Federated testing  
  
Testing done with the file federated baseline config search

at link <https://colab.research.google.com/drive/1KYtrSgZWEprGxEteMTSH-NNBpr56IO0Y#scrollTo=3rNf-hPZdWuf>

im testing all with 128 and 100 for last 2 values

"""# ── Define the values you want to test ────────────────────────────────────────────

learning\_rates = [0.001, 0.01, 0.1] # typical LR grid for CIFAR-100 :contentReference[oaicite:0]{index=0}

weight\_decays = [1e-5, 1e-4, 1e-3] # common L2 regs :contentReference[oaicite:1]{index=1}

batch\_sizes = [32, 64, 128] # fits 6 GB VRAM on Colab GPUs :contentReference[oaicite:2]{index=2}

epoch\_counts = [50, 100] # short vs full training"""

Results

{'lr': 0.001, 'weight\_decay': 1e-5, 'batch\_size': 128, 'epochs': 100},

Running config: lr=0.001, wd=1e-05, bs=128, epochs=100

Epoch 1/100 train\_acc=0.0126 val\_acc=0.0166

Epoch 2/100 train\_acc=0.0223 val\_acc=0.0364

Epoch 3/100 train\_acc=0.0433 val\_acc=0.0674

Epoch 4/100 train\_acc=0.0751 val\_acc=0.0842

Epoch 5/100 train\_acc=0.0945 val\_acc=0.0994

Epoch 6/100 train\_acc=0.1123 val\_acc=0.1198

Epoch 7/100 train\_acc=0.1315 val\_acc=0.1414

Epoch 8/100 train\_acc=0.1430 val\_acc=0.1466

Epoch 9/100 train\_acc=0.1550 val\_acc=0.1714

Epoch 10/100 train\_acc=0.1665 val\_acc=0.1704

Epoch 11/100 train\_acc=0.1788 val\_acc=0.1856

Epoch 12/100 train\_acc=0.1903 val\_acc=0.1942

Epoch 13/100 train\_acc=0.1982 val\_acc=0.2070

Epoch 14/100 train\_acc=0.2088 val\_acc=0.2226

Epoch 15/100 train\_acc=0.2201 val\_acc=0.2258

Epoch 16/100 train\_acc=0.2290 val\_acc=0.2198

Epoch 17/100 train\_acc=0.2356 val\_acc=0.2384

Epoch 18/100 train\_acc=0.2460 val\_acc=0.2444

Epoch 19/100 train\_acc=0.2564 val\_acc=0.2422

Epoch 20/100 train\_acc=0.2641 val\_acc=0.2554

Epoch 21/100 train\_acc=0.2680 val\_acc=0.2624

Epoch 22/100 train\_acc=0.2764 val\_acc=0.2760

Epoch 23/100 train\_acc=0.2823 val\_acc=0.2744

Epoch 24/100 train\_acc=0.2915 val\_acc=0.2838

Epoch 25/100 train\_acc=0.2962 val\_acc=0.2804

Epoch 26/100 train\_acc=0.3028 val\_acc=0.2924

Epoch 27/100 train\_acc=0.3090 val\_acc=0.2966

Epoch 28/100 train\_acc=0.3142 val\_acc=0.3004

Epoch 29/100 train\_acc=0.3225 val\_acc=0.3158

Epoch 30/100 train\_acc=0.3272 val\_acc=0.3084

Epoch 31/100 train\_acc=0.3324 val\_acc=0.3208

Epoch 32/100 train\_acc=0.3378 val\_acc=0.3254

Epoch 33/100 train\_acc=0.3453 val\_acc=0.3282

Epoch 34/100 train\_acc=0.3494 val\_acc=0.3262

Epoch 35/100 train\_acc=0.3560 val\_acc=0.3396

Epoch 36/100 train\_acc=0.3611 val\_acc=0.3228

Epoch 37/100 train\_acc=0.3633 val\_acc=0.3416

Epoch 38/100 train\_acc=0.3723 val\_acc=0.3418

Epoch 39/100 train\_acc=0.3736 val\_acc=0.3452

Epoch 40/100 train\_acc=0.3780 val\_acc=0.3450

Epoch 41/100 train\_acc=0.3827 val\_acc=0.3430

Epoch 42/100 train\_acc=0.3844 val\_acc=0.3444

Epoch 43/100 train\_acc=0.3921 val\_acc=0.3630

Epoch 44/100 train\_acc=0.3927 val\_acc=0.3624

Epoch 45/100 train\_acc=0.3968 val\_acc=0.3662

Epoch 46/100 train\_acc=0.4017 val\_acc=0.3582

Epoch 47/100 train\_acc=0.4047 val\_acc=0.3662

Epoch 48/100 train\_acc=0.4075 val\_acc=0.3588

Epoch 49/100 train\_acc=0.4132 val\_acc=0.3776

Epoch 50/100 train\_acc=0.4148 val\_acc=0.3754

Epoch 51/100 train\_acc=0.4159 val\_acc=0.3700

Epoch 52/100 train\_acc=0.4196 val\_acc=0.3744

Epoch 53/100 train\_acc=0.4208 val\_acc=0.3838

Epoch 54/100 train\_acc=0.4247 val\_acc=0.3852

Epoch 55/100 train\_acc=0.4307 val\_acc=0.3848

Epoch 56/100 train\_acc=0.4320 val\_acc=0.3800

Epoch 57/100 train\_acc=0.4345 val\_acc=0.3820

Epoch 58/100 train\_acc=0.4399 val\_acc=0.3920

Epoch 59/100 train\_acc=0.4382 val\_acc=0.3890

Epoch 60/100 train\_acc=0.4409 val\_acc=0.3938

Epoch 61/100 train\_acc=0.4426 val\_acc=0.3864

Epoch 62/100 train\_acc=0.4447 val\_acc=0.3912

Epoch 63/100 train\_acc=0.4480 val\_acc=0.3958

Epoch 64/100 train\_acc=0.4501 val\_acc=0.3956

Epoch 65/100 train\_acc=0.4533 val\_acc=0.3958

Epoch 66/100 train\_acc=0.4540 val\_acc=0.3920

Epoch 67/100 train\_acc=0.4589 val\_acc=0.3950

Epoch 68/100 train\_acc=0.4572 val\_acc=0.4008

Epoch 69/100 train\_acc=0.4607 val\_acc=0.3948

Epoch 70/100 train\_acc=0.4642 val\_acc=0.4008

Epoch 71/100 train\_acc=0.4625 val\_acc=0.3984

Epoch 72/100 train\_acc=0.4657 val\_acc=0.4000

Epoch 73/100 train\_acc=0.4676 val\_acc=0.4010

Epoch 74/100 train\_acc=0.4673 val\_acc=0.4052

Epoch 75/100 train\_acc=0.4730 val\_acc=0.4036

Epoch 76/100 train\_acc=0.4697 val\_acc=0.4044

Epoch 77/100 train\_acc=0.4712 val\_acc=0.4030

Epoch 78/100 train\_acc=0.4744 val\_acc=0.4070

Epoch 79/100 train\_acc=0.4746 val\_acc=0.4052

Epoch 80/100 train\_acc=0.4776 val\_acc=0.4070

Epoch 81/100 train\_acc=0.4760 val\_acc=0.4036

Epoch 82/100 train\_acc=0.4788 val\_acc=0.4022

Epoch 83/100 train\_acc=0.4804 val\_acc=0.4048

Epoch 84/100 train\_acc=0.4809 val\_acc=0.4102

Epoch 85/100 train\_acc=0.4828 val\_acc=0.4028

Epoch 86/100 train\_acc=0.4829 val\_acc=0.4046

Epoch 87/100 train\_acc=0.4843 val\_acc=0.4108

Epoch 88/100 train\_acc=0.4827 val\_acc=0.4192

Epoch 89/100 train\_acc=0.4846 val\_acc=0.4136

Epoch 90/100 train\_acc=0.4852 val\_acc=0.4158

Epoch 91/100 train\_acc=0.4847 val\_acc=0.4078

Epoch 92/100 train\_acc=0.4829 val\_acc=0.4094

Epoch 93/100 train\_acc=0.4862 val\_acc=0.4068

Epoch 94/100 train\_acc=0.4855 val\_acc=0.4114

Epoch 95/100 train\_acc=0.4867 val\_acc=0.4072

Epoch 96/100 train\_acc=0.4864 val\_acc=0.4098

Epoch 97/100 train\_acc=0.4863 val\_acc=0.4070

Epoch 98/100 train\_acc=0.4897 val\_acc=0.4118

Epoch 99/100 train\_acc=0.4883 val\_acc=0.4122

Epoch 100/100 train\_acc=0.4876 val\_acc=0.4158

Config {'lr': 0.001, 'weight\_decay': 1e-05, 'batch\_size': 128, 'epochs': 100} → best\_val\_acc=0.4192, test\_acc=0.4401

==================== RUN SUMMARY (2025-05-01 16:00:19) ====================

➜ Environment:

• python\_version: 3.11.12

• pytorch\_version: 2.6.0+cu124

• torchvision\_version: 0.21.0+cu124

• cuda\_available: True

• cuda\_device: Tesla T4

• device\_count: 1

• platform: Linux-6.1.123+-x86\_64-with-glibc2.35

• cwd: /content

➜ Data:

• train\_samples: 45000

• val\_samples: 5000

• test\_samples: 10000

• batch\_size: 128

• num\_batches\_train: 352

• num\_batches\_val: 40

• num\_batches\_test: 79

➜ Seed:

• seed: not set

➜ Hyperparameters:

