Run build_manifest_and_split.py

```
python build_manifest_and_split.py \
  --dataset_dir "$(pwd)" \
  --bias_train_to_many_images \
  --train_ratio 0.6 \
  --val_ratio 0.20 \
  --test_ratio 0.20 \
  --seed 123
```

build_roi_manifest.py

```
python build_roi_manifest.py --dataset_dir "$(pwd)"
```

dataset_multimodal.py

```
python dataset_multimodal.py --dataset_dir "$(pwd)"

python3 dataset_multimodal.py \
  --dataset_dir "$(pwd)" \
  --batch_size 32 \
  --num_workers 8 \
  --img_size 224
```

model_multimodal.py

```
python model_multimodal.py
```

RUN 1

Prompt convnext:

```
(base) bioinfo@pop-os:~/Downloads/Dataset$ python train_multimodal.py   --dataset_dir
"$(pwd)"   --epochs 20   --batch_size 16   --lr 1e-5   --weight_decay 1e-3   --device cuda
--freeze_backbone_epochs 10   --backbone_unfreeze_mode partial
--backbone_unfreeze_fraction 0.3   --backbone convnext_tiny   --dropout_p 0.5
--class_weights "2.11,4.60,4.98,2.0"
[INFO] Usando device: cuda
[INFO] Usando backbone: convnext_tiny
[INFO] Usando dropout_p: 0.5
[INFO] Épocas de treino: 20
[INFO] Batch size: 16
[INFO] Learning rate: 1e-05
[INFO] Weight decay: 0.001 (1e-03)
[INFO] Loss type: ce
[INFO] Focal gamma: 2.0
[INFO] freeze_backbone_epochs: 10
```

[INFO] backbone_unfreeze_mode: partial
[INFO] backbone_unfreeze_fraction: 0.3
[INFO] Pesos manuais (string): 2.11,4.60,4.98,2.0
[INFO] Criando DataLoaders...
[INFO] Usando PESOS MANUAIS para a loss:
  classe 0: peso = 2.1100
  classe 1: peso = 4.6000
  classe 2: peso = 4.9800
  classe 3: peso = 2.0000
[INFO] Congelando backbone nas primeiras 10 épocas.
[INFO] Usando CrossEntropyLoss ponderada.

Output:

(base) bioinfo@pop-os:~/Downloads/Dataset$ python train_multimodal.py --dataset_dir "$(pwd)" --epochs 20 --batch_size 16 --lr 1e-5 --weight_decay 1e-3 --device cuda --freeze_backbone_epochs 10 --backbone_unfreeze_mode partial --backbone_unfreeze_fraction 0.3 --backbone convnext_tiny --dropout_p 0.5 --class_weights "2.11,4.60,4.98,2.0"
[INFO] Usando device: cuda
[INFO] Usando backbone: convnext_tiny
[INFO] Usando dropout_p: 0.5
[INFO] Épocas de treino: 20
[INFO] Batch size: 16
[INFO] Learning rate: 1e-05
[INFO] Weight decay: 0.001
[INFO] Loss type: ce
[INFO] Focal gamma: 2.0
[INFO] freeze_backbone_epochs: 10
[INFO] backbone_unfreeze_mode: partial
[INFO] backbone_unfreeze_fraction: 0.3
[INFO] Pesos manuais (string): 2.11,4.60,4.98,2.0
[INFO] Criando DataLoaders...
[INFO] Usando PESOS MANUAIS para a loss:
  classe 0: peso = 2.1100
  classe 1: peso = 4.6000
  classe 2: peso = 4.9800
  classe 3: peso = 2.0000
[INFO] Congelando backbone nas primeiras 10 épocas.
[INFO] Usando CrossEntropyLoss ponderada.

========== Época 1/20 ==========
[TRAIN] loss=1.4623 acc=0.3109 macro_f1=0.3044 mcc=0.0859
[VAL] loss=1.3082 acc=0.4333 macro_f1=0.4187 mcc=0.3021
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.4187)

========== Época 2/20 ==========
[TRAIN] loss=1.3207 acc=0.4175 macro_f1=0.4126 mcc=0.2270
[VAL]   loss=1.2008 acc=0.6000 macro_f1=0.5373 mcc=0.4525
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.5373)

========== Época 3/20 ==========
[TRAIN] loss=1.2200 acc=0.4886 macro_f1=0.4866 mcc=0.3190
[VAL]   loss=1.1058 acc=0.6833 macro_f1=0.5868 mcc=0.5440
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.5868)

