Image-based porosity estimation in Cambrian sandstones from the Vilkyčiai-22 well

Nurlan Talibov, Donatas Kaminskas, Anna Cichon-Pupienis
nurlan.talibov@gamtc.lt

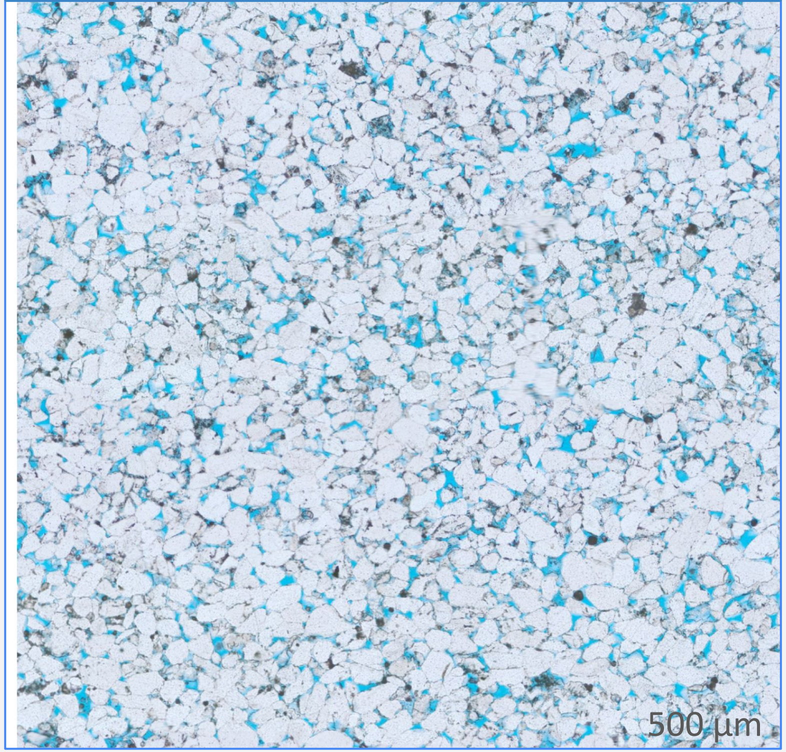
Research Question

Can Deep Learning Replace Manual Thresholding in Thin Sections?

Photography encouraged

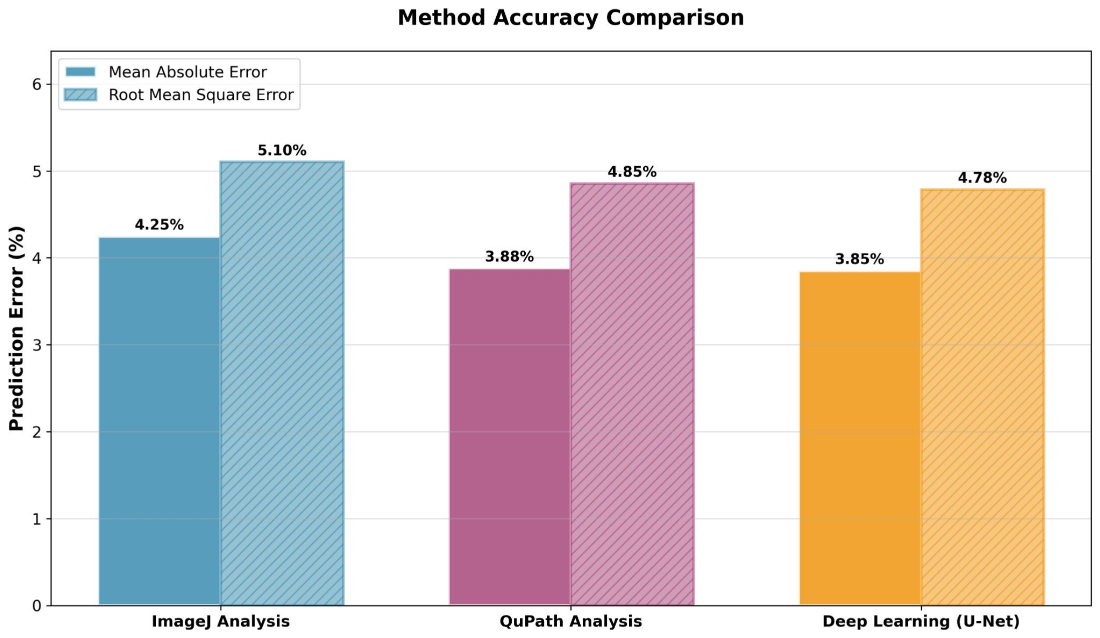
- 26 Cambrian sandstone core plugs - Vilkyčiai-22 well, Baltic Paleobasin, Western Lithuania
- Lab ground truth: porosity - 2.18% -15.14%.
- Three image-based methods: ImageJ (auto), QuPath (manual), U-Net (deep learning).
- Key question: can U-Net match expert accuracy without per-sample manual input?

