赫库兰尼姆纸草文献表面地形中的墨迹检测

Giorgio Angelot ti[1, *], **Federica Nicolardi**[1,2], **Paul Henderson**[1,3], **和 W. Brent Seales**[1, 4]

[1]维苏威挑战赛，美国[2]那不勒斯费德里科二世大学，意大利[3]格拉斯哥大学，苏格兰，英国[4]EduceLab，肯塔基大学，美国[*]
giorgio@scrollprize.org

arXiv:2603.27698v1 [cs.CV] 29 Mar 2026

摘要

赫库兰尼姆纸草文献难以阅读，因为卷轴和碳基墨迹都已碳化。在X射线射线照相术和断层成像中，墨迹检测通常依赖密度驱动或成分驱动的对比，但碳化纸草纸上的碳墨几乎无法提供衰减对比。基于形态学假设，我们表明，书写区域的表面形态包含足够的信号，能够将墨迹与纸草纸区分开来。为此，我们利用机械打开的赫库兰尼姆纸草文献的三维光学轮廓测量数据训练机器学习模型，以区分有墨迹和无墨迹区域。我们进一步量化了横向采样如何影响可学习性，以及原生分辨率模型在粗化输入上的表现。我们表明，仅凭高分辨率地形就包含可用于墨迹检测的信号。随着横向分辨率降低，分割性能随之下降，这揭示了要利用数据集中的形态信号，必须解析的特征空间尺度。这些发现有助于确定通过X射线断层成像读取封闭卷轴时，基于形态学的方法所需的空间分辨率目标。

引言

赫库兰尼姆纸草文献是数百卷碳化书籍，呈卷轴形态，于十八世纪下半叶在罗马城市赫库兰尼姆的遗址中被发现。公元79年，维苏威火山喷发，这座城市因此遭到毁灭并被掩埋。这批文献在碳化后又被火山灰掩埋，历经近两千年，是希腊罗马古代唯一留存下来的图书馆[1,2]。近三个世纪以来，尽管人们采用了多种方法打开和展开纸草文献，仍有数百卷未能释读[3,4]。超过800件已登记的藏品——包括完整卷轴和卷轴残片——仍处于卷曲状态；它们极其脆弱，因此要以非侵入方式读取，难度极大[5,6]。近期研究提出了针对密封卷轴X射线计算机断层成像（CT）扫描的虚拟展开流程[7–11]。这些方法通过几何方式展平纸草纸层，并结合能够产生墨迹对比的成像技术，从封闭卷轴中提取文字[7,8]。然而，对于碳化纸草纸上的碳墨，吸收对比本质上较弱，因此需要采用相衬断层成像和X射线荧光等先进成像手段[12,13,15]。即使是那些尽管脆弱、但能够机械展开的卷轴，阅读赫库兰尼姆纸草文献也颇具挑战。由于碳墨难以在可见光谱中辨识，红外成像等互补的表面敏感方法[16–18] 以及脉冲热成像近期已在赫库兰尼姆残片上显现出肉眼几乎不可见的文字，证实表面化学性质能够为碳墨提供可用信号[19]。这些技术共同展现出良好前景，可增强微弱的、基于密度的墨迹信号，从而提高开放式托盘中的文献以及封闭卷轴X射线CT成像结果的可读性。

与此同时，文化遗产研究日益采用光学轮廓测量和微观轮廓测量，以亚微米级横向采样和纳米级垂直重复性来表征表面地形[20]。轮廓测量已广泛应用于纸张科学和文件司法鉴定领域，在这些领域中，微地形记录了书写过程和压痕笔画[21,22]。这些观察结果为炭化纸草纸提供了一种互补的假设，通常称为形态学假设[14]，：墨水沉积会充分改变局部微地形，使纯形态学 信号能够区分有墨迹和无墨迹区域，而不受成分影响。与这一观点一致，对史密森学会卷轴的初步扫描电子显微镜观察显示，墨迹与基底之间的表面粗糙度存在差异，墨迹区域看起来更为光滑[13]。

在本研究中，我们在机械展开的炭化赫库兰尼姆纸草文献碎片上检验这一假设。我们使用处于共聚焦模式的三维光学轮廓仪（Sensofar S lynx 2），对单个字母采集经过校准的高度图。我们对高度图进行归一化并将其转换为图像，再使用与明场照片对齐的手工绘制标签训练深度学习分割模型，以区分墨迹和纸莎草纸。我们探讨三个问题：（i）仅凭地形，在高横向采样率下，是否足以检测墨迹？（ii）当相同的