• lr: 0.001

• weight\_decay: 1e-05

• batch\_size: 128

• epochs: 100

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

{'lr': 0.01, 'weight\_decay': 1e-5, 'batch\_size': 128, 'epochs': 100},

Running config: lr=0.01, wd=1e-05, bs=128, epochs=100

Epoch 1/100 train\_acc=0.0545 val\_acc=0.0982

Epoch 2/100 train\_acc=0.1315 val\_acc=0.1582

Epoch 3/100 train\_acc=0.1851 val\_acc=0.2126

Epoch 4/100 train\_acc=0.2242 val\_acc=0.2516

Epoch 5/100 train\_acc=0.2613 val\_acc=0.2692

Epoch 6/100 train\_acc=0.2981 val\_acc=0.3014

Epoch 7/100 train\_acc=0.3183 val\_acc=0.3326

Epoch 8/100 train\_acc=0.3427 val\_acc=0.3336

Epoch 9/100 train\_acc=0.3636 val\_acc=0.3440

Epoch 10/100 train\_acc=0.3854 val\_acc=0.3742

Epoch 11/100 train\_acc=0.4055 val\_acc=0.3826

Epoch 12/100 train\_acc=0.4172 val\_acc=0.3892

Epoch 13/100 train\_acc=0.4338 val\_acc=0.3964

Epoch 14/100 train\_acc=0.4471 val\_acc=0.4034

Epoch 15/100 train\_acc=0.4627 val\_acc=0.4088

Epoch 16/100 train\_acc=0.4712 val\_acc=0.3994

Epoch 17/100 train\_acc=0.4853 val\_acc=0.4156

Epoch 18/100 train\_acc=0.4938 val\_acc=0.4342

Epoch 19/100 train\_acc=0.5074 val\_acc=0.4288

Epoch 20/100 train\_acc=0.5139 val\_acc=0.4244

Epoch 21/100 train\_acc=0.5252 val\_acc=0.4410

Epoch 22/100 train\_acc=0.5355 val\_acc=0.4490

Epoch 23/100 train\_acc=0.5428 val\_acc=0.4504

Epoch 24/100 train\_acc=0.5512 val\_acc=0.4436

Epoch 25/100 train\_acc=0.5610 val\_acc=0.4494

Epoch 26/100 train\_acc=0.5675 val\_acc=0.4522

Epoch 27/100 train\_acc=0.5747 val\_acc=0.4596

Epoch 28/100 train\_acc=0.5830 val\_acc=0.4550

Epoch 29/100 train\_acc=0.5895 val\_acc=0.4616

Epoch 30/100 train\_acc=0.5995 val\_acc=0.4612

Epoch 31/100 train\_acc=0.6034 val\_acc=0.4496

Epoch 32/100 train\_acc=0.6147 val\_acc=0.4566

Epoch 33/100 train\_acc=0.6239 val\_acc=0.4656

Epoch 34/100 train\_acc=0.6254 val\_acc=0.4670

Epoch 35/100 train\_acc=0.6295 val\_acc=0.4646

Epoch 36/100 train\_acc=0.6381 val\_acc=0.4616

Epoch 37/100 train\_acc=0.6448 val\_acc=0.4708

Epoch 38/100 train\_acc=0.6506 val\_acc=0.4634

Epoch 39/100 train\_acc=0.6587 val\_acc=0.4784

Epoch 40/100 train\_acc=0.6664 val\_acc=0.4700

Epoch 41/100 train\_acc=0.6760 val\_acc=0.4678

Epoch 42/100 train\_acc=0.6783 val\_acc=0.4692

Epoch 43/100 train\_acc=0.6828 val\_acc=0.4770

Epoch 44/100 train\_acc=0.6908 val\_acc=0.4652

Epoch 45/100 train\_acc=0.6953 val\_acc=0.4774

Epoch 46/100 train\_acc=0.7044 val\_acc=0.4744

Epoch 47/100 train\_acc=0.7050 val\_acc=0.4642

Epoch 48/100 train\_acc=0.7156 val\_acc=0.4728

Epoch 49/100 train\_acc=0.7183 val\_acc=0.4798

Epoch 50/100 train\_acc=0.7253 val\_acc=0.4888

Epoch 51/100 train\_acc=0.7324 val\_acc=0.4730

Epoch 52/100 train\_acc=0.7370 val\_acc=0.4752

Epoch 53/100 train\_acc=0.7457 val\_acc=0.4784

Epoch 54/100 train\_acc=0.7474 val\_acc=0.4822

Epoch 55/100 train\_acc=0.7548 val\_acc=0.4858

Epoch 56/100 train\_acc=0.7638 val\_acc=0.4842

Epoch 57/100 train\_acc=0.7669 val\_acc=0.4778

Epoch 58/100 train\_acc=0.7722 val\_acc=0.4890

Epoch 59/100 train\_acc=0.7730 val\_acc=0.4788

Epoch 60/100 train\_acc=0.7793 val\_acc=0.4844

Epoch 61/100 train\_acc=0.7891 val\_acc=0.4776

Epoch 62/100 train\_acc=0.7933 val\_acc=0.4852

Epoch 63/100 train\_acc=0.7987 val\_acc=0.4980

Epoch 64/100 train\_acc=0.8020 val\_acc=0.4828

Epoch 65/100 train\_acc=0.8084 val\_acc=0.4918

Epoch 66/100 train\_acc=0.8140 val\_acc=0.4900

Epoch 67/100 train\_acc=0.8182 val\_acc=0.4934

Epoch 68/100 train\_acc=0.8224 val\_acc=0.4852

Epoch 69/100 train\_acc=0.8270 val\_acc=0.4986

Epoch 70/100 train\_acc=0.8276 val\_acc=0.4854

Epoch 71/100 train\_acc=0.8361 val\_acc=0.4828

Epoch 72/100 train\_acc=0.8386 val\_acc=0.4824

Epoch 73/100 train\_acc=0.8430 val\_acc=0.5000

Epoch 74/100 train\_acc=0.8453 val\_acc=0.4816

Epoch 75/100 train\_acc=0.8534 val\_acc=0.4950

Epoch 76/100 train\_acc=0.8553 val\_acc=0.4898

Epoch 77/100 train\_acc=0.8598 val\_acc=0.4858

Epoch 78/100 train\_acc=0.8621 val\_acc=0.4936

Epoch 79/100 train\_acc=0.8638 val\_acc=0.4920

Epoch 80/100 train\_acc=0.8662 val\_acc=0.4932

Epoch 81/100 train\_acc=0.8678 val\_acc=0.4928

Epoch 82/100 train\_acc=0.8752 val\_acc=0.4948

Epoch 83/100 train\_acc=0.8756 val\_acc=0.4962

Epoch 84/100 train\_acc=0.8777 val\_acc=0.4990

Epoch 85/100 train\_acc=0.8791 val\_acc=0.4962

Epoch 86/100 train\_acc=0.8787 val\_acc=0.4964

Epoch 87/100 train\_acc=0.8849 val\_acc=0.5000

Epoch 88/100 train\_acc=0.8854 val\_acc=0.4958

Epoch 89/100 train\_acc=0.8868 val\_acc=0.5044

Epoch 90/100 train\_acc=0.8880 val\_acc=0.4972

Epoch 91/100 train\_acc=0.8881 val\_acc=0.5016

Epoch 92/100 train\_acc=0.8906 val\_acc=0.4928

Epoch 93/100 train\_acc=0.8905 val\_acc=0.5018

Epoch 94/100 train\_acc=0.8911 val\_acc=0.5004

Epoch 95/100 train\_acc=0.8920 val\_acc=0.4972

Epoch 96/100 train\_acc=0.8928 val\_acc=0.5044

Epoch 97/100 train\_acc=0.8927 val\_acc=0.4950

Epoch 98/100 train\_acc=0.8936 val\_acc=0.4980

Epoch 99/100 train\_acc=0.8910 val\_acc=0.5018

Epoch 100/100 train\_acc=0.8948 val\_acc=0.5042

Config {'lr': 0.01, 'weight\_decay': 1e-05, 'batch\_size': 128, 'epochs': 100} → best\_val\_acc=0.5044, test\_acc=0.5234

==================== RUN SUMMARY (2025-05-01 16:43:23) ====================

➜ Environment:

• python\_version: 3.11.12

• pytorch\_version: 2.6.0+cu124

• torchvision\_version: 0.21.0+cu124

• cuda\_available: True

• cuda\_device: Tesla T4

• device\_count: 1

• platform: Linux-6.1.123+-x86\_64-with-glibc2.35

• cwd: /content

➜ Data:

• train\_samples: 45000

• val\_samples: 5000

• test\_samples: 10000

• batch\_size: 128

• num\_batches\_train: 352

• num\_batches\_val: 40

• num\_batches\_test: 79

➜ Seed:

• seed: not set

➜ Hyperparameters:

• lr: 0.01

• weight\_decay: 1e-05

• batch\_size: 128

• epochs: 100

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

{'lr': 0.1, 'weight\_decay': 1e-5, 'batch\_size': 128, 'epochs': 100},

Running config: lr=0.1, wd=1e-05, bs=128, epochs=100

Epoch 1/100 train\_acc=0.0553 val\_acc=0.0884

Epoch 2/100 train\_acc=0.0833 val\_acc=0.1024

Epoch 3/100 train\_acc=0.1078 val\_acc=0.1130

Epoch 4/100 train\_acc=0.1261 val\_acc=0.1310

Epoch 5/100 train\_acc=0.1350 val\_acc=0.1452

Epoch 6/100 train\_acc=0.1397 val\_acc=0.1556

Epoch 7/100 train\_acc=0.1498 val\_acc=0.1506

Epoch 8/100 train\_acc=0.1512 val\_acc=0.1460

Epoch 9/100 train\_acc=0.1585 val\_acc=0.1482

Epoch 10/100 train\_acc=0.1627 val\_acc=0.1598

Epoch 11/100 train\_acc=0.1688 val\_acc=0.1812

Epoch 12/100 train\_acc=0.1719 val\_acc=0.1588

Epoch 13/100 train\_acc=0.1763 val\_acc=0.1794

Epoch 14/100 train\_acc=0.1764 val\_acc=0.1742

Epoch 15/100 train\_acc=0.1773 val\_acc=0.1744

Epoch 16/100 train\_acc=0.1819 val\_acc=0.1754

Epoch 17/100 train\_acc=0.1858 val\_acc=0.1660

Epoch 18/100 train\_acc=0.1882 val\_acc=0.1814

Epoch 19/100 train\_acc=0.1838 val\_acc=0.1842

Epoch 20/100 train\_acc=0.1890 val\_acc=0.1724

Epoch 21/100 train\_acc=0.1912 val\_acc=0.1898

Epoch 22/100 train\_acc=0.1932 val\_acc=0.1852

Epoch 23/100 train\_acc=0.1966 val\_acc=0.1792

Epoch 24/100 train\_acc=0.1966 val\_acc=0.2096

Epoch 25/100 train\_acc=0.2022 val\_acc=0.1962

Epoch 26/100 train\_acc=0.2072 val\_acc=0.2074

Epoch 27/100 train\_acc=0.2153 val\_acc=0.2234

Epoch 28/100 train\_acc=0.2144 val\_acc=0.2006

Epoch 29/100 train\_acc=0.2148 val\_acc=0.2178

Epoch 30/100 train\_acc=0.2240 val\_acc=0.2146

Epoch 31/100 train\_acc=0.2264 val\_acc=0.2230

Epoch 32/100 train\_acc=0.2274 val\_acc=0.2222

Epoch 33/100 train\_acc=0.2317 val\_acc=0.2280

Epoch 34/100 train\_acc=0.2380 val\_acc=0.2228

Epoch 35/100 train\_acc=0.2466 val\_acc=0.2268

Epoch 36/100 train\_acc=0.2485 val\_acc=0.2310

Epoch 37/100 train\_acc=0.2569 val\_acc=0.2382

Epoch 38/100 train\_acc=0.2578 val\_acc=0.2606

Epoch 39/100 train\_acc=0.2644 val\_acc=0.2490

Epoch 40/100 train\_acc=0.2708 val\_acc=0.2592

Epoch 41/100 train\_acc=0.2749 val\_acc=0.2590

Epoch 42/100 train\_acc=0.2830 val\_acc=0.2664

Epoch 43/100 train\_acc=0.2870 val\_acc=0.2500

Epoch 44/100 train\_acc=0.2901 val\_acc=0.2708

Epoch 45/100 train\_acc=0.2951 val\_acc=0.2818

Epoch 46/100 train\_acc=0.3008 val\_acc=0.2740

Epoch 47/100 train\_acc=0.3066 val\_acc=0.2782

Epoch 48/100 train\_acc=0.3157 val\_acc=0.2946

Epoch 49/100 train\_acc=0.3223 val\_acc=0.2884

Epoch 50/100 train\_acc=0.3270 val\_acc=0.2950

Epoch 51/100 train\_acc=0.3334 val\_acc=0.2996

Epoch 52/100 train\_acc=0.3366 val\_acc=0.2984

Epoch 53/100 train\_acc=0.3462 val\_acc=0.3068

Epoch 54/100 train\_acc=0.3528 val\_acc=0.3106

Epoch 55/100 train\_acc=0.3560 val\_acc=0.3172

Epoch 56/100 train\_acc=0.3675 val\_acc=0.3168

Epoch 57/100 train\_acc=0.3750 val\_acc=0.3192

Epoch 58/100 train\_acc=0.3812 val\_acc=0.3400

Epoch 59/100 train\_acc=0.3879 val\_acc=0.3318

Epoch 60/100 train\_acc=0.3969 val\_acc=0.3368

Epoch 61/100 train\_acc=0.4025 val\_acc=0.3466

Epoch 62/100 train\_acc=0.4055 val\_acc=0.3496

Epoch 63/100 train\_acc=0.4183 val\_acc=0.3602

Epoch 64/100 train\_acc=0.4195 val\_acc=0.3574

Epoch 65/100 train\_acc=0.4292 val\_acc=0.3562

Epoch 66/100 train\_acc=0.4372 val\_acc=0.3618

Epoch 67/100 train\_acc=0.4425 val\_acc=0.3726

Epoch 68/100 train\_acc=0.4503 val\_acc=0.3678

Epoch 69/100 train\_acc=0.4525 val\_acc=0.3770

Epoch 70/100 train\_acc=0.4611 val\_acc=0.3774

Epoch 71/100 train\_acc=0.4692 val\_acc=0.3888

Epoch 72/100 train\_acc=0.4753 val\_acc=0.3826

Epoch 73/100 train\_acc=0.4782 val\_acc=0.3888

Epoch 74/100 train\_acc=0.4886 val\_acc=0.3866

Epoch 75/100 train\_acc=0.4926 val\_acc=0.3930

Epoch 76/100 train\_acc=0.4975 val\_acc=0.3954

Epoch 77/100 train\_acc=0.5024 val\_acc=0.3916

Epoch 78/100 train\_acc=0.5098 val\_acc=0.3960

Epoch 79/100 train\_acc=0.5162 val\_acc=0.4094

Epoch 80/100 train\_acc=0.5188 val\_acc=0.4016

Epoch 81/100 train\_acc=0.5275 val\_acc=0.4058

Epoch 82/100 train\_acc=0.5319 val\_acc=0.4072

Epoch 83/100 train\_acc=0.5375 val\_acc=0.4122

Epoch 84/100 train\_acc=0.5399 val\_acc=0.4174

Epoch 85/100 train\_acc=0.5411 val\_acc=0.4160

Epoch 86/100 train\_acc=0.5493 val\_acc=0.4186

Epoch 87/100 train\_acc=0.5536 val\_acc=0.4158

Epoch 88/100 train\_acc=0.5549 val\_acc=0.4188

Epoch 89/100 train\_acc=0.5582 val\_acc=0.4160

Epoch 90/100 train\_acc=0.5605 val\_acc=0.4214

Epoch 91/100 train\_acc=0.5665 val\_acc=0.4148

Epoch 92/100 train\_acc=0.5641 val\_acc=0.4262

Epoch 93/100 train\_acc=0.5705 val\_acc=0.4196

Epoch 94/100 train\_acc=0.5729 val\_acc=0.4170

Epoch 95/100 train\_acc=0.5721 val\_acc=0.4178

Epoch 96/100 train\_acc=0.5721 val\_acc=0.4182

Epoch 97/100 train\_acc=0.5745 val\_acc=0.4230

Epoch 98/100 train\_acc=0.5768 val\_acc=0.4280

Epoch 99/100 train\_acc=0.5780 val\_acc=0.4262

Epoch 100/100 train\_acc=0.5760 val\_acc=0.4230

Config {'lr': 0.1, 'weight\_decay': 1e-05, 'batch\_size': 128, 'epochs': 100} → best\_val\_acc=0.4280, test\_acc=0.4562

