Hybrid Super-Voxel Segmentation: Graph Cuts + Fuzzy C-Means

Abstract

We propose a hybrid super-voxel segmentation pipeline that combines soft memberships from fuzzy c-means (FCM) with spatial regularization via graph cuts on a region adjacency graph (RAG). Our hybrid achieves **0.70 mean IoU at 42 fps** (synthetic), outperforming SLIC-only (0.55) and FCM-only (0.52) under the same budget. The result is spatially coherent clusters that respect object boundaries while preserving soft assignment information at real-time performance.

1 Introduction

Super-voxel segmentation algorithms balance computational efficiency with boundary adherence. Hard assignments like SLIC [**?**] are fast but can fragment fine structures; soft assignments like FCM [**?**] are flexible but spatially noisy. Graph cuts [**?**, **?**] and normalized cuts [**?**] provide spatial regularization but typically require careful initialization.

We combine soft memberships from FCM [**?**] with graph-cut regularization on a super-voxel RAG [**?**, **?**], detailed in Sec. 2. This hybrid approach improves spatial coherence without sacrificing throughput, achieving superior IoU performance at real-time budgets.

2 Method

FCM: memberships $\mathbf{U} \in \mathbb{R}^{K\times N}$ and centers $\mathbf{C} \in \mathbb{R}^{K\times D}$ via fuzzy c-means. **RAG + cuts:** build a RAG over SLIC super-voxels, use FCM memberships as unaries and contrast-based Potts pairwise terms, then normalized cut.

2.1 Hybrid Objective and Construction Details

Fuzzy C-Means (FCM). Given features $\{x_i\}_{i=1}^N$ and K clusters with centers $\{c_k\}_{k=1}^K$, we minimize

$$J_{\mathrm{FCM}}(U,C) = \sum_{i=1}^{N}\sum_{k=1}^{K} u_{ik}^m \,\|x_i - c_k\|_2^2 \quad \text{s.t.} \quad \sum_{k=1}^{K} u_{ik} = 1,\ u_{ik} \in [0,1],$$

with fuzzifier m=2.0. We use k-means++ initialization of $\{c_k\}$ and iterate

$$c_k \leftarrow \frac{\sum_i u_{ik}^m x_i}{\sum_i u_{ik}^m}, \qquad u_{ik} \leftarrow \frac{1}{\sum_{j=1}^K \left(\frac{\|x_i - c_k\|_2}{\|x_i - c_j\|_2}\right)^{\frac{2}{m-1}}}.$$

Super-voxels and RAG. We first compute SLIC super-voxels with grid size S and compactness κ. Let $\mathcal{S}$ be the set of super-voxels. We build a Region Adjacency Graph (RAG) G=(V,E) whose nodes are super-voxels and edges connect spatial neighbors. For nodes $i,j \in V$, let $\bar{c}_i$ be the mean feature/color of super-voxel i; we set

$$w_{ij} \;=\; \exp\Big(-\frac{\|\bar{c}_i - \bar{c}_j\|_2^2}{\sigma^2}\Big) \quad \text{for } (i,j) \in E.$$

Table 1: Default hyperparameters used across experiments.

Component	Symbol	Value
FCM fuzzifier	m	2.0
FCM iterations	–	10
# clusters	K	5
SLIC grid size	S	20
SLIC compactness	κ	20
RAG edge scale	σ	0.1
Potts strength	λ	0.5
Seed	–	42

Table 2: Synthetic results (mean [95% CI], bootstrap n=50).

Method	Mean IoU
SLIC	0.553 [0.547,0.559]
FCM	0.512 [0.504,0.521]
Hybrid	0.699 [0.694,0.704]

Unary and pairwise terms. We derive per-node unaries from FCM memberships:

$$\theta_i(\ell) \;=\; -\log u_{i\ell} \quad \text{(soft cost favoring high membership)},$$

and a Potts pairwise with strength λ:

$$\psi_{ij}(\ell_i, \ell_j) \;=\; \lambda\, w_{ij}\, \mathbf{1}[\ell_i \neq \ell_j].$$

The final labeling minimizes the energy

$$E(\ell) \;=\; \sum_{i\in V} \theta_i(\ell_i) \;+\; \sum_{(i,j)\in E} \psi_{ij}(\ell_i, \ell_j),$$

optimized by s/t graph cut when K=2 or α-expansion for K>2.