样本被合成降采样至更粗的像素尺寸时，以及高分辨率模型在上采样后的退化输入上进行评估时，表观性能会如何变化？（iii）在一些带有墨迹的纸草文献上学到的形态学信号，能否泛化到未见过的样本？具体而言，本研究的贡献如下：

- **数据集**。 一套经过标注的光学轮廓测量数据集，涵盖 PHerc. 248、PHerc. 250 和 PHerc. 500P2 中的 14 个区域，共包含 16 个单独字母；采集时横向采样间距为 0.34 μm，垂直灵敏度为 8nm。
- **仅基于地形的分割**。 研究表明，仅凭表面地形即可在原生横向采样间距（0.34μm）下实现准确的墨迹分割，留出数据的 Dice 分数约为 0.88–0.90。
- **分辨率–性能趋势**。 在该数据集上，采样间距越大，性能呈单调下降趋势：中位数 Dice 分数从 0.34 μm 时的 $0.890\,[0.864, 0.905]$ 降至 10.88 μm 时的 $0.467\,[0.364, 0.524]$ （Friedman $\chi^2 = 88.34$，$p = 1.01 \times 10^{-15}$；Page's $L = 3848$，单侧 $p = 1.0 \times 10^{-4}$）。这一趋势仅反映我们的数据集，并不意味着它是普遍适用的界限。
- **跨分辨率推断**。 在 0.34 μm 分辨率下训练的模型，在上采样的粗分辨率输入上的性能会下降；像素尺寸达到 $\leq 2.04\,\mu$m 时，Dice系数缓慢下降，但到 $\geq 3.40\,\mu$m 时则骤降至接近零。这表明，在精细特征上训练的模型在像素尺寸达到 2.04μm 之前仍能捕捉到一些信号，但此后不久，这些信号便会完全消失。
- **各向同性体素模拟**。 对垂直离散化高度图（$\Delta z = \Delta x$）上的匹配分辨率模型进行评估，所得 Dice值与连续高度设置下的结果相当，这表明垂直量化并非该数据集的主要瓶颈。
- **跨纸草纸泛化**。 在原生采样分辨率（0.34μm）下进行留一纸草纸评估，以量化模型在不同手稿间的迁移效果。汇总各数据划分中的留出样本后，均值 Dice系数为 0.691；但各纸草文献的表现存在差异（表2）。

方法

样本。 我们分析了三份炭化纸草纸的残片：PHerc. 248、PHerc. 250 和 PHerc. 500P2。PHerc. 248（斐洛德摩斯，《论虔敬》）和 PHerc. 250（斐洛德摩斯，《论修辞学》I）是两份外层纸草纸现存的纸页。这些外层纸草纸是卷轴最外侧的部分，于 18 世纪 50 年代经纵向切割取下，并在 19 世纪上半叶被逐层刮除。PHerc. 500P2 是编入 PHerc. 500 号名下的散片之一，目前存放于*cassetto*28 中。我们根据字迹清晰度和物理损坏程度较轻的标准，从这些残片中选取了 16 个字母。每个字母区域的边长约为 1.5 毫米。

数据采集。 使用在共聚焦模式下运行的 Sensofar S lynx 2 光学轮廓仪测量表面地形。原生分辨率下的横向采样为 0.34μm，垂直灵敏度为 8 nm。针对每个字母，均采集了一张配准明场照片，以辅助人工标注。图1展示了 PHerc. 250 中的字母示例。预处理与合成下采样。原始轮廓测量导出数据偶尔会在共聚焦重建失败的像素处包含非有限高度值（NaN/Inf）。在进行任何重采样或归一化之前，我们记录二值有效性掩码 $M(x) = \mathbb{1}[\text{finite}(z(x))]$，并填充缺失像素，以获得适用于 nnU-Net 训练的稠密图。我们仅使用 Telea[25] 快速行进法（OpenCV 实现；半径为 $= 3$ 像素）对非有限像素进行图像修复，填补缺失值。图像修复在临时的 8位归一化图上进行，采用稳健线性缩放（0.5–99.5 百分位数），然后映射回原始高度范围。为确保测量值不变，图像修复后将所有原本有限的像素精确恢复。随后，将高度图（原生横向采样为 0.34 μm）归一化至范围 $[0,1]$，并从 64位浮点数转换为 16位无符号整数。为量化分割性能对横向采样的依赖程度，并模拟封闭卷轴体积成像中可能遇到的较粗有效采样，我们对每幅高度图进行合成下采样。通过使用尺寸为 $n \times n$ 像素的方形核进行分块平均，模拟较粗的有效横向分辨率，其中 n 的取值为 2 至 32。随后，使用双线性插值将下采样后的图重新上采样至原始网格，生成用于跨分辨率推断的成对输入。在补充的“各向同性体素”模拟中，我们还将高度值按物理单位（μm）离散化为均匀分箱，其宽度与各分辨率下的有效横向像素尺寸相匹配（$\Delta z = \Delta x$）。这些操作（下采样、上采样和 z 分箱）均在逐样本归一化和 uint16 转换之前执行。我们未进行平面扣除或倾斜校正。