==================== RUN SUMMARY (2025-05-01 17:14:14) ====================

➜ Environment:

• python\_version: 3.11.12

• pytorch\_version: 2.6.0+cu124

• torchvision\_version: 0.21.0+cu124

• cuda\_available: True

• cuda\_device: Tesla T4

• device\_count: 1

• platform: Linux-6.1.123+-x86\_64-with-glibc2.35

• cwd: /content

➜ Data:

• train\_samples: 45000

• val\_samples: 5000

• test\_samples: 10000

• batch\_size: 128

• num\_batches\_train: 352

• num\_batches\_val: 40

• num\_batches\_test: 79

➜ Seed:

• seed: not set

➜ Hyperparameters:

• lr: 0.1

• weight\_decay: 1e-05

• batch\_size: 128

• epochs: 100

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

{'lr': 0.001, 'weight\_decay': 1e-3, 'batch\_size': 128, 'epochs': 100},

Running config: lr=0.001, wd=0.001, bs=128, epochs=100

Epoch 1/100 train\_acc=0.0125 val\_acc=0.0162

Epoch 2/100 train\_acc=0.0215 val\_acc=0.0342

Epoch 3/100 train\_acc=0.0418 val\_acc=0.0648

Epoch 4/100 train\_acc=0.0741 val\_acc=0.0854

Epoch 5/100 train\_acc=0.0928 val\_acc=0.0978

Epoch 6/100 train\_acc=0.1106 val\_acc=0.1168

Epoch 7/100 train\_acc=0.1292 val\_acc=0.1394

Epoch 8/100 train\_acc=0.1407 val\_acc=0.1456

Epoch 9/100 train\_acc=0.1523 val\_acc=0.1698

Epoch 10/100 train\_acc=0.1638 val\_acc=0.1652

Epoch 11/100 train\_acc=0.1746 val\_acc=0.1834

Epoch 12/100 train\_acc=0.1861 val\_acc=0.1872

Epoch 13/100 train\_acc=0.1936 val\_acc=0.1992

Epoch 14/100 train\_acc=0.2056 val\_acc=0.2144

Epoch 15/100 train\_acc=0.2139 val\_acc=0.2218

Epoch 16/100 train\_acc=0.2230 val\_acc=0.2136

Epoch 17/100 train\_acc=0.2306 val\_acc=0.2316

Epoch 18/100 train\_acc=0.2398 val\_acc=0.2396

Epoch 19/100 train\_acc=0.2508 val\_acc=0.2382

Epoch 20/100 train\_acc=0.2558 val\_acc=0.2478

Epoch 21/100 train\_acc=0.2628 val\_acc=0.2566

Epoch 22/100 train\_acc=0.2711 val\_acc=0.2682

Epoch 23/100 train\_acc=0.2764 val\_acc=0.2712

Epoch 24/100 train\_acc=0.2825 val\_acc=0.2786

Epoch 25/100 train\_acc=0.2878 val\_acc=0.2734

Epoch 26/100 train\_acc=0.2947 val\_acc=0.2860

Epoch 27/100 train\_acc=0.3004 val\_acc=0.2870

Epoch 28/100 train\_acc=0.3046 val\_acc=0.2964

Epoch 29/100 train\_acc=0.3124 val\_acc=0.3038

Epoch 30/100 train\_acc=0.3181 val\_acc=0.3062

Epoch 31/100 train\_acc=0.3223 val\_acc=0.3080

Epoch 32/100 train\_acc=0.3276 val\_acc=0.3116

Epoch 33/100 train\_acc=0.3340 val\_acc=0.3182

Epoch 34/100 train\_acc=0.3380 val\_acc=0.3144

Epoch 35/100 train\_acc=0.3438 val\_acc=0.3264

Epoch 36/100 train\_acc=0.3485 val\_acc=0.3176

Epoch 37/100 train\_acc=0.3533 val\_acc=0.3330

Epoch 38/100 train\_acc=0.3590 val\_acc=0.3332

Epoch 39/100 train\_acc=0.3633 val\_acc=0.3398

Epoch 40/100 train\_acc=0.3673 val\_acc=0.3324

Epoch 41/100 train\_acc=0.3705 val\_acc=0.3338

Epoch 42/100 train\_acc=0.3719 val\_acc=0.3282

Epoch 43/100 train\_acc=0.3779 val\_acc=0.3544

Epoch 44/100 train\_acc=0.3820 val\_acc=0.3572

Epoch 45/100 train\_acc=0.3864 val\_acc=0.3588

Epoch 46/100 train\_acc=0.3885 val\_acc=0.3530

Epoch 47/100 train\_acc=0.3935 val\_acc=0.3640

Epoch 48/100 train\_acc=0.3958 val\_acc=0.3550

Epoch 49/100 train\_acc=0.4008 val\_acc=0.3658

Epoch 50/100 train\_acc=0.4032 val\_acc=0.3656

Epoch 51/100 train\_acc=0.4044 val\_acc=0.3656

Epoch 52/100 train\_acc=0.4076 val\_acc=0.3688

Epoch 53/100 train\_acc=0.4096 val\_acc=0.3746

Epoch 54/100 train\_acc=0.4131 val\_acc=0.3744

Epoch 55/100 train\_acc=0.4172 val\_acc=0.3758

Epoch 56/100 train\_acc=0.4201 val\_acc=0.3780

Epoch 57/100 train\_acc=0.4224 val\_acc=0.3752

Epoch 58/100 train\_acc=0.4258 val\_acc=0.3806

Epoch 59/100 train\_acc=0.4243 val\_acc=0.3862

Epoch 60/100 train\_acc=0.4272 val\_acc=0.3868

Epoch 61/100 train\_acc=0.4278 val\_acc=0.3794

Epoch 62/100 train\_acc=0.4298 val\_acc=0.3852

Epoch 63/100 train\_acc=0.4340 val\_acc=0.3866

Epoch 64/100 train\_acc=0.4360 val\_acc=0.3956

Epoch 65/100 train\_acc=0.4394 val\_acc=0.3876

Epoch 66/100 train\_acc=0.4400 val\_acc=0.3806

Epoch 67/100 train\_acc=0.4434 val\_acc=0.3868

Epoch 68/100 train\_acc=0.4420 val\_acc=0.3952

Epoch 69/100 train\_acc=0.4473 val\_acc=0.3920

Epoch 70/100 train\_acc=0.4498 val\_acc=0.3894

Epoch 71/100 train\_acc=0.4467 val\_acc=0.3930

Epoch 72/100 train\_acc=0.4520 val\_acc=0.3950

Epoch 73/100 train\_acc=0.4528 val\_acc=0.4008

Epoch 74/100 train\_acc=0.4528 val\_acc=0.3974

Epoch 75/100 train\_acc=0.4582 val\_acc=0.3922

Epoch 76/100 train\_acc=0.4544 val\_acc=0.3986

Epoch 77/100 train\_acc=0.4558 val\_acc=0.3974

Epoch 78/100 train\_acc=0.4582 val\_acc=0.3994

Epoch 79/100 train\_acc=0.4586 val\_acc=0.3938

Epoch 80/100 train\_acc=0.4643 val\_acc=0.4004

Epoch 81/100 train\_acc=0.4630 val\_acc=0.4018

Epoch 82/100 train\_acc=0.4656 val\_acc=0.3976

Epoch 83/100 train\_acc=0.4667 val\_acc=0.4020

Epoch 84/100 train\_acc=0.4665 val\_acc=0.4052

Epoch 85/100 train\_acc=0.4676 val\_acc=0.4000

Epoch 86/100 train\_acc=0.4667 val\_acc=0.4022

Epoch 87/100 train\_acc=0.4702 val\_acc=0.4022

Epoch 88/100 train\_acc=0.4686 val\_acc=0.4126

Epoch 89/100 train\_acc=0.4686 val\_acc=0.4034

Epoch 90/100 train\_acc=0.4692 val\_acc=0.4068

Epoch 91/100 train\_acc=0.4682 val\_acc=0.4036

Epoch 92/100 train\_acc=0.4703 val\_acc=0.4018

Epoch 93/100 train\_acc=0.4715 val\_acc=0.4000

Epoch 94/100 train\_acc=0.4703 val\_acc=0.4056

Epoch 95/100 train\_acc=0.4720 val\_acc=0.3970

Epoch 96/100 train\_acc=0.4724 val\_acc=0.4064

Epoch 97/100 train\_acc=0.4732 val\_acc=0.4088

Epoch 98/100 train\_acc=0.4729 val\_acc=0.4096

Epoch 99/100 train\_acc=0.4720 val\_acc=0.4032

Epoch 100/100 train\_acc=0.4738 val\_acc=0.4030

Config {'lr': 0.001, 'weight\_decay': 0.001, 'batch\_size': 128, 'epochs': 100} → best\_val\_acc=0.4126, test\_acc=0.4343

==================== RUN SUMMARY (2025-05-05 09:02:29) ====================

➜ Environment:

• python\_version: 3.11.12

• pytorch\_version: 2.6.0+cu124

• torchvision\_version: 0.21.0+cu124

• cuda\_available: True

• cuda\_device: Tesla T4

• device\_count: 1

• platform: Linux-6.1.123+-x86\_64-with-glibc2.35

• cwd: /content

➜ Data:

• train\_samples: 45000

• val\_samples: 5000

• test\_samples: 10000

• batch\_size: 128

• num\_batches\_train: 352

• num\_batches\_val: 40

• num\_batches\_test: 79

➜ Seed:

• seed: not set

➜ Hyperparameters:

• lr: 0.001

• weight\_decay: 0.001

• batch\_size: 128

• epochs: 100

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

{'lr': 0.01, 'weight\_decay': 1e-3, 'batch\_size': 128, 'epochs': 100},

Running config: lr=0.01, wd=0.001, bs=128, epochs=100

Epoch 1/100 train\_acc=0.0537 val\_acc=0.0946

Epoch 2/100 train\_acc=0.1279 val\_acc=0.1498

Epoch 3/100 train\_acc=0.1795 val\_acc=0.2064

Epoch 4/100 train\_acc=0.2191 val\_acc=0.2440

Epoch 5/100 train\_acc=0.2497 val\_acc=0.2498

Epoch 6/100 train\_acc=0.2828 val\_acc=0.2812

Epoch 7/100 train\_acc=0.3049 val\_acc=0.3196

Epoch 8/100 train\_acc=0.3270 val\_acc=0.3202

Epoch 9/100 train\_acc=0.3496 val\_acc=0.3350

Epoch 10/100 train\_acc=0.3710 val\_acc=0.3680

Epoch 11/100 train\_acc=0.3914 val\_acc=0.3732

Epoch 12/100 train\_acc=0.4011 val\_acc=0.3710

Epoch 13/100 train\_acc=0.4145 val\_acc=0.3820

Epoch 14/100 train\_acc=0.4295 val\_acc=0.3908

Epoch 15/100 train\_acc=0.4415 val\_acc=0.4034

Epoch 16/100 train\_acc=0.4529 val\_acc=0.3970

Epoch 17/100 train\_acc=0.4634 val\_acc=0.4122

Epoch 18/100 train\_acc=0.4728 val\_acc=0.4156

Epoch 19/100 train\_acc=0.4841 val\_acc=0.4316

Epoch 20/100 train\_acc=0.4918 val\_acc=0.4340

Epoch 21/100 train\_acc=0.5012 val\_acc=0.4332

Epoch 22/100 train\_acc=0.5085 val\_acc=0.4460

Epoch 23/100 train\_acc=0.5151 val\_acc=0.4474

Epoch 24/100 train\_acc=0.5275 val\_acc=0.4466

Epoch 25/100 train\_acc=0.5337 val\_acc=0.4498

Epoch 26/100 train\_acc=0.5462 val\_acc=0.4470

Epoch 27/100 train\_acc=0.5472 val\_acc=0.4632

Epoch 28/100 train\_acc=0.5522 val\_acc=0.4484

Epoch 29/100 train\_acc=0.5626 val\_acc=0.4586

Epoch 30/100 train\_acc=0.5655 val\_acc=0.4676

Epoch 31/100 train\_acc=0.5727 val\_acc=0.4724

Epoch 32/100 train\_acc=0.5822 val\_acc=0.4700

Epoch 33/100 train\_acc=0.5899 val\_acc=0.4754

Epoch 34/100 train\_acc=0.5928 val\_acc=0.4588

Epoch 35/100 train\_acc=0.5991 val\_acc=0.4668

Epoch 36/100 train\_acc=0.6032 val\_acc=0.4774

Epoch 37/100 train\_acc=0.6110 val\_acc=0.4780

Epoch 38/100 train\_acc=0.6152 val\_acc=0.4736

Epoch 39/100 train\_acc=0.6224 val\_acc=0.4860

Epoch 40/100 train\_acc=0.6299 val\_acc=0.4748

Epoch 41/100 train\_acc=0.6370 val\_acc=0.4812

Epoch 42/100 train\_acc=0.6392 val\_acc=0.4842

Epoch 43/100 train\_acc=0.6483 val\_acc=0.4798

Epoch 44/100 train\_acc=0.6479 val\_acc=0.5038

Epoch 45/100 train\_acc=0.6554 val\_acc=0.4914

Epoch 46/100 train\_acc=0.6632 val\_acc=0.4910

Epoch 47/100 train\_acc=0.6677 val\_acc=0.4980

Epoch 48/100 train\_acc=0.6740 val\_acc=0.5038

Epoch 49/100 train\_acc=0.6820 val\_acc=0.4934

Epoch 50/100 train\_acc=0.6848 val\_acc=0.5016

Epoch 51/100 train\_acc=0.6875 val\_acc=0.4906

Epoch 52/100 train\_acc=0.6957 val\_acc=0.4948

Epoch 53/100 train\_acc=0.7035 val\_acc=0.4998

Epoch 54/100 train\_acc=0.7042 val\_acc=0.5026

Epoch 55/100 train\_acc=0.7116 val\_acc=0.4958

Epoch 56/100 train\_acc=0.7181 val\_acc=0.5026

Epoch 57/100 train\_acc=0.7235 val\_acc=0.5042

Epoch 58/100 train\_acc=0.7332 val\_acc=0.4996

Epoch 59/100 train\_acc=0.7337 val\_acc=0.4978

Epoch 60/100 train\_acc=0.7431 val\_acc=0.5080

Epoch 61/100 train\_acc=0.7491 val\_acc=0.5070

Epoch 62/100 train\_acc=0.7530 val\_acc=0.4984

Epoch 63/100 train\_acc=0.7582 val\_acc=0.5086

Epoch 64/100 train\_acc=0.7649 val\_acc=0.5078

Epoch 65/100 train\_acc=0.7692 val\_acc=0.5144

Epoch 66/100 train\_acc=0.7733 val\_acc=0.5076

Epoch 67/100 train\_acc=0.7789 val\_acc=0.5198

Epoch 68/100 train\_acc=0.7829 val\_acc=0.5086

Epoch 69/100 train\_acc=0.7912 val\_acc=0.5138

Epoch 70/100 train\_acc=0.7943 val\_acc=0.5144

Epoch 71/100 train\_acc=0.8045 val\_acc=0.5104

Epoch 72/100 train\_acc=0.8053 val\_acc=0.5140

Epoch 73/100 train\_acc=0.8093 val\_acc=0.5206

Epoch 74/100 train\_acc=0.8145 val\_acc=0.5158

Epoch 75/100 train\_acc=0.8177 val\_acc=0.5116

Epoch 76/100 train\_acc=0.8240 val\_acc=0.5224

Epoch 77/100 train\_acc=0.8303 val\_acc=0.5204

Epoch 78/100 train\_acc=0.8302 val\_acc=0.5238

Epoch 79/100 train\_acc=0.8357 val\_acc=0.5160

Epoch 80/100 train\_acc=0.8394 val\_acc=0.5270

Epoch 81/100 train\_acc=0.8451 val\_acc=0.5308

Epoch 82/100 train\_acc=0.8481 val\_acc=0.5212

Epoch 83/100 train\_acc=0.8516 val\_acc=0.5182

Epoch 84/100 train\_acc=0.8552 val\_acc=0.5304

Epoch 85/100 train\_acc=0.8572 val\_acc=0.5286

Epoch 86/100 train\_acc=0.8604 val\_acc=0.5266

Epoch 87/100 train\_acc=0.8630 val\_acc=0.5350

Epoch 88/100 train\_acc=0.8645 val\_acc=0.5276

Epoch 89/100 train\_acc=0.8665 val\_acc=0.5360

Epoch 90/100 train\_acc=0.8729 val\_acc=0.5252

Epoch 91/100 train\_acc=0.8694 val\_acc=0.5250

Epoch 92/100 train\_acc=0.8732 val\_acc=0.5214

Epoch 93/100 train\_acc=0.8753 val\_acc=0.5312

Epoch 94/100 train\_acc=0.8760 val\_acc=0.5310

Epoch 95/100 train\_acc=0.8763 val\_acc=0.5272

Epoch 96/100 train\_acc=0.8765 val\_acc=0.5388

Epoch 97/100 train\_acc=0.8746 val\_acc=0.5244

Epoch 98/100 train\_acc=0.8778 val\_acc=0.5326

Epoch 99/100 train\_acc=0.8763 val\_acc=0.5294

Epoch 100/100 train\_acc=0.8779 val\_acc=0.5254

Config {'lr': 0.01, 'weight\_decay': 0.001, 'batch\_size': 128, 'epochs': 100} → best\_val\_acc=0.5388, test\_acc=0.5513

==================== RUN SUMMARY (2025-05-05 09:50:11) ====================

➜ Environment:

• python\_version: 3.11.12

• pytorch\_version: 2.6.0+cu124

• torchvision\_version: 0.21.0+cu124

• cuda\_available: True

• cuda\_device: Tesla T4

• device\_count: 1

• platform: Linux-6.1.123+-x86\_64-with-glibc2.35

• cwd: /content

➜ Data:

• train\_samples: 45000

• val\_samples: 5000

• test\_samples: 10000

• batch\_size: 128

• num\_batches\_train: 352

• num\_batches\_val: 40

• num\_batches\_test: 79

➜ Seed:

• seed: not set

➜ Hyperparameters:

• lr: 0.01

• weight\_decay: 0.001

• batch\_size: 128

• epochs: 100

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

{'lr': 0.1, 'weight\_decay': 1e-3, 'batch\_size': 128, 'epochs': 100},

Running config: lr=0.1, wd=0.001, bs=128, epochs=100

Epoch 1/100 train\_acc=0.0581 val\_acc=0.0806

Epoch 2/100 train\_acc=0.0985 val\_acc=0.1086

Epoch 3/100 train\_acc=0.1244 val\_acc=0.1416

Epoch 4/100 train\_acc=0.1352 val\_acc=0.1372

Epoch 5/100 train\_acc=0.1482 val\_acc=0.1766

Epoch 6/100 train\_acc=0.1565 val\_acc=0.1718

Epoch 7/100 train\_acc=0.1688 val\_acc=0.1540

Epoch 8/100 train\_acc=0.1787 val\_acc=0.1716

Epoch 9/100 train\_acc=0.1858 val\_acc=0.1950

Epoch 10/100 train\_acc=0.1835 val\_acc=0.1748

Epoch 11/100 train\_acc=0.1887 val\_acc=0.1978

Epoch 12/100 train\_acc=0.1957 val\_acc=0.1852

Epoch 13/100 train\_acc=0.2004 val\_acc=0.2160

Epoch 14/100 train\_acc=0.2035 val\_acc=0.2194

Epoch 15/100 train\_acc=0.2105 val\_acc=0.2072

Epoch 16/100 train\_acc=0.2148 val\_acc=0.2162

Epoch 17/100 train\_acc=0.2169 val\_acc=0.2090

Epoch 18/100 train\_acc=0.2248 val\_acc=0.2062

Epoch 19/100 train\_acc=0.2241 val\_acc=0.2278

Epoch 20/100 train\_acc=0.2313 val\_acc=0.2324

Epoch 21/100 train\_acc=0.2339 val\_acc=0.2144

Epoch 22/100 train\_acc=0.2302 val\_acc=0.2404

Epoch 23/100 train\_acc=0.2420 val\_acc=0.2322

Epoch 24/100 train\_acc=0.2448 val\_acc=0.2244

Epoch 25/100 train\_acc=0.2461 val\_acc=0.2454

Epoch 26/100 train\_acc=0.2501 val\_acc=0.2560

Epoch 27/100 train\_acc=0.2554 val\_acc=0.2376

Epoch 28/100 train\_acc=0.2592 val\_acc=0.2248

Epoch 29/100 train\_acc=0.2606 val\_acc=0.2604

Epoch 30/100 train\_acc=0.2744 val\_acc=0.2626

Epoch 31/100 train\_acc=0.2711 val\_acc=0.2846

Epoch 32/100 train\_acc=0.2766 val\_acc=0.2762

Epoch 33/100 train\_acc=0.2786 val\_acc=0.2676

Epoch 34/100 train\_acc=0.2812 val\_acc=0.2826

Epoch 35/100 train\_acc=0.2849 val\_acc=0.2680

Epoch 36/100 train\_acc=0.2887 val\_acc=0.2928

Epoch 37/100 train\_acc=0.2936 val\_acc=0.2710

Epoch 38/100 train\_acc=0.3017 val\_acc=0.2758

Epoch 39/100 train\_acc=0.3090 val\_acc=0.3084

Epoch 40/100 train\_acc=0.3138 val\_acc=0.3096

Epoch 41/100 train\_acc=0.3211 val\_acc=0.3126

Epoch 42/100 train\_acc=0.3227 val\_acc=0.3206

Epoch 43/100 train\_acc=0.3266 val\_acc=0.3110

Epoch 44/100 train\_acc=0.3310 val\_acc=0.3344

Epoch 45/100 train\_acc=0.3361 val\_acc=0.3214

Epoch 46/100 train\_acc=0.3397 val\_acc=0.3292

Epoch 47/100 train\_acc=0.3549 val\_acc=0.3290

Epoch 48/100 train\_acc=0.3565 val\_acc=0.3398

Epoch 49/100 train\_acc=0.3650 val\_acc=0.3424

Epoch 50/100 train\_acc=0.3684 val\_acc=0.3354

Epoch 51/100 train\_acc=0.3742 val\_acc=0.3516

Epoch 52/100 train\_acc=0.3859 val\_acc=0.3502

Epoch 53/100 train\_acc=0.3894 val\_acc=0.3536

Epoch 54/100 train\_acc=0.3974 val\_acc=0.3484

Epoch 55/100 train\_acc=0.4006 val\_acc=0.3662

Epoch 56/100 train\_acc=0.4114 val\_acc=0.3952

Epoch 57/100 train\_acc=0.4190 val\_acc=0.3822

Epoch 58/100 train\_acc=0.4221 val\_acc=0.3856

Epoch 59/100 train\_acc=0.4336 val\_acc=0.3856

Epoch 60/100 train\_acc=0.4392 val\_acc=0.3918

Epoch 61/100 train\_acc=0.4435 val\_acc=0.4042

Epoch 62/100 train\_acc=0.4534 val\_acc=0.3906

Epoch 63/100 train\_acc=0.4606 val\_acc=0.4034

Epoch 64/100 train\_acc=0.4721 val\_acc=0.4320

Epoch 65/100 train\_acc=0.4763 val\_acc=0.4268

Epoch 66/100 train\_acc=0.4843 val\_acc=0.4290

Epoch 67/100 train\_acc=0.4953 val\_acc=0.4258

Epoch 68/100 train\_acc=0.5056 val\_acc=0.4548

Epoch 69/100 train\_acc=0.5146 val\_acc=0.4482

Epoch 70/100 train\_acc=0.5224 val\_acc=0.4486

Epoch 71/100 train\_acc=0.5292 val\_acc=0.4504

Epoch 72/100 train\_acc=0.5400 val\_acc=0.4568

Epoch 73/100 train\_acc=0.5476 val\_acc=0.4616

Epoch 74/100 train\_acc=0.5597 val\_acc=0.4594

Epoch 75/100 train\_acc=0.5659 val\_acc=0.4780

Epoch 76/100 train\_acc=0.5718 val\_acc=0.4878

Epoch 77/100 train\_acc=0.5841 val\_acc=0.4840

Epoch 78/100 train\_acc=0.5980 val\_acc=0.4844

Epoch 79/100 train\_acc=0.6114 val\_acc=0.4996

Epoch 80/100 train\_acc=0.6194 val\_acc=0.4936

Epoch 81/100 train\_acc=0.6291 val\_acc=0.4946

Epoch 82/100 train\_acc=0.6406 val\_acc=0.4954

Epoch 83/100 train\_acc=0.6497 val\_acc=0.5076

Epoch 84/100 train\_acc=0.6602 val\_acc=0.5058

Epoch 85/100 train\_acc=0.6692 val\_acc=0.5088

Epoch 86/100 train\_acc=0.6787 val\_acc=0.5116

Epoch 87/100 train\_acc=0.6878 val\_acc=0.5172

Epoch 88/100 train\_acc=0.6983 val\_acc=0.5194

Epoch 89/100 train\_acc=0.7076 val\_acc=0.5216

Epoch 90/100 train\_acc=0.7158 val\_acc=0.5294

Epoch 91/100 train\_acc=0.7270 val\_acc=0.5276

Epoch 92/100 train\_acc=0.7321 val\_acc=0.5334

Epoch 93/100 train\_acc=0.7420 val\_acc=0.5298

Epoch 94/100 train\_acc=0.7449 val\_acc=0.5350

Epoch 95/100 train\_acc=0.7506 val\_acc=0.5332

Epoch 96/100 train\_acc=0.7575 val\_acc=0.5370

Epoch 97/100 train\_acc=0.7578 val\_acc=0.5404

Epoch 98/100 train\_acc=0.7626 val\_acc=0.5396

Epoch 99/100 train\_acc=0.7646 val\_acc=0.5390

Epoch 100/100 train\_acc=0.7638 val\_acc=0.5344

Config {'lr': 0.1, 'weight\_decay': 0.001, 'batch\_size': 128, 'epochs': 100} → best\_val\_acc=0.5404, test\_acc=0.5729

