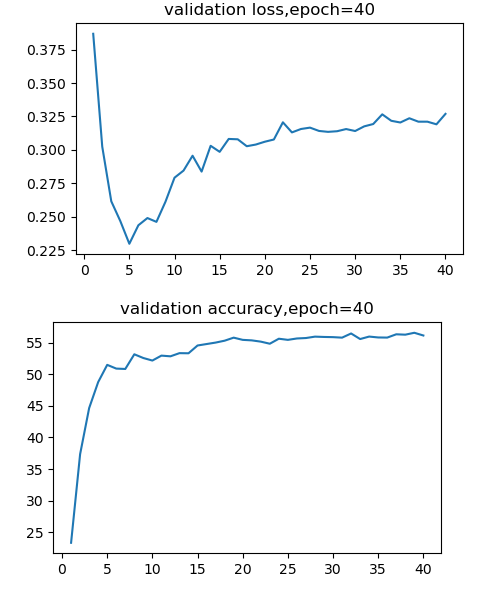
Validation set: Average loss: 0.3191, Accuracy: 5654/10000 (57%)

[40, 2000] loss: 0.005,Train accuracy:99.92

[40, 4000] loss: 0.004,Train accuracy:99.96

[40, 6000] loss: 0.006,Train accuracy:99.90

Validation set: Average loss: 0.3269, Accuracy: 5611/10000 (56%)



CA

epochs =40

batch\_size = 8

lr=0.0015

momentum=0.9

[1, 2000] loss: 4.379,Train accuracy:5.13

[1, 4000] loss: 3.887,Train accuracy:9.88

[1, 6000] loss: 3.592,Train accuracy:14.53

Validation set: Average loss: 0.4208, Accuracy: 1804/10000 (18%)

[2, 2000] loss: 3.349,Train accuracy:18.60

[2, 4000] loss: 3.132,Train accuracy:22.16

[2, 6000] loss: 2.949,Train accuracy:25.64

Validation set: Average loss: 1.1157, Accuracy: 3118/10000 (31%)

[3, 2000] loss: 2.663,Train accuracy:31.25

[3, 4000] loss: 2.566,Train accuracy:33.48

[3, 6000] loss: 2.451,Train accuracy:35.44

Validation set: Average loss: 0.2874, Accuracy: 3973/10000 (40%)

[4, 2000] loss: 2.199,Train accuracy:40.88

[4, 4000] loss: 2.156,Train accuracy:42.29

[4, 6000] loss: 2.122,Train accuracy:42.60

Validation set: Average loss: 0.3211, Accuracy: 4388/10000 (44%)

[5, 2000] loss: 1.850,Train accuracy:49.19

[5, 4000] loss: 1.852,Train accuracy:48.59

[5, 6000] loss: 1.875,Train accuracy:48.32

Validation set: Average loss: 0.2489, Accuracy: 4728/10000 (47%)

[6, 2000] loss: 1.558,Train accuracy:55.53

[6, 4000] loss: 1.602,Train accuracy:55.03

[6, 6000] loss: 1.613,Train accuracy:54.47

Validation set: Average loss: 0.2627, Accuracy: 4955/10000 (50%)

[7, 2000] loss: 1.268,Train accuracy:62.67

[7, 4000] loss: 1.356,Train accuracy:60.41

[7, 6000] loss: 1.396,Train accuracy:59.48

Validation set: Average loss: 0.4495, Accuracy: 4724/10000 (47%)

[8, 2000] loss: 1.050,Train accuracy:68.38

[8, 4000] loss: 1.112,Train accuracy:66.84

[8, 6000] loss: 1.183,Train accuracy:65.45

Validation set: Average loss: 0.2394, Accuracy: 5209/10000 (52%)

[9, 2000] loss: 0.828,Train accuracy:74.51

[9, 4000] loss: 0.926,Train accuracy:71.48

[9, 6000] loss: 0.981,Train accuracy:70.20

Validation set: Average loss: 0.2414, Accuracy: 5295/10000 (53%)

[10, 2000] loss: 0.600,Train accuracy:81.36

[10, 4000] loss: 0.739,Train accuracy:76.64

[10, 6000] loss: 0.808,Train accuracy:75.20

Validation set: Average loss: 0.2681, Accuracy: 5181/10000 (52%)

[11, 2000] loss: 0.461,Train accuracy:85.17

[11, 4000] loss: 0.550,Train accuracy:82.91

[11, 6000] loss: 0.638,Train accuracy:79.98

Validation set: Average loss: 0.2784, Accuracy: 5206/10000 (52%)

[12, 2000] loss: 0.343,Train accuracy:89.21

[12, 4000] loss: 0.403,Train accuracy:87.02

[12, 6000] loss: 0.487,Train accuracy:84.33

Validation set: Average loss: 0.2752, Accuracy: 5343/10000 (53%)

[13, 2000] loss: 0.271,Train accuracy:91.74

[13, 4000] loss: 0.320,Train accuracy:89.74

[13, 6000] loss: 0.386,Train accuracy:87.55

Validation set: Average loss: 0.3064, Accuracy: 5266/10000 (53%)

[14, 2000] loss: 0.216,Train accuracy:93.27

[14, 4000] loss: 0.255,Train accuracy:91.71

[14, 6000] loss: 0.297,Train accuracy:90.62

Validation set: Average loss: 0.3082, Accuracy: 5284/10000 (53%)

[15, 2000] loss: 0.167,Train accuracy:94.88

[15, 4000] loss: 0.179,Train accuracy:94.33

[15, 6000] loss: 0.229,Train accuracy:92.61

Validation set: Average loss: 0.3171, Accuracy: 5275/10000 (53%)

[16, 2000] loss: 0.134,Train accuracy:95.84

[16, 4000] loss: 0.164,Train accuracy:94.75

[16, 6000] loss: 0.182,Train accuracy:94.31

Validation set: Average loss: 0.3245, Accuracy: 5324/10000 (53%)

[17, 2000] loss: 0.118,Train accuracy:96.53

[17, 4000] loss: 0.133,Train accuracy:95.84

[17, 6000] loss: 0.148,Train accuracy:95.38

Validation set: Average loss: 0.3341, Accuracy: 5229/10000 (52%)

[18, 2000] loss: 0.103,Train accuracy:96.61

[18, 6000] loss: 0.133,Train accuracy:95.78

Validation set: Average loss: 0.3290, Accuracy: 5340/10000 (53%)

[19, 2000] loss: 0.086,Train accuracy:97.47

[19, 4000] loss: 0.103,Train accuracy:96.85

[19, 6000] loss: 0.104,Train accuracy:96.77

Validation set: Average loss: 0.3388, Accuracy: 5306/10000 (53%)

[20, 2000] loss: 0.083,Train accuracy:97.67

[20, 6000] loss: 0.085,Train accuracy:97.53

Validation set: Average loss: 0.3316, Accuracy: 5458/10000 (55%)

[21, 2000] loss: 0.059,Train accuracy:98.27

[21, 4000] loss: 0.066,Train accuracy:97.94

[21, 6000] loss: 0.077,Train accuracy:97.79

Validation set: Average loss: 0.3379, Accuracy: 5488/10000 (55%)

[22, 2000] loss: 0.062,Train accuracy:98.14

[22, 4000] loss: 0.061,Train accuracy:98.22

[22, 6000] loss: 0.064,Train accuracy:98.03

Validation set: Average loss: 0.3393, Accuracy: 5500/10000 (55%)

[23, 2000] loss: 0.044,Train accuracy:98.75

[23, 4000] loss: 0.051,Train accuracy:98.46

[23, 6000] loss: 0.070,Train accuracy:97.91

Validation set: Average loss: 0.3325, Accuracy: 5534/10000 (55%)

[24, 2000] loss: 0.047,Train accuracy:98.70

[24, 4000] loss: 0.041,Train accuracy:98.88

[24, 6000] loss: 0.055,Train accuracy:98.44

Validation set: Average loss: 0.3414, Accuracy: 5535/10000 (55%)

