Cascade Fusion Segmentation Model

TTG Input
B×176×1024×1024

FeatureLevel 0 — 512×512×64
FeatureLevel 1 — 256×256×128
FeatureLevel 2 — 128×128×250
FeatureLevel 3 — 64×64×512

DCT Input
B×292×128×128

Resize →128×128×292

Key Innovation: Adaptive Feature Fusion
Leverages SMP Pre-trained Models
Configurable: ttg_backbone (TIMM models), Unet_model (SMP variants), in_channels (TTG: 176, DCT: 292), _loss (BCE/Dice weighting)

Adaptive Fusion Strategy

FusionLayerDetection
_detect_fusion_layer()
Auto-detect 128×128

FeatureConcatenation
(Channel Expansion)
= 128×128×548

Spatial Matching: 128×128

SMP Encoder(ResNet) — Multi-scale
Bottleneck — Deepest features
SMP Decoder(U-N...) — Multi-scale
SegmentationOutput — B×K×1024×1024

U-Net Skip Connections

BCE Loss — weight
Dice Loss
CombinedLoss — ·BCE + (1-)·Dice

SegmentationMetrics
IoU, Dice, per-class

Legend

Input/Output	Fusion (orange arrow)
Feature Extractor	Resize (red dashed arrow)
Fusion Module	Forward (black arrow)
Processing	