python -u engine.py --corpus=swda --mode=train --gpu=0,1 --batch\_size=2 --batch\_size\_val=2 --epochs=100 --lr=0.0001 --nlayer=2 --chunk\_size=196 --dropout=0.5 --nfinetune=1 --speaker\_info=emb\_cls --topic\_info=none --nclass=43 --emb\_batch=0

Namespace(corpus='swda', mode='train', nclass=43, batch\_size=2, batch\_size\_val=2, emb\_batch=0, epochs=100, gpu='0,1', lr=0.0001, nlayer=2, chunk\_size=196, dropout=0.5, speaker\_info='emb\_cls', topic\_info='none', nfinetune=1, seed=0)

Tokenizing train....

Done

Tokenizing val....

Done

Tokenizing test....

Done

Done

Let's use 1 GPUs!

Initializing model....

The cache for model files in Transformers v4.22.0 has been updated. Migrating your old cache. This is a one-time only operation. You can interrupt this and resume the migration later on by calling `transformers.utils.move\_cache()`.

0it [00:00, ?it/s]

/usr/local/lib/python3.11/dist-packages/huggingface\_hub/file\_download.py:896: FutureWarning: `resume\_download` is deprecated and will be removed in version 1.0.0. Downloads always resume when possible. If you want to force a new download, use `force\_download=True`.

warnings.warn(

Some weights of the model checkpoint at roberta-base were not used when initializing RobertaModel: ['lm\_head.bias', 'lm\_head.layer\_norm.weight', 'lm\_head.layer\_norm.bias', 'lm\_head.dense.bias', 'lm\_head.dense.weight']

- This IS expected if you are initializing RobertaModel from the checkpoint of a model trained on another task or with another architecture (e.g. initializing a BertForSequenceClassification model from a BertForPreTraining model).

- This IS NOT expected if you are initializing RobertaModel from the checkpoint of a model that you expect to be exactly identical (initializing a BertForSequenceClassification model from a BertForSequenceClassification model).

2025-04-16 03:12:51.750644: E external/local\_xla/xla/stream\_executor/cuda/cuda\_fft.cc:477] Unable to register cuFFT factory: Attempting to register factory for plugin cuFFT when one has already been registered

WARNING: All log messages before absl::InitializeLog() is called are written to STDERR

E0000 00:00:1744773171.768780 9452 cuda\_dnn.cc:8310] Unable to register cuDNN factory: Attempting to register factory for plugin cuDNN when one has already been registered

E0000 00:00:1744773171.774066 9452 cuda\_blas.cc:1418] Unable to register cuBLAS factory: Attempting to register factory for plugin cuBLAS when one has already been registered

/usr/local/lib/python3.11/dist-packages/transformers/optimization.py:391: FutureWarning: This implementation of AdamW is deprecated and will be removed in a future version. Use the PyTorch implementation torch.optim.AdamW instead, or set `no\_deprecation\_warning=True` to disable this warning

warnings.warn(

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 1\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 3.769 loss\_act:3.769

Batch: 36/714 loss: 2.961 loss\_act:2.961

Batch: 71/714 loss: 2.611 loss\_act:2.611

Batch: 106/714 loss: 2.089 loss\_act:2.089

Batch: 141/714 loss: 1.581 loss\_act:1.581

Batch: 176/714 loss: 1.342 loss\_act:1.342

Batch: 211/714 loss: 1.328 loss\_act:1.328

Batch: 246/714 loss: 1.200 loss\_act:1.200

Batch: 281/714 loss: 1.102 loss\_act:1.102

Batch: 316/714 loss: 1.198 loss\_act:1.198

Batch: 351/714 loss: 0.960 loss\_act:0.960

Batch: 386/714 loss: 0.867 loss\_act:0.867

Batch: 421/714 loss: 0.930 loss\_act:0.930

Batch: 456/714 loss: 1.213 loss\_act:1.213

Batch: 491/714 loss: 0.967 loss\_act:0.967

Batch: 526/714 loss: 0.853 loss\_act:0.853

Batch: 561/714 loss: 0.804 loss\_act:0.804

Batch: 596/714 loss: 0.826 loss\_act:0.826

Batch: 631/714 loss: 1.144 loss\_act:1.144

Batch: 666/714 loss: 0.645 loss\_act:0.645

Batch: 701/714 loss: 0.745 loss\_act:0.745

Batch: 714/714 loss: 0.633 loss\_act:0.633

Epoch 1 Train Loss: 1.241 Val Acc: 0.779 Test Acc: 0.765

Best Epoch: 1 Best Epoch Val Acc: 0.779 Best Epoch Test Acc: 0.765, Best Test Acc: 0.765

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 2\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.560 loss\_act:0.560