[25, 2000] loss: 0.041,Train accuracy:98.82

[25, 4000] loss: 0.035,Train accuracy:99.06

[25, 6000] loss: 0.038,Train accuracy:98.99

Validation set: Average loss: 0.3534, Accuracy: 5521/10000 (55%)

[26, 2000] loss: 0.030,Train accuracy:99.14

[26, 4000] loss: 0.031,Train accuracy:99.19

[26, 6000] loss: 0.041,Train accuracy:98.74

Validation set: Average loss: 0.3448, Accuracy: 5518/10000 (55%)

[27, 2000] loss: 0.033,Train accuracy:99.08

[27, 4000] loss: 0.031,Train accuracy:99.23

[27, 6000] loss: 0.034,Train accuracy:99.06

Validation set: Average loss: 0.3471, Accuracy: 5545/10000 (55%)

[28, 2000] loss: 0.027,Train accuracy:99.24

[28, 4000] loss: 0.025,Train accuracy:99.41

[28, 6000] loss: 0.032,Train accuracy:99.19

Validation set: Average loss: 0.3566, Accuracy: 5544/10000 (55%)

[29, 2000] loss: 0.021,Train accuracy:99.42

[29, 4000] loss: 0.024,Train accuracy:99.48

[29, 6000] loss: 0.029,Train accuracy:99.19

Validation set: Average loss: 0.3574, Accuracy: 5520/10000 (55%)

[30, 2000] loss: 0.022,Train accuracy:99.41

[30, 4000] loss: 0.020,Train accuracy:99.49

[30, 6000] loss: 0.023,Train accuracy:99.43

Validation set: Average loss: 0.3598, Accuracy: 5487/10000 (55%)

[31, 2000] loss: 0.022,Train accuracy:99.51

[31, 4000] loss: 0.018,Train accuracy:99.51

[31, 6000] loss: 0.025,Train accuracy:99.36

Validation set: Average loss: 0.3504, Accuracy: 5596/10000 (56%)

[32, 2000] loss: 0.014,Train accuracy:99.63

[32, 4000] loss: 0.019,Train accuracy:99.59

[32, 6000] loss: 0.019,Train accuracy:99.58

Validation set: Average loss: 0.3542, Accuracy: 5545/10000 (55%)

[33, 2000] loss: 0.014,Train accuracy:99.58

[33, 4000] loss: 0.015,Train accuracy:99.59

[33, 6000] loss: 0.015,Train accuracy:99.62

Validation set: Average loss: 0.3625, Accuracy: 5520/10000 (55%)

[34, 2000] loss: 0.013,Train accuracy:99.72

[34, 4000] loss: 0.017,Train accuracy:99.61

[34, 6000] loss: 0.016,Train accuracy:99.54

Validation set: Average loss: 0.3636, Accuracy: 5578/10000 (56%)

[35, 2000] loss: 0.015,Train accuracy:99.65

[35, 4000] loss: 0.014,Train accuracy:99.65

[35, 6000] loss: 0.015,Train accuracy:99.66

Validation set: Average loss: 0.3658, Accuracy: 5547/10000 (55%)

[36, 2000] loss: 0.013,Train accuracy:99.73

[36, 4000] loss: 0.012,Train accuracy:99.69

[36, 6000] loss: 0.015,Train accuracy:99.58

Validation set: Average loss: 0.3574, Accuracy: 5538/10000 (55%)

[37, 2000] loss: 0.011,Train accuracy:99.76

[37, 4000] loss: 0.015,Train accuracy:99.58

[37, 6000] loss: 0.017,Train accuracy:99.55

Validation set: Average loss: 0.3590, Accuracy: 5514/10000 (55%)

[38, 2000] loss: 0.011,Train accuracy:99.72

[38, 4000] loss: 0.012,Train accuracy:99.71

[38, 6000] loss: 0.015,Train accuracy:99.64

Validation set: Average loss: 0.3643, Accuracy: 5475/10000 (55%)

[39, 2000] loss: 0.010,Train accuracy:99.73

[39, 4000] loss: 0.011,Train accuracy:99.76

[39, 6000] loss: 0.013,Train accuracy:99.64

Validation set: Average loss: 0.3680, Accuracy: 5562/10000 (56%)

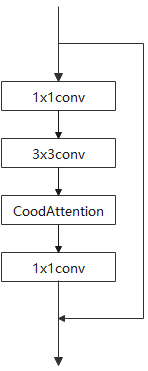
[40, 2000] loss: 0.007,Train accuracy:99.83

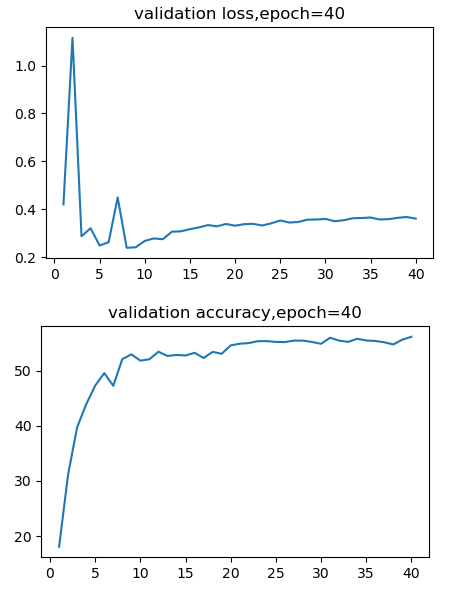
[40, 4000] loss: 0.010,Train accuracy:99.83

[40, 6000] loss: 0.013,Train accuracy:99.66

Validation set: Average loss: 0.3610, Accuracy: 5614/10000 (56%)

