

Synthetic gastric H&E images

A technical account of tissue modelling, pathology diffusion and measurement-driven evaluation.

Grail Computer Pte Ltd | 9 September 2026
SGH / IMDA Open Innovation Platform

We combined a procedural model of gastric mucosa, a pathology-trained diffusion generator and a calibrated evaluation harness to investigate synthetic histopathology for anomaly-detection training. The work separates two questions: how to specify tissue architecture, and how to render its cellular appearance.

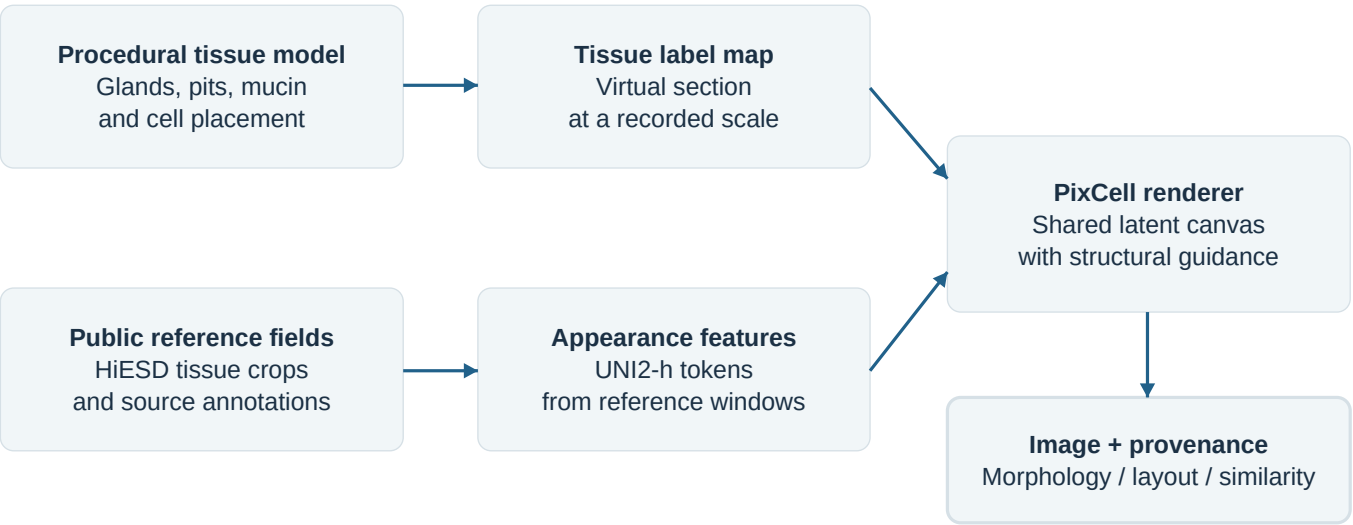


Figure 1. The two conditioning streams: procedural structure and reference appearance. The structure adapter supplies direct map guidance; reference-matching experiments supply appearance features without the adapter.

Output specification

Property	Implementation
Image raster	4096 × 2048 pixels, RGB PNG
Physical sampling	0.25 µm/px; a 1.024 × 0.512 mm field of view
Intended categories	H. pylori gastritis; intestinal metaplasia; co-occurrence; normal
Normal presets	Body-type and antrum-type gastric mucosa
Record per image	Generation route, map, sources, settings, measurements and checksum

The key experimental result is that appearance and layout respond differently to conditioning. Matching reference windows improved measured morphology; a direct structure adapter produced substantially stronger control of gland layout. Removing explicit nucleus channels from that adapter improved nuclear count and size while retaining much of its layout control.

Research record: sdh-challenge.grail.computer. This report describes recorded experiments; the linked reports contain measurements, comparisons and provenance.

Building the tissue architecture

A three-dimensional construction produces a two-dimensional section and an explicit label map.

The tissue model represents hollow glands beneath a mucosal surface, with a pit region, epithelial walls and lamina propria. Presets control gland dimensions, pit depth, lumen shape, goblet-cell mucin, stromal nuclei and inflammatory follicles. Section angle and depth determine whether glands appear as rings, oblique profiles or longitudinal channels.