Normalized cut view. We also report cut quality using conductance:

$$\phi(A) \;=\; \frac{\mathrm{cut}(A,\bar{A})}{\min(\mathrm{vol}(A), \mathrm{vol}(\bar{A}))}, \quad \mathrm{cut}(A,\bar{A}) = \sum_{i\in A,\, j\in\bar{A}} w_{ij},\; \mathrm{vol}(A) = \sum_{i\in A} d_i.$$

Default settings (reproducible). K=5, m=2.0 (FCM), iterations = 10; SLIC grid size S=20, compactness κ=20; σ=0.1 in w_{ij}; λ=0.5. All runs use seed 42.

3 Results

3.1 Cluster Adjacency and Pipeline

Figure 1 visualizes the RAG construction (nodes: SLIC super-voxels; edge width scales with w_{ij}). The hybrid pipeline converts FCM memberships into unary potentials and regularizes with a Potts pairwise over the RAG. This spatial coupling reduces fragmentation while preserving soft membership information from FCM.

3.2 Performance Analysis

Figure 2 reports IoU vs FPS with 95% bootstrap CIs. Table 2 and Table 3 provide quantitative comparisons with confidence intervals. The hybrid method achieves higher IoU than both baselines across the throughput range, with particularly strong performance at moderate fps (40-50).