标注与数据划分。 我们使用 GIMP 在对齐的明场照片上手动描绘二值墨迹掩码。16 个字母被划分为 14 个独立样本。大多数样本只包含一个字母。在所有实验中，我们对这 14 个样本进行了五折交叉验证。每一折中，模型使用约 80% 的样本进行训练，并在留出的 20% 样本上进行评估；随后汇总结果，确保每个样本在每种条件下恰好贡献一个留出评分。在匹配分辨率实验中，我们针对每种横向采样分别训练了独立的 2D nnU-Net 模型。对于跨分辨率

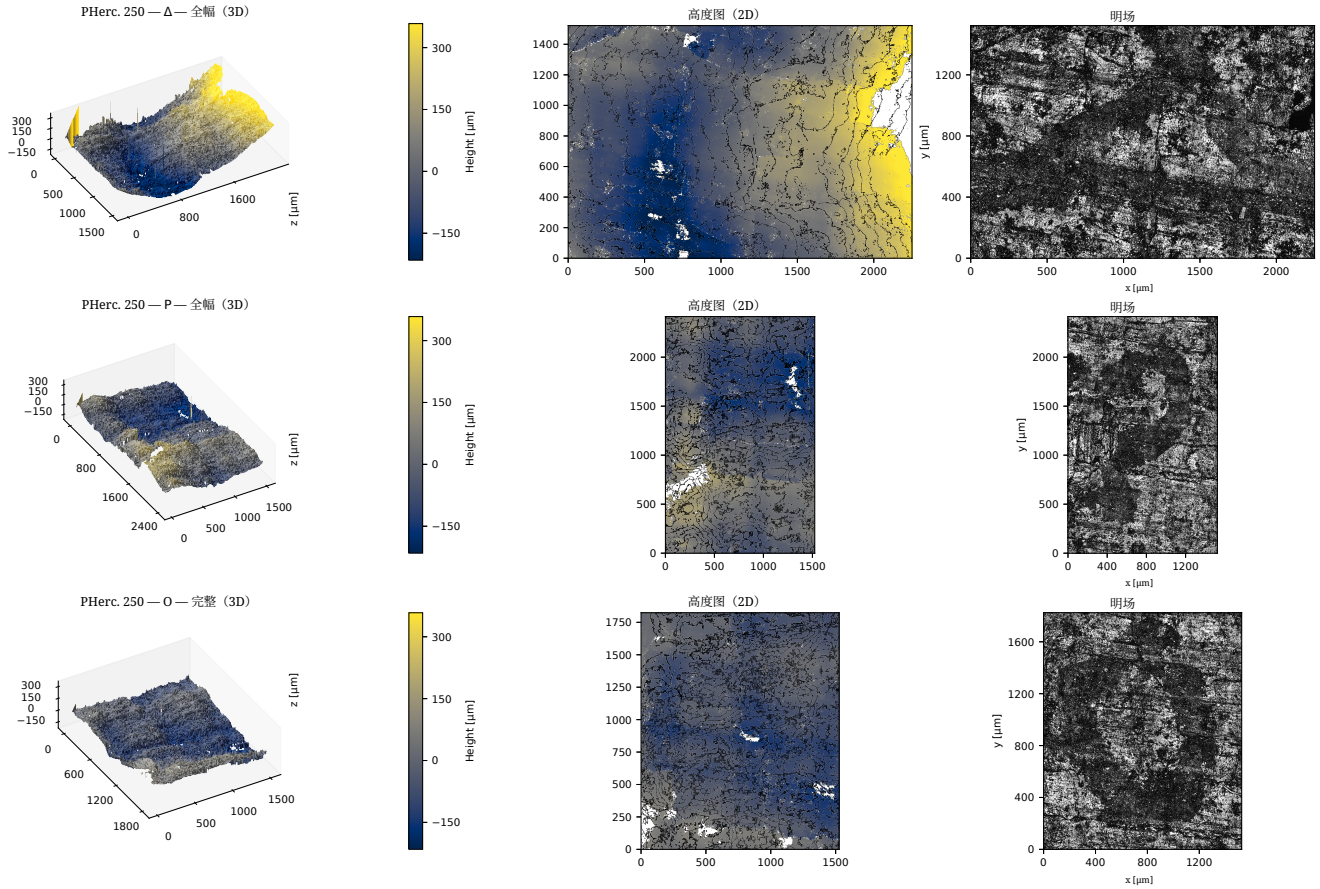

图1. PHerc. 250 中三个样本的代表性轮廓测量概述。每一行展示一个不同的字母。从左到右依次为：(i) 测得地形的完整三维渲染；(ii) 带有地形等高线的二维高度图；(iii) 与之对齐且可见墨迹的明场图像。坐标轴单位为 µm；每行中的高度色标采用相同的比例尺。

推断，我们使用在 0.34µm 下训练的模型，对留出输入的降质版本进行评估。对于 z 分箱，我们使用基于连续高度图训练的相同逐折模型，对留出输入经过 z 分箱处理的版本进行评估。为评估跨纸草纸泛化，我们还在 0.34µm 下进行了留一纸草纸评估：每次选取一份纸莎草纸，用另外两份纸莎草纸的所有样本进行训练，并在留出的纸莎草纸的所有样本上进行评估。

模型与训练。 我们使用了 nnU-Net 框架[23] 采用类 U-Net 编码器–解码器架构和默认超参数。训练使用 Dice 加交叉熵损失、Adam 优化器，学习率为 10^{-3}，并采用随机翻转、旋转、强度缩放和弹性形变进行数据增强。模型通过提前停止训练，在验证损失收敛时停止。

