========== Training ==========

Epoch: [1][1/384] Data 0.854 (0.854) Elapsed 0m 0s (remain 5m 59s) Loss: 0.6539(0.6539) Grad: 37.0500

Epoch: [1][51/384] Data 0.000 (0.069) Elapsed 0m 6s (remain 0m 41s) Loss: 0.3489(0.5665) Grad: 27.2237

Epoch: [1][101/384] Data 0.173 (0.062) Elapsed 0m 11s (remain 0m 33s) Loss: 0.5882(0.5585) Grad: 30.8909

Epoch: [1][151/384] Data 0.000 (0.056) Elapsed 0m 16s (remain 0m 25s) Loss: 0.4974(0.5340) Grad: 36.5368

Epoch: [1][201/384] Data 0.291 (0.055) Elapsed 0m 22s (remain 0m 20s) Loss: 0.5779(0.5186) Grad: 34.1893

Epoch: [1][251/384] Data 0.000 (0.055) Elapsed 0m 27s (remain 0m 14s) Loss: 0.3990(0.5091) Grad: 31.1121

Epoch: [1][301/384] Data 0.223 (0.056) Elapsed 0m 33s (remain 0m 9s) Loss: 0.2693(0.4960) Grad: 24.8492

Epoch: [1][351/384] Data 0.000 (0.053) Elapsed 0m 38s (remain 0m 3s) Loss: 0.1823(0.4852) Grad: 22.1760

Epoch: [1][384/384] Data 0.000 (0.053) Elapsed 0m 41s (remain 0m 0s) Loss: 0.3251(0.4797) Grad: 25.8214

EVAL: [1/97] Data 0.576 (0.576) Elapsed 0m 0s (remain 0m 56s) Loss: 0.5758(0.5758)

EVAL: [51/97] Data 0.000 (0.088) Elapsed 0m 5s (remain 0m 4s) Loss: 0.2708(0.4205)

EVAL: [97/97] Data 0.000 (0.084) Elapsed 0m 9s (remain 0m 0s) Loss: 0.0947(0.3958)

Epoch {1} - avg\_train\_loss: {0.4797} avg\_val\_loss: {0.3958} time: {51}s

Epoch {1} - Accuracy(validation): {0.84590}

Epoch {1} - Save Best Score: {0.8459} Model

Epoch: [2][1/384] Data 0.265 (0.265) Elapsed 0m 0s (remain 2m 7s) Loss: 0.4126(0.4126) Grad: 19.9719

Epoch: [2][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 20s) Loss: 0.2502(0.3848) Grad: 12.6762

Epoch: [2][101/384] Data 0.000 (0.003) Elapsed 0m 5s (remain 0m 16s) Loss: 0.4241(0.3782) Grad: 24.9954

Epoch: [2][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.2644(0.3727) Grad: 23.0005

Epoch: [2][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.4105(0.3674) Grad: 28.9922

Epoch: [2][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.2650(0.3613) Grad: 19.4480

Epoch: [2][301/384] Data 0.000 (0.001) Elapsed 0m 16s (remain 0m 4s) Loss: 0.1587(0.3564) Grad: 15.4089

Epoch: [2][351/384] Data 0.000 (0.001) Elapsed 0m 19s (remain 0m 1s) Loss: 0.1474(0.3488) Grad: 17.2613

Epoch: [2][384/384] Data 0.000 (0.001) Elapsed 0m 21s (remain 0m 0s) Loss: 0.1499(0.3435) Grad: 17.0750

EVAL: [1/97] Data 0.267 (0.267) Elapsed 0m 0s (remain 0m 27s) Loss: 0.6175(0.6175)

EVAL: [51/97] Data 0.000 (0.014) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1679(0.4002)

EVAL: [97/97] Data 0.000 (0.012) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0303(0.3672)

Epoch {2} - avg\_train\_loss: {0.3435} avg\_val\_loss: {0.3672} time: {24}s

Epoch {2} - Accuracy(validation): {0.85436}

Epoch {2} - Save Best Score: {0.8544} Model

Epoch: [3][1/384] Data 0.272 (0.272) Elapsed 0m 0s (remain 2m 5s) Loss: 0.3159(0.3159) Grad: 27.0951

