Improved Weakly Supervised Semantic Segmentation with a Relationally Optimized Prototype Memory Bank Framework

Xiang Zhong, Qiudan Zhang, Jianmin Jiang

深圳大学 SHENZHEN UNIVERSITY

NEURAL INFORMATION PROCESSING SYSTEMS

Background:

Weakly Supervised Semantic Segmentation (WSSS):

Learn pixel-level segmentation from image-level labels. Class Activation Maps (CAMs) provide localization cues for pseudo-label generation.

The prototype bottleneck:

→ Single prototypes miss diverse object appearances.

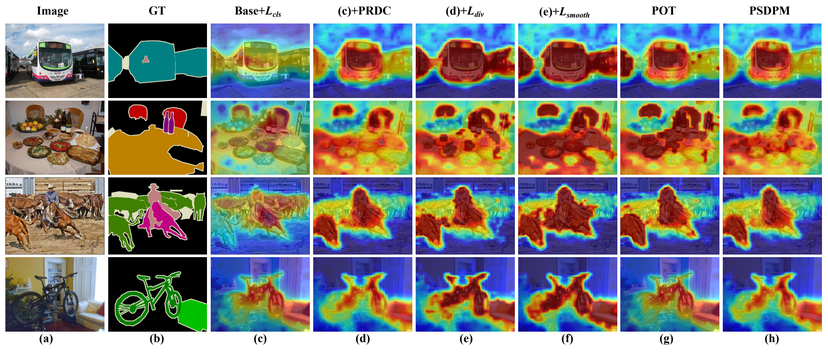
→ Independent prototypes overlook supportive semantic relations across classes.

→ Incomplete CAMs propagate errors to pseudo-labels.

Our Key Idea:

Maintain multiple global prototypes and connect them through a positive relation graph. Refine local prototypes using context-aware global knowledge.

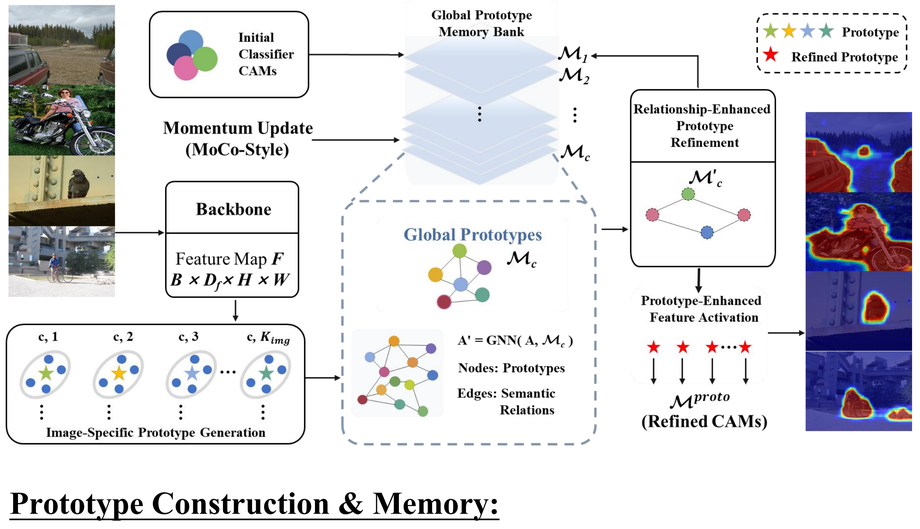
Qualitative CAM Analysis::



From sparse activation to full objects:

Relational propagation improves activation completeness. Diversity and smoothness further support coherent localization.

Method:



Prototype Construction & Memory:

K-means extracts image-specific prototypes. Per-class global memory captures dataset-level variation and evolves through momentum updates.

Relational Propagation (PRDC):

Visual similarity x label co-occurrence defines positive graph edges. A single-layer GCN enriches global prototypes with semantic context.

Context-Aware Refinement:

Similarity-weighted global prototypes refine local representations for CAM generation. Training combines classification, local-global consistency, diversity and graph smoothness.

Segmentation Examples:

Experiments:

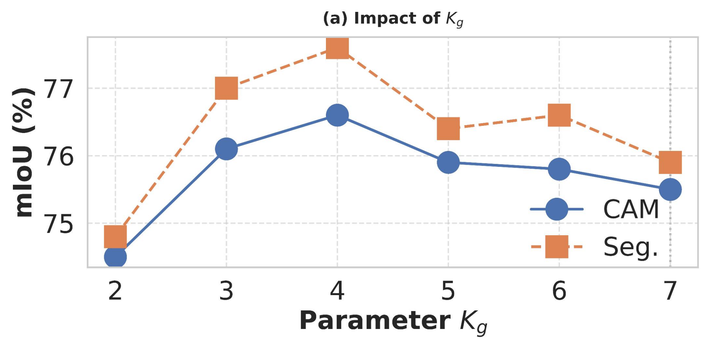
Final Semantic Segmentation:

Methods	VOC test	VOC val	COCO val
PCRE	75.5	75.9	47.2
POT	76.1	76.7	47.9
Our	**77.6**	**77.8**	**48.8**

CAM & Pseudo-Mask Quality:

Methods	CAMs	Mask
MuP-VSS	71.7	74.1
POT	75.0	79.3
Our	**76.6**	**80.4**

Ablation of Global Prototypes per Class K_g :



Ablation of the Image-Specific Prototypes K_{img} :