training finished！





----------------------------------------------------------------

Layer (type) Output Shape Param #

================================================================

Conv2d-1 [-1, 64, 32, 32] 1,728

BatchNorm2d-2 [-1, 64, 32, 32] 128

Conv2d-3 [-1, 64, 32, 32] 4,096

BatchNorm2d-4 [-1, 64, 32, 32] 128

ReLU-5 [-1, 64, 32, 32] 0

Conv2d-6 [-1, 64, 32, 32] 36,864

AdaptiveAvgPool2d-7 [-1, 64, 32, 1] 0

AdaptiveAvgPool2d-8 [-1, 64, 1, 32] 0

Conv2d-9 [-1, 8, 64, 1] 520

BatchNorm2d-10 [-1, 8, 64, 1] 16

ReLU6-11 [-1, 8, 64, 1] 0

h\_sigmoid-12 [-1, 8, 64, 1] 0

h\_swish-13 [-1, 8, 64, 1] 0

Conv2d-14 [-1, 64, 32, 1] 576

Conv2d-15 [-1, 64, 1, 32] 576

CoordAtt-16 [-1, 64, 32, 32] 0

BatchNorm2d-17 [-1, 64, 32, 32] 128

ReLU-18 [-1, 64, 32, 32] 0

Conv2d-19 [-1, 256, 32, 32] 16,384

BatchNorm2d-20 [-1, 256, 32, 32] 512

Conv2d-21 [-1, 256, 32, 32] 16,384

BatchNorm2d-22 [-1, 256, 32, 32] 512

ReLU-23 [-1, 256, 32, 32] 0

Bottleneck-24 [-1, 256, 32, 32] 0

Conv2d-25 [-1, 64, 32, 32] 16,384

BatchNorm2d-26 [-1, 64, 32, 32] 128

ReLU-27 [-1, 64, 32, 32] 0

Conv2d-28 [-1, 64, 32, 32] 36,864

AdaptiveAvgPool2d-29 [-1, 64, 32, 1] 0

AdaptiveAvgPool2d-30 [-1, 64, 1, 32] 0

Conv2d-31 [-1, 8, 64, 1] 520

BatchNorm2d-32 [-1, 8, 64, 1] 16

ReLU6-33 [-1, 8, 64, 1] 0

h\_sigmoid-34 [-1, 8, 64, 1] 0

h\_swish-35 [-1, 8, 64, 1] 0

Conv2d-36 [-1, 64, 32, 1] 576

Conv2d-37 [-1, 64, 1, 32] 576

CoordAtt-38 [-1, 64, 32, 32] 0

BatchNorm2d-39 [-1, 64, 32, 32] 128

ReLU-40 [-1, 64, 32, 32] 0

Conv2d-41 [-1, 256, 32, 32] 16,384

BatchNorm2d-42 [-1, 256, 32, 32] 512

ReLU-43 [-1, 256, 32, 32] 0

Bottleneck-44 [-1, 256, 32, 32] 0

Conv2d-45 [-1, 128, 32, 32] 32,768

BatchNorm2d-46 [-1, 128, 32, 32] 256

ReLU-47 [-1, 128, 32, 32] 0

Conv2d-48 [-1, 128, 16, 16] 147,456

AdaptiveAvgPool2d-49 [-1, 128, 16, 1] 0

AdaptiveAvgPool2d-50 [-1, 128, 1, 16] 0

Conv2d-51 [-1, 8, 32, 1] 1,032

BatchNorm2d-52 [-1, 8, 32, 1] 16

ReLU6-53 [-1, 8, 32, 1] 0

h\_sigmoid-54 [-1, 8, 32, 1] 0

h\_swish-55 [-1, 8, 32, 1] 0

Conv2d-56 [-1, 128, 16, 1] 1,152

Conv2d-57 [-1, 128, 1, 16] 1,152

CoordAtt-58 [-1, 128, 16, 16] 0

BatchNorm2d-59 [-1, 128, 16, 16] 256

ReLU-60 [-1, 128, 16, 16] 0

Conv2d-61 [-1, 512, 16, 16] 65,536

BatchNorm2d-62 [-1, 512, 16, 16] 1,024

Conv2d-63 [-1, 512, 16, 16] 131,072

BatchNorm2d-64 [-1, 512, 16, 16] 1,024

ReLU-65 [-1, 512, 16, 16] 0

Bottleneck-66 [-1, 512, 16, 16] 0

Conv2d-67 [-1, 128, 16, 16] 65,536

BatchNorm2d-68 [-1, 128, 16, 16] 256

ReLU-69 [-1, 128, 16, 16] 0

Conv2d-70 [-1, 128, 16, 16] 147,456

AdaptiveAvgPool2d-71 [-1, 128, 16, 1] 0

AdaptiveAvgPool2d-72 [-1, 128, 1, 16] 0

Conv2d-73 [-1, 8, 32, 1] 1,032

BatchNorm2d-74 [-1, 8, 32, 1] 16

ReLU6-75 [-1, 8, 32, 1] 0

h\_sigmoid-76 [-1, 8, 32, 1] 0

h\_swish-77 [-1, 8, 32, 1] 0

Conv2d-78 [-1, 128, 16, 1] 1,152

Conv2d-79 [-1, 128, 1, 16] 1,152

CoordAtt-80 [-1, 128, 16, 16] 0

BatchNorm2d-81 [-1, 128, 16, 16] 256

ReLU-82 [-1, 128, 16, 16] 0

Conv2d-83 [-1, 512, 16, 16] 65,536

BatchNorm2d-84 [-1, 512, 16, 16] 1,024

ReLU-85 [-1, 512, 16, 16] 0

Bottleneck-86 [-1, 512, 16, 16] 0

Conv2d-87 [-1, 256, 16, 16] 131,072

BatchNorm2d-88 [-1, 256, 16, 16] 512

ReLU-89 [-1, 256, 16, 16] 0

Conv2d-90 [-1, 256, 8, 8] 589,824

AdaptiveAvgPool2d-91 [-1, 256, 8, 1] 0

AdaptiveAvgPool2d-92 [-1, 256, 1, 8] 0

Conv2d-93 [-1, 8, 16, 1] 2,056

BatchNorm2d-94 [-1, 8, 16, 1] 16

ReLU6-95 [-1, 8, 16, 1] 0

h\_sigmoid-96 [-1, 8, 16, 1] 0

h\_swish-97 [-1, 8, 16, 1] 0

Conv2d-98 [-1, 256, 8, 1] 2,304

Conv2d-99 [-1, 256, 1, 8] 2,304

CoordAtt-100 [-1, 256, 8, 8] 0

BatchNorm2d-101 [-1, 256, 8, 8] 512

ReLU-102 [-1, 256, 8, 8] 0

Conv2d-103 [-1, 1024, 8, 8] 262,144

BatchNorm2d-104 [-1, 1024, 8, 8] 2,048

Conv2d-105 [-1, 1024, 8, 8] 524,288

BatchNorm2d-106 [-1, 1024, 8, 8] 2,048

ReLU-107 [-1, 1024, 8, 8] 0

Bottleneck-108 [-1, 1024, 8, 8] 0

Conv2d-109 [-1, 256, 8, 8] 262,144

BatchNorm2d-110 [-1, 256, 8, 8] 512

ReLU-111 [-1, 256, 8, 8] 0

Conv2d-112 [-1, 256, 8, 8] 589,824

AdaptiveAvgPool2d-113 [-1, 256, 8, 1] 0

AdaptiveAvgPool2d-114 [-1, 256, 1, 8] 0

Conv2d-115 [-1, 8, 16, 1] 2,056

BatchNorm2d-116 [-1, 8, 16, 1] 16

ReLU6-117 [-1, 8, 16, 1] 0

h\_sigmoid-118 [-1, 8, 16, 1] 0

h\_swish-119 [-1, 8, 16, 1] 0

Conv2d-120 [-1, 256, 8, 1] 2,304

Conv2d-121 [-1, 256, 1, 8] 2,304

CoordAtt-122 [-1, 256, 8, 8] 0

BatchNorm2d-123 [-1, 256, 8, 8] 512

ReLU-124 [-1, 256, 8, 8] 0

Conv2d-125 [-1, 1024, 8, 8] 262,144

BatchNorm2d-126 [-1, 1024, 8, 8] 2,048

ReLU-127 [-1, 1024, 8, 8] 0

Bottleneck-128 [-1, 1024, 8, 8] 0

Conv2d-129 [-1, 512, 8, 8] 524,288

BatchNorm2d-130 [-1, 512, 8, 8] 1,024

ReLU-131 [-1, 512, 8, 8] 0

Conv2d-132 [-1, 512, 4, 4] 2,359,296

AdaptiveAvgPool2d-133 [-1, 512, 4, 1] 0

AdaptiveAvgPool2d-134 [-1, 512, 1, 4] 0

Conv2d-135 [-1, 16, 8, 1] 8,208

BatchNorm2d-136 [-1, 16, 8, 1] 32

ReLU6-137 [-1, 16, 8, 1] 0

h\_sigmoid-138 [-1, 16, 8, 1] 0

h\_swish-139 [-1, 16, 8, 1] 0

Conv2d-140 [-1, 512, 4, 1] 8,704

Conv2d-141 [-1, 512, 1, 4] 8,704

CoordAtt-142 [-1, 512, 4, 4] 0

BatchNorm2d-143 [-1, 512, 4, 4] 1,024

ReLU-144 [-1, 512, 4, 4] 0

Conv2d-145 [-1, 2048, 4, 4] 1,048,576

BatchNorm2d-146 [-1, 2048, 4, 4] 4,096

Conv2d-147 [-1, 2048, 4, 4] 2,097,152

BatchNorm2d-148 [-1, 2048, 4, 4] 4,096

ReLU-149 [-1, 2048, 4, 4] 0

Bottleneck-150 [-1, 2048, 4, 4] 0

Conv2d-151 [-1, 512, 4, 4] 1,048,576

BatchNorm2d-152 [-1, 512, 4, 4] 1,024

ReLU-153 [-1, 512, 4, 4] 0

Conv2d-154 [-1, 512, 4, 4] 2,359,296

AdaptiveAvgPool2d-155 [-1, 512, 4, 1] 0

AdaptiveAvgPool2d-156 [-1, 512, 1, 4] 0

Conv2d-157 [-1, 16, 8, 1] 8,208

BatchNorm2d-158 [-1, 16, 8, 1] 32

ReLU6-159 [-1, 16, 8, 1] 0