Epoch: [3][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 20s) Loss: 0.1838(0.2755) Grad: 15.8821

Epoch: [3][101/384] Data 0.000 (0.003) Elapsed 0m 5s (remain 0m 16s) Loss: 0.2965(0.2582) Grad: 23.7920

Epoch: [3][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.1190(0.2475) Grad: 12.1172

Epoch: [3][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.2927(0.2411) Grad: 39.0243

Epoch: [3][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.1627(0.2343) Grad: 19.6258

Epoch: [3][301/384] Data 0.000 (0.001) Elapsed 0m 16s (remain 0m 4s) Loss: 0.0769(0.2284) Grad: 8.3787

Epoch: [3][351/384] Data 0.000 (0.001) Elapsed 0m 19s (remain 0m 1s) Loss: 0.0703(0.2188) Grad: 9.5818

Epoch: [3][384/384] Data 0.000 (0.001) Elapsed 0m 21s (remain 0m 0s) Loss: 0.0771(0.2122) Grad: 15.9209

EVAL: [1/97] Data 0.277 (0.277) Elapsed 0m 0s (remain 0m 28s) Loss: 0.8037(0.8037)

EVAL: [51/97] Data 0.000 (0.014) Elapsed 0m 1s (remain 0m 1s) Loss: 0.0811(0.4099)

EVAL: [97/97] Data 0.000 (0.011) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0129(0.3748)

Epoch {3} - avg\_train\_loss: {0.2122} avg\_val\_loss: {0.3748} time: {24}s

Epoch {3} - Accuracy(validation): {0.85956}

Epoch {3} - Save Best Score: {0.8596} Model

Epoch: [4][1/384] Data 0.259 (0.259) Elapsed 0m 0s (remain 2m 4s) Loss: 0.1591(0.1591) Grad: 18.4934

Epoch: [4][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 20s) Loss: 0.1161(0.1585) Grad: 15.0249

Epoch: [4][101/384] Data 0.000 (0.003) Elapsed 0m 5s (remain 0m 16s) Loss: 0.2062(0.1472) Grad: 24.1750

Epoch: [4][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0504(0.1410) Grad: 13.8438

Epoch: [4][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.1146(0.1328) Grad: 15.6024

Epoch: [4][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0494(0.1250) Grad: 8.2484

Epoch: [4][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0410(0.1206) Grad: 7.2590

Epoch: [4][351/384] Data 0.000 (0.001) Elapsed 0m 19s (remain 0m 1s) Loss: 0.0387(0.1151) Grad: 8.8754

Epoch: [4][384/384] Data 0.000 (0.001) Elapsed 0m 21s (remain 0m 0s) Loss: 0.0515(0.1104) Grad: 11.2418

EVAL: [1/97] Data 0.259 (0.259) Elapsed 0m 0s (remain 0m 26s) Loss: 0.8502(0.8502)

EVAL: [51/97] Data 0.000 (0.013) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1374(0.4401)

EVAL: [97/97] Data 0.000 (0.011) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0061(0.4000)

Epoch {4} - avg\_train\_loss: {0.1104} avg\_val\_loss: {0.4000} time: {24}s

Epoch {4} - Accuracy(validation): {0.85371}

Epoch: [5][1/384] Data 0.282 (0.282) Elapsed 0m 0s (remain 2m 13s) Loss: 0.0971(0.0971) Grad: 17.1453

Epoch: [5][51/384] Data 0.000 (0.006) Elapsed 0m 3s (remain 0m 20s) Loss: 0.0784(0.1050) Grad: 12.3765

Epoch: [5][101/384] Data 0.000 (0.003) Elapsed 0m 5s (remain 0m 16s) Loss: 0.1333(0.1070) Grad: 21.2819

Epoch: [5][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0499(0.1183) Grad: 11.9689

Epoch: [5][201/384] Data 0.000 (0.002) Elapsed 0m 11s (remain 0m 10s) Loss: 0.1732(0.1198) Grad: 37.2388

Epoch: [5][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0552(0.1159) Grad: 11.9107

Epoch: [5][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0401(0.1090) Grad: 8.1699