Sample 009 - Lab porosity - 11.18%



Key Results

Three Method Accuracy vs Helium Porosimetry



ImageJ

MAE : 4.25%

Otsu’s algorithm overestimates pores systematically

QuPath

MAE : 3.88%

Strong MAE with expert threshold adjustment; operator dependent.

U-Net

MAE : 3.85%

Lowest MAE; fully automated at inference; inherits QuPath bias.

Sample Material

26 core plugs from Vilkyčiai-22 well; standard petrographic thin sections, GSL, 2014

Property	Value/Detail
Number of samples	26 core plugs
Lithology	Cambrian sandstone
Depth range	2003.5 - 2201.4 m
Lab. porosity range	2.18% - 15.14%
Thin section thickness	~30 µm (standard petrographic)
Impregnation	Blue-dyed epoxy resin (pore-space filler)
Microscope	Nikon Eclipse
Camera	Nikon DS-Fi3
Magnification / mode	10x objective, plane-polarized light (PPL)

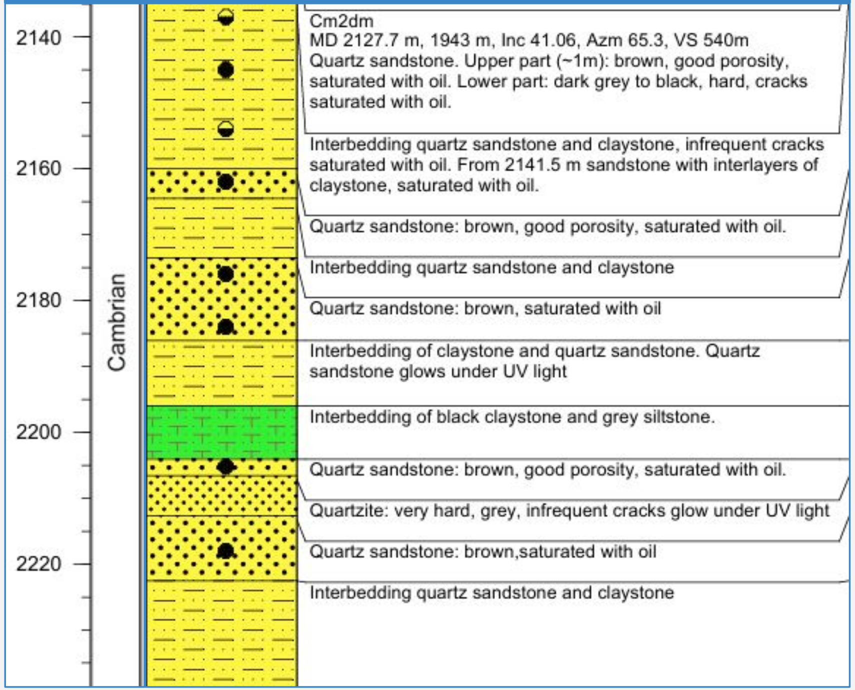
Core plug sample from Vilkyčiai-22 well, GSL core repository, Vievis