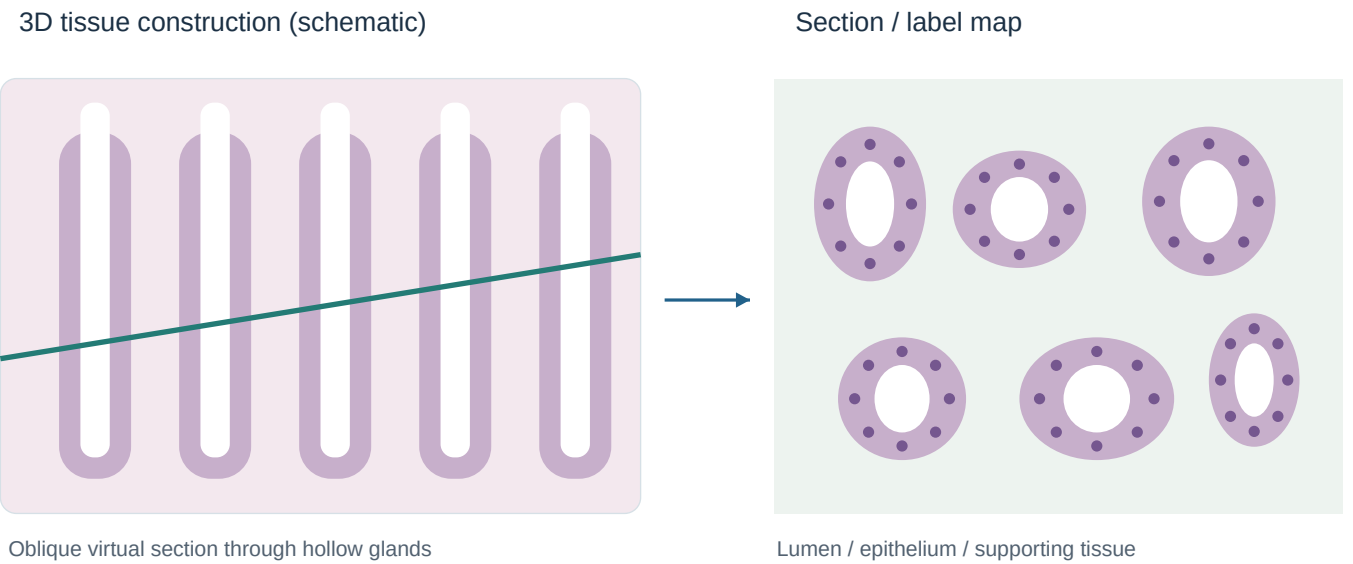


Figure 2. Conceptual geometry, not a microscopy image. The section inherits its spatial organisation from the procedural tissue model.

Recorded v2 preset parameters

Parameter	Body	Antrum	Gastritis	IM	Mixed
Outer diameter (µm)	44-60	46-66	52-74	58-84	56-80
Wall thickness (µm)	14-22	16-24	18-25	20-29	19-28
Pit depth / slab depth	0.22	0.42	0.50	0.28	0.38
Goblet fraction	0	0	0	0.36	0.30
Stromal nuclei / mm²*	12,500	10,500	17,000	10,500	15,000

*Procedural placement settings, not measured output densities. IM = intestinal metaplasia. These are implementation parameters, not universal histological reference ranges.

The nine labels are background, lumen, epithelial cytoplasm, goblet mucin, epithelial nucleus, stroma, stromal nucleus, vessel and red cell. The v2 renderer added epithelial granularity and a minimum mucin-to-lumen clearance so the gland wall did not collapse into a ring of separate mucin wedges.

Across the v2 cartoon evaluation, 68 of 72 category × cut × metric summaries fell inside the real reference bands, versus 50 of 72 for the earlier construction. This measures the drawing before diffusion rendering. [Source: A1-V2 report.](#)

Rendering a full-size field

Reference conditioning and overlapping-window diffusion on one shared canvas.

PixCell-1024 supplies a pathology-trained diffusion transformer. UNI2-h supplies 16 appearance tokens per 1024-pixel reference window. At inference, 21 overlapping windows cover the 4096 × 2048 canvas; denoising predictions are combined on a shared latent representation before decoding.