评估与统计。 使用留出样本上的 Dice 系数对分割性能进行量化。对于匹配分辨率模型，我们报告了 14 个留出样本的 Dice 分数，并对交叉验证折的结果进行了汇总。分析针对不同像素尺寸下相互关联的样本进行。由于无法假设数据符合正态分布，我们采用了非参数检验：Friedman 检验（总体检验，双尾）和 Page L 检验（有序趋势，单尾）。除非另有说明，我们采用显著性阈值 $\alpha = 0.05$。事后两两比较采用 Wilcoxon 符号秩检验（双尾），并使用 Holm 校正控制所有像素尺寸对比的族错误率。我们使用中位数（[第一四分位数，Q3]）和均值±标准差汇总数据。为评估缺失值模式是否会干扰学习，我们还量化了墨迹和纸莎草纸区域内原始（填补前）高度图中非有限像素的比例，并使用 Wilcoxon 符号秩检验比较配对比例；此外，我们报告缺失掩膜与墨迹标注之间的 Dice 重叠，作为简单的仅掩膜基线。

结果

概述

我们总共分析了 16 个字母，分属三份纸莎草纸中的 14 个样本。分割实验的主要结果见图 2 和 3、表 1，以及表 2 中的留一纸草文献泛化汇总。高分辨率地形支持准确的墨迹分割，留出样本的 Dice 分数中位数约为 0.89。通过分块平均合成降低横向采样精度时，性能呈单调下降；使用原生采样训练的模型仅在有效像素尺寸约为 1µm 以内时仍保持合理的准确度。将较粗分辨率的输入交给高分辨率模型时，横向采样精度达到 1.36µm 之前，性能仍然较高，但在更粗的分辨率下则有所下降。为便于直观参考，我们展示 Dice $= 0.70$，将其作为分割中常用的参考重叠阈值[24]。在原生采样下，留一纸草纸评估在所有留出样本上的合并平均Dice为 0.691，但不同纸莎草纸之间存在显著差异（表 2）。