h\_sigmoid-160 [-1, 16, 8, 1] 0

h\_swish-161 [-1, 16, 8, 1] 0

Conv2d-162 [-1, 512, 4, 1] 8,704

Conv2d-163 [-1, 512, 1, 4] 8,704

CoordAtt-164 [-1, 512, 4, 4] 0

BatchNorm2d-165 [-1, 512, 4, 4] 1,024

ReLU-166 [-1, 512, 4, 4] 0

Conv2d-167 [-1, 2048, 4, 4] 1,048,576

BatchNorm2d-168 [-1, 2048, 4, 4] 4,096

ReLU-169 [-1, 2048, 4, 4] 0

Bottleneck-170 [-1, 2048, 4, 4] 0

Linear-171 [-1, 100] 204,900

================================================================

Total params: 14,218,132

Trainable params: 14,218,132

Non-trainable params: 0

----------------------------------------------------------------

Input size (MB): 0.01

Forward/backward pass size (MB): 56.14

Params size (MB): 54.24

Estimated Total Size (MB): 110.39

----------------------------------------------------------------

FLOPs = 0.737591296G

Params = 14.218132M

CA+Channel

lr=0.0015

[1, 2000] loss: 4.385,Train accuracy:5.38

[1, 4000] loss: 3.896,Train accuracy:9.69

[1, 6000] loss: 3.614,Train accuracy:14.10

Validation set: Average loss: 1.2624, Accuracy: 1704/10000 (17%)

[2, 2000] loss: 3.349,Train accuracy:18.53

[2, 4000] loss: 3.175,Train accuracy:21.33

[2, 6000] loss: 3.001,Train accuracy:24.81

Validation set: Average loss: 1.9177, Accuracy: 2869/10000 (29%)

[3, 2000] loss: 2.730,Train accuracy:29.41

[3, 4000] loss: 2.605,Train accuracy:32.68

[3, 6000] loss: 2.474,Train accuracy:35.21

Validation set: Average loss: 0.4079, Accuracy: 3693/10000 (37%)

[4, 2000] loss: 2.231,Train accuracy:40.54

[4, 4000] loss: 2.190,Train accuracy:40.88

[4, 6000] loss: 2.166,Train accuracy:41.96

Validation set: Average loss: 0.2852, Accuracy: 4495/10000 (45%)

[5, 2000] loss: 1.889,Train accuracy:47.84

[5, 4000] loss: 1.930,Train accuracy:47.73

[5, 6000] loss: 1.845,Train accuracy:49.18

Validation set: Average loss: 0.3083, Accuracy: 4900/10000 (49%)

[6, 2000] loss: 1.574,Train accuracy:55.63

[6, 4000] loss: 1.644,Train accuracy:54.44

[6, 6000] loss: 1.630,Train accuracy:54.04

Validation set: Average loss: 0.2280, Accuracy: 5146/10000 (51%)

[7, 2000] loss: 1.299,Train accuracy:62.59

[7, 4000] loss: 1.386,Train accuracy:59.59

[7, 6000] loss: 1.428,Train accuracy:59.17

Validation set: Average loss: 0.2693, Accuracy: 5102/10000 (51%)

[8, 2000] loss: 1.066,Train accuracy:67.88

[8, 4000] loss: 1.173,Train accuracy:65.66

[8, 6000] loss: 1.194,Train accuracy:65.31

Validation set: Average loss: 0.2313, Accuracy: 5360/10000 (54%)

[9, 2000] loss: 0.832,Train accuracy:74.94

[9, 4000] loss: 0.933,Train accuracy:71.64

[9, 6000] loss: 0.996,Train accuracy:69.63

Validation set: Average loss: 0.2409, Accuracy: 5426/10000 (54%)

[10, 2000] loss: 0.636,Train accuracy:80.27

[10, 4000] loss: 0.742,Train accuracy:76.63

[10, 6000] loss: 0.820,Train accuracy:74.88

Validation set: Average loss: 0.2451, Accuracy: 5398/10000 (54%)

[11, 2000] loss: 0.461,Train accuracy:85.29

[11, 4000] loss: 0.556,Train accuracy:82.44

[11, 6000] loss: 0.653,Train accuracy:79.70

Validation set: Average loss: 0.2681, Accuracy: 5331/10000 (53%)

[12, 2000] loss: 0.365,Train accuracy:88.28

[12, 4000] loss: 0.427,Train accuracy:86.83

[12, 6000] loss: 0.480,Train accuracy:84.79

Validation set: Average loss: 0.2811, Accuracy: 5322/10000 (53%)

[13, 2000] loss: 0.268,Train accuracy:91.58

[13, 4000] loss: 0.322,Train accuracy:89.68

[13, 6000] loss: 0.389,Train accuracy:87.59

Validation set: Average loss: 0.2925, Accuracy: 5272/10000 (53%)

[14, 2000] loss: 0.205,Train accuracy:93.56

[14, 4000] loss: 0.239,Train accuracy:92.65

[14, 6000] loss: 0.276,Train accuracy:91.52

Validation set: Average loss: 0.2956, Accuracy: 5489/10000 (55%)

[15, 2000] loss: 0.158,Train accuracy:94.99

[15, 4000] loss: 0.194,Train accuracy:93.91

[15, 6000] loss: 0.210,Train accuracy:93.48

Validation set: Average loss: 0.3075, Accuracy: 5430/10000 (54%)

[16, 2000] loss: 0.133,Train accuracy:95.71

[16, 4000] loss: 0.151,Train accuracy:95.34

[16, 6000] loss: 0.172,Train accuracy:94.69

Validation set: Average loss: 0.3119, Accuracy: 5447/10000 (54%)

[17, 2000] loss: 0.113,Train accuracy:96.55

[17, 4000] loss: 0.120,Train accuracy:96.30

[17, 6000] loss: 0.139,Train accuracy:95.64

Validation set: Average loss: 0.3180, Accuracy: 5511/10000 (55%)

[18, 2000] loss: 0.095,Train accuracy:97.04

[18, 4000] loss: 0.108,Train accuracy:96.86

[18, 6000] loss: 0.120,Train accuracy:96.19

Validation set: Average loss: 0.3207, Accuracy: 5443/10000 (54%)

[19, 2000] loss: 0.088,Train accuracy:97.52

[19, 4000] loss: 0.088,Train accuracy:97.46

[19, 6000] loss: 0.095,Train accuracy:97.04

Validation set: Average loss: 0.3224, Accuracy: 5484/10000 (55%)

[20, 2000] loss: 0.066,Train accuracy:98.07

[20, 4000] loss: 0.078,Train accuracy:97.68

[20, 6000] loss: 0.086,Train accuracy:97.24

Validation set: Average loss: 0.3180, Accuracy: 5531/10000 (55%)

[21, 2000] loss: 0.066,Train accuracy:98.01

[21, 4000] loss: 0.070,Train accuracy:97.78

[21, 6000] loss: 0.081,Train accuracy:97.34

Validation set: Average loss: 0.3325, Accuracy: 5510/10000 (55%)

[22, 2000] loss: 0.057,Train accuracy:98.27

[22, 4000] loss: 0.060,Train accuracy:98.38

[22, 6000] loss: 0.069,Train accuracy:97.78

Validation set: Average loss: 0.3297, Accuracy: 5582/10000 (56%)

[23, 2000] loss: 0.047,Train accuracy:98.63

[23, 4000] loss: 0.055,Train accuracy:98.44

[23, 6000] loss: 0.055,Train accuracy:98.30

Validation set: Average loss: 0.3305, Accuracy: 5554/10000 (56%)

[24, 2000] loss: 0.042,Train accuracy:98.86

[24, 4000] loss: 0.046,Train accuracy:98.73

[24, 6000] loss: 0.050,Train accuracy:98.51

Validation set: Average loss: 0.3377, Accuracy: 5566/10000 (56%)