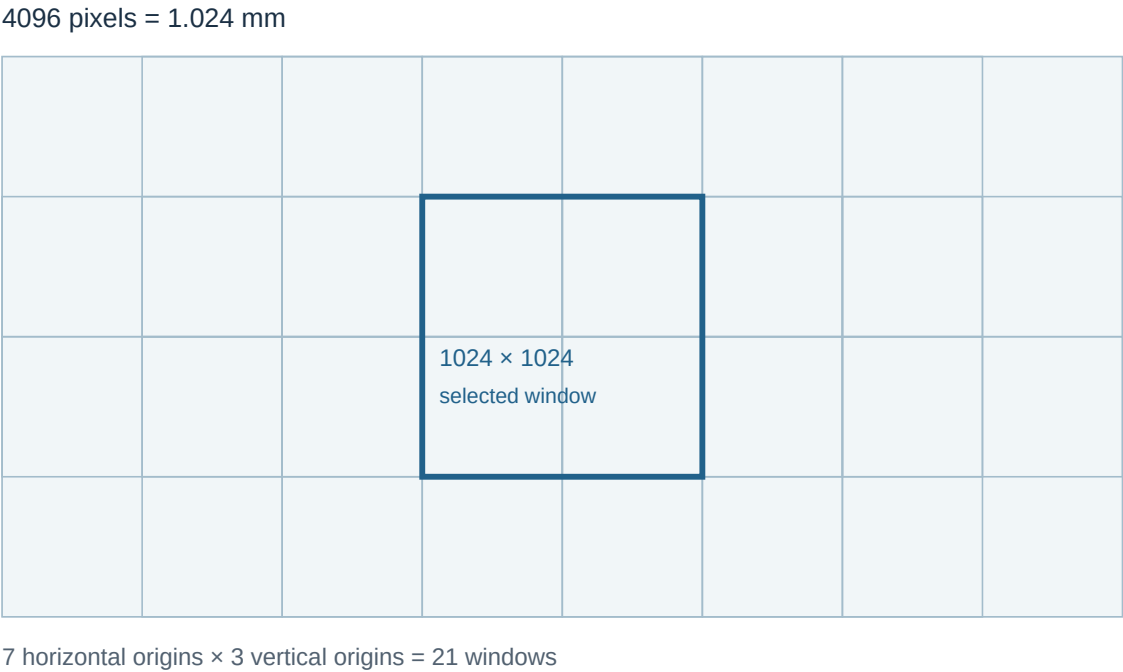


Figure 3. Window geometry: 1024-pixel windows with 512-pixel spacing. Overlap is part of generation, rather than a final collage of independently rendered fields.

Three conditioning routes tested

Route	What supplied the appearance / structure
Single reference	Tokens from one real field; repainting used as a baseline. Reference architecture can persist.
Matched mosaic (A5)	Each canvas window receives tokens from the real window best matching its synthetic lumen and epithelial masks.
Structure adapter (F2/F3)	An additional trained branch reads the label map; reference tokens still supply appearance.

Matched retrieval. For each output window, the search compared 24 same-category training fields × 21 window origins × 8 rotations/reflections: 4,032 candidates. The score was $\text{IoU}(\text{lumen}) + 0.5 \times \text{IoU}(\text{epithelium})$, evaluated at 1/8 resolution using block-mean masks. A canvas used a median of 11 distinct reference fields (range 5-14).

Diffusion settings. The adapter inference runs used a 20-step DPMSolver++ schedule and classifier-free guidance of 1.5. From-noise generation and image-to-image repaint were tested separately. Repaint start indices 6 and 3 execute 14 and 17 remaining steps respectively. [Sources: A5; F3.](#)

Teaching the renderer to read structure

A small trainable branch adds spatial guidance to a frozen pathology generator.

