Dec 22, 2025

Cell and Nucleus Segmentation in the Aegle Pipeline

DOI
https://dx.doi.org/10.17504/protocols.io.8epv55614v1b/v1

Da Kuang[1]

[1]University of Pennsylvania

Da Kuang

Create & collaborate more with a free account

Edit and publish protocols, collaborate in communities, share insights through comments, and track progress with run records.

Create free account

OPEN ACCESS

DOI: https://dx.doi.org/10.17504/protocols.io.8epv55614v1b/v1

Protocol Citation: Da Kuang 2025. Cell and Nucleus Segmentation in the Aegle Pipeline. **protocols.io** https://dx.doi.org/10.17504/protocols.io.8epv55614v1b/v1

License: This is an open access protocol distributed under the terms of the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited

Protocol status: Working
We use this protocol and it's working

Created: December 22, 2025

Last Modified: December 22, 2025

Protocol Integer ID: 235653

Keywords: nucleus segmentation in the aegle pipeline, hubmap female reproductive system imaging data, nucleus segmentation, aegle pipeline, level segmentation output, automated segmentation, tissue region, image preprocessing, cell

Abstract

This protocol documents the computational workflow used to generate cell- and nucleus-level segmentation outputs for HuBMAP Female Reproductive System imaging data. Tissue regions are manually annotated, followed by image preprocessing and automated segmentation using the Aegle pipeline.

Overview

1 This protocol documents the computational segmentation procedure used to generate cell- and nucleus-level segmentation masks and quantitative cell profiles for HuBMAP Female Reproductive System (FRS) imaging data.

The protocol records:

- how tissue regions are defined,
- how raw images are preprocessed,
- how cell and nucleus segmentation is performed,
- and what artifacts are produced.

This protocol is intended for **data provenance, reproducibility, and auditability** in HuBMAP data submissions.
It does **not** describe algorithmic implementation details beyond what is required to reproduce outputs.

2 Software and Resources:

- Pipeline: Aegle image analysis pipeline
- GitHub: https://github.com/kimpenn/aegle
- Annotation Tool: QuPath v0.6.0
- Segmentation Model: DeepCell Mesmer (via Aegle)

Input Data:

- PhenoCycler / CODEX whole-slide images (.qptiff)
- Manually annotated tissue boundaries (.geojson)

———

Manual tissue annotation

3 Purpose
Define biologically relevant tissue regions to restrict downstream processing.

Procedure

1. Open the raw .qptiff image in QuPath (v0.6.0).
2. Manually delineate tissue boundaries corresponding to the target organ region.
3. Export annotations as GeoJSON files.

Outputs

- One GeoJSON file per scan defining tissue boundaries.

File Naming Convention
<sample_id>_<organ>_<region>_<scan_id>.tissue-boundary.geojson

Examples
- 10RTAMP5_FT_Ampula_Scan1.tissue-boundary.geojson
- 10RTF5_FT_Fimbriae_Scan1.tissue-boundary.geojson

Segmentation

4 Configure segmentation experiment

Purpose
Specify channels and parameters for cell and nucleus segmentation.

Procedure

Select the appropriate segmentation configuration:
- exps/configs/main/preprocess_ft_hb
- exps/configs/main/preprocess_uterus_hb
- exps/configs/main/preprocess_ovary_hb
- exps/configs/main/main_ft_hb
- exps/configs/main/main_uterus_hb
- exps/configs/main/main_ovary_hb

Configurations define:
- nuclear and membrane channels,
- segmentation resolution,
- post-processing rules,
- feature extraction settings.

5 Run preprocessing

Procedure

Execute the preprocessing launcher corresponding to the organ:

```sh
bash launcher/run_preprocess_ft.sh
bash launcher/run_preprocess_uterus.sh
bash launcher/run_preprocess_ovary.sh
```

Outputs

Preprocessing outputs are written to the same directory as the input image:

- extras/antibodies.tsv
- processed_hubmap/manual_polygon_overlay.png
- Downsampled overview images (.jpg)
- Tissue-restricted OME-TIFF image (.ome.tiff)

These outputs serve as inputs for segmentation.

6 Run cell and nucleus segmentation

Procedure

Execute the segmentation launcher:

```
bash launcher/run_main_ft.sh
bash launcher/run_main_uterus.sh
bash launcher/run_main_ovary.sh
```

The pipeline:

1. Applies DeepCell Mesmer for initial nucleus and cell boundary prediction.
2. Performs nucleus–cell matching and morphological refinement.
3. Extracts per-cell and per-nucleus quantitative features.

7 Generated outputs

Segmentation Masks

- <exp_id>.cell_matched_mask.segmentations.ome.tiff
- <exp_id>.nucleus_matched_mask.segmentations.ome.tiff

Quantitative Profiles

- cell_overview.csv — per-cell spatial, morphological, and intensity features
- nucleus_overview.csv — per-nucleus features

These files constitute the final segmentation products submitted to HuBMAP.