定性示例

图2 展示了三个有代表性的字母： ε 来自 PHerc. 248（赫库兰尼姆纸草文献编号）， κ 来自 PHerc. 250（赫库兰尼姆纸草文献编号），以及 η 来自 PHerc. 500P2。从左至右，各列依次展示转换后的地形16位无符号整数图像、预测的墨迹掩膜，以及配准明场照片。地形图像中无法直接看到墨迹。在完整数据集中，高度图似乎主要受表面的倾斜度和曲率，以及构成纸莎草纸页的纤维形状影响；表面并非完全平坦。事实上，表面可能存在与纸莎草纤维交错分布的凹谷或凸丘。尽管如此，训练后的机器学习模型仍能紧密追踪笔画内部，并在高度起伏趋近背景微观粗糙度的羽化边界处逐渐变细。

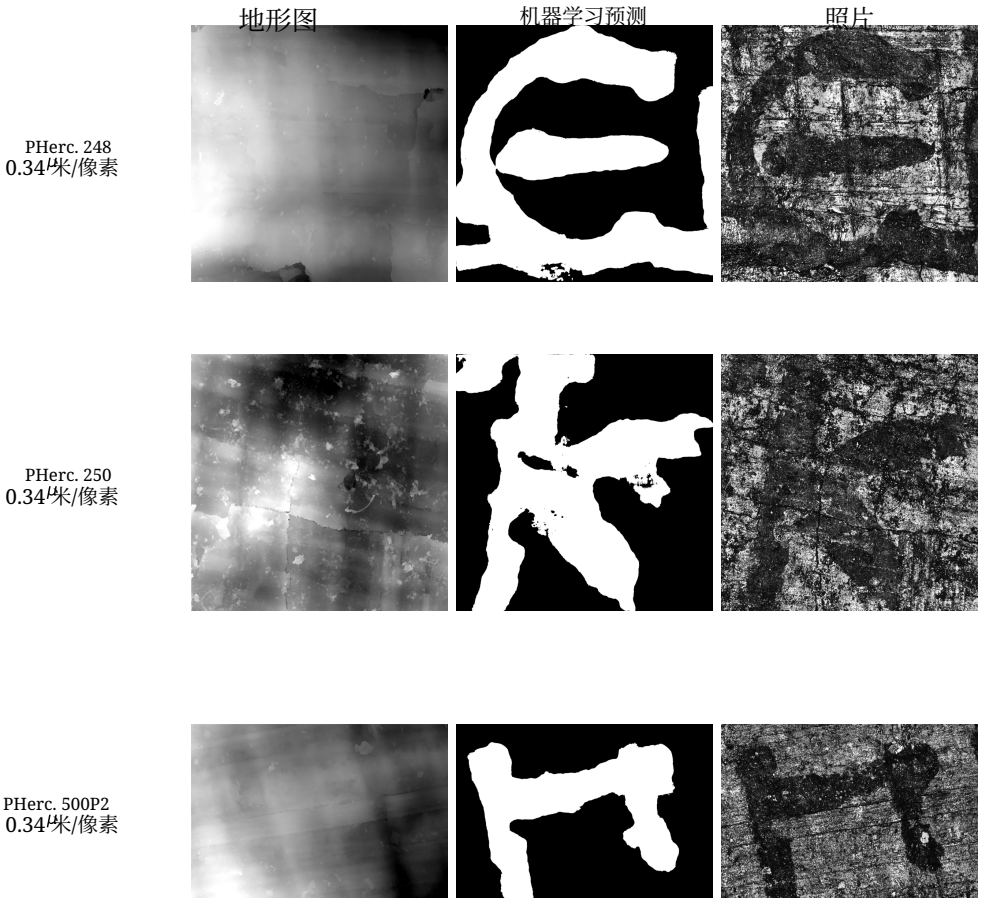

图2。三个纸草文献中的定性示例。从上到下各行依次为 ε PHerc.248（赫库兰尼姆纸草文献编号）上的字母、 κ PHerc.250（赫库兰尼姆纸草文献编号）上的字母，以及 η PHerc. 500P2上的字母。由左至右各列依次为：（i）转换为uint16图像的地形图，（ii）nnU-Net模型的预测结果，以及（iii）明场照片。高度图中可见的网格结构是构成纸莎草基底的纤维格栅。

仅从地形中学习。在 0.34 µm 的原生采样下，各折的训练均平稳收敛。留出集的 Dice

分数集中在 0.88–0.90 之间，中位数为 $0.890\,[0.864, 0.905]$，支持了这一核心论点：仅凭表面形态即可为墨迹提供判别信号，并验证了形态学假设。