========== Época 4/20 ==========
[TRAIN] loss=1.1436 acc=0.5436 macro_f1=0.5440 mcc=0.3938
[VAL]   loss=1.0352 acc=0.7000 macro_f1=0.6033 mcc=0.5657
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.6033)

========== Época 5/20 ==========
[TRAIN] loss=1.0607 acc=0.6097 macro_f1=0.6085 mcc=0.4788
[VAL]   loss=0.9919 acc=0.7183 macro_f1=0.6219 mcc=0.5926
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.6219)

========== Época 6/20 ==========
[TRAIN] loss=1.0206 acc=0.6252 macro_f1=0.6178 mcc=0.5056
[VAL]   loss=0.9278 acc=0.7167 macro_f1=0.6188 mcc=0.5874

========== Época 7/20 ==========
[TRAIN] loss=0.9322 acc=0.6641 macro_f1=0.6542 mcc=0.5553
[VAL]   loss=0.8896 acc=0.7383 macro_f1=0.6274 mcc=0.6211
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.6274)

========== Época 8/20 ==========
[TRAIN] loss=0.9120 acc=0.7029 macro_f1=0.7073 mcc=0.6043
[VAL]   loss=0.8907 acc=0.7417 macro_f1=0.6704 mcc=0.6309
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.6704)

========== Época 9/20 ==========
[TRAIN] loss=0.8862 acc=0.7024 macro_f1=0.6994 mcc=0.6064
[VAL]   loss=0.8380 acc=0.7500 macro_f1=0.6585 mcc=0.6388

========== Época 10/20 ==========
[TRAIN] loss=0.8526 acc=0.7268 macro_f1=0.7239 mcc=0.6357
[VAL]   loss=0.8102 acc=0.7450 macro_f1=0.6392 mcc=0.6307

========== Época 11/20 ==========
[INFO] Descongelando PARCIALMENTE o backbone (fração = 0.30).
[TRAIN] loss=0.7271 acc=0.7746 macro_f1=0.7780 mcc=0.7012
[VAL]   loss=0.6654 acc=0.7867 macro_f1=0.7027 mcc=0.6933
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.7027)

========== Época 12/20 ==========
[TRAIN] loss=0.6307 acc=0.8273 macro_f1=0.8275 mcc=0.7704
[VAL]   loss=0.6054 acc=0.7933 macro_f1=0.7021 mcc=0.6998

========== Época 13/20 ==========
[TRAIN] loss=0.5266 acc=0.8828 macro_f1=0.8870 mcc=0.8438
[VAL]   loss=0.5684 acc=0.8150 macro_f1=0.7110 mcc=0.7295
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.7110)

========== Época 14/20 ==========
[TRAIN] loss=0.4668 acc=0.9039 macro_f1=0.9058 mcc=0.8720
[VAL]   loss=0.5596 acc=0.8050 macro_f1=0.7368 mcc=0.7161
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.7368)

========== Época 15/20 ==========
[TRAIN] loss=0.4210 acc=0.9178 macro_f1=0.9185 mcc=0.8905
[VAL]   loss=0.5275 acc=0.8117 macro_f1=0.7311 mcc=0.7250

========== Época 16/20 ==========
[TRAIN] loss=0.3774 acc=0.9339 macro_f1=0.9355 mcc=0.9119
[VAL]   loss=0.5471 acc=0.7950 macro_f1=0.7087 mcc=0.7087

========== Época 17/20 ==========
[TRAIN] loss=0.3215 acc=0.9489 macro_f1=0.9493 mcc=0.9319
[VAL]   loss=0.5775 acc=0.7933 macro_f1=0.6449 mcc=0.7032

========== Época 18/20 ==========
[TRAIN] loss=0.2990 acc=0.9500 macro_f1=0.9494 mcc=0.9334
[VAL]   loss=0.6386 acc=0.7667 macro_f1=0.6606 mcc=0.6737

========== Época 19/20 ==========
[TRAIN] loss=0.2455 acc=0.9706 macro_f1=0.9709 mcc=0.9608
[VAL]   loss=0.5863 acc=0.7917 macro_f1=0.6780 mcc=0.7075

========== Época 20/20 ==========
[TRAIN] loss=0.2466 acc=0.9706 macro_f1=0.9704 mcc=0.9608
[VAL]   loss=0.6652 acc=0.7767 macro_f1=0.6447 mcc=0.6965

[OK] Histórico de treino salvo em: /home/bioinfo/Downloads/Dataset/training_history.csv

[INFO] Carregando melhor modelo para avaliação no TEST...