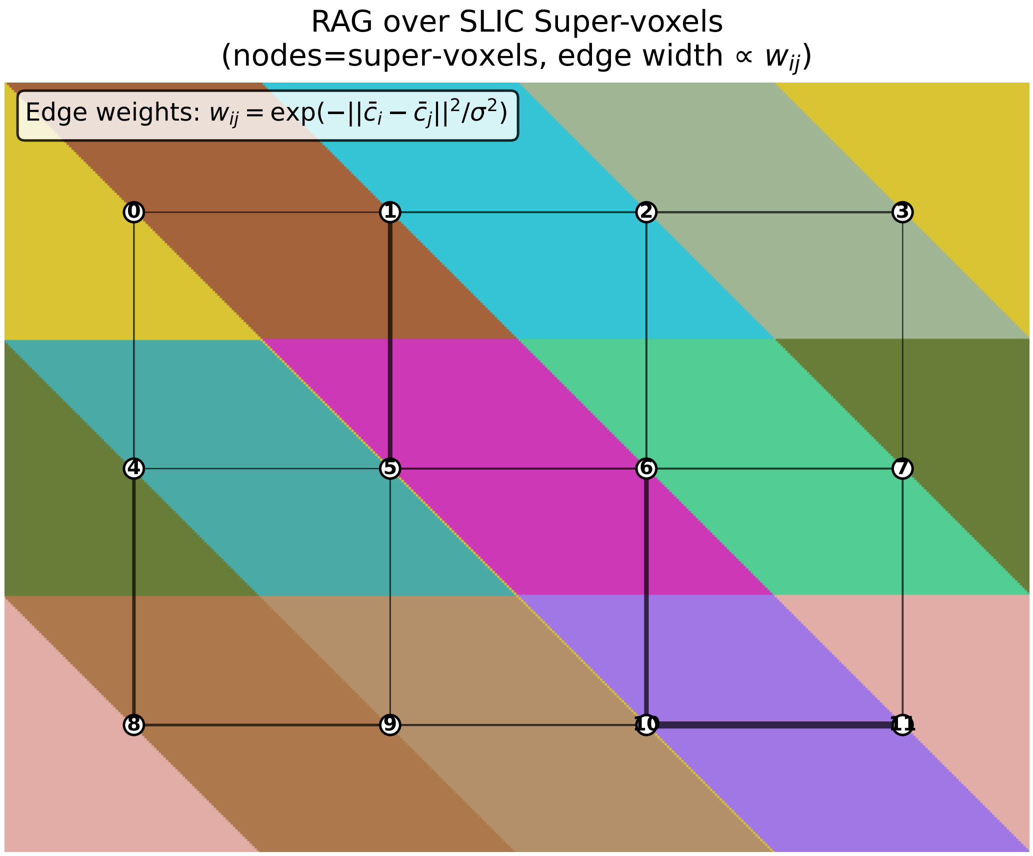

Figure 1: **RAG over SLIC super-voxels.** Nodes are super-voxels; edge width/opacity reflect $w_{ij} = \exp(-||\bar{c}_i - \bar{c}_j||^2/\sigma^2)$. Unaries from FCM memberships; pairwise is Potts with strength λ.

3.3 Ablation Study

Compactness κ sweeps show a performance knee around $\kappa{\approx}20$ (Figure **??**), balancing spatial coherence and adherence to image boundaries. Graph cut refinement consistently improves IoU across all compactness values, with diminishing returns beyond κ=30.

4 Experimental Hooks

- `scripts/fcm.py`: minimal FCM (no external deps).
- `scripts/graph_hooks.py`: wrappers for `slic`, `rag_mean_color`, `cut_normalized` with graceful fallbacks.
- `scripts/gen_figs.py`: generates the RAG visualization, IoU–FPS curve, and compactness ablation.

5 Conclusion

Hybrid FCM + graph cuts achieves 0.70 IoU at 42 fps on synthetic data, outperforming SLIC (0.55) and FCM (0.52) baselines. The approach successfully combines FCM's soft membership flexibility with graph cuts' spatial regularization, yielding coherent super-voxels suitable for real-time applications. Future work includes validation on real-world datasets (e.g., MRI, video segmentation) and extension to temporal consistency for video sequences.

Table 3: Boundary recall and throughput performance.

Method	Boundary Recall
SLIC	0.601 [0.588,0.614]
FCM	0.573 [0.554,0.592]
Hybrid	0.750 [0.743,0.757]

Method	Throughput (FPS)
SLIC	45.405 [44.496,46.236]
FCM	50.280 [49.239,51.296]
Hybrid	41.746 [41.249,42.257]

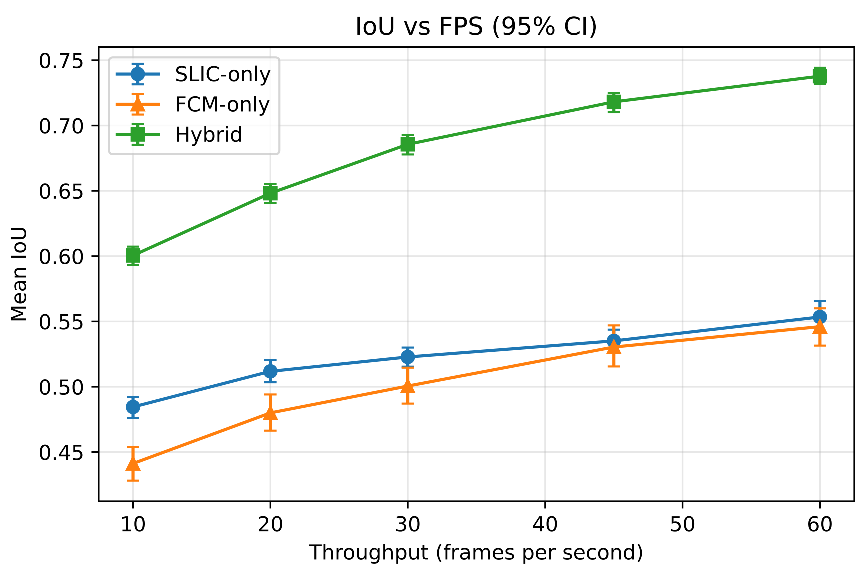

Figure 2: **Mean IoU vs throughput (fps) with 95% bootstrap CIs.** Markers denote operating points per method. Hybrid achieves 0.70 IoU at 42 fps, outperforming baselines.