Batch: 36/714 loss: 0.724 loss\_act:0.724

Batch: 71/714 loss: 0.455 loss\_act:0.455

Batch: 106/714 loss: 0.623 loss\_act:0.623

Batch: 141/714 loss: 0.718 loss\_act:0.718

Batch: 176/714 loss: 0.395 loss\_act:0.395

Batch: 211/714 loss: 0.528 loss\_act:0.528

Batch: 246/714 loss: 0.825 loss\_act:0.825

Batch: 281/714 loss: 0.764 loss\_act:0.764

Batch: 316/714 loss: 0.772 loss\_act:0.772

Batch: 351/714 loss: 0.559 loss\_act:0.559

Batch: 386/714 loss: 0.780 loss\_act:0.780

Batch: 421/714 loss: 0.677 loss\_act:0.677

Batch: 456/714 loss: 0.752 loss\_act:0.752

Batch: 491/714 loss: 0.581 loss\_act:0.581

Batch: 526/714 loss: 0.741 loss\_act:0.741

Batch: 561/714 loss: 0.615 loss\_act:0.615

Batch: 596/714 loss: 0.913 loss\_act:0.913

Batch: 631/714 loss: 0.567 loss\_act:0.567

Batch: 666/714 loss: 0.939 loss\_act:0.939

Batch: 701/714 loss: 0.721 loss\_act:0.721

Batch: 714/714 loss: 0.562 loss\_act:0.562

Epoch 2 Train Loss: 0.709 Val Acc: 0.808 Test Acc: 0.802

Best Epoch: 2 Best Epoch Val Acc: 0.808 Best Epoch Test Acc: 0.802, Best Test Acc: 0.802

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 3\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.512 loss\_act:0.512

Batch: 36/714 loss: 0.605 loss\_act:0.605

Batch: 71/714 loss: 0.567 loss\_act:0.567

Batch: 106/714 loss: 0.524 loss\_act:0.524

Batch: 141/714 loss: 0.630 loss\_act:0.630

Batch: 176/714 loss: 0.715 loss\_act:0.715

Batch: 211/714 loss: 0.539 loss\_act:0.539

Batch: 246/714 loss: 0.697 loss\_act:0.697

Batch: 281/714 loss: 0.479 loss\_act:0.479

Batch: 316/714 loss: 0.765 loss\_act:0.765

Batch: 351/714 loss: 0.653 loss\_act:0.653

Batch: 386/714 loss: 0.721 loss\_act:0.721

Batch: 421/714 loss: 0.606 loss\_act:0.606

Batch: 456/714 loss: 0.467 loss\_act:0.467

Batch: 491/714 loss: 0.593 loss\_act:0.593

Batch: 526/714 loss: 0.618 loss\_act:0.618

Batch: 561/714 loss: 0.454 loss\_act:0.454

Batch: 596/714 loss: 0.586 loss\_act:0.586

Batch: 631/714 loss: 0.715 loss\_act:0.715

Batch: 666/714 loss: 0.644 loss\_act:0.644

Batch: 701/714 loss: 0.474 loss\_act:0.474

Batch: 714/714 loss: 0.605 loss\_act:0.605

Epoch 3 Train Loss: 0.645 Val Acc: 0.817 Test Acc: 0.808

Best Epoch: 3 Best Epoch Val Acc: 0.817 Best Epoch Test Acc: 0.808, Best Test Acc: 0.808

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 4\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.400 loss\_act:0.400

Batch: 36/714 loss: 0.629 loss\_act:0.629

Batch: 71/714 loss: 0.678 loss\_act:0.678

Batch: 106/714 loss: 0.584 loss\_act:0.584

Batch: 141/714 loss: 0.541 loss\_act:0.541

Batch: 176/714 loss: 0.764 loss\_act:0.764

Batch: 211/714 loss: 0.496 loss\_act:0.496

Batch: 246/714 loss: 0.692 loss\_act:0.692

Batch: 281/714 loss: 0.622 loss\_act:0.622

Batch: 316/714 loss: 0.486 loss\_act:0.486

Batch: 351/714 loss: 0.498 loss\_act:0.498

Batch: 386/714 loss: 0.498 loss\_act:0.498

Batch: 421/714 loss: 0.686 loss\_act:0.686

Batch: 456/714 loss: 0.768 loss\_act:0.768

Batch: 491/714 loss: 0.440 loss\_act:0.440

Batch: 526/714 loss: 0.502 loss\_act:0.502

Batch: 561/714 loss: 0.653 loss\_act:0.653

Batch: 596/714 loss: 0.678 loss\_act:0.678

Batch: 631/714 loss: 0.568 loss\_act:0.568

Batch: 666/714 loss: 1.055 loss\_act:1.055

Batch: 701/714 loss: 0.532 loss\_act:0.532

Batch: 714/714 loss: 0.485 loss\_act:0.485

Epoch 4 Train Loss: 0.602 Val Acc: 0.823 Test Acc: 0.808

Best Epoch: 4 Best Epoch Val Acc: 0.823 Best Epoch Test Acc: 0.808, Best Test Acc: 0.808

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 5\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.468 loss\_act:0.468