缺失数据伪影对照。 在原生采样的$n = 14$ 个样本中，原始（填补前）高度图仅含少量非有限像素（中位数为 1.665% [1.250% 2.352%]）。与纸莎草纸区域相比，墨迹区域的缺失率并未系统性升高（墨迹：1.653% [1.443% 2.357%]；纸莎草纸：1.768% [1.075% 2.379%]；配对 Wilcoxon 符号秩检验：$W = 52$，双侧 $p = 1.0$）。缺失掩膜与墨迹标注的直接重叠可以忽略不计（Dice系数$= 0.032\,[0.028, 0.044]$），这表明在该数据集中，非有限像素的丢失模式并未与墨迹区域简单重合。

匹配训练与测试条件下的分辨率敏感性。图3中的橙色方框以及表1 的第一列分区总结了横向采样逐渐变粗时的性能表现。随着像素尺寸增大，Dice系数总体呈下降趋势，尽管不同分辨率之间存在轻微的非单调差异。Friedman 检验结果高度显著（$\chi^2 = 88.34$，$p = 1.01 \times 10^{-15}$），Page's L 证实了有序下降趋势（$L = 3848$，单侧 $p = 1.0 \times 10^{-4}$）。涉及 0.34 µm 的 Wilcoxon 符号秩两两比较经 Holm校正后仍具有显著性（校正后 $p \leq 0.0044$），相对于 0 -0.086，Dice系数的配对变化中位数范围为 -0.417 至。经校正后，约 2 µm 及以上的像素尺寸之间无法区分（校正后 $p \geq 0.245$）。我们将这些视为数据集层面的观察，而非普适阈值，因为表观界限可能随数据集规模、多样性和机器学习模型架构而变化。