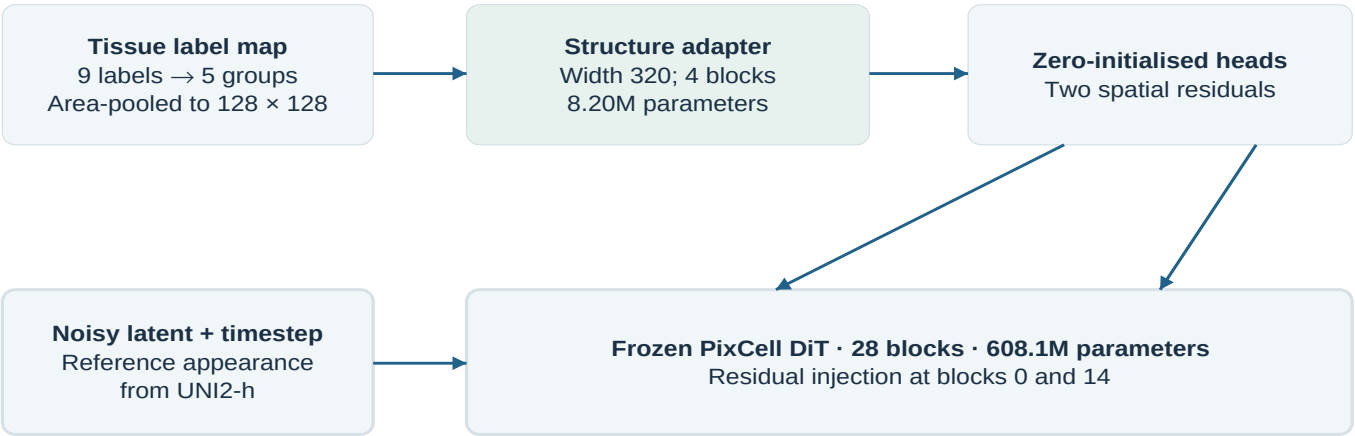


Figure 4. F3 coarse adapter. The trainable branch changes spatial conditioning while the PixCell backbone remains frozen. A zero adapter scale adds no residual and provides the paired control.

Coarse conditioning vocabulary

Channel	Merged tissue labels
0 / background	Background above the surface
1 / lumen	Lumen
2 / epithelium	Cytoplasm + goblet mucin + epithelial nucleus
3 / stroma	Stroma + stromal nucleus
4 / vascular	Vessel + red cell

The F2 adapter used all nine channels (8.25M parameters). F3 removed separate nucleus channels by merging them into their tissue compartments. Each latent cell retains a soft partition whose channels sum to one. The group definition is stored in the checkpoint and reconstructed at inference.

Training configuration

The 96-field development pool was split again by slide into 64 training and 32 validation fields. This internal adapter split is distinct from the 22-field external morphometric calibration set. Training pairs used automatically segmented real images, rather than pathologist-drawn masks.

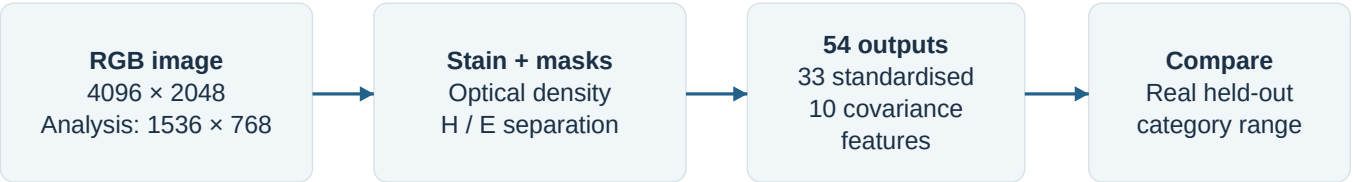
Both adapters used 1024-pixel crops, batch size 4, 3,000 steps, learning rate 10^{-4} with cosine decay and 100 warmup steps, weight decay 0.01, gradient clipping at 1.0, classifier-free and conditioning dropout of 0.10, bf16 and gradient checkpointing. Checkpoints were saved every 500 steps; validation ran every 250 steps.

The coarse run completed 3,000 steps in 8,198 seconds. Model, reference-token and label-map caching allowed a single inference process to run many jobs. Sources: [F2 training](#); [F3](#).

Measuring morphological similarity

A calibrated multi-feature instrument separates gland topology from broader tissue appearance.

Reference fields came from public HiESD slides. Native sampling of approximately 0.2458 $\mu\text{m}/\text{px}$ was resampled to 0.25 $\mu\text{m}/\text{px}$. The morphometric calibration used 96 training fields (24 per category) and 22 fields held out by slide: four normal fields and six for each other category.



Nuclei · epithelial spacing · stroma · stain · lumen geometry

Figure 5. Feature extraction and calibration. The 54 reported values are not all dimensions of the main distance score.

