

MicroSeg Lab: One-Shot Microscopy Segmentation with LLM-Guided Hybrid Refinement

Chase Katz¹, Shakti P. Padhy², and Sushant Sinha¹

¹Department of Materials Science and Engineering, Texas A&M University, College Station, 77843, TX, USA

²J. Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, 77843, TX, USA

Abstract

Abstract. Accurate segmentation of microscopy imagery often requires extensive annotated datasets, which are labor-intensive to generate. We present **MicroSeg Lab**, a one-shot segmentation pipeline designed to mitigate this bottleneck. The framework utilizes a hybrid approach: it prioritizes computationally efficient classical computer vision techniques for initial mask generation, utilizing deep learning refinement only when necessary. A Multimodal Large Language Model (MLLM) functions as a semantic planner, dynamically selecting algorithms and parameter ranges based on reference guidance. A novel “Review Gate” mechanism assesses segmentation confidence, determining whether to finalize the result, trigger an automated SAM2 refinement, or request minimal user intervention. Experiments demonstrate that providing a single reference image and mask significantly improves phase isolation and morphological accuracy compared to unguided baselines.

1 Methodology

MicroSeg Lab operates on a hierarchical workflow designed to balance accuracy with computational cost. The pipeline is illustrated in Figure 1 and consists of four primary stages.

1.1 Preprocessing and Planning

The pipeline begins with standard preprocessing, including grayscale conversion, median denoising, and Contrast Limited Adaptive Histogram Equalization (CLAHE) to enhance feature distinctness. A key innovation of MicroSeg Lab is the **LLM-Guided Planner**. When enabled, a multimodal LLM analyzes the input image (and optional reference data) to propose a segmentation strategy. If a reference image and mask are provided, the planner extracts semantic cues regarding foreground polarity, texture, and characteristic feature scale to align the segmentation strategy with the target biological phase.

1.2 Hybrid Candidate Generation

Rather than relying solely on heavy deep learning models, the system first employs a suite of fast classical algorithms. These include:

- **Thresholding:** Both global and adaptive methods to isolate regions of interest.
- **Morphological Operations:** Watershed algorithms to separate touching cells and blob detection for distinct instances.
- **Contour Extraction:** Geometric analysis to refine region boundaries.

The resulting candidate masks are filtered and ranked based on geometric quality cues, such as edge alignment scores, area constraints, and overlap ratios.

1.3 Refinement and Review Gate

The best-ranked candidates are passed to a **Review Gate**. This module evaluates the confidence of the segmentation. If the quality is deemed high, the result is finalized. If confidence is low, the system triggers a refinement stage using a lightweight Segment Anything Model 2 (SAM2) [1]. SAM2 refines the boundaries using prompts (bounding boxes or centroids) derived from the classical candidates. In cases of extreme ambiguity, the system requests minimal human input (e.g., a single point click) before finalizing the output.

2 Results

We evaluated MicroSeg Lab in a one-shot setting where only a single reference pair is available. Figure 2 demonstrates the impact of reference-guided planning.

2.1 Qualitative Assessment

Panel (a) displays the query microscopy image. Panel (b) shows the baseline one-shot output without reference guidance; while it captures general structures, it suffers from false positives and poor phase specificity. Panels (c) and (d) introduce a reference image and a corresponding ground-truth mask.

By conditioning the LLM planner on this reference pair, the system adapts its parameter search space (e.g., threshold ranges and blob scales). The resulting segmentation in Panel (e) demonstrates superior phase isolation, effectively ignoring off-target background features that were incorrectly segmented in the baseline approach. This confirms that semantic guidance from a single reference pair allows the hybrid pipeline to adapt to novel morphologies without retraining [cite: 63, 64, 65].