[25, 2000] loss: 0.037,Train accuracy:99.01

[25, 4000] loss: 0.039,Train accuracy:98.88

[25, 6000] loss: 0.042,Train accuracy:98.74

Validation set: Average loss: 0.3404, Accuracy: 5559/10000 (56%)

[26, 2000] loss: 0.028,Train accuracy:99.30

[26, 4000] loss: 0.032,Train accuracy:99.13

[26, 6000] loss: 0.037,Train accuracy:98.96

Validation set: Average loss: 0.3350, Accuracy: 5629/10000 (56%)

[27, 2000] loss: 0.025,Train accuracy:99.24

[27, 4000] loss: 0.028,Train accuracy:99.24

[27, 6000] loss: 0.033,Train accuracy:99.10

Validation set: Average loss: 0.3332, Accuracy: 5646/10000 (56%)

[28, 2000] loss: 0.024,Train accuracy:99.36

[28, 4000] loss: 0.026,Train accuracy:99.36

[28, 6000] loss: 0.027,Train accuracy:99.21

Validation set: Average loss: 0.3359, Accuracy: 5782/10000 (58%)

[29, 2000] loss: 0.021,Train accuracy:99.44

[29, 4000] loss: 0.025,Train accuracy:99.34

[29, 6000] loss: 0.023,Train accuracy:99.34

Validation set: Average loss: 0.3426, Accuracy: 5646/10000 (56%)

[30, 2000] loss: 0.018,Train accuracy:99.55

[30, 4000] loss: 0.021,Train accuracy:99.42

[30, 6000] loss: 0.023,Train accuracy:99.48

Validation set: Average loss: 0.3458, Accuracy: 5637/10000 (56%)

[31, 2000] loss: 0.022,Train accuracy:99.48

[31, 4000] loss: 0.016,Train accuracy:99.62

[31, 6000] loss: 0.016,Train accuracy:99.62

Validation set: Average loss: 0.3411, Accuracy: 5643/10000 (56%)

[32, 2000] loss: 0.015,Train accuracy:99.61

[32, 4000] loss: 0.021,Train accuracy:99.44

[32, 6000] loss: 0.015,Train accuracy:99.66

Validation set: Average loss: 0.3380, Accuracy: 5710/10000 (57%)

[33, 2000] loss: 0.015,Train accuracy:99.59

[33, 4000] loss: 0.014,Train accuracy:99.64

[33, 6000] loss: 0.016,Train accuracy:99.59

Validation set: Average loss: 0.3442, Accuracy: 5664/10000 (57%)

[34, 2000] loss: 0.012,Train accuracy:99.74

[34, 4000] loss: 0.015,Train accuracy:99.55

[34, 6000] loss: 0.015,Train accuracy:99.61

Validation set: Average loss: 0.3429, Accuracy: 5642/10000 (56%)

[35, 2000] loss: 0.009,Train accuracy:99.79

[35, 4000] loss: 0.012,Train accuracy:99.76

[35, 6000] loss: 0.011,Train accuracy:99.77

Validation set: Average loss: 0.3456, Accuracy: 5738/10000 (57%)

[36, 2000] loss: 0.008,Train accuracy:99.88

[36, 4000] loss: 0.009,Train accuracy:99.79

[36, 6000] loss: 0.012,Train accuracy:99.73

Validation set: Average loss: 0.3444, Accuracy: 5678/10000 (57%)

[37, 2000] loss: 0.010,Train accuracy:99.78

[37, 4000] loss: 0.008,Train accuracy:99.85

[37, 6000] loss: 0.011,Train accuracy:99.73

Validation set: Average loss: 0.3429, Accuracy: 5667/10000 (57%)

[38, 2000] loss: 0.008,Train accuracy:99.83

[38, 4000] loss: 0.010,Train accuracy:99.81

[38, 6000] loss: 0.010,Train accuracy:99.76

Validation set: Average loss: 0.3467, Accuracy: 5694/10000 (57%)

[39, 2000] loss: 0.007,Train accuracy:99.81

[39, 4000] loss: 0.008,Train accuracy:99.80

[39, 6000] loss: 0.010,Train accuracy:99.74

Validation set: Average loss: 0.3460, Accuracy: 5638/10000 (56%)

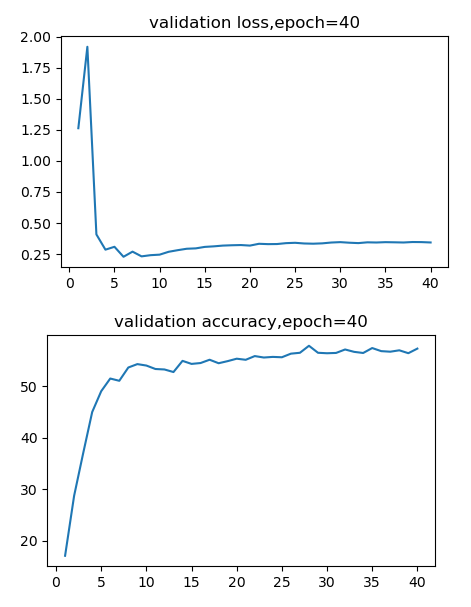
[40, 2000] loss: 0.007,Train accuracy:99.88

[40, 4000] loss: 0.007,Train accuracy:99.90

[40, 6000] loss: 0.009,Train accuracy:99.84

Validation set: Average loss: 0.3431, Accuracy: 5727/10000 (57%)

training finished！



moment=0.95

[1, 2000] loss: 4.434,Train accuracy:4.37

[1, 4000] loss: 3.972,Train accuracy:8.37

[1, 6000] loss: 3.722,Train accuracy:12.11

Validation set: Average loss: 1.1845, Accuracy: 1549/10000 (15%)

[2, 2000] loss: 3.477,Train accuracy:16.48

[2, 4000] loss: 3.361,Train accuracy:18.13

[2, 6000] loss: 3.168,Train accuracy:21.73

Validation set: Average loss: 1.0719, Accuracy: 2602/10000 (26%)

Number of parameter: 18.40M

[3, 2000] loss: 2.909,Train accuracy:26.48

[3, 4000] loss: 2.817,Train accuracy:28.74

[3, 6000] loss: 2.689,Train accuracy:31.08

Validation set: Average loss: 0.5398, Accuracy: 3471/10000 (35%)

[4, 2000] loss: 2.456,Train accuracy:35.84

[4, 4000] loss: 2.422,Train accuracy:36.70

[4, 6000] loss: 2.315,Train accuracy:38.73

Validation set: Average loss: 3.4685, Accuracy: 4036/10000 (40%)

[5, 2000] loss: 2.102,Train accuracy:43.33

[5, 4000] loss: 2.065,Train accuracy:43.93

[5, 6000] loss: 2.041,Train accuracy:44.81

Validation set: Average loss: 0.3348, Accuracy: 4576/10000 (46%)

[6, 2000] loss: 1.779,Train accuracy:50.63

[6, 4000] loss: 1.809,Train accuracy:49.39

[6, 6000] loss: 1.793,Train accuracy:50.16

Validation set: Average loss: 0.2464, Accuracy: 4841/10000 (48%)

[7, 2000] loss: 1.519,Train accuracy:56.95

[7, 4000] loss: 1.550,Train accuracy:56.34

[7, 6000] loss: 1.563,Train accuracy:56.13

Validation set: Average loss: 0.2577, Accuracy: 5050/10000 (50%)

[8, 2000] loss: 1.260,Train accuracy:63.51

[8, 4000] loss: 1.337,Train accuracy:60.91

[8, 6000] loss: 1.359,Train accuracy:60.97

Validation set: Average loss: 0.2647, Accuracy: 5151/10000 (52%)

[9, 2000] loss: 1.017,Train accuracy:69.24

[9, 4000] loss: 1.117,Train accuracy:66.51

[9, 6000] loss: 1.166,Train accuracy:65.91

Validation set: Average loss: 0.2531, Accuracy: 5241/10000 (52%)