Batch: 36/714 loss: 0.527 loss\_act:0.527

Batch: 71/714 loss: 0.637 loss\_act:0.637

Batch: 106/714 loss: 0.625 loss\_act:0.625

Batch: 141/714 loss: 0.663 loss\_act:0.663

Batch: 176/714 loss: 0.421 loss\_act:0.421

Batch: 211/714 loss: 0.705 loss\_act:0.705

Batch: 246/714 loss: 0.697 loss\_act:0.697

Batch: 281/714 loss: 0.706 loss\_act:0.706

Batch: 316/714 loss: 0.484 loss\_act:0.484

Batch: 351/714 loss: 0.690 loss\_act:0.690

Batch: 386/714 loss: 0.538 loss\_act:0.538

Batch: 421/714 loss: 0.422 loss\_act:0.422

Batch: 456/714 loss: 0.515 loss\_act:0.515

Batch: 491/714 loss: 0.382 loss\_act:0.382

Batch: 526/714 loss: 0.593 loss\_act:0.593

Batch: 561/714 loss: 0.928 loss\_act:0.928

Batch: 596/714 loss: 0.518 loss\_act:0.518

Batch: 631/714 loss: 0.326 loss\_act:0.326

Batch: 666/714 loss: 0.422 loss\_act:0.422

Batch: 701/714 loss: 0.567 loss\_act:0.567

Batch: 714/714 loss: 0.679 loss\_act:0.679

Epoch 5 Train Loss: 0.577 Val Acc: 0.821 Test Acc: 0.804

Best Epoch: 4 Best Epoch Val Acc: 0.823 Best Epoch Test Acc: 0.808, Best Test Acc: 0.808

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 6\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.654 loss\_act:0.654

Batch: 36/714 loss: 0.668 loss\_act:0.668

Batch: 71/714 loss: 0.647 loss\_act:0.647

Batch: 106/714 loss: 0.435 loss\_act:0.435

Batch: 141/714 loss: 0.557 loss\_act:0.557

Batch: 176/714 loss: 0.453 loss\_act:0.453

Batch: 211/714 loss: 0.746 loss\_act:0.746

Batch: 246/714 loss: 0.638 loss\_act:0.638

Batch: 281/714 loss: 0.651 loss\_act:0.651

Batch: 316/714 loss: 0.569 loss\_act:0.569

Batch: 351/714 loss: 0.355 loss\_act:0.355

Batch: 386/714 loss: 0.728 loss\_act:0.728

Batch: 421/714 loss: 0.500 loss\_act:0.500

Batch: 456/714 loss: 0.443 loss\_act:0.443

Batch: 491/714 loss: 0.636 loss\_act:0.636

Batch: 526/714 loss: 0.393 loss\_act:0.393

Batch: 561/714 loss: 0.522 loss\_act:0.522

Batch: 596/714 loss: 0.541 loss\_act:0.541

Batch: 631/714 loss: 0.584 loss\_act:0.584

Batch: 666/714 loss: 0.621 loss\_act:0.621

Batch: 701/714 loss: 0.564 loss\_act:0.564

Batch: 714/714 loss: 0.596 loss\_act:0.596

Epoch 6 Train Loss: 0.554 Val Acc: 0.821 Test Acc: 0.800

Best Epoch: 4 Best Epoch Val Acc: 0.823 Best Epoch Test Acc: 0.808, Best Test Acc: 0.808

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 7\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.466 loss\_act:0.466

Batch: 36/714 loss: 0.460 loss\_act:0.460

Batch: 71/714 loss: 0.473 loss\_act:0.473

Batch: 106/714 loss: 0.645 loss\_act:0.645

Batch: 141/714 loss: 0.434 loss\_act:0.434

Batch: 176/714 loss: 0.610 loss\_act:0.610

Batch: 211/714 loss: 0.438 loss\_act:0.438

Batch: 246/714 loss: 0.579 loss\_act:0.579

Batch: 281/714 loss: 0.390 loss\_act:0.390

Batch: 316/714 loss: 0.446 loss\_act:0.446

Batch: 351/714 loss: 0.691 loss\_act:0.691

Batch: 386/714 loss: 0.625 loss\_act:0.625

Batch: 421/714 loss: 0.575 loss\_act:0.575

Batch: 456/714 loss: 0.553 loss\_act:0.553

Batch: 491/714 loss: 0.621 loss\_act:0.621

Batch: 526/714 loss: 0.527 loss\_act:0.527

Batch: 561/714 loss: 0.481 loss\_act:0.481

Batch: 596/714 loss: 0.335 loss\_act:0.335

Batch: 631/714 loss: 0.651 loss\_act:0.651

Batch: 666/714 loss: 0.470 loss\_act:0.470

Batch: 701/714 loss: 0.585 loss\_act:0.585

Batch: 714/714 loss: 0.474 loss\_act:0.474

Epoch 7 Train Loss: 0.537 Val Acc: 0.825 Test Acc: 0.820

Best Epoch: 7 Best Epoch Val Acc: 0.825 Best Epoch Test Acc: 0.820, Best Test Acc: 0.820

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 8\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.453 loss\_act:0.453