========== RESULTADOS FINAIS (TEST) ==========
[TEST] acc=0.8074 macro_f1=0.7402 mcc=0.7146

Matriz de confusão (linhas = verdade, colunas = predito):
[[140   0   0   0]
 [0 119 53   4]
 [0 33 211   5]
 [0 16   4 12]]

RUN 2

Prompt

```
(base) bioinfo@pop-os:~/Downloads/Dataset$ python train_multimodal.py \
  --dataset_dir "$(pwd)" \
  --epochs 20 \
  --batch_size 16 \
  --device cuda \
  --backbone convnext_tiny \
  --dropout_p 0.4 \
  \
  --backbone_unfreeze_mode partial \
  --backbone_unfreeze_fraction 0.7 \
  \
  --freeze_epoch_start 11 \
  --freeze_epoch_end 20 \
  \
  --phase_boundary_epoch 10 \
  --lr_phase1 1e-5 \
  --lr_phase2 1e-4 \
  --weight_decay 1e-4 \
  \
  --loss_type ce  --class_weights "2.0,6.0,6.0,4.0"
[INFO] Usando device: cuda
[INFO] Usando backbone: convnext_tiny
[INFO] Usando dropout_p: 0.4
[INFO] Épocas de treino: 20
[INFO] Batch size: 16
[INFO] LR base (--lr): 0.0001
[INFO] Fases de LR: Fase 1 (<= epoch 10) lr=1e-05, Fase 2 (> epoch 10) lr=0.0001
[INFO] Weight decay: 0.0001
[INFO] Loss type: ce
[INFO] Focal gamma: 2.0
```

[INFO] freeze_backbone_epochs (modo antigo): 0
[INFO] freeze_epoch_start=11, freeze_epoch_end=20 (janela explícita de freeze; 0=desativado)
[INFO] backbone_unfreeze_mode=partial, backbone_unfreeze_fraction=0.7
[INFO] Pesos manuais (string): 2.0,6.0,6.0,4.0
[INFO] Criando DataLoaders...
[INFO] Usando PESOS MANUAIS para a loss:
  classe 0: peso = 2.0000
  classe 1: peso = 6.0000
  classe 2: peso = 6.0000
  classe 3: peso = 4.0000
[INFO] Usando CrossEntropyLoss ponderada.

Output

========== Época 1/20 ==========
[INFO] Ajustando LR para 1.00e-05 (época 1)
[INFO] Estado do backbone na época 1: unfrozen_last_fraction=0.70
[TRAIN] loss=1.1070 acc=0.5986 macro_f1=0.5949 mcc=0.4714
[VAL]   loss=0.8588 acc=0.7500 macro_f1=0.6832 mcc=0.6372
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.6832)

========== Época 2/20 ==========
[TRAIN] loss=0.7212 acc=0.8223 macro_f1=0.8172 mcc=0.7637
[VAL]   loss=0.7352 acc=0.7667 macro_f1=0.7232 mcc=0.6652
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.7232)

========== Época 3/20 ==========
[TRAIN] loss=0.5480 acc=0.8862 macro_f1=0.8871 mcc=0.8485
[VAL]   loss=0.6346 acc=0.7983 macro_f1=0.7403 mcc=0.7072
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.7403)

========== Época 4/20 ==========
[TRAIN] loss=0.4241 acc=0.9311 macro_f1=0.9318 mcc=0.9082
[VAL]   loss=0.6566 acc=0.7733 macro_f1=0.7058 mcc=0.6895

========== Época 5/20 ==========
[TRAIN] loss=0.3563 acc=0.9439 macro_f1=0.9450 mcc=0.9253
[VAL]   loss=0.5853 acc=0.7983 macro_f1=0.7200 mcc=0.7065

========== Época 6/20 ==========

[TRAIN] loss=0.3001 acc=0.9606 macro_f1=0.9591 mcc=0.9474
[VAL]   loss=0.5778 acc=0.8150 macro_f1=0.7138 mcc=0.7384