Epoch: [5][351/384] Data 0.000 (0.001) Elapsed 0m 19s (remain 0m 1s) Loss: 0.0350(0.1048) Grad: 6.6974

Epoch: [5][384/384] Data 0.000 (0.001) Elapsed 0m 21s (remain 0m 0s) Loss: 0.0421(0.1006) Grad: 10.8379

EVAL: [1/97] Data 0.274 (0.274) Elapsed 0m 0s (remain 0m 27s) Loss: 0.8623(0.8623)

EVAL: [51/97] Data 0.001 (0.011) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1177(0.4551)

EVAL: [97/97] Data 0.000 (0.009) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0041(0.4063)

Epoch {5} - avg\_train\_loss: {0.1006} avg\_val\_loss: {0.4063} time: {24}s

Epoch {5} - Accuracy(validation): {0.85696}

Epoch: [6][1/384] Data 0.285 (0.285) Elapsed 0m 0s (remain 2m 11s) Loss: 0.0537(0.0537) Grad: 9.0710

Epoch: [6][51/384] Data 0.000 (0.006) Elapsed 0m 3s (remain 0m 20s) Loss: 0.0451(0.0580) Grad: 8.9660

Epoch: [6][101/384] Data 0.000 (0.003) Elapsed 0m 5s (remain 0m 16s) Loss: 0.0560(0.0509) Grad: 10.2717

Epoch: [6][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0076(0.0468) Grad: 2.5455

Epoch: [6][201/384] Data 0.000 (0.002) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0264(0.0431) Grad: 6.3256

Epoch: [6][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0123(0.0402) Grad: 3.2726

Epoch: [6][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0110(0.0376) Grad: 3.0974

Epoch: [6][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0113(0.0352) Grad: 5.8941

Epoch: [6][384/384] Data 0.000 (0.001) Elapsed 0m 21s (remain 0m 0s) Loss: 0.0048(0.0334) Grad: 1.6976

EVAL: [1/97] Data 0.273 (0.273) Elapsed 0m 0s (remain 0m 28s) Loss: 1.0909(1.0909)

EVAL: [51/97] Data 0.000 (0.013) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1198(0.5830)

EVAL: [97/97] Data 0.000 (0.011) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0008(0.5249)

Epoch {6} - avg\_train\_loss: {0.0334} avg\_val\_loss: {0.5249} time: {25}s

Epoch {6} - Accuracy(validation): {0.85241}

Epoch: [7][1/384] Data 0.277 (0.277) Elapsed 0m 0s (remain 2m 11s) Loss: 0.0059(0.0059) Grad: 3.3229

Epoch: [7][51/384] Data 0.000 (0.006) Elapsed 0m 3s (remain 0m 20s) Loss: 0.0122(0.0115) Grad: 3.5837

Epoch: [7][101/384] Data 0.000 (0.003) Elapsed 0m 6s (remain 0m 16s) Loss: 0.0086(0.0112) Grad: 2.6473

Epoch: [7][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0015(0.0106) Grad: 0.8989

Epoch: [7][201/384] Data 0.000 (0.002) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0143(0.0103) Grad: 10.0777

Epoch: [7][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0036(0.0098) Grad: 1.8043

Epoch: [7][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0032(0.0094) Grad: 1.1332

Epoch: [7][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0031(0.0088) Grad: 1.2838

Epoch: [7][384/384] Data 0.000 (0.001) Elapsed 0m 22s (remain 0m 0s) Loss: 0.0015(0.0085) Grad: 0.6132

EVAL: [1/97] Data 0.288 (0.288) Elapsed 0m 0s (remain 0m 29s) Loss: 1.3831(1.3831)

EVAL: [51/97] Data 0.000 (0.012) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1886(0.6580)

EVAL: [97/97] Data 0.000 (0.009) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0002(0.6009)

Epoch {7} - avg\_train\_loss: {0.0085} avg\_val\_loss: {0.6009} time: {25}s

Epoch {7} - Accuracy(validation): {0.84915}

Epoch: [8][1/384] Data 0.272 (0.272) Elapsed 0m 0s (remain 2m 7s) Loss: 0.0011(0.0011) Grad: 0.3834