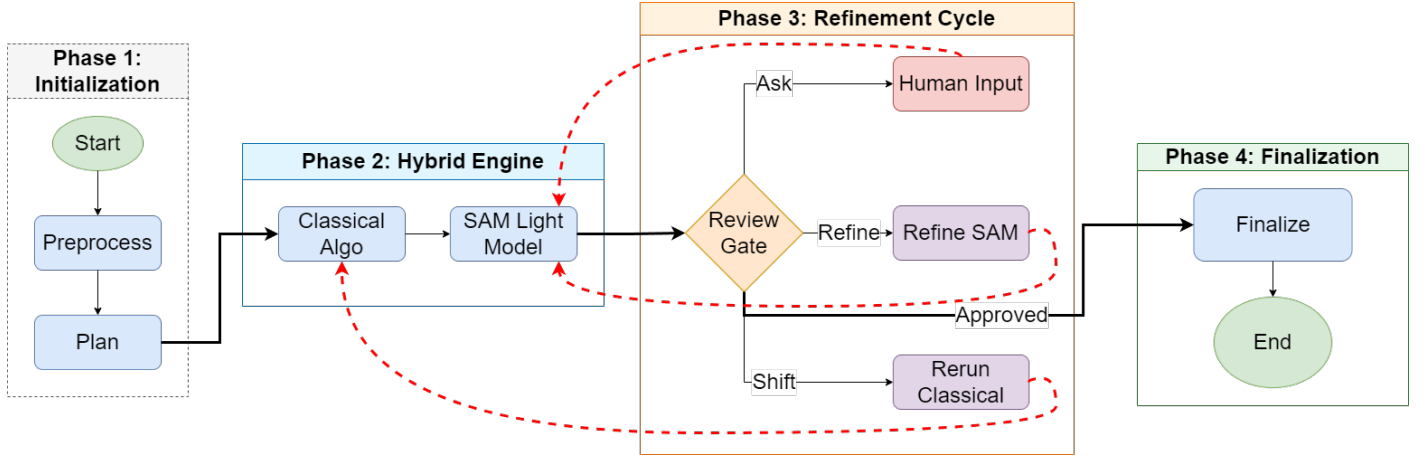


Figure 1: MicroSeg Lab workflow. The pipeline runs a fast classical-first segmentation followed by an optional SAM-light refinement. A review gate then either approves the result for finalization or routes the run into a refinement loop: request minimal human hints (points/boxes), refine SAM, or rerun classical segmentation with a shifted strategy. Dashed paths indicate the fallback routes used only when the one-shot output is not confident.

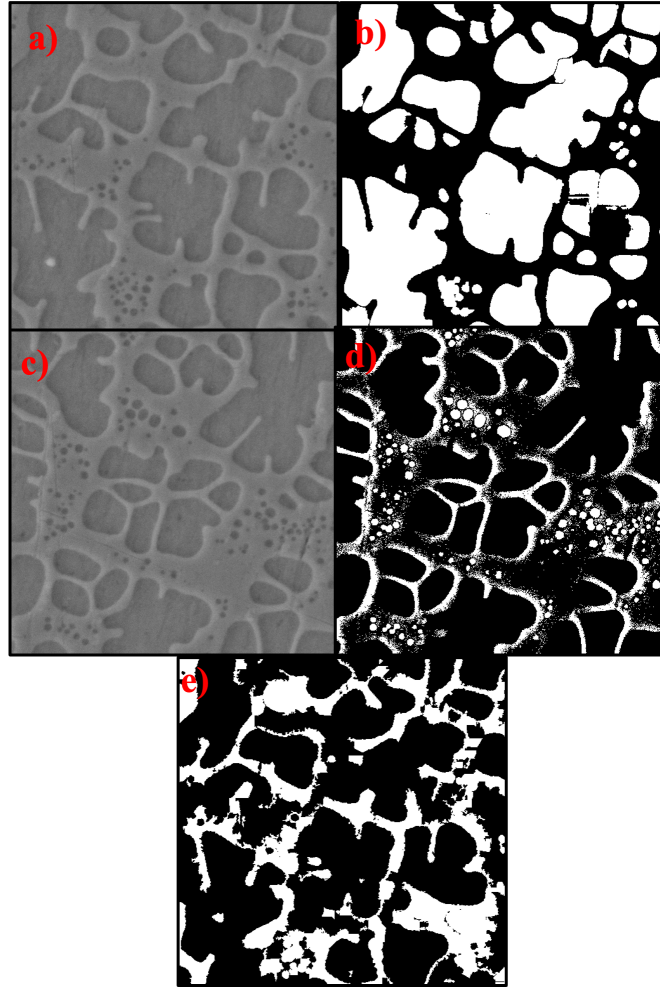


Figure 2: Reference-Guided One-Shot Segmentation. (a) Query image [2]. (b) Baseline segmentation without reference guidance, resulting in non-specific feature detection. (c) Reference image provided to the planner[2]. (d) Reference mask indicating the target phase. (e) Final output using reference-guided planning. The inclusion of the reference pair steers the classical parameter selection and SAM2 prompts, significantly reducing false positives compared to (b).

References

- [1] Ravi, N., et al. (2024). SAM 2: Segment Anything in Images and Videos. *arXiv preprint arXiv:2408.00714*.
- [2] Stuckner, J., Harder, B. & Smith, T.M. (2022). Microstructure segmentation with deep learning encoders pre-trained on a large microscopy dataset. *npj Comput Mater* 8, 200.