

Technical Summary

The project focuses on image-based classification for two distinct tasks – Binary gender classification and Multi-class facial identity recognition.

It leverages EfficientNet family backbones (efficientnet_v2_s and efficientnet_b0) pretrained on ImageNet, fine-tuned on custom datasets using PyTorch.

Key Innovations

- SWA + Focal Loss + TTA: includes a robust combination to counteract class imbalance and improve generalization.
- Mixed Precision Training across both tasks.
- Dynamic label extraction and distortion-based data enrichment.
- Flexible K-Fold Cross-Validation for reproducibility.

Model Architecture

- Task A
 - EfficientNetV2-S backbone:
Replaces final classifier with:
Linear(in_features → 512) → ReLU → Dropout(0.5) → Linear(512 → 2)
 - Dynamic face detection & cropping using facenet_pytorch's MTCNN.
 - Albumenations augmentations.
 - Balanced sampling using WeightedRandomSampler to address class imbalance.
 - Stratified K-Fold cross-validation for robust performance estimation.
 - Focal Loss to address class imbalance.
 - Stochastic Weight Averaging (SWA) after epoch 5 to improve generalization.
 - Automatic Mixed Precision (AMP) (torch.cuda.amp) for faster training.
 - Test-Time Augmentation (TTA) by horizontal flipping during validation.
- Task B
 - EfficientNet-B0 backbone:
Final classifier adapted dynamically to the number of identities.
 - Dataset already containing Distorted images for robustness.
 - Resize to 224x224.
 - Normalization.
 - ReduceLROnPlateau scheduler.
 - AMP for efficiency.
 - Early stopping after 5 epochs without improvement.
 - Training combines original and distorted images.