Batch: 36/714 loss: 0.617 loss\_act:0.617

Batch: 71/714 loss: 0.526 loss\_act:0.526

Batch: 106/714 loss: 0.340 loss\_act:0.340

Batch: 141/714 loss: 0.529 loss\_act:0.529

Batch: 176/714 loss: 0.494 loss\_act:0.494

Batch: 211/714 loss: 0.409 loss\_act:0.409

Batch: 246/714 loss: 0.541 loss\_act:0.541

Batch: 281/714 loss: 0.566 loss\_act:0.566

Batch: 316/714 loss: 0.739 loss\_act:0.739

Batch: 351/714 loss: 0.465 loss\_act:0.465

Batch: 386/714 loss: 0.505 loss\_act:0.505

Batch: 421/714 loss: 0.511 loss\_act:0.511

Batch: 456/714 loss: 0.515 loss\_act:0.515

Batch: 491/714 loss: 0.562 loss\_act:0.562

Batch: 526/714 loss: 0.644 loss\_act:0.644

Batch: 561/714 loss: 0.369 loss\_act:0.369

Batch: 596/714 loss: 0.453 loss\_act:0.453

Batch: 631/714 loss: 0.671 loss\_act:0.671

Batch: 666/714 loss: 0.554 loss\_act:0.554

Batch: 701/714 loss: 0.534 loss\_act:0.534

Batch: 714/714 loss: 0.748 loss\_act:0.748

Epoch 8 Train Loss: 0.522 Val Acc: 0.826 Test Acc: 0.819

Best Epoch: 8 Best Epoch Val Acc: 0.826 Best Epoch Test Acc: 0.819, Best Test Acc: 0.820

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 9\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.426 loss\_act:0.426

Batch: 36/714 loss: 0.450 loss\_act:0.450

Batch: 71/714 loss: 0.443 loss\_act:0.443

Batch: 106/714 loss: 0.408 loss\_act:0.408

Batch: 141/714 loss: 0.361 loss\_act:0.361

Batch: 176/714 loss: 0.433 loss\_act:0.433

Batch: 211/714 loss: 0.362 loss\_act:0.362

Batch: 246/714 loss: 0.370 loss\_act:0.370

Batch: 281/714 loss: 0.484 loss\_act:0.484

Batch: 316/714 loss: 0.726 loss\_act:0.726

Batch: 351/714 loss: 0.461 loss\_act:0.461

Batch: 386/714 loss: 0.612 loss\_act:0.612

Batch: 421/714 loss: 0.420 loss\_act:0.420

Batch: 456/714 loss: 0.535 loss\_act:0.535

Batch: 491/714 loss: 0.501 loss\_act:0.501

Batch: 526/714 loss: 0.530 loss\_act:0.530

Batch: 561/714 loss: 0.496 loss\_act:0.496

Batch: 596/714 loss: 0.442 loss\_act:0.442

Batch: 631/714 loss: 0.583 loss\_act:0.583

Batch: 666/714 loss: 0.433 loss\_act:0.433

Batch: 701/714 loss: 0.534 loss\_act:0.534

Batch: 714/714 loss: 0.695 loss\_act:0.695

Epoch 9 Train Loss: 0.504 Val Acc: 0.824 Test Acc: 0.816

Best Epoch: 8 Best Epoch Val Acc: 0.826 Best Epoch Test Acc: 0.819, Best Test Acc: 0.820

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 10\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.465 loss\_act:0.465

Batch: 36/714 loss: 0.755 loss\_act:0.755

Batch: 71/714 loss: 0.370 loss\_act:0.370

Batch: 106/714 loss: 0.505 loss\_act:0.505

Batch: 141/714 loss: 0.346 loss\_act:0.346

Batch: 176/714 loss: 0.499 loss\_act:0.499

Batch: 211/714 loss: 0.381 loss\_act:0.381

Batch: 246/714 loss: 0.646 loss\_act:0.646

Batch: 281/714 loss: 0.485 loss\_act:0.485

Batch: 316/714 loss: 0.373 loss\_act:0.373

Batch: 351/714 loss: 0.413 loss\_act:0.413

Batch: 386/714 loss: 0.337 loss\_act:0.337

Batch: 421/714 loss: 0.480 loss\_act:0.480

Batch: 456/714 loss: 0.456 loss\_act:0.456

Batch: 491/714 loss: 0.626 loss\_act:0.626

Batch: 526/714 loss: 0.461 loss\_act:0.461

Batch: 561/714 loss: 0.440 loss\_act:0.440

Batch: 596/714 loss: 0.383 loss\_act:0.383

Batch: 631/714 loss: 0.457 loss\_act:0.457

Batch: 666/714 loss: 0.360 loss\_act:0.360

Batch: 701/714 loss: 0.475 loss\_act:0.475

Batch: 714/714 loss: 0.345 loss\_act:0.345

Epoch 10 Train Loss: 0.491 Val Acc: 0.826 Test Acc: 0.812

Best Epoch: 8 Best Epoch Val Acc: 0.826 Best Epoch Test Acc: 0.819, Best Test Acc: 0.820

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 11\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.651 loss\_act:0.651