==================== RUN SUMMARY (2025-05-05 10:42:32) ====================

➜ Environment:

• python\_version: 3.11.12

• pytorch\_version: 2.6.0+cu124

• torchvision\_version: 0.21.0+cu124

• cuda\_available: True

• cuda\_device: Tesla T4

• device\_count: 1

• platform: Linux-6.1.123+-x86\_64-with-glibc2.35

• cwd: /content

➜ Data:

• train\_samples: 45000

• val\_samples: 5000

• test\_samples: 10000

• batch\_size: 128

• num\_batches\_train: 352

• num\_batches\_val: 40

• num\_batches\_test: 79

➜ Seed:

• seed: not set

➜ Hyperparameters:

• lr: 0.1

• weight\_decay: 0.001

• batch\_size: 128

• epochs: 100

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

{'lr': 0.001, 'weight\_decay': 1e-4, 'batch\_size': 128, 'epochs': 100},

Running config: lr=0.001, wd=0.0001, bs=128, epochs=100

Epoch 1/100 train\_acc=0.0175 val\_acc=0.0238

Epoch 2/100 train\_acc=0.0227 val\_acc=0.0326

Epoch 3/100 train\_acc=0.0565 val\_acc=0.0748

Epoch 4/100 train\_acc=0.0829 val\_acc=0.1014

Epoch 5/100 train\_acc=0.1003 val\_acc=0.1140

Epoch 6/100 train\_acc=0.1195 val\_acc=0.1366

Epoch 7/100 train\_acc=0.1326 val\_acc=0.1374

Epoch 8/100 train\_acc=0.1440 val\_acc=0.1536

Epoch 9/100 train\_acc=0.1518 val\_acc=0.1656

Epoch 10/100 train\_acc=0.1646 val\_acc=0.1648

Epoch 11/100 train\_acc=0.1767 val\_acc=0.1746

Epoch 12/100 train\_acc=0.1822 val\_acc=0.1880

Epoch 13/100 train\_acc=0.1952 val\_acc=0.2022

Epoch 14/100 train\_acc=0.2057 val\_acc=0.2024

Epoch 15/100 train\_acc=0.2152 val\_acc=0.2116

Epoch 16/100 train\_acc=0.2250 val\_acc=0.2258

Epoch 17/100 train\_acc=0.2312 val\_acc=0.2366

Epoch 18/100 train\_acc=0.2405 val\_acc=0.2318

Epoch 19/100 train\_acc=0.2474 val\_acc=0.2492

Epoch 20/100 train\_acc=0.2546 val\_acc=0.2446

Epoch 21/100 train\_acc=0.2633 val\_acc=0.2536

Epoch 22/100 train\_acc=0.2675 val\_acc=0.2590

Epoch 23/100 train\_acc=0.2739 val\_acc=0.2732

Epoch 24/100 train\_acc=0.2828 val\_acc=0.2754

Epoch 25/100 train\_acc=0.2874 val\_acc=0.2674

Epoch 26/100 train\_acc=0.2944 val\_acc=0.2792

Epoch 27/100 train\_acc=0.2986 val\_acc=0.2836

Epoch 28/100 train\_acc=0.3057 val\_acc=0.2882

Epoch 29/100 train\_acc=0.3090 val\_acc=0.2918

Epoch 30/100 train\_acc=0.3151 val\_acc=0.2964

Epoch 31/100 train\_acc=0.3214 val\_acc=0.3024

Epoch 32/100 train\_acc=0.3247 val\_acc=0.3136

Epoch 33/100 train\_acc=0.3295 val\_acc=0.3090

Epoch 34/100 train\_acc=0.3360 val\_acc=0.3114

Epoch 35/100 train\_acc=0.3389 val\_acc=0.3170

Epoch 36/100 train\_acc=0.3463 val\_acc=0.3212

Epoch 37/100 train\_acc=0.3468 val\_acc=0.3210

Epoch 38/100 train\_acc=0.3536 val\_acc=0.3248

Epoch 39/100 train\_acc=0.3590 val\_acc=0.3328

Epoch 40/100 train\_acc=0.3606 val\_acc=0.3258

Epoch 41/100 train\_acc=0.3674 val\_acc=0.3394

Epoch 42/100 train\_acc=0.3687 val\_acc=0.3378

Epoch 43/100 train\_acc=0.3751 val\_acc=0.3474

Epoch 44/100 train\_acc=0.3799 val\_acc=0.3454

Epoch 45/100 train\_acc=0.3825 val\_acc=0.3504

Epoch 46/100 train\_acc=0.3860 val\_acc=0.3478

Epoch 47/100 train\_acc=0.3884 val\_acc=0.3534

Epoch 48/100 train\_acc=0.3958 val\_acc=0.3552

Epoch 49/100 train\_acc=0.3973 val\_acc=0.3604

Epoch 50/100 train\_acc=0.4008 val\_acc=0.3552

Epoch 51/100 train\_acc=0.4046 val\_acc=0.3606

Epoch 52/100 train\_acc=0.4078 val\_acc=0.3690

Epoch 53/100 train\_acc=0.4142 val\_acc=0.3704

Epoch 54/100 train\_acc=0.4126 val\_acc=0.3698

Epoch 55/100 train\_acc=0.4164 val\_acc=0.3778

Epoch 56/100 train\_acc=0.4190 val\_acc=0.3690

Epoch 57/100 train\_acc=0.4206 val\_acc=0.3758

Epoch 58/100 train\_acc=0.4252 val\_acc=0.3730

Epoch 59/100 train\_acc=0.4265 val\_acc=0.3746

Epoch 60/100 train\_acc=0.4310 val\_acc=0.3768

Epoch 61/100 train\_acc=0.4346 val\_acc=0.3844

Epoch 62/100 train\_acc=0.4349 val\_acc=0.3732

Epoch 63/100 train\_acc=0.4354 val\_acc=0.3840

Epoch 64/100 train\_acc=0.4396 val\_acc=0.3810

Epoch 65/100 train\_acc=0.4434 val\_acc=0.3852

Epoch 66/100 train\_acc=0.4448 val\_acc=0.3864

Epoch 67/100 train\_acc=0.4454 val\_acc=0.3940

Epoch 68/100 train\_acc=0.4509 val\_acc=0.3900

Epoch 69/100 train\_acc=0.4516 val\_acc=0.3884

Epoch 70/100 train\_acc=0.4518 val\_acc=0.3824

Epoch 71/100 train\_acc=0.4534 val\_acc=0.3968

Epoch 72/100 train\_acc=0.4574 val\_acc=0.3968

Epoch 73/100 train\_acc=0.4588 val\_acc=0.3942

Epoch 74/100 train\_acc=0.4566 val\_acc=0.3954

Epoch 75/100 train\_acc=0.4590 val\_acc=0.3978

Epoch 76/100 train\_acc=0.4626 val\_acc=0.3966

Epoch 77/100 train\_acc=0.4643 val\_acc=0.4000

Epoch 78/100 train\_acc=0.4639 val\_acc=0.4014

Epoch 79/100 train\_acc=0.4648 val\_acc=0.4046

Epoch 80/100 train\_acc=0.4678 val\_acc=0.3982

Epoch 81/100 train\_acc=0.4659 val\_acc=0.4020

Epoch 82/100 train\_acc=0.4689 val\_acc=0.4012

Epoch 83/100 train\_acc=0.4708 val\_acc=0.3978

Epoch 84/100 train\_acc=0.4713 val\_acc=0.4038

Epoch 85/100 train\_acc=0.4721 val\_acc=0.4030

Epoch 86/100 train\_acc=0.4723 val\_acc=0.4018

Epoch 87/100 train\_acc=0.4732 val\_acc=0.3982

Epoch 88/100 train\_acc=0.4737 val\_acc=0.4050

Epoch 89/100 train\_acc=0.4763 val\_acc=0.3984

Epoch 90/100 train\_acc=0.4762 val\_acc=0.4040

Epoch 91/100 train\_acc=0.4771 val\_acc=0.4090

Epoch 92/100 train\_acc=0.4790 val\_acc=0.4044

Epoch 93/100 train\_acc=0.4762 val\_acc=0.4080

Epoch 94/100 train\_acc=0.4757 val\_acc=0.4096

Epoch 95/100 train\_acc=0.4801 val\_acc=0.4028

Epoch 96/100 train\_acc=0.4765 val\_acc=0.4080

Epoch 97/100 train\_acc=0.4780 val\_acc=0.4098

Epoch 98/100 train\_acc=0.4786 val\_acc=0.4026

Epoch 99/100 train\_acc=0.4782 val\_acc=0.3996

Epoch 100/100 train\_acc=0.4760 val\_acc=0.4036

Config {'lr': 0.001, 'weight\_decay': 0.0001, 'batch\_size': 128, 'epochs': 100} → best\_val\_acc=0.4098, test\_acc=0.4359

==================== RUN SUMMARY (2025-05-05 11:14:46) ====================

➜ Environment:

• python\_version: 3.11.12

• pytorch\_version: 2.6.0+cu124

• torchvision\_version: 0.21.0+cu124

• cuda\_available: True

• cuda\_device: Tesla T4

• device\_count: 1

• platform: Linux-6.1.123+-x86\_64-with-glibc2.35

• cwd: /content

➜ Data:

• train\_samples: 45000

• val\_samples: 5000

• test\_samples: 10000

• batch\_size: 128

• num\_batches\_train: 352

• num\_batches\_val: 40

• num\_batches\_test: 79

➜ Seed:

• seed: not set

➜ Hyperparameters:

• lr: 0.001

• weight\_decay: 0.0001

• batch\_size: 128

• epochs: 100

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

{'lr': 0.01, 'weight\_decay': 1e-4, 'batch\_size': 128, 'epochs': 100},

Running config: lr=0.01, wd=0.0001, bs=128, epochs=100

Epoch 1/100 train\_acc=0.0542 val\_acc=0.1038

Epoch 2/100 train\_acc=0.1318 val\_acc=0.1640

Epoch 3/100 train\_acc=0.1821 val\_acc=0.2130

Epoch 4/100 train\_acc=0.2240 val\_acc=0.2422

Epoch 5/100 train\_acc=0.2624 val\_acc=0.2838

Epoch 6/100 train\_acc=0.2938 val\_acc=0.2874

Epoch 7/100 train\_acc=0.3231 val\_acc=0.3296

Epoch 8/100 train\_acc=0.3437 val\_acc=0.3440

Epoch 9/100 train\_acc=0.3663 val\_acc=0.3488

Epoch 10/100 train\_acc=0.3853 val\_acc=0.3608

Epoch 11/100 train\_acc=0.3984 val\_acc=0.3660

Epoch 12/100 train\_acc=0.4213 val\_acc=0.3952

Epoch 13/100 train\_acc=0.4325 val\_acc=0.3838

Epoch 14/100 train\_acc=0.4467 val\_acc=0.4072

Epoch 15/100 train\_acc=0.4616 val\_acc=0.3968

Epoch 16/100 train\_acc=0.4704 val\_acc=0.4228

Epoch 17/100 train\_acc=0.4815 val\_acc=0.4164

Epoch 18/100 train\_acc=0.4869 val\_acc=0.4260

Epoch 19/100 train\_acc=0.5036 val\_acc=0.4190

Epoch 20/100 train\_acc=0.5126 val\_acc=0.4238

Epoch 21/100 train\_acc=0.5233 val\_acc=0.4362

Epoch 22/100 train\_acc=0.5321 val\_acc=0.4422

Epoch 23/100 train\_acc=0.5371 val\_acc=0.4234

Epoch 24/100 train\_acc=0.5480 val\_acc=0.4492

Epoch 25/100 train\_acc=0.5555 val\_acc=0.4598

Epoch 26/100 train\_acc=0.5626 val\_acc=0.4440

Epoch 27/100 train\_acc=0.5728 val\_acc=0.4518

Epoch 28/100 train\_acc=0.5783 val\_acc=0.4540

Epoch 29/100 train\_acc=0.5870 val\_acc=0.4636

Epoch 30/100 train\_acc=0.5981 val\_acc=0.4520

Epoch 31/100 train\_acc=0.6008 val\_acc=0.4588

Epoch 32/100 train\_acc=0.6074 val\_acc=0.4632

Epoch 33/100 train\_acc=0.6161 val\_acc=0.4610

Epoch 34/100 train\_acc=0.6220 val\_acc=0.4596

Epoch 35/100 train\_acc=0.6293 val\_acc=0.4710

Epoch 36/100 train\_acc=0.6343 val\_acc=0.4710

Epoch 37/100 train\_acc=0.6402 val\_acc=0.4652

Epoch 38/100 train\_acc=0.6484 val\_acc=0.4680

Epoch 39/100 train\_acc=0.6543 val\_acc=0.4648

Epoch 40/100 train\_acc=0.6654 val\_acc=0.4704

Epoch 41/100 train\_acc=0.6702 val\_acc=0.4636

Epoch 42/100 train\_acc=0.6752 val\_acc=0.4612

Epoch 43/100 train\_acc=0.6797 val\_acc=0.4766

Epoch 44/100 train\_acc=0.6868 val\_acc=0.4716

Epoch 45/100 train\_acc=0.6908 val\_acc=0.4718

Epoch 46/100 train\_acc=0.6968 val\_acc=0.4772

Epoch 47/100 train\_acc=0.7065 val\_acc=0.4752

Epoch 48/100 train\_acc=0.7102 val\_acc=0.4820

Epoch 49/100 train\_acc=0.7159 val\_acc=0.4878

Epoch 50/100 train\_acc=0.7190 val\_acc=0.4694

Epoch 51/100 train\_acc=0.7300 val\_acc=0.4808

Epoch 52/100 train\_acc=0.7337 val\_acc=0.4814

Epoch 53/100 train\_acc=0.7380 val\_acc=0.4852

Epoch 54/100 train\_acc=0.7428 val\_acc=0.4802

Epoch 55/100 train\_acc=0.7511 val\_acc=0.4826

Epoch 56/100 train\_acc=0.7595 val\_acc=0.4810

Epoch 57/100 train\_acc=0.7599 val\_acc=0.4870

Epoch 58/100 train\_acc=0.7685 val\_acc=0.4842

Epoch 59/100 train\_acc=0.7731 val\_acc=0.4870

Epoch 60/100 train\_acc=0.7754 val\_acc=0.4868

Epoch 61/100 train\_acc=0.7832 val\_acc=0.4878

Epoch 62/100 train\_acc=0.7895 val\_acc=0.4932

Epoch 63/100 train\_acc=0.7886 val\_acc=0.4846

Epoch 64/100 train\_acc=0.7983 val\_acc=0.4898

Epoch 65/100 train\_acc=0.8056 val\_acc=0.4976

Epoch 66/100 train\_acc=0.8086 val\_acc=0.4970

Epoch 67/100 train\_acc=0.8123 val\_acc=0.4884

Epoch 68/100 train\_acc=0.8177 val\_acc=0.5026

Epoch 69/100 train\_acc=0.8203 val\_acc=0.4984

Epoch 70/100 train\_acc=0.8300 val\_acc=0.4894

Epoch 71/100 train\_acc=0.8295 val\_acc=0.4920

Epoch 72/100 train\_acc=0.8341 val\_acc=0.4972

Epoch 73/100 train\_acc=0.8382 val\_acc=0.4838

Epoch 74/100 train\_acc=0.8452 val\_acc=0.4992

Epoch 75/100 train\_acc=0.8490 val\_acc=0.4970

Epoch 76/100 train\_acc=0.8506 val\_acc=0.4956

Epoch 77/100 train\_acc=0.8543 val\_acc=0.5022

Epoch 78/100 train\_acc=0.8603 val\_acc=0.5052

Epoch 79/100 train\_acc=0.8626 val\_acc=0.5046

Epoch 80/100 train\_acc=0.8668 val\_acc=0.5038

Epoch 81/100 train\_acc=0.8665 val\_acc=0.5034

Epoch 82/100 train\_acc=0.8703 val\_acc=0.4934

Epoch 83/100 train\_acc=0.8722 val\_acc=0.5050

Epoch 84/100 train\_acc=0.8736 val\_acc=0.4994

Epoch 85/100 train\_acc=0.8762 val\_acc=0.5008

Epoch 86/100 train\_acc=0.8802 val\_acc=0.4990

Epoch 87/100 train\_acc=0.8820 val\_acc=0.5024

Epoch 88/100 train\_acc=0.8819 val\_acc=0.5122

Epoch 89/100 train\_acc=0.8857 val\_acc=0.5108

Epoch 90/100 train\_acc=0.8890 val\_acc=0.5096

Epoch 91/100 train\_acc=0.8860 val\_acc=0.5054

Epoch 92/100 train\_acc=0.8872 val\_acc=0.4948

Epoch 93/100 train\_acc=0.8898 val\_acc=0.5062

Epoch 94/100 train\_acc=0.8879 val\_acc=0.5030

Epoch 95/100 train\_acc=0.8904 val\_acc=0.5026

Epoch 96/100 train\_acc=0.8898 val\_acc=0.5042

Epoch 97/100 train\_acc=0.8912 val\_acc=0.5062

Epoch 98/100 train\_acc=0.8902 val\_acc=0.4990

Epoch 99/100 train\_acc=0.8901 val\_acc=0.5076

Epoch 100/100 train\_acc=0.8915 val\_acc=0.4992

Config {'lr': 0.01, 'weight\_decay': 0.0001, 'batch\_size': 128, 'epochs': 100} → best\_val\_acc=0.5122, test\_acc=0.5243

==================== RUN SUMMARY (2025-05-05 11:49:02) ====================

➜ Environment:

• python\_version: 3.11.12

• pytorch\_version: 2.6.0+cu124

• torchvision\_version: 0.21.0+cu124

• cuda\_available: True

• cuda\_device: Tesla T4

• device\_count: 1

• platform: Linux-6.1.123+-x86\_64-with-glibc2.35

• cwd: /content

➜ Data:

• train\_samples: 45000

• val\_samples: 5000

• test\_samples: 10000

• batch\_size: 128

• num\_batches\_train: 352

• num\_batches\_val: 40

• num\_batches\_test: 79

➜ Seed:

• seed: not set

➜ Hyperparameters:

• lr: 0.01

• weight\_decay: 0.0001

• batch\_size: 128

• epochs: 100

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

{'lr': 0.1, 'weight\_decay': 1e-4, 'batch\_size': 128, 'epochs': 100},

Running config: lr=0.1, wd=0.0001, bs=128, epochs=100

Epoch 1/100 train\_acc=0.0585 val\_acc=0.0946

Epoch 2/100 train\_acc=0.0948 val\_acc=0.1114

Epoch 3/100 train\_acc=0.1170 val\_acc=0.1356

Epoch 4/100 train\_acc=0.1274 val\_acc=0.1544

Epoch 5/100 train\_acc=0.1391 val\_acc=0.1342

Epoch 6/100 train\_acc=0.1454 val\_acc=0.1486

Epoch 7/100 train\_acc=0.1493 val\_acc=0.1660

Epoch 8/100 train\_acc=0.1546 val\_acc=0.1618

Epoch 9/100 train\_acc=0.1605 val\_acc=0.1618

Epoch 10/100 train\_acc=0.1667 val\_acc=0.1742

Epoch 11/100 train\_acc=0.1731 val\_acc=0.1782

Epoch 12/100 train\_acc=0.1724 val\_acc=0.1686

Epoch 13/100 train\_acc=0.1808 val\_acc=0.1716

Epoch 14/100 train\_acc=0.1865 val\_acc=0.1852

Epoch 15/100 train\_acc=0.1902 val\_acc=0.1936

Epoch 16/100 train\_acc=0.1949 val\_acc=0.1958

Epoch 17/100 train\_acc=0.1988 val\_acc=0.2070

Epoch 18/100 train\_acc=0.2018 val\_acc=0.2000

Epoch 19/100 train\_acc=0.2093 val\_acc=0.1882

Epoch 20/100 train\_acc=0.2128 val\_acc=0.2130

Epoch 21/100 train\_acc=0.2118 val\_acc=0.2012

Epoch 22/100 train\_acc=0.2098 val\_acc=0.2030

Epoch 23/100 train\_acc=0.2237 val\_acc=0.2126

Epoch 24/100 train\_acc=0.2259 val\_acc=0.2056

Epoch 25/100 train\_acc=0.2285 val\_acc=0.2214

Epoch 26/100 train\_acc=0.2357 val\_acc=0.2292

Epoch 27/100 train\_acc=0.2348 val\_acc=0.2270

Epoch 28/100 train\_acc=0.2402 val\_acc=0.2306

Epoch 29/100 train\_acc=0.2472 val\_acc=0.2138

Epoch 30/100 train\_acc=0.2519 val\_acc=0.2392

Epoch 31/100 train\_acc=0.2581 val\_acc=0.2504

Epoch 32/100 train\_acc=0.2598 val\_acc=0.2610

Epoch 33/100 train\_acc=0.2710 val\_acc=0.2430

Epoch 34/100 train\_acc=0.2715 val\_acc=0.2620

Epoch 35/100 train\_acc=0.2780 val\_acc=0.2580

Epoch 36/100 train\_acc=0.2859 val\_acc=0.2710

Epoch 37/100 train\_acc=0.2890 val\_acc=0.2808

Epoch 38/100 train\_acc=0.2964 val\_acc=0.2840

Epoch 39/100 train\_acc=0.3006 val\_acc=0.2656

Epoch 40/100 train\_acc=0.3082 val\_acc=0.2790

Epoch 41/100 train\_acc=0.3147 val\_acc=0.3022

Epoch 42/100 train\_acc=0.3204 val\_acc=0.2914

Epoch 43/100 train\_acc=0.3275 val\_acc=0.2966

Epoch 44/100 train\_acc=0.3346 val\_acc=0.2986

Epoch 45/100 train\_acc=0.3401 val\_acc=0.3026

Epoch 46/100 train\_acc=0.3470 val\_acc=0.3222

Epoch 47/100 train\_acc=0.3593 val\_acc=0.3238

Epoch 48/100 train\_acc=0.3596 val\_acc=0.3240

Epoch 49/100 train\_acc=0.3702 val\_acc=0.3490

Epoch 50/100 train\_acc=0.3813 val\_acc=0.3428

Epoch 51/100 train\_acc=0.3880 val\_acc=0.3302

Epoch 52/100 train\_acc=0.3956 val\_acc=0.3380

Epoch 53/100 train\_acc=0.3985 val\_acc=0.3512

Epoch 54/100 train\_acc=0.4091 val\_acc=0.3576

Epoch 55/100 train\_acc=0.4168 val\_acc=0.3544

Epoch 56/100 train\_acc=0.4241 val\_acc=0.3676

Epoch 57/100 train\_acc=0.4310 val\_acc=0.3658

Epoch 58/100 train\_acc=0.4435 val\_acc=0.3656

Epoch 59/100 train\_acc=0.4472 val\_acc=0.3798

Epoch 60/100 train\_acc=0.4594 val\_acc=0.3752

Epoch 61/100 train\_acc=0.4657 val\_acc=0.3772

Epoch 62/100 train\_acc=0.4770 val\_acc=0.3926

Epoch 63/100 train\_acc=0.4850 val\_acc=0.4046

Epoch 64/100 train\_acc=0.4891 val\_acc=0.4104

Epoch 65/100 train\_acc=0.5006 val\_acc=0.4036

Epoch 66/100 train\_acc=0.5065 val\_acc=0.4118

Epoch 67/100 train\_acc=0.5133 val\_acc=0.4228

Epoch 68/100 train\_acc=0.5256 val\_acc=0.4242

Epoch 69/100 train\_acc=0.5356 val\_acc=0.4236

Epoch 70/100 train\_acc=0.5448 val\_acc=0.4334

Epoch 71/100 train\_acc=0.5520 val\_acc=0.4328

Epoch 72/100 train\_acc=0.5611 val\_acc=0.4326

Epoch 73/100 train\_acc=0.5686 val\_acc=0.4364

Epoch 74/100 train\_acc=0.5772 val\_acc=0.4408

Epoch 75/100 train\_acc=0.5896 val\_acc=0.4270

Epoch 76/100 train\_acc=0.5938 val\_acc=0.4406

Epoch 77/100 train\_acc=0.6045 val\_acc=0.4476

Epoch 78/100 train\_acc=0.6099 val\_acc=0.4470

Epoch 79/100 train\_acc=0.6177 val\_acc=0.4506

Epoch 80/100 train\_acc=0.6268 val\_acc=0.4594

Epoch 81/100 train\_acc=0.6358 val\_acc=0.4578

Epoch 82/100 train\_acc=0.6400 val\_acc=0.4596

Epoch 83/100 train\_acc=0.6449 val\_acc=0.4644

Epoch 84/100 train\_acc=0.6523 val\_acc=0.4670

Epoch 85/100 train\_acc=0.6618 val\_acc=0.4722

Epoch 86/100 train\_acc=0.6689 val\_acc=0.4668

Epoch 87/100 train\_acc=0.6738 val\_acc=0.4658

Epoch 88/100 train\_acc=0.6762 val\_acc=0.4644

Epoch 89/100 train\_acc=0.6819 val\_acc=0.4772

Epoch 90/100 train\_acc=0.6858 val\_acc=0.4706

Epoch 91/100 train\_acc=0.6945 val\_acc=0.4798

Epoch 92/100 train\_acc=0.6954 val\_acc=0.4722

Epoch 93/100 train\_acc=0.6978 val\_acc=0.4760

Epoch 94/100 train\_acc=0.6995 val\_acc=0.4908

Epoch 95/100 train\_acc=0.7047 val\_acc=0.4758

Epoch 96/100 train\_acc=0.7031 val\_acc=0.4750

Epoch 97/100 train\_acc=0.7064 val\_acc=0.4744

Epoch 98/100 train\_acc=0.7044 val\_acc=0.4710

Epoch 99/100 train\_acc=0.7074 val\_acc=0.4780

Epoch 100/100 train\_acc=0.7091 val\_acc=0.4840

Config {'lr': 0.1, 'weight\_decay': 0.0001, 'batch\_size': 128, 'epochs': 100} → best\_val\_acc=0.4908, test\_acc=0.5101