========== Época 7/20 ==========
[TRAIN] loss=0.2358 acc=0.9750 macro_f1=0.9740 mcc=0.9667
[VAL]   loss=0.5338 acc=0.8267 macro_f1=0.7527 mcc=0.7488
[INFO] Novo melhor modelo salvo em /home/bioinfo/Downloads/Dataset/best_model.pt (val_macro_f1=0.7527)

========== Época 8/20 ==========
[TRAIN] loss=0.2202 acc=0.9761 macro_f1=0.9770 mcc=0.9682
[VAL]   loss=0.6294 acc=0.7883 macro_f1=0.6753 mcc=0.7052

========== Época 9/20 ==========
[TRAIN] loss=0.1794 acc=0.9867 macro_f1=0.9864 mcc=0.9823
[VAL]   loss=0.5692 acc=0.8050 macro_f1=0.7161 mcc=0.7171

========== Época 10/20 ==========
[TRAIN] loss=0.1604 acc=0.9872 macro_f1=0.9872 mcc=0.9830
[VAL]   loss=0.5897 acc=0.8183 macro_f1=0.7117 mcc=0.7413

========== Época 11/20 ==========
[INFO] Ajustando LR para 1.00e-04 (época 11)
[INFO] Estado do backbone na época 11: frozen (explicit_window)
[TRAIN] loss=0.1149 acc=0.9922 macro_f1=0.9922 mcc=0.9896
[VAL]   loss=0.5752 acc=0.8317 macro_f1=0.7243 mcc=0.7557

========== Época 12/20 ==========
[TRAIN] loss=0.0968 acc=0.9895 macro_f1=0.9893 mcc=0.9859
[VAL]   loss=0.5993 acc=0.8317 macro_f1=0.7239 mcc=0.7564

========== Época 13/20 ==========
[TRAIN] loss=0.0787 acc=0.9889 macro_f1=0.9891 mcc=0.9852
[VAL]   loss=0.5647 acc=0.8367 macro_f1=0.7282 mcc=0.7624

========== Época 14/20 ==========
[TRAIN] loss=0.0637 acc=0.9895 macro_f1=0.9897 mcc=0.9859
[VAL]   loss=0.5971 acc=0.8333 macro_f1=0.7248 mcc=0.7570

========== Época 15/20 ==========
[TRAIN] loss=0.0569 acc=0.9939 macro_f1=0.9939 mcc=0.9919
[VAL]   loss=0.5863 acc=0.8367 macro_f1=0.7283 mcc=0.7620

========== Época 16/20 ==========
[TRAIN] loss=0.0586 acc=0.9922 macro_f1=0.9924 mcc=0.9896
[VAL]   loss=0.6720 acc=0.8267 macro_f1=0.7018 mcc=0.7495

========== Época 17/20 ==========

```
[TRAIN] loss=0.0442  acc=0.9939  macro_f1=0.9939  mcc=0.9919
[VAL]   loss=0.6948  acc=0.8267  macro_f1=0.7107  mcc=0.7507

========== Época 18/20 ==========
[TRAIN] loss=0.0470  acc=0.9917  macro_f1=0.9915  mcc=0.9889
[VAL]   loss=0.6408  acc=0.8333  macro_f1=0.7169  mcc=0.7580

========== Época 19/20 ==========
[TRAIN] loss=0.0323  acc=0.9956  macro_f1=0.9956  mcc=0.9941
[VAL]   loss=0.6456  acc=0.8300  macro_f1=0.7230  mcc=0.7534

========== Época 20/20 ==========
[TRAIN] loss=0.0298  acc=0.9967  macro_f1=0.9966  mcc=0.9956
[VAL]   loss=0.7394  acc=0.8233  macro_f1=0.6985  mcc=0.7463

[OK] Histórico de treino salvo em: /home/bioinfo/Downloads/Dataset/training_history.csv

[INFO] Carregando melhor modelo para avaliação no TEST...

========== RESULTADOS FINAIS (TEST) ==========
[TEST] acc=0.8157  macro_f1=0.7695  mcc=0.7273

Matriz de confusão (linhas = verdade, colunas = predito):
[[140   0   0   0]
 [  0 130  45   1]
 [  0  44 204   1]
 [  0  16   3  13]]
```

plot_training_curves.py

```
python plot_training_curves.py \
  --history_path "/home/bioinfo/Downloads/Dataset/training_history.csv” \
  --dataset_dir "$(pwd)” \
  --out_dir "./plots" \
  --backbone "convnext_tiny" \
  --img_size 224 \
  --dropout_p 0.4
```