Photography encouraged

Study Context: Vilkyčiai-22 Well

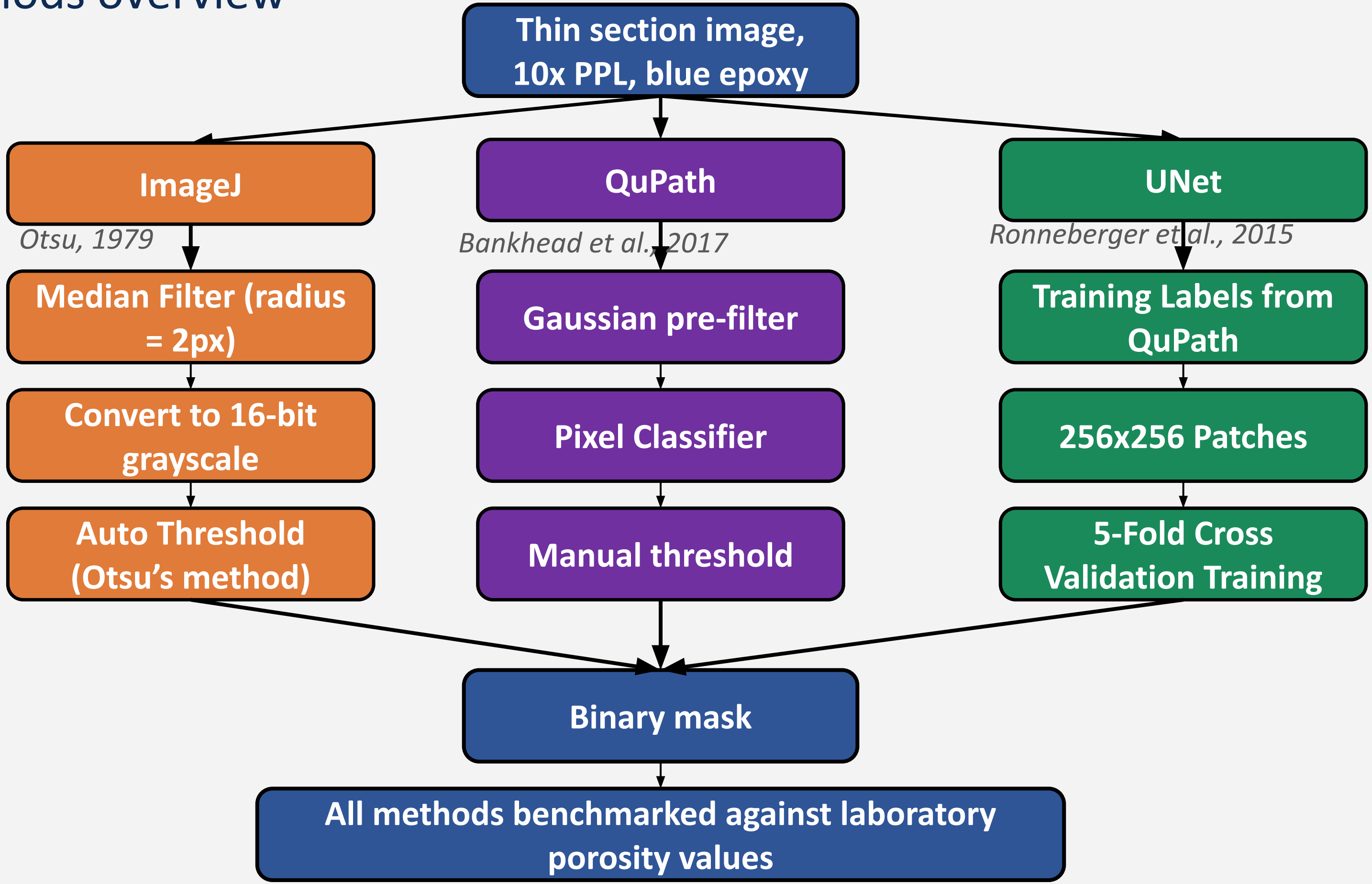
Location	Vilkyčiai-22 well, western Lithuania, Baltic Basin – a Vendian–Phanerozoic pericratonic sedimentary basin
Lithology	Cambrian sandstones with usually fine to medium grain size and cementation; detrital quartz-dominated framework
Depth interval	2003.5 – 2201.4 m; representative of hydrocarbon reservoirs
Core material	26 core plugs; helium porosimetry: 2.18% – 15.14%