Epoch: [8][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 21s) Loss: 0.0031(0.0042) Grad: 1.0053

Epoch: [8][101/384] Data 0.000 (0.003) Elapsed 0m 6s (remain 0m 17s) Loss: 0.0036(0.0043) Grad: 1.6090

Epoch: [8][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0034(0.0082) Grad: 5.3129

Epoch: [8][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0188(0.0088) Grad: 23.7881

Epoch: [8][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0059(0.0084) Grad: 7.3699

Epoch: [8][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0045(0.0081) Grad: 4.0763

Epoch: [8][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0035(0.0076) Grad: 1.9704

Epoch: [8][384/384] Data 0.000 (0.001) Elapsed 0m 22s (remain 0m 0s) Loss: 0.0027(0.0073) Grad: 1.5300

EVAL: [1/97] Data 0.272 (0.272) Elapsed 0m 0s (remain 0m 27s) Loss: 1.3447(1.3447)

EVAL: [51/97] Data 0.000 (0.013) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1421(0.7371)

EVAL: [97/97] Data 0.000 (0.010) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0001(0.6633)

Epoch {8} - avg\_train\_loss: {0.0073} avg\_val\_loss: {0.6633} time: {25}s

Epoch {8} - Accuracy(validation): {0.85436}

Epoch: [9][1/384] Data 0.280 (0.280) Elapsed 0m 0s (remain 2m 18s) Loss: 0.0005(0.0005) Grad: 0.1532

Epoch: [9][51/384] Data 0.000 (0.006) Elapsed 0m 3s (remain 0m 21s) Loss: 0.0051(0.0038) Grad: 4.2530

Epoch: [9][101/384] Data 0.000 (0.003) Elapsed 0m 6s (remain 0m 17s) Loss: 0.0041(0.0048) Grad: 2.4447

Epoch: [9][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0040(0.0056) Grad: 3.9533

Epoch: [9][201/384] Data 0.000 (0.002) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0042(0.0072) Grad: 2.7602

Epoch: [9][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0023(0.0070) Grad: 1.2570

Epoch: [9][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0027(0.0067) Grad: 2.9816

Epoch: [9][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0015(0.0064) Grad: 1.1878

Epoch: [9][384/384] Data 0.000 (0.001) Elapsed 0m 22s (remain 0m 0s) Loss: 0.0007(0.0061) Grad: 0.2933

EVAL: [1/97] Data 0.288 (0.288) Elapsed 0m 0s (remain 0m 29s) Loss: 1.2071(1.2071)

EVAL: [51/97] Data 0.000 (0.012) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1448(0.7750)

EVAL: [97/97] Data 0.000 (0.009) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0001(0.6938)

Epoch {9} - avg\_train\_loss: {0.0061} avg\_val\_loss: {0.6938} time: {25}s

Epoch {9} - Accuracy(validation): {0.85371}

Epoch: [10][1/384] Data 0.261 (0.261) Elapsed 0m 0s (remain 2m 6s) Loss: 0.0003(0.0003) Grad: 0.1332

Epoch: [10][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 21s) Loss: 0.0026(0.0035) Grad: 0.9730

Epoch: [10][101/384] Data 0.000 (0.003) Elapsed 0m 6s (remain 0m 16s) Loss: 0.0032(0.0040) Grad: 2.4645

Epoch: [10][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0012(0.0049) Grad: 0.8192

Epoch: [10][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0039(0.0055) Grad: 2.0187

Epoch: [10][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0067(0.0056) Grad: 7.7881

Epoch: [10][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0015(0.0053) Grad: 1.2606

Epoch: [10][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0033(0.0052) Grad: 3.6761

Epoch: [10][384/384] Data 0.000 (0.001) Elapsed 0m 22s (remain 0m 0s) Loss: 0.0006(0.0050) Grad: 0.6589

EVAL: [1/97] Data 0.263 (0.263) Elapsed 0m 0s (remain 0m 26s) Loss: 1.2808(1.2808)

EVAL: [51/97] Data 0.000 (0.013) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1233(0.7803)

EVAL: [97/97] Data 0.000 (0.010) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0001(0.6903)