Batch: 36/714 loss: 0.413 loss\_act:0.413

Batch: 71/714 loss: 0.332 loss\_act:0.332

Batch: 106/714 loss: 0.440 loss\_act:0.440

Batch: 141/714 loss: 0.438 loss\_act:0.438

Batch: 176/714 loss: 0.702 loss\_act:0.702

Batch: 211/714 loss: 0.504 loss\_act:0.504

Batch: 246/714 loss: 0.550 loss\_act:0.550

Batch: 281/714 loss: 0.510 loss\_act:0.510

Batch: 316/714 loss: 0.574 loss\_act:0.574

Batch: 351/714 loss: 0.481 loss\_act:0.481

Batch: 386/714 loss: 0.352 loss\_act:0.352

Batch: 421/714 loss: 0.636 loss\_act:0.636

Batch: 456/714 loss: 0.426 loss\_act:0.426

Batch: 491/714 loss: 0.482 loss\_act:0.482

Batch: 526/714 loss: 0.539 loss\_act:0.539

Batch: 561/714 loss: 0.396 loss\_act:0.396

Batch: 596/714 loss: 0.570 loss\_act:0.570

Batch: 631/714 loss: 0.399 loss\_act:0.399

Batch: 666/714 loss: 0.522 loss\_act:0.522

Batch: 701/714 loss: 0.710 loss\_act:0.710

Batch: 714/714 loss: 0.412 loss\_act:0.412

Epoch 11 Train Loss: 0.475 Val Acc: 0.831 Test Acc: 0.818

Best Epoch: 11 Best Epoch Val Acc: 0.831 Best Epoch Test Acc: 0.818, Best Test Acc: 0.820

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 12\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.373 loss\_act:0.373

Batch: 36/714 loss: 0.467 loss\_act:0.467

Batch: 71/714 loss: 0.302 loss\_act:0.302

Batch: 106/714 loss: 0.448 loss\_act:0.448

Batch: 141/714 loss: 0.663 loss\_act:0.663

Batch: 176/714 loss: 0.379 loss\_act:0.379

Batch: 211/714 loss: 0.352 loss\_act:0.352

Batch: 246/714 loss: 0.394 loss\_act:0.394

Batch: 281/714 loss: 0.353 loss\_act:0.353

Batch: 316/714 loss: 0.472 loss\_act:0.472

Batch: 351/714 loss: 0.534 loss\_act:0.534

Batch: 386/714 loss: 0.535 loss\_act:0.535

Batch: 421/714 loss: 0.485 loss\_act:0.485

Batch: 456/714 loss: 0.622 loss\_act:0.622

Batch: 491/714 loss: 0.564 loss\_act:0.564

Batch: 526/714 loss: 0.543 loss\_act:0.543

Batch: 561/714 loss: 0.501 loss\_act:0.501

Batch: 596/714 loss: 0.517 loss\_act:0.517

Batch: 631/714 loss: 0.609 loss\_act:0.609

Batch: 666/714 loss: 0.313 loss\_act:0.313

Batch: 701/714 loss: 0.294 loss\_act:0.294

Batch: 714/714 loss: 0.762 loss\_act:0.762

Epoch 12 Train Loss: 0.462 Val Acc: 0.830 Test Acc: 0.821

Best Epoch: 11 Best Epoch Val Acc: 0.831 Best Epoch Test Acc: 0.818, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 13\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.532 loss\_act:0.532

Batch: 36/714 loss: 0.450 loss\_act:0.450

Batch: 71/714 loss: 0.454 loss\_act:0.454

Batch: 106/714 loss: 0.343 loss\_act:0.343

Batch: 141/714 loss: 0.347 loss\_act:0.347

Batch: 176/714 loss: 0.246 loss\_act:0.246

Batch: 211/714 loss: 0.355 loss\_act:0.355

Batch: 246/714 loss: 0.709 loss\_act:0.709

Batch: 281/714 loss: 0.560 loss\_act:0.560

Batch: 316/714 loss: 0.377 loss\_act:0.377

Batch: 351/714 loss: 0.449 loss\_act:0.449

Batch: 386/714 loss: 0.523 loss\_act:0.523

Batch: 421/714 loss: 0.405 loss\_act:0.405

Batch: 456/714 loss: 0.485 loss\_act:0.485

Batch: 491/714 loss: 0.336 loss\_act:0.336

Batch: 526/714 loss: 0.415 loss\_act:0.415

Batch: 561/714 loss: 0.549 loss\_act:0.549

Batch: 596/714 loss: 0.506 loss\_act:0.506

Batch: 631/714 loss: 0.574 loss\_act:0.574

Batch: 666/714 loss: 0.307 loss\_act:0.307

Batch: 701/714 loss: 0.498 loss\_act:0.498

Batch: 714/714 loss: 0.360 loss\_act:0.360

Epoch 13 Train Loss: 0.448 Val Acc: 0.832 Test Acc: 0.817

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 14\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.407 loss\_act:0.407