Lithological log - Vilkyčiai-22



Methods Overview

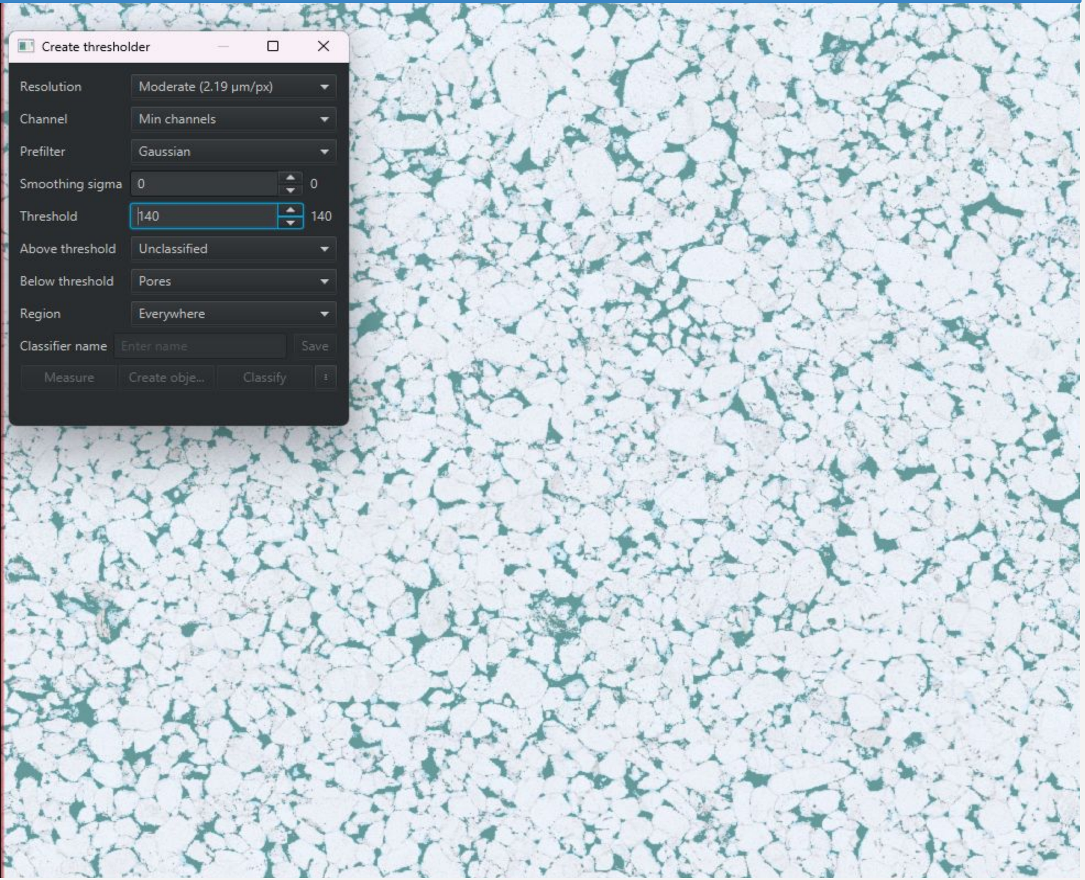
Three methods overview



QuPath in Practice: The Operator's Decision

Photography encouraged

QuPath Interface - live threshold preview



Key operator decisions

Threshold = 140

Set manually per sample. Dark blue-epoxy pixels (pores) fall below this intensity value.

Channel: Min channels

Uses the minimum value across RGB channels - maximises contrast between blue epoxy (low min) and mineral grains (high min).