Feature construction and distances

The analysis raster has effective sampling of 0.6667 $\mu\text{m}/\text{px}$. Fixed H&E stain vectors and optical density $OD = -\log((RGB + 1)/256)$ preserve stain shifts instead of normalising them away. Nuclear candidates are split by watershed and filtered to 5-180 μm^2 . The epithelial proxy uses a 6 μm morphological closing of the haematoxylin-rich mask.

The instrument reports 54 values across gland topology, nuclear morphology, epithelial spacing, stromal texture, colour and global tissue fractions. For 33 selected features, robust standardisation uses $z = (x - \text{median})/(1.4826 \times \text{MAD})$, with standard-deviation fallback. Their RMS provides a diagonal summary.

The main distance D is a robust Mahalanobis distance in **10 core features**: gland-hole fraction, ring density, nuclear density and median area, epithelial nearest-neighbour distance and its coefficient of variation, stromal orientation coherence and eosin variance, and mean H and E optical density. MinCovDet uses support fraction 0.75, with a ridge of $0.10 \times \text{mean covariance diagonal}$.

Held-out calibration by programme category

Category	Train / held out	Median D	Maximum D
Normal	24 / 4	3.33	4.34
Gastritis	24 / 6	4.19	7.97
Intestinal metaplasia	24 / 6	2.58	4.50
Mixed	24 / 6	3.66	4.38

“Inside the band” means at or below the category-specific held-out maximum. It is a measured morphology criterion, not a disease classifier. The small held-out set produces coarse ranges. HiESD chronic-gastritis labels do not establish H. pylori positivity. [Source: C1 instrument report.](#)

Controlled experiments and results

Paired comparisons identify what each conditioning change actually contributed.

A5: morphology distance D (lower is closer)



Within category band: 14/40 baseline → 35/40 matched mosaic

Figure 6. A5 medians over 40 paired drawings (four categories, five seeds, two cuts). The real comparator contains 22 held-out fields. Matching improved D on 37 of 40 pairs, but lumen IoU remained low: 0.0835 versus 0.0452.

F3: remove explicit nucleus conditioning

Arm / comparator	D	Lumen IoU	Nuclei / mm²	Nuclear area µm²
Real held-out (n=22)	3.51	-	6,376	23.1
Two-pass si3	5.77	0.051	5,935	29.6
F2 nine-channel, noise	9.03	0.466	10,070	17.6
F3 coarse, noise	4.60	0.395	6,655	24.2
F3 coarse, repaint si3	4.40	0.357	6,351	25.8

Medians over the same 20 across-section drawings for generated arms; values are from the F3 comparison, not the A5 population. IoU = intersection over union with the intended mask.

Merging the nucleus channels reduced density error from approximately +58% to +4% and median area error from -24% to +5%, relative to the real comparator. Lumen IoU decreased from 0.466 to 0.395: the coarse adapter traded some fine layout adherence for substantially better nuclear statistics.

Adapter-off control isolates the layout effect

A separate F3 test held the drawing, starting image, donor, seed and 17-step repaint schedule fixed and changed only the nine-channel adapter scale. Across 20 pairs, enabling the adapter increased lumen IoU by a median 0.434 and ring IoU by 0.271. The corresponding adapter-off medians were 0.051 and 0.045. Retaining layout carried a median cost of +1.30 D in that comparison.

The evidence supports two distinct findings: reference matching improved appearance, while direct label conditioning controlled layout. These scores do not establish that the generated categories are diagnostically correct. Sources: A5; F3.

Similarity screens and provenance

Record source influence explicitly and assess image similarity at two different levels.

The evaluation harness combined a pathology-encoder screen with targeted pixel search. The results below describe the screened image set after procedural organism rendering. They are checks against reference similarity, not a proof that no source architecture was inherited.

Screen	Coverage	Recorded result
Embedding similarity	80 candidate tiles against 944 tiles from 118 reference fields; phikon-v2 encoder	Maximum cosine 0.8995; flag threshold 0.95; no flagged tiles.
Pixel-level NCC	83 candidate/reference pairs; 10,568 search windows; 64-pixel patches, stride 256	Maximum NCC 0.7889; threshold 0.92; no flagged windows.