Batch: 36/714 loss: 0.279 loss\_act:0.279

Batch: 71/714 loss: 0.404 loss\_act:0.404

Batch: 106/714 loss: 0.412 loss\_act:0.412

Batch: 141/714 loss: 0.388 loss\_act:0.388

Batch: 176/714 loss: 0.489 loss\_act:0.489

Batch: 211/714 loss: 0.530 loss\_act:0.530

Batch: 246/714 loss: 0.484 loss\_act:0.484

Batch: 281/714 loss: 0.276 loss\_act:0.276

Batch: 316/714 loss: 0.410 loss\_act:0.410

Batch: 351/714 loss: 0.469 loss\_act:0.469

Batch: 386/714 loss: 0.452 loss\_act:0.452

Batch: 421/714 loss: 0.509 loss\_act:0.509

Batch: 456/714 loss: 0.635 loss\_act:0.635

Batch: 491/714 loss: 0.436 loss\_act:0.436

Batch: 526/714 loss: 0.456 loss\_act:0.456

Batch: 561/714 loss: 0.501 loss\_act:0.501

Batch: 596/714 loss: 0.459 loss\_act:0.459

Batch: 631/714 loss: 0.501 loss\_act:0.501

Batch: 666/714 loss: 0.546 loss\_act:0.546

Batch: 701/714 loss: 0.227 loss\_act:0.227

Batch: 714/714 loss: 0.373 loss\_act:0.373

Epoch 14 Train Loss: 0.436 Val Acc: 0.826 Test Acc: 0.817

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 15\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.401 loss\_act:0.401

Batch: 36/714 loss: 0.566 loss\_act:0.566

Batch: 71/714 loss: 0.364 loss\_act:0.364

Batch: 106/714 loss: 0.318 loss\_act:0.318

Batch: 141/714 loss: 0.293 loss\_act:0.293

Batch: 176/714 loss: 0.386 loss\_act:0.386

Batch: 211/714 loss: 0.323 loss\_act:0.323

Batch: 246/714 loss: 0.490 loss\_act:0.490

Batch: 281/714 loss: 0.456 loss\_act:0.456

Batch: 316/714 loss: 0.427 loss\_act:0.427

Batch: 351/714 loss: 0.442 loss\_act:0.442

Batch: 386/714 loss: 0.466 loss\_act:0.466

Batch: 421/714 loss: 0.314 loss\_act:0.314

Batch: 456/714 loss: 0.420 loss\_act:0.420

Batch: 491/714 loss: 0.429 loss\_act:0.429

Batch: 526/714 loss: 0.561 loss\_act:0.561

Batch: 561/714 loss: 0.377 loss\_act:0.377

Batch: 596/714 loss: 0.487 loss\_act:0.487

Batch: 631/714 loss: 0.369 loss\_act:0.369

Batch: 666/714 loss: 0.439 loss\_act:0.439

Batch: 701/714 loss: 0.353 loss\_act:0.353

Batch: 714/714 loss: 0.475 loss\_act:0.475

Epoch 15 Train Loss: 0.423 Val Acc: 0.828 Test Acc: 0.821

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 16\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.400 loss\_act:0.400

Batch: 36/714 loss: 0.401 loss\_act:0.401

Batch: 71/714 loss: 0.393 loss\_act:0.393

Batch: 106/714 loss: 0.275 loss\_act:0.275

Batch: 141/714 loss: 0.216 loss\_act:0.216

Batch: 176/714 loss: 0.307 loss\_act:0.307

Batch: 211/714 loss: 0.416 loss\_act:0.416

Batch: 246/714 loss: 0.578 loss\_act:0.578

Batch: 281/714 loss: 0.387 loss\_act:0.387

Batch: 316/714 loss: 0.440 loss\_act:0.440

Batch: 351/714 loss: 0.501 loss\_act:0.501

Batch: 386/714 loss: 0.363 loss\_act:0.363

Batch: 421/714 loss: 0.403 loss\_act:0.403

Batch: 456/714 loss: 0.354 loss\_act:0.354

Batch: 491/714 loss: 0.568 loss\_act:0.568

Batch: 526/714 loss: 0.358 loss\_act:0.358

Batch: 561/714 loss: 0.275 loss\_act:0.275

Batch: 596/714 loss: 0.302 loss\_act:0.302

Batch: 631/714 loss: 0.579 loss\_act:0.579

Batch: 666/714 loss: 0.623 loss\_act:0.623

Batch: 701/714 loss: 0.591 loss\_act:0.591

Batch: 714/714 loss: 0.388 loss\_act:0.388

Epoch 16 Train Loss: 0.409 Val Acc: 0.831 Test Acc: 0.812

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 17\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.386 loss\_act:0.386

Batch: 36/714 loss: 0.274 loss\_act:0.274

Batch: 71/714 loss: 0.324 loss\_act:0.324

Batch: 106/714 loss: 0.406 loss\_act:0.406

Batch: 141/714 loss: 0.643 loss\_act:0.643

Batch: 176/714 loss: 0.252 loss\_act:0.252

Batch: 211/714 loss: 0.257 loss\_act:0.257

Batch: 246/714 loss: 0.272 loss\_act:0.272

Batch: 281/714 loss: 0.401 loss\_act:0.401

Batch: 316/714 loss: 0.364 loss\_act:0.364

Batch: 351/714 loss: 0.213 loss\_act:0.213

Batch: 386/714 loss: 0.361 loss\_act:0.361

Batch: 421/714 loss: 0.272 loss\_act:0.272

Batch: 456/714 loss: 0.337 loss\_act:0.337

Batch: 491/714 loss: 0.528 loss\_act:0.528

Batch: 526/714 loss: 0.386 loss\_act:0.386

Batch: 561/714 loss: 0.395 loss\_act:0.395

Batch: 596/714 loss: 0.491 loss\_act:0.491

Batch: 631/714 loss: 0.334 loss\_act:0.334

Batch: 666/714 loss: 0.238 loss\_act:0.238

Batch: 701/714 loss: 0.475 loss\_act:0.475

Batch: 714/714 loss: 0.456 loss\_act:0.456

Epoch 17 Train Loss: 0.396 Val Acc: 0.830 Test Acc: 0.821

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 18\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.329 loss\_act:0.329