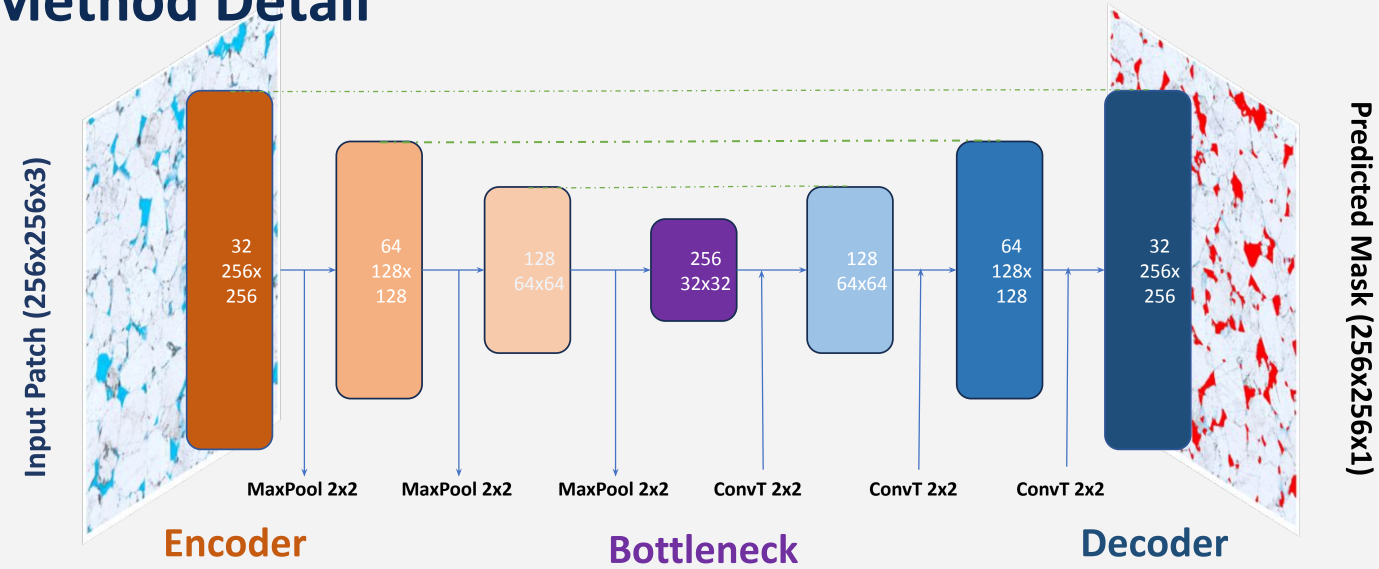
Below threshold -> Pores

Classification direction: pixels darker than threshold are labelled as pore space.

Live preview

Operator sees segmentation update in real time while adjusting threshold.

UNet Method Detail



Photography encouraged

Patch size 256 x 256 px	Patches / image 300–500 non-overlapping
Loss function Focal (α=0.75, γ=2.0) + Tversky (α=0.7, β=0.3)	Optimizer / LR Adam · 1×10^{-4}
Augmentation Flips, 90° rotations, brightness/contrast jitter	Validation 5-fold cross-validation, stratified by porosity
Threshold 0.9 per fold (optimised on validation set)	Framework TensorFlow 2.10, RTX 3050 Ti GPU

Segmentation Results

UNet benchmarked against ground truth QuPath

Photography encouraged

Metric	Mean ± SD	Range
Dice Coefficient	0.8554 ± 0.0424	0.6884 - 0.9042
IoU (Jaccard)	0.7495 ± 0.0602	0.5249 - 0.8251
Precision	0.8295 ± 0.0513	0.6126 - 0.8942
Recall	0.8844 ± 0.0429	0.7655 - 0.9529
F1-Score	0.8554 ± 0.0424	0.6884 - 0.9042
Pixel Accuracy	0.9846 ± 0.0067	0.9662 - 0.9949
Specificity	0.9899 ± 0.0047	0.9777 - 0.9964
Porosity MAE (%)	0.3856 ± 0.3145	0.0236 - 1.2110

Table: Segmentation Performance Metrics Summary: U-Net vs. QuPath Reference Masks