Epoch {10} - avg\_train\_loss: {0.0050} avg\_val\_loss: {0.6903} time: {25}s

Epoch {10} - Accuracy(validation): {0.85306}

Epoch: [11][1/384] Data 0.248 (0.248) Elapsed 0m 0s (remain 1m 59s) Loss: 0.0007(0.0007) Grad: 0.3810

Epoch: [11][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 21s) Loss: 0.0013(0.0018) Grad: 0.6250

Epoch: [11][101/384] Data 0.000 (0.003) Elapsed 0m 6s (remain 0m 17s) Loss: 0.0010(0.0016) Grad: 0.6810

Epoch: [11][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0005(0.0016) Grad: 0.3480

Epoch: [11][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0015(0.0017) Grad: 0.6429

Epoch: [11][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0009(0.0017) Grad: 0.9225

Epoch: [11][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0009(0.0016) Grad: 0.6524

Epoch: [11][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0005(0.0016) Grad: 0.2904

Epoch: [11][384/384] Data 0.000 (0.001) Elapsed 0m 22s (remain 0m 0s) Loss: 0.0002(0.0015) Grad: 0.0636

EVAL: [1/97] Data 0.259 (0.259) Elapsed 0m 0s (remain 0m 27s) Loss: 1.3754(1.3754)

EVAL: [51/97] Data 0.016 (0.011) Elapsed 0m 1s (remain 0m 1s) Loss: 0.2017(0.8495)

EVAL: [97/97] Data 0.000 (0.009) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0000(0.7492)

Epoch {11} - avg\_train\_loss: {0.0015} avg\_val\_loss: {0.7492} time: {25}s

Epoch {11} - Accuracy(validation): {0.85826}

Epoch: [12][1/384] Data 0.269 (0.269) Elapsed 0m 0s (remain 2m 6s) Loss: 0.0001(0.0001) Grad: 0.0393

Epoch: [12][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 21s) Loss: 0.0008(0.0006) Grad: 0.6299

Epoch: [12][101/384] Data 0.000 (0.003) Elapsed 0m 6s (remain 0m 17s) Loss: 0.0005(0.0006) Grad: 0.3052

Epoch: [12][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0004(0.0006) Grad: 0.4844

Epoch: [12][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0011(0.0007) Grad: 1.2115

Epoch: [12][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0004(0.0007) Grad: 0.2738

Epoch: [12][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0004(0.0007) Grad: 0.2770

Epoch: [12][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0003(0.0007) Grad: 0.2518

Epoch: [12][384/384] Data 0.000 (0.001) Elapsed 0m 22s (remain 0m 0s) Loss: 0.0001(0.0007) Grad: 0.0698

EVAL: [1/97] Data 0.255 (0.255) Elapsed 0m 0s (remain 0m 26s) Loss: 1.4865(1.4865)

EVAL: [51/97] Data 0.000 (0.014) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1657(0.9115)

EVAL: [97/97] Data 0.000 (0.011) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0000(0.8054)

Epoch {12} - avg\_train\_loss: {0.0007} avg\_val\_loss: {0.8054} time: {25}s

Epoch {12} - Accuracy(validation): {0.85566}

Epoch: [13][1/384] Data 0.274 (0.274) Elapsed 0m 0s (remain 2m 9s) Loss: 0.0001(0.0001) Grad: 0.0431

Epoch: [13][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 20s) Loss: 0.0004(0.0004) Grad: 0.1496

Epoch: [13][101/384] Data 0.000 (0.003) Elapsed 0m 6s (remain 0m 17s) Loss: 0.0003(0.0004) Grad: 0.3530

Epoch: [13][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0003(0.0004) Grad: 0.3794

Epoch: [13][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0004(0.0004) Grad: 0.2922

Epoch: [13][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0004(0.0005) Grad: 0.5001

Epoch: [13][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0002(0.0005) Grad: 0.1029

Epoch: [13][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0003(0.0004) Grad: 0.3198

Epoch: [13][384/384] Data 0.000 (0.001) Elapsed 0m 22s (remain 0m 0s) Loss: 0.0001(0.0004) Grad: 0.0810

EVAL: [1/97] Data 0.268 (0.268) Elapsed 0m 0s (remain 0m 27s) Loss: 1.7819(1.7819)