Pixel search included token-source fields and nearest-neighbour fields identified by the embedding screen. It was targeted rather than an exhaustive comparison with every reference. Matching or near-matching tissue arrangement is a separate question from matching pixels.

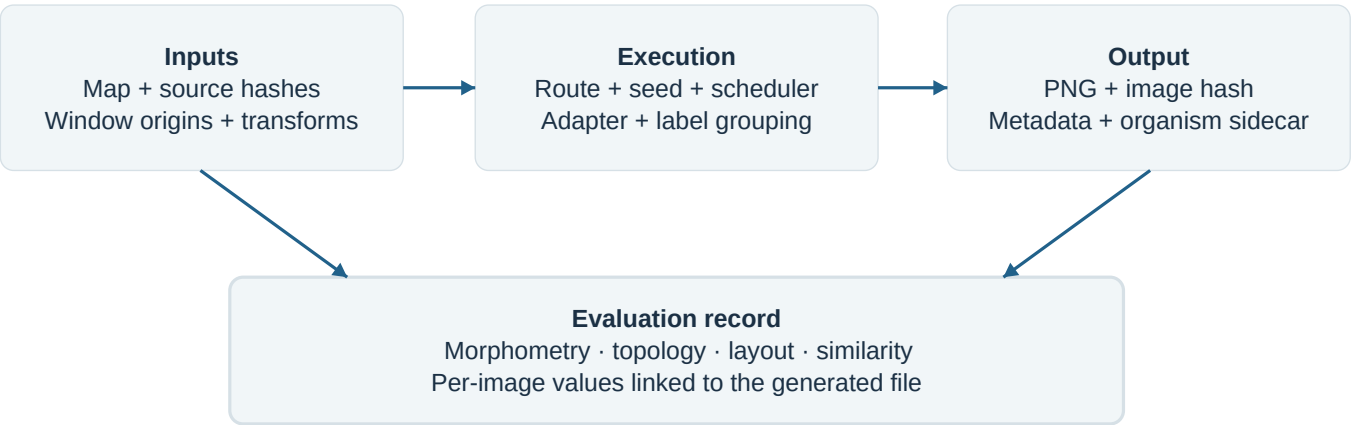


Figure 7. Provenance links the generated file to its inputs, execution settings and measurements. Checksums establish file identity and transfer integrity.

Procedural *H. pylori* rendering

The organism layer uses faint curved rods placed in surface mucus and pit or gland lumina, with the rendering route disclosed. These are procedural additions, not organisms learned from verified *H. pylori*-positive annotations. The sidecar records placement so that organism rendering can be distinguished from the underlying diffusion output.

Interpretation of the evidence

The measurements establish engineering behaviour: morphology relative to one public dataset, responses to controlled conditioning changes and similarity within the searched references. They do not establish clinical validity, downstream anomaly-detection benefit or complete independence from donor layout. No pathologist review was part of these experiments.

The 1536-pixel analysis raster is insensitive to some loss of fine detail, so it cannot certify authentic 0.25 µm/px information. Selection also included an explicit gland-hole-score exception (0.985 versus the held-out normal maximum of 0.976); the record therefore does not support a universal all-thresholds-pass claim. [Sources: similarity screens; selection record.](#)

Sources and evidence

Experiment reports, data attribution and model references.

Public evidence library

sdh-challenge.grail.computer contains the methods, comparison figures, source records and downloadable measurements. The experiments cited in this report can be accessed directly:

[A1-V2: tissue construction and preset evaluation](#)

[A5: matched-reference retrieval and paired comparisons](#)

[F2: nine-channel adapter training](#)

[F3: coarse adapter and adapter-off controls](#)

[C1: morphometric instrument and calibration](#)

[Similarity screens: coverage and per-image results](#)

Data and model attribution

HiESD: JSGe-AI / dataset authors, *A fully annotated pathology slide dataset for early gastric cancer and precancerous lesions*. doi:10.6084/m9.figshare.28919840. The dataset record states CC BY 4.0. Source categories and slide-level splits are retained in the evidence record.

PixCell: [author repository](#) and [PixCell-1024 model card](#). **UNI2-h:** [Mahmood Lab model card](#). **phikon-v2:** [Owkin model card](#).

Company: Grail Computer Pte Ltd | grail.computer

Contact: Yash Agarwal