==================== RUN SUMMARY (2025-05-06 08:40:24) ====================

➜ Environment:

• python\_version: 3.11.12

• pytorch\_version: 2.6.0+cu124

• torchvision\_version: 0.21.0+cu124

• cuda\_available: True

• cuda\_device: Tesla T4

• device\_count: 1

• platform: Linux-6.1.123+-x86\_64-with-glibc2.35

• cwd: /content

➜ Data:

• train\_samples: 45000

• val\_samples: 5000

• test\_samples: 10000

• batch\_size: 128

• num\_batches\_train: 352

• num\_batches\_val: 40

• num\_batches\_test: 79

➜ Seed:

• seed: not set

➜ Hyperparameters:

• lr: 0.1

• weight\_decay: 0.0001

• batch\_size: 128

• epochs: 100

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

SHAKESPEARE

BATCH\_SIZE = 1024 # doubled batch size

SEQ\_LEN = 80

EMBED\_DIM = 128

HIDDEN\_DIM = 256

NUM\_LAYERS = 2

NUM\_EPOCHS = 100

LR = 0.5

MOMENTUM = 0.9

small = full.select(range(1\_000\_000))

raw = small

tv = raw.train\_test\_split(test\_size=0.2, seed=42)

train\_raw, hold = tv["train"], tv["test"]

vt = hold.train\_test\_split(test\_size=0.5, seed=42)

val\_raw, test\_raw = vt["train"], vt["test"]

Epoch 1 → Train Acc: 0.4288, Val Acc: 0.5025

/tmp/ipython-input-2-233185215.py:128: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.

with autocast():

Epoch 2 → Train Acc: 0.5165, Val Acc: 0.5346

Epoch 3 → Train Acc: 0.5411, Val Acc: 0.5513

Epoch 4 → Train Acc: 0.5556, Val Acc: 0.5599

Epoch 5 → Train Acc: 0.5663, Val Acc: 0.5662

Epoch 6 → Train Acc: 0.5738, Val Acc: 0.5710

Epoch 7 → Train Acc: 0.5810, Val Acc: 0.5724

Epoch 8 → Train Acc: 0.5863, Val Acc: 0.5781

Epoch 9 → Train Acc: 0.5918, Val Acc: 0.5803

Epoch 10 → Train Acc: 0.5958, Val Acc: 0.5809

Epoch 11 → Train Acc: 0.5998, Val Acc: 0.5808

Epoch 12 → Train Acc: 0.6035, Val Acc: 0.5807

Epoch 13 → Train Acc: 0.6069, Val Acc: 0.5799

Epoch 14 → Train Acc: 0.6104, Val Acc: 0.5825

Epoch 15 → Train Acc: 0.6134, Val Acc: 0.5825

Epoch 16 → Train Acc: 0.6169, Val Acc: 0.5832

Epoch 17 → Train Acc: 0.6195, Val Acc: 0.5801

Epoch 18 → Train Acc: 0.6225, Val Acc: 0.5826

Epoch 19 → Train Acc: 0.6245, Val Acc: 0.5804

Epoch 20 → Train Acc: 0.6271, Val Acc: 0.5807

Epoch 21 → Train Acc: 0.6296, Val Acc: 0.5782

Epoch 22 → Train Acc: 0.6317, Val Acc: 0.5802

Epoch 23 → Train Acc: 0.6339, Val Acc: 0.5808

Epoch 24 → Train Acc: 0.6357, Val Acc: 0.5784

Epoch 25 → Train Acc: 0.6383, Val Acc: 0.5773

Epoch 26 → Train Acc: 0.6399, Val Acc: 0.5802

Epoch 27 → Train Acc: 0.6418, Val Acc: 0.5785

Epoch 28 → Train Acc: 0.6427, Val Acc: 0.5768

Epoch 29 → Train Acc: 0.6452, Val Acc: 0.5764

Epoch 30 → Train Acc: 0.6465, Val Acc: 0.5772

Epoch 31 → Train Acc: 0.6482, Val Acc: 0.5773

Epoch 32 → Train Acc: 0.6496, Val Acc: 0.5744

Epoch 33 → Train Acc: 0.6518, Val Acc: 0.5754

Epoch 34 → Train Acc: 0.6524, Val Acc: 0.5758

Epoch 35 → Train Acc: 0.6544, Val Acc: 0.5739

Epoch 36 → Train Acc: 0.6565, Val Acc: 0.5757

Epoch 37 → Train Acc: 0.6582, Val Acc: 0.5740

Epoch 38 → Train Acc: 0.6594, Val Acc: 0.5747

Epoch 39 → Train Acc: 0.6610, Val Acc: 0.5750

Epoch 40 → Train Acc: 0.6624, Val Acc: 0.5734

Epoch 41 → Train Acc: 0.6642, Val Acc: 0.5733

Epoch 42 → Train Acc: 0.6657, Val Acc: 0.5736

Epoch 43 → Train Acc: 0.6670, Val Acc: 0.5738

Epoch 44 → Train Acc: 0.6692, Val Acc: 0.5696

Epoch 45 → Train Acc: 0.6710, Val Acc: 0.5711

Epoch 46 → Train Acc: 0.6725, Val Acc: 0.5710

Epoch 47 → Train Acc: 0.6743, Val Acc: 0.5694

Epoch 48 → Train Acc: 0.6754, Val Acc: 0.5685

Epoch 49 → Train Acc: 0.6777, Val Acc: 0.5713

Epoch 50 → Train Acc: 0.6798, Val Acc: 0.5701

Epoch 51 → Train Acc: 0.6821, Val Acc: 0.5685

Epoch 52 → Train Acc: 0.6832, Val Acc: 0.5693

Epoch 53 → Train Acc: 0.6855, Val Acc: 0.5681

Epoch 54 → Train Acc: 0.6874, Val Acc: 0.5688

Epoch 55 → Train Acc: 0.6890, Val Acc: 0.5689

Epoch 56 → Train Acc: 0.6917, Val Acc: 0.5683

Epoch 57 → Train Acc: 0.6932, Val Acc: 0.5683

Epoch 58 → Train Acc: 0.6959, Val Acc: 0.5694

Epoch 59 → Train Acc: 0.6977, Val Acc: 0.5675

Epoch 60 → Train Acc: 0.6994, Val Acc: 0.5676

Epoch 61 → Train Acc: 0.7016, Val Acc: 0.5667

Epoch 62 → Train Acc: 0.7035, Val Acc: 0.5677

Epoch 63 → Train Acc: 0.7063, Val Acc: 0.5671

Epoch 64 → Train Acc: 0.7087, Val Acc: 0.5644

Epoch 65 → Train Acc: 0.7110, Val Acc: 0.5655

Epoch 66 → Train Acc: 0.7130, Val Acc: 0.5653

Epoch 67 → Train Acc: 0.7147, Val Acc: 0.5646

Epoch 68 → Train Acc: 0.7181, Val Acc: 0.5658

Epoch 69 → Train Acc: 0.7200, Val Acc: 0.5644

Epoch 70 → Train Acc: 0.7224, Val Acc: 0.5640

Epoch 71 → Train Acc: 0.7245, Val Acc: 0.5630

Epoch 72 → Train Acc: 0.7273, Val Acc: 0.5642

Epoch 73 → Train Acc: 0.7290, Val Acc: 0.5634

Epoch 74 → Train Acc: 0.7311, Val Acc: 0.5642

Epoch 75 → Train Acc: 0.7335, Val Acc: 0.5636

Epoch 76 → Train Acc: 0.7355, Val Acc: 0.5633

Epoch 77 → Train Acc: 0.7371, Val Acc: 0.5610

Epoch 78 → Train Acc: 0.7394, Val Acc: 0.5624

Epoch 79 → Train Acc: 0.7415, Val Acc: 0.5626

Epoch 80 → Train Acc: 0.7434, Val Acc: 0.5614

Epoch 81 → Train Acc: 0.7456, Val Acc: 0.5625

Epoch 82 → Train Acc: 0.7465, Val Acc: 0.5616

Epoch 83 → Train Acc: 0.7486, Val Acc: 0.5609

Epoch 84 → Train Acc: 0.7508, Val Acc: 0.5601

Epoch 85 → Train Acc: 0.7522, Val Acc: 0.5613

Epoch 86 → Train Acc: 0.7533, Val Acc: 0.5609

Epoch 87 → Train Acc: 0.7549, Val Acc: 0.5612

Epoch 88 → Train Acc: 0.7560, Val Acc: 0.5613

Epoch 89 → Train Acc: 0.7573, Val Acc: 0.5607

Epoch 90 → Train Acc: 0.7585, Val Acc: 0.5598

Epoch 91 → Train Acc: 0.7592, Val Acc: 0.5605

Epoch 92 → Train Acc: 0.7602, Val Acc: 0.5602

Epoch 93 → Train Acc: 0.7607, Val Acc: 0.5600

Epoch 94 → Train Acc: 0.7610, Val Acc: 0.5602

Epoch 95 → Train Acc: 0.7621, Val Acc: 0.5602

Epoch 96 → Train Acc: 0.7626, Val Acc: 0.5602

Epoch 97 → Train Acc: 0.7630, Val Acc: 0.5603

Epoch 98 → Train Acc: 0.7630, Val Acc: 0.5605

Epoch 99 → Train Acc: 0.7636, Val Acc: 0.5606

Epoch 100 → Train Acc: 0.7634, Val Acc: 0.5604

Test → Loss: 1.7746, Acc: 0.5593

another one

# Hyperparameters

BATCH\_SIZE = 1024 # doubled batch size

SEQ\_LEN = 80

EMBED\_DIM = 128

HIDDEN\_DIM = 256

NUM\_LAYERS = 2

NUM\_EPOCHS = 50

LR = 0.5

MOMENTUM = 0.9

…

for EMB in [8]:

for LR in [0.1]:

for MOM in [0.0]:

[emb8\_lr0.1\_mom0.0] Epoch 1 → Train Acc: 0.2096, Val Acc: 0.2344

[emb8\_lr0.1\_mom0.0] Epoch 2 → Train Acc: 0.2899, Val Acc: 0.3326

[emb8\_lr0.1\_mom0.0] Epoch 3 → Train Acc: 0.3585, Val Acc: 0.3851

[emb8\_lr0.1\_mom0.0] Epoch 4 → Train Acc: 0.3990, Val Acc: 0.4169

[emb8\_lr0.1\_mom0.0] Epoch 5 → Train Acc: 0.4272, Val Acc: 0.4386

[emb8\_lr0.1\_mom0.0] Epoch 6 → Train Acc: 0.4463, Val Acc: 0.4598

[emb8\_lr0.1\_mom0.0] Epoch 7 → Train Acc: 0.4630, Val Acc: 0.4752

[emb8\_lr0.1\_mom0.0] Epoch 8 → Train Acc: 0.4759, Val Acc: 0.4855

[emb8\_lr0.1\_mom0.0] Epoch 9 → Train Acc: 0.4861, Val Acc: 0.4958

[emb8\_lr0.1\_mom0.0] Epoch 10 → Train Acc: 0.4945, Val Acc: 0.5046

[emb8\_lr0.1\_mom0.0] Epoch 11 → Train Acc: 0.5012, Val Acc: 0.5085

[emb8\_lr0.1\_mom0.0] Epoch 12 → Train Acc: 0.5069, Val Acc: 0.5160

[emb8\_lr0.1\_mom0.0] Epoch 13 → Train Acc: 0.5121, Val Acc: 0.5210

[emb8\_lr0.1\_mom0.0] Epoch 14 → Train Acc: 0.5164, Val Acc: 0.5256

[emb8\_lr0.1\_mom0.0] Epoch 15 → Train Acc: 0.5204, Val Acc: 0.5287

[emb8\_lr0.1\_mom0.0] Epoch 16 → Train Acc: 0.5238, Val Acc: 0.5324

[emb8\_lr0.1\_mom0.0] Epoch 17 → Train Acc: 0.5270, Val Acc: 0.5357

[emb8\_lr0.1\_mom0.0] Epoch 18 → Train Acc: 0.5297, Val Acc: 0.5379

[emb8\_lr0.1\_mom0.0] Epoch 19 → Train Acc: 0.5321, Val Acc: 0.5402

[emb8\_lr0.1\_mom0.0] Epoch 20 → Train Acc: 0.5343, Val Acc: 0.5436

[emb8\_lr0.1\_mom0.0] Epoch 21 → Train Acc: 0.5367, Val Acc: 0.5448

[emb8\_lr0.1\_mom0.0] Epoch 22 → Train Acc: 0.5382, Val Acc: 0.5470

[emb8\_lr0.1\_mom0.0] Epoch 23 → Train Acc: 0.5400, Val Acc: 0.5473

[emb8\_lr0.1\_mom0.0] Epoch 24 → Train Acc: 0.5415, Val Acc: 0.5500

[emb8\_lr0.1\_mom0.0] Epoch 25 → Train Acc: 0.5430, Val Acc: 0.5510

[emb8\_lr0.1\_mom0.0] Epoch 26 → Train Acc: 0.5445, Val Acc: 0.5523

/usr/local/lib/python3.11/dist-packages/torch/nn/modules/rnn.py:1124: UserWarning: RNN module weights are not part of single contiguous chunk of memory. This means they need to be compacted at every call, possibly greatly increasing memory usage. To compact weights again call flatten\_parameters(). (Triggered internally at /pytorch/aten/src/ATen/native/cudnn/RNN.cpp:1412.)