EVAL: [51/97] Data 0.000 (0.011) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1336(0.9565)

EVAL: [97/97] Data 0.000 (0.009) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0000(0.8465)

Epoch {13} - avg\_train\_loss: {0.0004} avg\_val\_loss: {0.8465} time: {25}s

Epoch {13} - Accuracy(validation): {0.85501}

Epoch: [14][1/384] Data 0.256 (0.256) Elapsed 0m 0s (remain 2m 1s) Loss: 0.0001(0.0001) Grad: 0.0708

Epoch: [14][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 20s) Loss: 0.0003(0.0004) Grad: 0.1689

Epoch: [14][101/384] Data 0.000 (0.003) Elapsed 0m 6s (remain 0m 16s) Loss: 0.0003(0.0004) Grad: 0.2579

Epoch: [14][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0001(0.0003) Grad: 0.1128

Epoch: [14][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0004(0.0003) Grad: 0.2536

Epoch: [14][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0004(0.0004) Grad: 0.5105

Epoch: [14][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0002(0.0004) Grad: 0.1191

Epoch: [14][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0001(0.0003) Grad: 0.0744

Epoch: [14][384/384] Data 0.000 (0.001) Elapsed 0m 22s (remain 0m 0s) Loss: 0.0001(0.0003) Grad: 0.0794

EVAL: [1/97] Data 0.257 (0.257) Elapsed 0m 0s (remain 0m 26s) Loss: 1.7784(1.7784)

EVAL: [51/97] Data 0.000 (0.014) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1258(0.9782)

EVAL: [97/97] Data 0.000 (0.011) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0000(0.8611)

Epoch {14} - avg\_train\_loss: {0.0003} avg\_val\_loss: {0.8611} time: {25}s

Epoch {14} - Accuracy(validation): {0.85436}

Epoch: [15][1/384] Data 0.266 (0.266) Elapsed 0m 0s (remain 2m 8s) Loss: 0.0002(0.0002) Grad: 0.2144

Epoch: [15][51/384] Data 0.000 (0.005) Elapsed 0m 3s (remain 0m 20s) Loss: 0.0003(0.0004) Grad: 0.1838

Epoch: [15][101/384] Data 0.000 (0.003) Elapsed 0m 6s (remain 0m 16s) Loss: 0.0003(0.0003) Grad: 0.2401

Epoch: [15][151/384] Data 0.000 (0.002) Elapsed 0m 8s (remain 0m 13s) Loss: 0.0001(0.0003) Grad: 0.1272

Epoch: [15][201/384] Data 0.000 (0.001) Elapsed 0m 11s (remain 0m 10s) Loss: 0.0002(0.0003) Grad: 0.1046

Epoch: [15][251/384] Data 0.000 (0.001) Elapsed 0m 14s (remain 0m 7s) Loss: 0.0002(0.0003) Grad: 0.1861

Epoch: [15][301/384] Data 0.000 (0.001) Elapsed 0m 17s (remain 0m 4s) Loss: 0.0002(0.0003) Grad: 0.1102

Epoch: [15][351/384] Data 0.000 (0.001) Elapsed 0m 20s (remain 0m 1s) Loss: 0.0001(0.0003) Grad: 0.0691

Epoch: [15][384/384] Data 0.000 (0.001) Elapsed 0m 22s (remain 0m 0s) Loss: 0.0001(0.0002) Grad: 0.0445

EVAL: [1/97] Data 0.258 (0.258) Elapsed 0m 0s (remain 0m 26s) Loss: 1.7141(1.7141)

EVAL: [51/97] Data 0.000 (0.011) Elapsed 0m 1s (remain 0m 1s) Loss: 0.1386(0.9853)

EVAL: [97/97] Data 0.000 (0.009) Elapsed 0m 2s (remain 0m 0s) Loss: 0.0000(0.8640)

Epoch {15} - avg\_train\_loss: {0.0002} avg\_val\_loss: {0.8640} time: {25}s

Epoch {15} - Accuracy(validation): {0.85241}

========== CV Result ==========

Score: 0.85956}

TN 1116

FP 45

FN 171

TP 206