Batch: 36/714 loss: 0.361 loss\_act:0.361

Batch: 71/714 loss: 0.346 loss\_act:0.346

Batch: 106/714 loss: 0.412 loss\_act:0.412

Batch: 141/714 loss: 0.400 loss\_act:0.400

Batch: 176/714 loss: 0.311 loss\_act:0.311

Batch: 211/714 loss: 0.412 loss\_act:0.412

Batch: 246/714 loss: 0.380 loss\_act:0.380

Batch: 281/714 loss: 0.329 loss\_act:0.329

Batch: 316/714 loss: 0.300 loss\_act:0.300

Batch: 351/714 loss: 0.188 loss\_act:0.188

Batch: 386/714 loss: 0.252 loss\_act:0.252

Batch: 421/714 loss: 0.355 loss\_act:0.355

Batch: 456/714 loss: 0.375 loss\_act:0.375

Batch: 491/714 loss: 0.485 loss\_act:0.485

Batch: 526/714 loss: 0.348 loss\_act:0.348

Batch: 561/714 loss: 0.506 loss\_act:0.506

Batch: 596/714 loss: 0.364 loss\_act:0.364

Batch: 631/714 loss: 0.325 loss\_act:0.325

Batch: 666/714 loss: 0.314 loss\_act:0.314

Batch: 701/714 loss: 0.510 loss\_act:0.510

Batch: 714/714 loss: 0.506 loss\_act:0.506

Epoch 18 Train Loss: 0.384 Val Acc: 0.826 Test Acc: 0.821

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 19\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.269 loss\_act:0.269

Batch: 36/714 loss: 0.209 loss\_act:0.209

Batch: 71/714 loss: 0.180 loss\_act:0.180

Batch: 106/714 loss: 0.375 loss\_act:0.375

Batch: 141/714 loss: 0.415 loss\_act:0.415

Batch: 176/714 loss: 0.294 loss\_act:0.294

Batch: 211/714 loss: 0.284 loss\_act:0.284

Batch: 246/714 loss: 0.283 loss\_act:0.283

Batch: 281/714 loss: 0.443 loss\_act:0.443

Batch: 316/714 loss: 0.432 loss\_act:0.432

Batch: 351/714 loss: 0.257 loss\_act:0.257

Batch: 386/714 loss: 0.322 loss\_act:0.322

Batch: 421/714 loss: 0.302 loss\_act:0.302

Batch: 456/714 loss: 0.432 loss\_act:0.432

Batch: 491/714 loss: 0.463 loss\_act:0.463

Batch: 526/714 loss: 0.373 loss\_act:0.373

Batch: 561/714 loss: 0.301 loss\_act:0.301

Batch: 596/714 loss: 0.400 loss\_act:0.400

Batch: 631/714 loss: 0.389 loss\_act:0.389

Batch: 666/714 loss: 0.337 loss\_act:0.337

Batch: 701/714 loss: 0.432 loss\_act:0.432

Batch: 714/714 loss: 0.541 loss\_act:0.541

Epoch 19 Train Loss: 0.371 Val Acc: 0.828 Test Acc: 0.817

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 20\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.267 loss\_act:0.267

Batch: 36/714 loss: 0.292 loss\_act:0.292

Batch: 71/714 loss: 0.388 loss\_act:0.388

Batch: 106/714 loss: 0.339 loss\_act:0.339

Batch: 141/714 loss: 0.371 loss\_act:0.371

Batch: 176/714 loss: 0.499 loss\_act:0.499

Batch: 211/714 loss: 0.251 loss\_act:0.251

Batch: 246/714 loss: 0.337 loss\_act:0.337

Batch: 281/714 loss: 0.301 loss\_act:0.301

Batch: 316/714 loss: 0.317 loss\_act:0.317

Batch: 351/714 loss: 0.361 loss\_act:0.361

Batch: 386/714 loss: 0.362 loss\_act:0.362

Batch: 421/714 loss: 0.235 loss\_act:0.235

Batch: 456/714 loss: 0.479 loss\_act:0.479

Batch: 491/714 loss: 0.350 loss\_act:0.350

Batch: 526/714 loss: 0.233 loss\_act:0.233

Batch: 561/714 loss: 0.455 loss\_act:0.455

Batch: 596/714 loss: 0.346 loss\_act:0.346

Batch: 631/714 loss: 0.395 loss\_act:0.395

Batch: 666/714 loss: 0.262 loss\_act:0.262

Batch: 701/714 loss: 0.260 loss\_act:0.260

Batch: 714/714 loss: 0.259 loss\_act:0.259

Epoch 20 Train Loss: 0.357 Val Acc: 0.828 Test Acc: 0.814

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 21\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.326 loss\_act:0.326