result = \_VF.lstm(

[emb8\_lr0.1\_mom0.0] Epoch 27 → Train Acc: 0.5454, Val Acc: 0.5537

/tmp/ipython-input-4-2835131870.py:154: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.

with autocast():

[emb8\_lr0.1\_mom0.0] Epoch 28 → Train Acc: 0.5469, Val Acc: 0.5550

[emb8\_lr0.1\_mom0.0] Epoch 29 → Train Acc: 0.5476, Val Acc: 0.5557

[emb8\_lr0.1\_mom0.0] Epoch 30 → Train Acc: 0.5487, Val Acc: 0.5556

[emb8\_lr0.1\_mom0.0] Epoch 31 → Train Acc: 0.5493, Val Acc: 0.5567

[emb8\_lr0.1\_mom0.0] Epoch 32 → Train Acc: 0.5502, Val Acc: 0.5574

[emb8\_lr0.1\_mom0.0] Epoch 33 → Train Acc: 0.5508, Val Acc: 0.5582

[emb8\_lr0.1\_mom0.0] Epoch 34 → Train Acc: 0.5515, Val Acc: 0.5587

[emb8\_lr0.1\_mom0.0] Epoch 35 → Train Acc: 0.5520, Val Acc: 0.5591

[emb8\_lr0.1\_mom0.0] Epoch 36 → Train Acc: 0.5526, Val Acc: 0.5596

[emb8\_lr0.1\_mom0.0] Epoch 37 → Train Acc: 0.5531, Val Acc: 0.5598

[emb8\_lr0.1\_mom0.0] Epoch 38 → Train Acc: 0.5531, Val Acc: 0.5601

[emb8\_lr0.1\_mom0.0] Epoch 39 → Train Acc: 0.5537, Val Acc: 0.5605

[emb8\_lr0.1\_mom0.0] Epoch 40 → Train Acc: 0.5540, Val Acc: 0.5608

[emb8\_lr0.1\_mom0.0] Epoch 41 → Train Acc: 0.5543, Val Acc: 0.5610

[emb8\_lr0.1\_mom0.0] Epoch 42 → Train Acc: 0.5545, Val Acc: 0.5610

[emb8\_lr0.1\_mom0.0] Epoch 43 → Train Acc: 0.5548, Val Acc: 0.5613

[emb8\_lr0.1\_mom0.0] Epoch 44 → Train Acc: 0.5551, Val Acc: 0.5615

[emb8\_lr0.1\_mom0.0] Epoch 45 → Train Acc: 0.5549, Val Acc: 0.5615

[emb8\_lr0.1\_mom0.0] Epoch 46 → Train Acc: 0.5551, Val Acc: 0.5616

[emb8\_lr0.1\_mom0.0] Epoch 47 → Train Acc: 0.5551, Val Acc: 0.5616

[emb8\_lr0.1\_mom0.0] Epoch 48 → Train Acc: 0.5552, Val Acc: 0.5615

[emb8\_lr0.1\_mom0.0] Epoch 49 → Train Acc: 0.5551, Val Acc: 0.5616

[emb8\_lr0.1\_mom0.0] Epoch 50 → Train Acc: 0.5551, Val Acc: 0.5616

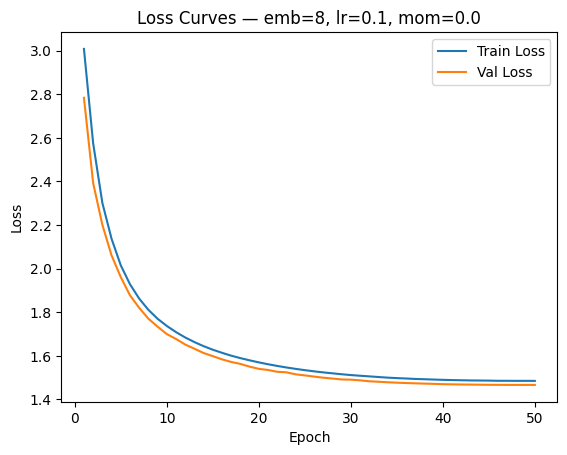
=== Best config ===

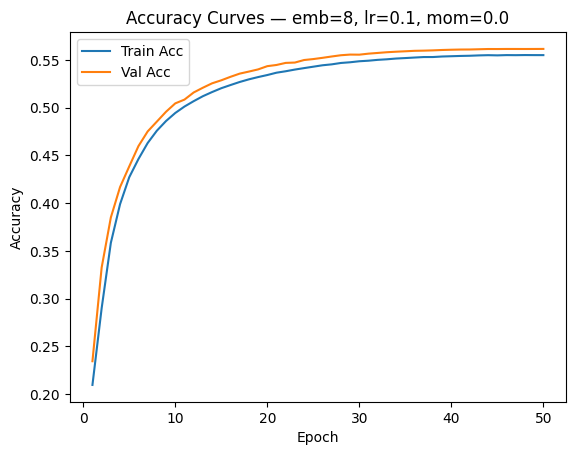
{'embed': 8, 'lr': 0.1, 'momentum': 0.0} → Val Acc: 0.5616152437206264

/usr/local/lib/python3.11/dist-packages/torch/utils/data/dataloader.py:624: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

warnings.warn(

Test → Loss: 1.4640, Acc: 0.5624





# Hyperparameters

BATCH\_SIZE = 1024 # doubled batch size

SEQ\_LEN = 80

EMBED\_DIM = 128

HIDDEN\_DIM = 256

NUM\_LAYERS = 2

NUM\_EPOCHS = 50

LR = 0.5

MOMENTUM = 0.9

for EMB in [8]:

for LR in [0.1]:

for MOM in [0.9]:

[emb8\_lr0.1\_mom0.9] Epoch 1 → Train Acc: 0.4001, Val Acc: 0.4930

/tmp/ipython-input-2-634397099.py:154: FutureWarning: `torch.cuda.amp.autocast(args...)` is deprecated. Please use `torch.amp.autocast('cuda', args...)` instead.

with autocast():

[emb8\_lr0.1\_mom0.9] Epoch 2 → Train Acc: 0.5132, Val Acc: 0.5360

[emb8\_lr0.1\_mom0.9] Epoch 3 → Train Acc: 0.5417, Val Acc: 0.5555

[emb8\_lr0.1\_mom0.9] Epoch 4 → Train Acc: 0.5565, Val Acc: 0.5651

[emb8\_lr0.1\_mom0.9] Epoch 5 → Train Acc: 0.5663, Val Acc: 0.5726

[emb8\_lr0.1\_mom0.9] Epoch 6 → Train Acc: 0.5735, Val Acc: 0.5791

[emb8\_lr0.1\_mom0.9] Epoch 7 → Train Acc: 0.5788, Val Acc: 0.5830

[emb8\_lr0.1\_mom0.9] Epoch 8 → Train Acc: 0.5832, Val Acc: 0.5861

[emb8\_lr0.1\_mom0.9] Epoch 9 → Train Acc: 0.5868, Val Acc: 0.5888

[emb8\_lr0.1\_mom0.9] Epoch 10 → Train Acc: 0.5899, Val Acc: 0.5908

[emb8\_lr0.1\_mom0.9] Epoch 11 → Train Acc: 0.5926, Val Acc: 0.5930

[emb8\_lr0.1\_mom0.9] Epoch 12 → Train Acc: 0.5949, Val Acc: 0.5942

[emb8\_lr0.1\_mom0.9] Epoch 13 → Train Acc: 0.5973, Val Acc: 0.5961

[emb8\_lr0.1\_mom0.9] Epoch 14 → Train Acc: 0.5993, Val Acc: 0.5973

[emb8\_lr0.1\_mom0.9] Epoch 15 → Train Acc: 0.6008, Val Acc: 0.5984

[emb8\_lr0.1\_mom0.9] Epoch 16 → Train Acc: 0.6029, Val Acc: 0.6005

[emb8\_lr0.1\_mom0.9] Epoch 17 → Train Acc: 0.6041, Val Acc: 0.6008

[emb8\_lr0.1\_mom0.9] Epoch 18 → Train Acc: 0.6057, Val Acc: 0.6015

[emb8\_lr0.1\_mom0.9] Epoch 19 → Train Acc: 0.6072, Val Acc: 0.6024

[emb8\_lr0.1\_mom0.9] Epoch 20 → Train Acc: 0.6087, Val Acc: 0.6032

[emb8\_lr0.1\_mom0.9] Epoch 21 → Train Acc: 0.6100, Val Acc: 0.6039

[emb8\_lr0.1\_mom0.9] Epoch 22 → Train Acc: 0.6114, Val Acc: 0.6047

[emb8\_lr0.1\_mom0.9] Epoch 23 → Train Acc: 0.6124, Val Acc: 0.6056

[emb8\_lr0.1\_mom0.9] Epoch 24 → Train Acc: 0.6133, Val Acc: 0.6063

[emb8\_lr0.1\_mom0.9] Epoch 25 → Train Acc: 0.6145, Val Acc: 0.6052

[emb8\_lr0.1\_mom0.9] Epoch 26 → Train Acc: 0.6158, Val Acc: 0.6066

[emb8\_lr0.1\_mom0.9] Epoch 27 → Train Acc: 0.6169, Val Acc: 0.6071

[emb8\_lr0.1\_mom0.9] Epoch 28 → Train Acc: 0.6174, Val Acc: 0.6074

[emb8\_lr0.1\_mom0.9] Epoch 29 → Train Acc: 0.6187, Val Acc: 0.6078

[emb8\_lr0.1\_mom0.9] Epoch 30 → Train Acc: 0.6192, Val Acc: 0.6088

[emb8\_lr0.1\_mom0.9] Epoch 31 → Train Acc: 0.6202, Val Acc: 0.6095

[emb8\_lr0.1\_mom0.9] Epoch 32 → Train Acc: 0.6212, Val Acc: 0.6095

[emb8\_lr0.1\_mom0.9] Epoch 33 → Train Acc: 0.6219, Val Acc: 0.6094

[emb8\_lr0.1\_mom0.9] Epoch 34 → Train Acc: 0.6228, Val Acc: 0.6102

[emb8\_lr0.1\_mom0.9] Epoch 35 → Train Acc: 0.6234, Val Acc: 0.6105

[emb8\_lr0.1\_mom0.9] Epoch 36 → Train Acc: 0.6241, Val Acc: 0.6105

[emb8\_lr0.1\_mom0.9] Epoch 37 → Train Acc: 0.6247, Val Acc: 0.6110

[emb8\_lr0.1\_mom0.9] Epoch 38 → Train Acc: 0.6255, Val Acc: 0.6112

[emb8\_lr0.1\_mom0.9] Epoch 39 → Train Acc: 0.6258, Val Acc: 0.6109

[emb8\_lr0.1\_mom0.9] Epoch 40 → Train Acc: 0.6265, Val Acc: 0.6111

[emb8\_lr0.1\_mom0.9] Epoch 41 → Train Acc: 0.6269, Val Acc: 0.6114

[emb8\_lr0.1\_mom0.9] Epoch 42 → Train Acc: 0.6273, Val Acc: 0.6115

[emb8\_lr0.1\_mom0.9] Epoch 43 → Train Acc: 0.6276, Val Acc: 0.6117

[emb8\_lr0.1\_mom0.9] Epoch 44 → Train Acc: 0.6281, Val Acc: 0.6115

[emb8\_lr0.1\_mom0.9] Epoch 45 → Train Acc: 0.6281, Val Acc: 0.6121

[emb8\_lr0.1\_mom0.9] Epoch 46 → Train Acc: 0.6287, Val Acc: 0.6118

[emb8\_lr0.1\_mom0.9] Epoch 47 → Train Acc: 0.6289, Val Acc: 0.6119

[emb8\_lr0.1\_mom0.9] Epoch 48 → Train Acc: 0.6289, Val Acc: 0.6120

[emb8\_lr0.1\_mom0.9] Epoch 49 → Train Acc: 0.6293, Val Acc: 0.6120

[emb8\_lr0.1\_mom0.9] Epoch 50 → Train Acc: 0.6291, Val Acc: 0.6121

=== Best config ===

{'embed': 8, 'lr': 0.1, 'momentum': 0.9} → Val Acc: 0.6121294731810286

/usr/local/lib/python3.11/dist-packages/torch/utils/data/dataloader.py:624: UserWarning: This DataLoader will create 8 worker processes in total. Our suggested max number of worker in current system is 2, which is smaller than what this DataLoader is going to create. Please be aware that excessive worker creation might get DataLoader running slow or even freeze, lower the worker number to avoid potential slowness/freeze if necessary.

warnings.warn(

Test → Loss: 1.2596, Acc: 0.6128