0.86 **Mean Dice**

strong segmentation fidelity

92% **samples with porosity error < 1%**

porosity MAE vs QuPath 0.39%

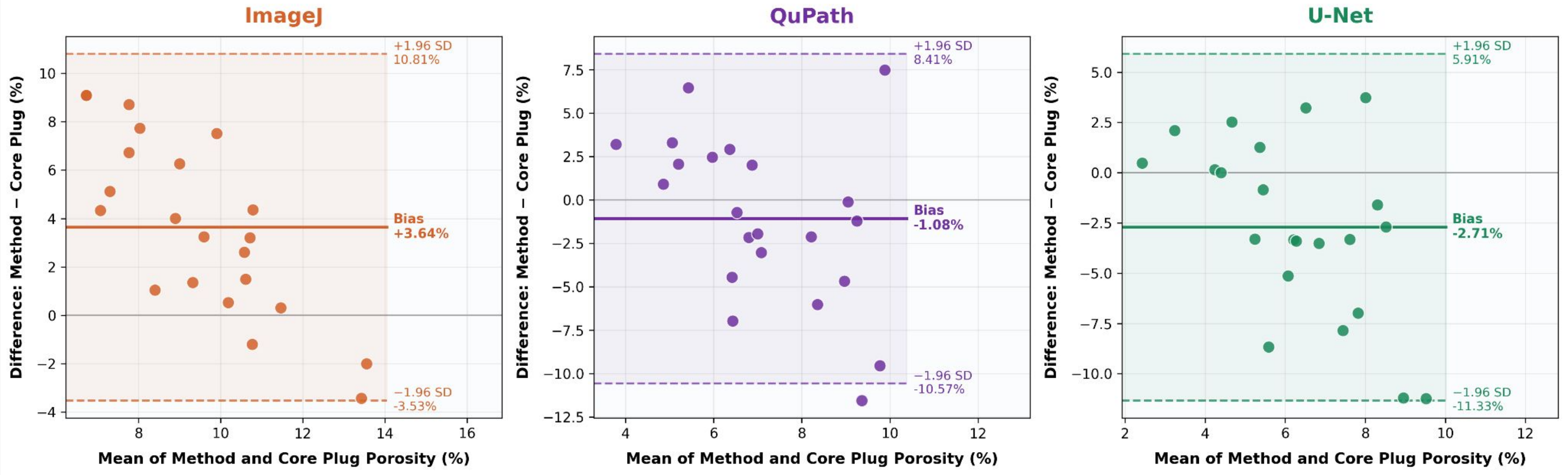
0.90 **Best Dice (sample 022)**

highest-performing sample

Method Agreement

Systematic Bias Patterns — Bland-Altman Analysis

Bland-Altman plots showing agreement between each image-based method and laboratory porosity. Solid lines = mean bias. Dashed lines = ±1.96 SD limits of agreement.



Zero-bias line remains within ±1.96 SD limits for all three method, no statistically fixed offset from helium porosimetry confirmed.

Bland & Altman, 1986

Discussion

Image-Based vs. Laboratory Porosity: Expected Discrepancies

Photography encouraged

Measurement Principle Comparison

Aspect	Lab. Porosity	Thin Section Image
Dimension	3D Volumetric	2D areal only
Sample volume	~10-40cm3 (core plug)	~0.036 cm3 (thin section)
Pore size range	All pores down to nano-scale	Resolution-limited
Measurement	Physical - Boyle's Law	Optical - pixel segmentation
Key Limitation	No spatial information	2D sampling bias

Discussion

What do the Segmentation Metrics tell us ?

Photography encouraged

U-Net predictions benchmarked pixel-wise against QuPath reference masks (n=26)

0.86
Mean Dice
strong segmentation fidelity

92%
Samples porosity error < 1%
vs QuPath reference · MAE 0.39%

0.90
Best Dice
Sample 022 · highest-performing

High Dice (0.86 ± 0.04) — what it means

86% pixel-level overlap on average, strong for a model trained on only 26 samples with 5-fold cross-validation.

Lowest performer — Sample 017 (Dice = 0.69)

Finest grain size (52 µm, near silt–sand boundary) — unusual texture the model saw less of during training.

Recall (0.88) > Precision (0.83) — slight over-detection

Missing pores is worse than over-detecting, acceptable trade-off for porosity estimation.

Important point — metrics reflect QuPath, not ground truth (Lab.)

MAE vs helium porosimetry (3.85%) is the more independent assessment of accuracy.

Practical Implications & Limitations

Photography encouraged

	ImageJ	QuPath	U-Net
Best for:	Quick overview · no setup Large automated pipelines	Small datasets · immediate results needed	Large datasets · scalable workflows
Time/sample:	~5 sec / sample	~2–3 min / sample	Seconds (post-training)
Scalability:	High	Low	High
Limitation:	No sensitivity to sample variation	Subjective · operator dependent	Needs training data · single well tested

- All three methods show comparable MAE against lab porosity (3.85–4.25%)
- ImageJ overestimates (+3.64%) — Otsu threshold converges across samples
- QuPath achieves smallest bias (−1.08%) but requires manual per-sample input
- U-Net matches QuPath accuracy (MAE 3.85%) with no per-sample operator input — scalable

References

Bankhead, P., Loughrey, M. B., Fernández, J. A., Dombrowski, Y., McArt, D. G., Dunne, P. D., McQuaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., James, J. A., Salto-Tellez, M., & Hamilton, P. W. (2017). QuPath: Open source software for digital pathology image analysis. *Scientific Reports*, 7, 16878. https://doi.org/10.1038/s41598-017-17204-5

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307–310. https://doi.org/10.1016/S0140-6736(86)90837-8

Lithuanian Geological Survey. (2014). Vilkyčiai-22 well report DK-4100-181/2014-07. Unpublished internal report.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 9351, 234–241. https://doi.org/10.1007/978-3-319-24574-4_28