[10, 2000] loss: 0.817,Train accuracy:75.01

[10, 4000] loss: 0.932,Train accuracy:71.29

[10, 6000] loss: 0.957,Train accuracy:71.24

Validation set: Average loss: 0.2615, Accuracy: 5344/10000 (53%)

[11, 2000] loss: 0.637,Train accuracy:79.78

[11, 4000] loss: 0.748,Train accuracy:76.78

[11, 6000] loss: 0.773,Train accuracy:75.93

Validation set: Average loss: 0.2467, Accuracy: 5356/10000 (54%)

[12, 2000] loss: 0.498,Train accuracy:84.11

[12, 4000] loss: 0.582,Train accuracy:81.46

[12, 6000] loss: 0.611,Train accuracy:80.42

Validation set: Average loss: 0.2726, Accuracy: 5283/10000 (53%)

[13, 2000] loss: 0.375,Train accuracy:87.87

[13, 4000] loss: 0.448,Train accuracy:85.60

[13, 6000] loss: 0.488,Train accuracy:84.37

Validation set: Average loss: 0.2695, Accuracy: 5399/10000 (54%)

[14, 2000] loss: 0.290,Train accuracy:90.70

[14, 4000] loss: 0.332,Train accuracy:89.29

[14, 6000] loss: 0.378,Train accuracy:87.88

Validation set: Average loss: 0.2905, Accuracy: 5356/10000 (54%)

[15, 2000] loss: 0.220,Train accuracy:92.82

[15, 4000] loss: 0.261,Train accuracy:91.69

[15, 6000] loss: 0.298,Train accuracy:90.39

Validation set: Average loss: 0.3296, Accuracy: 5250/10000 (52%)

[16, 2000] loss: 0.171,Train accuracy:94.77

[16, 4000] loss: 0.202,Train accuracy:93.49

[16, 6000] loss: 0.237,Train accuracy:92.44

Validation set: Average loss: 0.3255, Accuracy: 5227/10000 (52%)

[17, 2000] loss: 0.153,Train accuracy:95.08

[17, 4000] loss: 0.173,Train accuracy:94.34

[17, 6000] loss: 0.188,Train accuracy:93.84

Validation set: Average loss: 0.3157, Accuracy: 5273/10000 (53%)

[18, 2000] loss: 0.130,Train accuracy:95.85

[18, 4000] loss: 0.145,Train accuracy:95.29

[18, 6000] loss: 0.150,Train accuracy:95.29

Validation set: Average loss: 0.3579, Accuracy: 5322/10000 (53%)

[19, 2000] loss: 0.099,Train accuracy:97.08

[19, 4000] loss: 0.117,Train accuracy:96.26

[19, 6000] loss: 0.116,Train accuracy:96.33

Validation set: Average loss: 0.3417, Accuracy: 5336/10000 (53%)

[20, 2000] loss: 0.068,Train accuracy:97.98

[20, 4000] loss: 0.081,Train accuracy:97.38

[20, 6000] loss: 0.100,Train accuracy:96.91

Validation set: Average loss: 0.3376, Accuracy: 5395/10000 (54%)

[21, 2000] loss: 0.065,Train accuracy:98.03

[21, 4000] loss: 0.079,Train accuracy:97.42

[21, 6000] loss: 0.087,Train accuracy:97.28

Validation set: Average loss: 0.3624, Accuracy: 5355/10000 (54%)

[22, 2000] loss: 0.054,Train accuracy:98.46

[22, 4000] loss: 0.062,Train accuracy:98.27

[22, 6000] loss: 0.070,Train accuracy:97.96

Validation set: Average loss: 0.3633, Accuracy: 5366/10000 (54%)

[23, 2000] loss: 0.048,Train accuracy:98.66

[23, 4000] loss: 0.051,Train accuracy:98.39

[23, 6000] loss: 0.056,Train accuracy:98.26

Validation set: Average loss: 0.3516, Accuracy: 5375/10000 (54%)

[24, 2000] loss: 0.046,Train accuracy:98.77

[24, 4000] loss: 0.047,Train accuracy:98.62

[24, 6000] loss: 0.052,Train accuracy:98.59

Validation set: Average loss: 0.3476, Accuracy: 5536/10000 (55%)

[25, 2000] loss: 0.037,Train accuracy:98.96

[25, 4000] loss: 0.046,Train accuracy:98.52

[25, 6000] loss: 0.045,Train accuracy:98.61

Validation set: Average loss: 0.3513, Accuracy: 5462/10000 (55%)

[26, 2000] loss: 0.043,Train accuracy:98.64

[26, 4000] loss: 0.041,Train accuracy:98.66

[26, 6000] loss: 0.044,Train accuracy:98.80

Validation set: Average loss: 0.3727, Accuracy: 5360/10000 (54%)

[27, 2000] loss: 0.029,Train accuracy:99.19

[27, 4000] loss: 0.034,Train accuracy:99.13

[27, 6000] loss: 0.034,Train accuracy:99.06

Validation set: Average loss: 0.3665, Accuracy: 5527/10000 (55%)

[28, 2000] loss: 0.026,Train accuracy:99.18

[28, 4000] loss: 0.027,Train accuracy:99.25

[28, 6000] loss: 0.031,Train accuracy:99.19

Validation set: Average loss: 0.3885, Accuracy: 5462/10000 (55%)

[29, 2000] loss: 0.028,Train accuracy:99.16

[29, 4000] loss: 0.027,Train accuracy:99.19

[29, 6000] loss: 0.032,Train accuracy:99.04

Validation set: Average loss: 0.3822, Accuracy: 5411/10000 (54%)

[30, 2000] loss: 0.021,Train accuracy:99.41

[30, 4000] loss: 0.025,Train accuracy:99.29

[30, 6000] loss: 0.029,Train accuracy:99.08

Validation set: Average loss: 0.3744, Accuracy: 5462/10000 (55%)

[31, 2000] loss: 0.021,Train accuracy:99.45

[31, 4000] loss: 0.024,Train accuracy:99.38

[31, 6000] loss: 0.024,Train accuracy:99.35

Validation set: Average loss: 0.3621, Accuracy: 5472/10000 (55%)

[32, 2000] loss: 0.023,Train accuracy:99.41

[32, 4000] loss: 0.017,Train accuracy:99.57

[32, 6000] loss: 0.019,Train accuracy:99.47

Validation set: Average loss: 0.4053, Accuracy: 5416/10000 (54%)

[33, 2000] loss: 0.020,Train accuracy:99.48

[33, 4000] loss: 0.022,Train accuracy:99.41

[33, 6000] loss: 0.020,Train accuracy:99.49

Validation set: Average loss: 0.3743, Accuracy: 5484/10000 (55%)

[34, 2000] loss: 0.015,Train accuracy:99.63

[34, 4000] loss: 0.017,Train accuracy:99.51

[34, 6000] loss: 0.017,Train accuracy:99.58

Validation set: Average loss: 0.3973, Accuracy: 5476/10000 (55%)

[35, 2000] loss: 0.016,Train accuracy:99.56

[35, 4000] loss: 0.015,Train accuracy:99.56

[35, 6000] loss: 0.016,Train accuracy:99.66

Validation set: Average loss: 0.3716, Accuracy: 5553/10000 (56%)

[36, 2000] loss: 0.010,Train accuracy:99.74

[36, 4000] loss: 0.013,Train accuracy:99.63

[36, 6000] loss: 0.017,Train accuracy:99.56

Validation set: Average loss: 0.3785, Accuracy: 5459/10000 (55%)

[37, 2000] loss: 0.013,Train accuracy:99.68

[37, 4000] loss: 0.011,Train accuracy:99.75

[37, 6000] loss: 0.016,Train accuracy:99.60

Validation set: Average loss: 0.3937, Accuracy: 5498/10000 (55%)