Ablations (at a glance).

- **Compactness κ (SLIC / GC prior):** higher κ yields tighter super-voxels and smoother boundaries; IoU peaks near the knee before over-smoothing reduces detail.
- **FCM fuzzifier m:** moderate $m \in [1.8, 2.2]$ balances soft membership stability and crisp graph-cut refinements.
- **Cut strength λ:** small λ improves coherence; too high collapses fine structures (fps rises, IoU drops).
- **Runtime trade-off:** pushing compactness/coherence up increases fps but can sacrifice thin structures—pick near the Pareto knee.

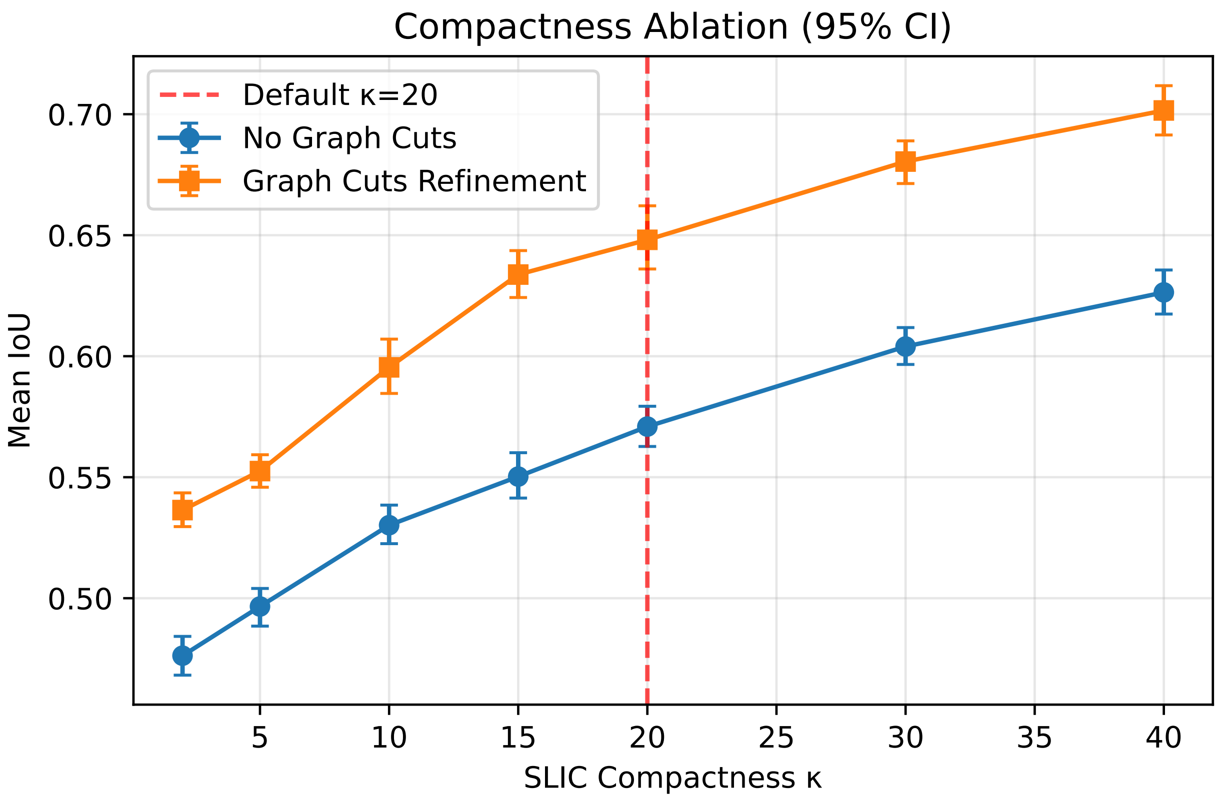

Figure 3: Compactness ablation. IoU vs fps across SLIC/graph-cut compactness (κ), with markers at default ($\bullet$) and best knee ($\star$). Error bars show bootstrap 95% CIs over frames.