Batch: 36/714 loss: 0.277 loss\_act:0.277

Batch: 71/714 loss: 0.320 loss\_act:0.320

Batch: 106/714 loss: 0.440 loss\_act:0.440

Batch: 141/714 loss: 0.519 loss\_act:0.519

Batch: 176/714 loss: 0.263 loss\_act:0.263

Batch: 211/714 loss: 0.340 loss\_act:0.340

Batch: 246/714 loss: 0.326 loss\_act:0.326

Batch: 281/714 loss: 0.380 loss\_act:0.380

Batch: 316/714 loss: 0.420 loss\_act:0.420

Batch: 351/714 loss: 0.330 loss\_act:0.330

Batch: 386/714 loss: 0.407 loss\_act:0.407

Batch: 421/714 loss: 0.437 loss\_act:0.437

Batch: 456/714 loss: 0.327 loss\_act:0.327

Batch: 491/714 loss: 0.361 loss\_act:0.361

Batch: 526/714 loss: 0.301 loss\_act:0.301

Batch: 561/714 loss: 0.483 loss\_act:0.483

Batch: 596/714 loss: 0.245 loss\_act:0.245

Batch: 631/714 loss: 0.250 loss\_act:0.250

Batch: 666/714 loss: 0.301 loss\_act:0.301

Batch: 701/714 loss: 0.198 loss\_act:0.198

Batch: 714/714 loss: 0.291 loss\_act:0.291

Epoch 21 Train Loss: 0.345 Val Acc: 0.825 Test Acc: 0.818

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 22\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.320 loss\_act:0.320

Batch: 36/714 loss: 0.381 loss\_act:0.381

Batch: 71/714 loss: 0.252 loss\_act:0.252

Batch: 106/714 loss: 0.357 loss\_act:0.357

Batch: 141/714 loss: 0.298 loss\_act:0.298

Batch: 176/714 loss: 0.445 loss\_act:0.445

Batch: 211/714 loss: 0.178 loss\_act:0.178

Batch: 246/714 loss: 0.188 loss\_act:0.188

Batch: 281/714 loss: 0.327 loss\_act:0.327

Batch: 316/714 loss: 0.271 loss\_act:0.271

Batch: 351/714 loss: 0.386 loss\_act:0.386

Batch: 386/714 loss: 0.171 loss\_act:0.171

Batch: 421/714 loss: 0.336 loss\_act:0.336

Batch: 456/714 loss: 0.390 loss\_act:0.390

Batch: 491/714 loss: 0.301 loss\_act:0.301

Batch: 526/714 loss: 0.303 loss\_act:0.303

Batch: 561/714 loss: 0.376 loss\_act:0.376

Batch: 596/714 loss: 0.382 loss\_act:0.382

Batch: 631/714 loss: 0.245 loss\_act:0.245

Batch: 666/714 loss: 0.309 loss\_act:0.309

Batch: 701/714 loss: 0.284 loss\_act:0.284

Batch: 714/714 loss: 0.406 loss\_act:0.406

Epoch 22 Train Loss: 0.330 Val Acc: 0.827 Test Acc: 0.810

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*Epoch: 23\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Batch: 1/714 loss: 0.366 loss\_act:0.366

Batch: 36/714 loss: 0.233 loss\_act:0.233

Batch: 71/714 loss: 0.262 loss\_act:0.262

Batch: 106/714 loss: 0.321 loss\_act:0.321

Batch: 141/714 loss: 0.352 loss\_act:0.352

Batch: 176/714 loss: 0.281 loss\_act:0.281

Batch: 211/714 loss: 0.215 loss\_act:0.215

Batch: 246/714 loss: 0.337 loss\_act:0.337

Batch: 281/714 loss: 0.275 loss\_act:0.275

Batch: 316/714 loss: 0.213 loss\_act:0.213

Batch: 351/714 loss: 0.227 loss\_act:0.227

Batch: 386/714 loss: 0.299 loss\_act:0.299

Batch: 421/714 loss: 0.367 loss\_act:0.367

Batch: 456/714 loss: 0.415 loss\_act:0.415

Batch: 491/714 loss: 0.352 loss\_act:0.352

Batch: 526/714 loss: 0.438 loss\_act:0.438

Batch: 561/714 loss: 0.290 loss\_act:0.290

Batch: 596/714 loss: 0.343 loss\_act:0.343

Batch: 631/714 loss: 0.367 loss\_act:0.367

Batch: 666/714 loss: 0.345 loss\_act:0.345

Batch: 701/714 loss: 0.308 loss\_act:0.308

Batch: 714/714 loss: 0.428 loss\_act:0.428

Epoch 23 Train Loss: 0.317 Val Acc: 0.825 Test Acc: 0.811

Best Epoch: 13 Best Epoch Val Acc: 0.832 Best Epoch Test Acc: 0.817, Best Test Acc: 0.821

Saving the best checkpoint....

Test Acc: 0.817