[38, 2000] loss: 0.010,Train accuracy:99.73

[38, 4000] loss: 0.013,Train accuracy:99.68

[38, 6000] loss: 0.016,Train accuracy:99.54

Validation set: Average loss: 0.3724, Accuracy: 5561/10000 (56%)

[39, 2000] loss: 0.009,Train accuracy:99.83

[39, 4000] loss: 0.011,Train accuracy:99.79

[39, 6000] loss: 0.013,Train accuracy:99.64

Validation set: Average loss: 0.3678, Accuracy: 5582/10000 (56%)

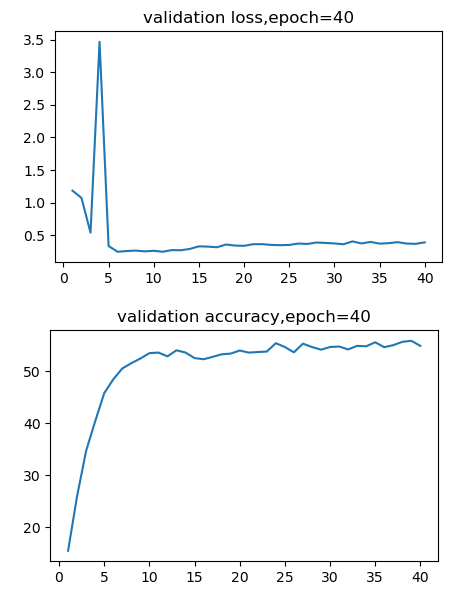
[40, 2000] loss: 0.010,Train accuracy:99.69

[40, 4000] loss: 0.010,Train accuracy:99.75

[40, 6000] loss: 0.012,Train accuracy:99.64

Validation set: Average loss: 0.3913, Accuracy: 5484/10000 (55%)

