

HistoSeg: One-Page App Summary

Evidence scope: README.md, setup.py, pyproject.toml, and src/histoseg/*.

What It Is

HistoSeg is a Python toolkit for spatial transcriptomics segmentation and geometry extraction. In this repo, the main implemented workflow generates Pattern1 isoline contours (default level 0.5) from clustered cell data.

Who It Is For

Primary persona: computational biology and bioinformatics users working with 10x Xenium and GraphClust outputs. Explicit persona statement: Not found in repo (persona inferred from README/tutorial inputs).

What It Does

- Aligns clusters.csv with cells.parquet and auto-detects join id and x/y coordinate columns.
- Selects target Pattern1 clusters and builds background points from non-target cells.
- Optionally generates synthetic background points from tissue_boundary.csv bounding box.
- Fits a distance-weighted KNN regressor, predicts on a mesh, and smooths probabilities with a Gaussian filter.
- Masks non-tissue grid regions, extracts isolines, and filters loops by minimum cells-inside threshold.
- Writes reproducible outputs: params.json, contour .npz arrays, and preview .png image.
- Provides optional Hugging Face dataset download helper and an experimental GUI entry point (histoseg-gui).

How It Works (Architecture)

- Inputs: clusters.csv + cells.parquet (+ optional tissue_boundary.csv) from local files or HF dataset helper.
- Core compute: run_pattern1_isoline in src/histoseg/contours/pattern1_isoline.py.
- I/O service: src/histoseg/io/huggingface.py downloads required dataset files for HF wrapper runs.
- UI surface: tkinter GUI in src/histoseg/gui/gui_app.py, exposed by console script histoseg-gui in setup.py.
- Data flow: ingest and align -> choose target/background -> KNN field -> smoothing/mask -> contour extraction -> saved artifacts.
- Backend API service or deployed web service: Not found in repo.

How To Run (Minimal)

- Install: pip install -U histoseg (or from source: pip install -e .)
- Prepare inputs: clusters.csv, cells.parquet, and optional tissue_boundary.csv.
- Run local API: Pattern1IsolineConfig(...) + run_pattern1_isoline(...).
- Alternative: run_pattern1_isoline_from_hf(repo_id=...) to pull inputs from a Hugging Face dataset repo.
- Optional GUI: run histoseg-gui (README marks this path experimental).