Related Work and State-of-the-Art: Ink Detection from Herculaneum CT

Context. Virtual unwrapping showed that volumetric X-ray CT plus computation can recover text without opening fragile scrolls. The En-Gedi study established the first fully non-invasive reading using metal-based ink with strong CT contrast [1]. For Herculaneum, carbon-based ink is near-isodense with papyrus. Parker et al. overturned the “what you see is what you get” assumption by showing that ink induces subtle **micro-structural**cues—surface smoothing, thickness change, crack morphology—detectable by learning rather than raw intensity [2]. To make progress reproducible, Parsons et al. released **EduceLab-Scrolls**, aligning micro-CT fragments with infrared ink maps and formalizing benchmark tasks, metrics, and a public challenge framework (Vesuvius Challenge) [3].

Methods and architectures. Early baselines used 3-D CNN/U-Net encoders trained on small CT patches to predict ink probability on the surface. Two architectural lines dominate. (i) **Frequency-augmented CNNs.** Chi et al. introduced **Fast Fourier Convolution (FFC)** to mix local spatial and global frequency reasoning efficiently [6], underpinned by the FFT of Cooley–Tukey [8]. Quattrini et al. extended FFC to **volumetric FFC (vFFC)** inside a U-Net-like 3-D network, improving detection of pseudo-periodic papyrus textures and faint ink signatures; code and protocol made it a strong, reproducible baseline [4]. (ii) **Residual/attention hybrids.** Modern variants adopt ResNet-style encoders for stable deep training [10] and V-Net/Dice-loss formulations for severe class imbalance [7]. Recent academic follow-ups report hybrid 3-D feature extractors with 2-D refinement heads and depth-invariance training, surpassing vFFC on the fragment-1 benchmark.

Datasets used and why. The field standard is **EduceLab-Scrolls fragments**: CT volumes with **IR-verified 2-D ink labels** (train on fragments 2–3, evaluate on fragment 1). This pairing uniquely enables **supervised** learning for carbon-ink in

CT. Classical virtual-unwrapping works provide geometric context for sheet extraction and flattening but lack supervised ink labels [5,9].

Benchmarks and evaluation; limitations. Common metrics are **$F\beta$ with β = 0.5**(precision-weighted), **pseudo-F-measure (pFM)**, and **PSNR**. F0.5 penalizes false positives strongly, matching curatorial needs; pFM reflects binarization quality; PSNR offers signal similarity. Limitations: $F\beta$ depends on threshold calibration and class balance; PSNR is weakly perceptual for sparse strokes; leaderboard scores can be swayed by ensembling and test-time augmentation rather than core model quality.

Current shortcomings. (1) **Depth misalignment / sheet-switching** lowers recall when ink lies off the expected surface slice. (2) **Over-mixing** in spectral bottlenecks can hallucinate fiber-parallel false positives. (3) **Weak geometry priors**: few models inject sheetness, local thickness, or surface-normal cues known to correlate with ink-altered micro-structure. (4) **Ensemble dependence**: top results often rely on many models; single-model clarity and ablations are under-reported.

How works relate. En-Gedi proved feasibility with metal ink [1]. Parker et al. established **carbon-ink detectability via structure**, motivating ML for Herculaneum [2]. Parsons et al. supplied data, labels, and metrics to make results verifiable [3]. vFFC added **global frequency reasoning** to 3-D CNNs and set a reproducible benchmark [4,6,8]. Later hybrids leverage **ResNet/V-Net backbones, Dice-based losses, depth-invariance augmentation, and 2-D refinement heads** to close the remaining gaps [7,10].

How this project builds and differs. This work isolates **ink detection** on fragments and targets **single-model** improvements over vFFC without large ensembles. It injects **geometry-aware inputs** (sheetness, thickness, normals), applies **depth-invariance training** and **calibrated thresholds**, and reports **clean ablations** to attribute gains.

Research questions. RQ1. Do **geometry-aware channels** combined with a vFFC-style network improve F0.5 on fragment 1 versus the vanilla vFFC baseline? **RQ2.** Do **depth-roll augmentation** and **per-fragment threshold calibration** reduce fiber-parallel false positives without harming recall on fragment 1?

References

[1] W. B. Seales, C. S. Parker, M. Segal, E. Tov, P. Shor, and Y. Porath, “From damage to discovery via virtual unwrapping: Reading the scroll from En-Gedi,” *Science Advances*, vol. 2, no. 9, e1601247, 2016.

[2] C. S. Parker, S. Parsons, J. Bandy, C. Chapman, F. Coppens, and W. B. Seales, “From invisibility to readability: Recovering the ink of Herculaneum,” *PLOS ONE*, vol. 14, no. 5, e0215775, 2019.

[3] S. Parsons, C. S. Parker, C. Chapman, M. Hayashida, and W. B. Seales, “EduceLab-Scrolls: Verifiable recovery of text from Herculaneum papyri using X-ray CT,” arXiv:2304.02084, 2023.

[4] F. Quattrini, V. Pippi, S. Cascianelli, and R. Cucchiara, “Volumetric Fast Fourier Convolution for Detecting Ink on the Carbonized Herculaneum Papyri,” in *Proc. ICCV 2023 Workshops (e-Heritage)*, 2023.

[5] O. Samko, Y.-K. Lai, D. Marshall, and P. L. Rosin, “Virtual unrolling and information recovery from scanned scrolled historical documents,” *Pattern Recognition*, vol. 47, no. 1, pp. 248–259, 2014.

[6] L. Chi, B. Jiang, and Y. Mu, “Fast Fourier Convolution,” in *Advances in Neural Information Processing Systems 33 (NeurIPS 2020)*, pp. 4479–4488. [7] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation,” in *Proc. 3DV*, pp. 565–571, 2016.

[8] J. W. Cooley and J. W. Tukey, “An algorithm for the machine calculation of complex Fourier series,” *Mathematics of Computation*, vol. 19, no. 90, pp. 297–301, 1965.

[9] P. L. Rosin, Y.-K. Lai, C. Liu, G. R. Davis, D. Mills, G. Tuson, and Y. Russell, “Virtual recovery of content from X-ray micro-tomography scans of damaged

historic scrolls," *Scientific Reports*, vol. 8, no. 1, p. 11901, 2018.
[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. CVPR*, pp. 770–778, 2016.