training finished！



----------------------------------------------------------------

Layer (type) Output Shape Param #

================================================================

Conv2d-1 [-1, 64, 32, 32] 1,728

BatchNorm2d-2 [-1, 64, 32, 32] 128

Conv2d-3 [-1, 64, 32, 32] 4,096

BatchNorm2d-4 [-1, 64, 32, 32] 128

ReLU-5 [-1, 64, 32, 32] 0

Conv2d-6 [-1, 64, 32, 32] 36,864

AdaptiveAvgPool2d-7 [-1, 64, 32, 1] 0

AdaptiveAvgPool2d-8 [-1, 64, 1, 32] 0

Conv2d-9 [-1, 8, 64, 1] 520

BatchNorm2d-10 [-1, 8, 64, 1] 16

ReLU6-11 [-1, 8, 64, 1] 0

h\_sigmoid-12 [-1, 8, 64, 1] 0

h\_swish-13 [-1, 8, 64, 1] 0

Conv2d-14 [-1, 64, 32, 1] 576

Conv2d-15 [-1, 64, 1, 32] 576

AdaptiveAvgPool2d-16 [-1, 64, 1, 1] 0

Linear-17 [-1, 2] 130

ReLU-18 [-1, 2] 0

Linear-19 [-1, 64] 192

Sigmoid-20 [-1, 64] 0

ChannelAtt-21 [-1, 64, 32, 32] 0

CoordAtt-22 [-1, 64, 32, 32] 0

BatchNorm2d-23 [-1, 64, 32, 32] 128

ReLU-24 [-1, 64, 32, 32] 0

Conv2d-25 [-1, 256, 32, 32] 16,384

BatchNorm2d-26 [-1, 256, 32, 32] 512

Conv2d-27 [-1, 256, 32, 32] 16,384

BatchNorm2d-28 [-1, 256, 32, 32] 512

ReLU-29 [-1, 256, 32, 32] 0

Bottleneck-30 [-1, 256, 32, 32] 0

Conv2d-31 [-1, 64, 32, 32] 16,384

BatchNorm2d-32 [-1, 64, 32, 32] 128

ReLU-33 [-1, 64, 32, 32] 0

Conv2d-34 [-1, 64, 32, 32] 36,864

AdaptiveAvgPool2d-35 [-1, 64, 32, 1] 0

AdaptiveAvgPool2d-36 [-1, 64, 1, 32] 0

Conv2d-37 [-1, 8, 64, 1] 520

BatchNorm2d-38 [-1, 8, 64, 1] 16

ReLU6-39 [-1, 8, 64, 1] 0

h\_sigmoid-40 [-1, 8, 64, 1] 0

h\_swish-41 [-1, 8, 64, 1] 0

Conv2d-42 [-1, 64, 32, 1] 576

Conv2d-43 [-1, 64, 1, 32] 576

AdaptiveAvgPool2d-44 [-1, 64, 1, 1] 0

Linear-45 [-1, 2] 130

ReLU-46 [-1, 2] 0

Linear-47 [-1, 64] 192

Sigmoid-48 [-1, 64] 0

ChannelAtt-49 [-1, 64, 32, 32] 0

CoordAtt-50 [-1, 64, 32, 32] 0

BatchNorm2d-51 [-1, 64, 32, 32] 128

ReLU-52 [-1, 64, 32, 32] 0

Conv2d-53 [-1, 256, 32, 32] 16,384

BatchNorm2d-54 [-1, 256, 32, 32] 512

ReLU-55 [-1, 256, 32, 32] 0

Bottleneck-56 [-1, 256, 32, 32] 0

Conv2d-57 [-1, 128, 32, 32] 32,768

BatchNorm2d-58 [-1, 128, 32, 32] 256

ReLU-59 [-1, 128, 32, 32] 0

Conv2d-60 [-1, 128, 16, 16] 147,456

AdaptiveAvgPool2d-61 [-1, 128, 16, 1] 0

AdaptiveAvgPool2d-62 [-1, 128, 1, 16] 0

Conv2d-63 [-1, 8, 32, 1] 1,032

BatchNorm2d-64 [-1, 8, 32, 1] 16

ReLU6-65 [-1, 8, 32, 1] 0

h\_sigmoid-66 [-1, 8, 32, 1] 0

h\_swish-67 [-1, 8, 32, 1] 0

Conv2d-68 [-1, 128, 16, 1] 1,152

Conv2d-69 [-1, 128, 1, 16] 1,152

AdaptiveAvgPool2d-70 [-1, 128, 1, 1] 0

Linear-71 [-1, 4] 516

ReLU-72 [-1, 4] 0

Linear-73 [-1, 128] 640

Sigmoid-74 [-1, 128] 0

ChannelAtt-75 [-1, 128, 16, 16] 0

CoordAtt-76 [-1, 128, 16, 16] 0

BatchNorm2d-77 [-1, 128, 16, 16] 256

ReLU-78 [-1, 128, 16, 16] 0

Conv2d-79 [-1, 512, 16, 16] 65,536

BatchNorm2d-80 [-1, 512, 16, 16] 1,024

Conv2d-81 [-1, 512, 16, 16] 131,072

BatchNorm2d-82 [-1, 512, 16, 16] 1,024

ReLU-83 [-1, 512, 16, 16] 0

Bottleneck-84 [-1, 512, 16, 16] 0

Conv2d-85 [-1, 128, 16, 16] 65,536

BatchNorm2d-86 [-1, 128, 16, 16] 256

ReLU-87 [-1, 128, 16, 16] 0

Conv2d-88 [-1, 128, 16, 16] 147,456

AdaptiveAvgPool2d-89 [-1, 128, 16, 1] 0

AdaptiveAvgPool2d-90 [-1, 128, 1, 16] 0

Conv2d-91 [-1, 8, 32, 1] 1,032

BatchNorm2d-92 [-1, 8, 32, 1] 16

ReLU6-93 [-1, 8, 32, 1] 0

h\_sigmoid-94 [-1, 8, 32, 1] 0

h\_swish-95 [-1, 8, 32, 1] 0

Conv2d-96 [-1, 128, 16, 1] 1,152

Conv2d-97 [-1, 128, 1, 16] 1,152

AdaptiveAvgPool2d-98 [-1, 128, 1, 1] 0

Linear-99 [-1, 4] 516

ReLU-100 [-1, 4] 0

Linear-101 [-1, 128] 640

Sigmoid-102 [-1, 128] 0

ChannelAtt-103 [-1, 128, 16, 16] 0

CoordAtt-104 [-1, 128, 16, 16] 0

BatchNorm2d-105 [-1, 128, 16, 16] 256

ReLU-106 [-1, 128, 16, 16] 0

Conv2d-107 [-1, 512, 16, 16] 65,536

BatchNorm2d-108 [-1, 512, 16, 16] 1,024

ReLU-109 [-1, 512, 16, 16] 0

Bottleneck-110 [-1, 512, 16, 16] 0

Conv2d-111 [-1, 256, 16, 16] 131,072

BatchNorm2d-112 [-1, 256, 16, 16] 512

ReLU-113 [-1, 256, 16, 16] 0

Conv2d-114 [-1, 256, 8, 8] 589,824

AdaptiveAvgPool2d-115 [-1, 256, 8, 1] 0

AdaptiveAvgPool2d-116 [-1, 256, 1, 8] 0

Conv2d-117 [-1, 8, 16, 1] 2,056

BatchNorm2d-118 [-1, 8, 16, 1] 16

ReLU6-119 [-1, 8, 16, 1] 0

h\_sigmoid-120 [-1, 8, 16, 1] 0

h\_swish-121 [-1, 8, 16, 1] 0

Conv2d-122 [-1, 256, 8, 1] 2,304

Conv2d-123 [-1, 256, 1, 8] 2,304

AdaptiveAvgPool2d-124 [-1, 256, 1, 1] 0

Linear-125 [-1, 8] 2,056

ReLU-126 [-1, 8] 0

Linear-127 [-1, 256] 2,304

Sigmoid-128 [-1, 256] 0

ChannelAtt-129 [-1, 256, 8, 8] 0

CoordAtt-130 [-1, 256, 8, 8] 0

BatchNorm2d-131 [-1, 256, 8, 8] 512

ReLU-132 [-1, 256, 8, 8] 0

Conv2d-133 [-1, 1024, 8, 8] 262,144

BatchNorm2d-134 [-1, 1024, 8, 8] 2,048

Conv2d-135 [-1, 1024, 8, 8] 524,288

BatchNorm2d-136 [-1, 1024, 8, 8] 2,048

ReLU-137 [-1, 1024, 8, 8] 0

Bottleneck-138 [-1, 1024, 8, 8] 0

Conv2d-139 [-1, 256, 8, 8] 262,144

BatchNorm2d-140 [-1, 256, 8, 8] 512

ReLU-141 [-1, 256, 8, 8] 0

Conv2d-142 [-1, 256, 8, 8] 589,824

AdaptiveAvgPool2d-143 [-1, 256, 8, 1] 0

AdaptiveAvgPool2d-144 [-1, 256, 1, 8] 0

Conv2d-145 [-1, 8, 16, 1] 2,056

BatchNorm2d-146 [-1, 8, 16, 1] 16

ReLU6-147 [-1, 8, 16, 1] 0

h\_sigmoid-148 [-1, 8, 16, 1] 0

h\_swish-149 [-1, 8, 16, 1] 0

Conv2d-150 [-1, 256, 8, 1] 2,304

Conv2d-151 [-1, 256, 1, 8] 2,304

AdaptiveAvgPool2d-152 [-1, 256, 1, 1] 0

Linear-153 [-1, 8] 2,056

ReLU-154 [-1, 8] 0

Linear-155 [-1, 256] 2,304

Sigmoid-156 [-1, 256] 0

ChannelAtt-157 [-1, 256, 8, 8] 0

CoordAtt-158 [-1, 256, 8, 8] 0

BatchNorm2d-159 [-1, 256, 8, 8] 512

ReLU-160 [-1, 256, 8, 8] 0

Conv2d-161 [-1, 1024, 8, 8] 262,144

BatchNorm2d-162 [-1, 1024, 8, 8] 2,048

ReLU-163 [-1, 1024, 8, 8] 0

Bottleneck-164 [-1, 1024, 8, 8] 0

Conv2d-165 [-1, 512, 8, 8] 524,288

BatchNorm2d-166 [-1, 512, 8, 8] 1,024

ReLU-167 [-1, 512, 8, 8] 0

Conv2d-168 [-1, 512, 4, 4] 2,359,296

AdaptiveAvgPool2d-169 [-1, 512, 4, 1] 0

AdaptiveAvgPool2d-170 [-1, 512, 1, 4] 0

Conv2d-171 [-1, 16, 8, 1] 8,208

BatchNorm2d-172 [-1, 16, 8, 1] 32

ReLU6-173 [-1, 16, 8, 1] 0

h\_sigmoid-174 [-1, 16, 8, 1] 0

h\_swish-175 [-1, 16, 8, 1] 0

Conv2d-176 [-1, 512, 4, 1] 8,704

Conv2d-177 [-1, 512, 1, 4] 8,704

AdaptiveAvgPool2d-178 [-1, 512, 1, 1] 0

Linear-179 [-1, 16] 8,208

ReLU-180 [-1, 16] 0

Linear-181 [-1, 512] 8,704

Sigmoid-182 [-1, 512] 0

ChannelAtt-183 [-1, 512, 4, 4] 0

CoordAtt-184 [-1, 512, 4, 4] 0

BatchNorm2d-185 [-1, 512, 4, 4] 1,024

ReLU-186 [-1, 512, 4, 4] 0

Conv2d-187 [-1, 2048, 4, 4] 1,048,576

BatchNorm2d-188 [-1, 2048, 4, 4] 4,096

Conv2d-189 [-1, 2048, 4, 4] 2,097,152

BatchNorm2d-190 [-1, 2048, 4, 4] 4,096

ReLU-191 [-1, 2048, 4, 4] 0

Bottleneck-192 [-1, 2048, 4, 4] 0

Conv2d-193 [-1, 512, 4, 4] 1,048,576

BatchNorm2d-194 [-1, 512, 4, 4] 1,024

ReLU-195 [-1, 512, 4, 4] 0

Conv2d-196 [-1, 512, 4, 4] 2,359,296

AdaptiveAvgPool2d-197 [-1, 512, 4, 1] 0

AdaptiveAvgPool2d-198 [-1, 512, 1, 4] 0

Conv2d-199 [-1, 16, 8, 1] 8,208

BatchNorm2d-200 [-1, 16, 8, 1] 32

ReLU6-201 [-1, 16, 8, 1] 0

h\_sigmoid-202 [-1, 16, 8, 1] 0

h\_swish-203 [-1, 16, 8, 1] 0

Conv2d-204 [-1, 512, 4, 1] 8,704

Conv2d-205 [-1, 512, 1, 4] 8,704

AdaptiveAvgPool2d-206 [-1, 512, 1, 1] 0

Linear-207 [-1, 16] 8,208

ReLU-208 [-1, 16] 0

Linear-209 [-1, 512] 8,704

Sigmoid-210 [-1, 512] 0

ChannelAtt-211 [-1, 512, 4, 4] 0

CoordAtt-212 [-1, 512, 4, 4] 0

BatchNorm2d-213 [-1, 512, 4, 4] 1,024

ReLU-214 [-1, 512, 4, 4] 0

Conv2d-215 [-1, 2048, 4, 4] 1,048,576

BatchNorm2d-216 [-1, 2048, 4, 4] 4,096

ReLU-217 [-1, 2048, 4, 4] 0

Bottleneck-218 [-1, 2048, 4, 4] 0

Linear-219 [-1, 100] 204,900

================================================================

Total params: 14,263,632

Trainable params: 14,263,632

Non-trainable params: 0

----------------------------------------------------------------

Input size (MB): 0.01

Forward/backward pass size (MB): 58.06

Params size (MB): 54.41

Estimated Total Size (MB): 112.48

----------------------------------------------------------------

FLOPs = 8.116707456G

Params = 14.263632M