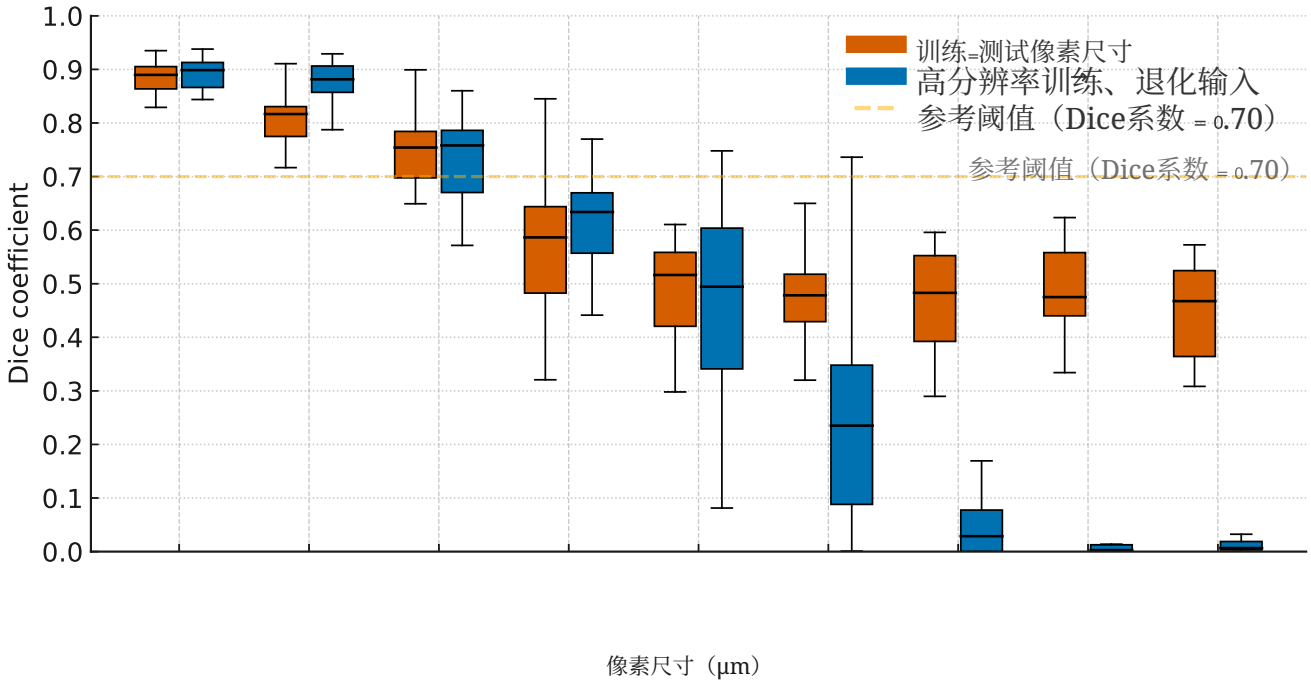

图3. 分割性能（Dice系数）随有效横向采样（像素尺寸，µm）的变化。橙色箱线图：在匹配分辨率下训练和评估的 nnU-Net 模型。蓝色箱线图：在 0.34 µm 下训练的单个 nnU-Net，在降采样至较粗像素尺寸并上采样至原生网格的输入上进行评估。箱体表示 $n = 14$ 个案例的中位数和四分位距；须线延伸至最极端的非离群值（Tukey，1.5 ×IQR；省略离群值）。水平虚线表示 Dice $= 0.70$ 参考重叠阈值。

通过 z 分箱模拟各向同性体素。 为近似具有各向同性体素的体积扫描，我们将高度值离散为多个区间，每个区间的宽度等于各分辨率下的有效横向像素尺寸（$\Delta z = \Delta x$）。表 1 最右侧列组汇总了匹配分辨率模型在经过 z 分箱的留出测试输入上的评估性能。不同分辨率下，z 分箱得到的 Dice值与连续高度设置（最左侧列组）接近，中位 Dice系数的绝对变化最多为 0.026。随着像素尺寸增大，性能下降的趋势依然存在。

固定高分辨率模型的跨分辨率泛化。我们在 0.34 µm 分辨率下训练了一个二维 nnU-Net 模型；测试时，先通过块平均将输入退化到更粗的有效分辨率，再上采样回原生网格。图3 中的蓝色框以及表1 的第二列分区总结了结果。在 ∼1µm 时，性能仍然很高；到 2.04 µm 时，性能有所下降；而到 3.40 µm 及更粗时，性能则降至接近零。与训练和测试分辨率匹配的设置相比，性能下降更为陡峭，这与高频形态学信息的丢失以及分布偏移相符

表1. 三种设置下 Dice系数与横向像素尺寸的关系：(i) 像素尺寸敏感性（nnU-Net 在匹配分辨率下训练和评估）；(ii) 在 0.34μm 数据上训练的模型在退化输入上的评估；以及 (iii) 在 z 分箱高度图（$\Delta z = \Delta x$）上进行匹配分辨率评估，以模拟各向同性体素。数值为中位数 [第一四分位数、Q3] 和均值 ± 标准差。只有分辨率为 1.02μm 或更精细的模型超过 Dice系数参考阈值。

横向像素尺寸	像素尺寸敏感性（n=14）		在 0.34μm 下训练的模型		Z 分箱测试输入（$\Delta z = \Delta x$，n=14）	
	中位数 [第一四分位数、Q3]	均值 ± 标准差	中位数 [第一四分位数，Q3]	均值 ± 标准差	中位数 [第一四分位数，Q3]	均值 ±标准差
0。34 μm	0.890 [0.864，0.905]	0.881 ± 0.039	0.899 [0.866，0.913]	0.889 ± 0.038	0.864 [0.808，0.900]	0.839 ± 0.075
0.68 μm	0.817 [0.775，0.831]	0.803 ± 0.068	0.881 [0.857，0.906]	0.877 ± 0.042	0.817 [0.755, 0.843]	0.792 ± 0.085
1。02 μm	0.754 [0.698，0.784]	0.748 ± 0.068	0.758 [0.670，0.786]	0.736 ± 0.091	0.742 [0.701, 0.787]	0.744 ± 0.080
1。36 μm	0.586 [0.483, 0.644]	0.570 ± 0.143	0.634 [0.557，0.670]	0.611 ± 0.104	0.596 [0.507, 0.645]	0.586 ± 0.139
2.04 μm	0.516 [0.421, 0.559]	0.493 ± 0.093	0.494 [0.341，0.604]	0.455 ± 0.208	0.499 [0.451，0.541]	0.486 ± 0.083
2.72 μm	0.478 [0.429, 0.518]	0.480 ± 0.096	0.235 [0.088，0.348]	0.264 ± 0.234	0.486 [0.426，0.515]	0.473 ± 0.085
3.40 μm	0.483 [0.392, 0.552]	0.460 ± 0.106	0.029 [0.000，0.078]	0.053 ± 0.063	0.480 [0.381, 0.516]	0.451 ± 0.106
5.44 μm	0.475 [0.440，0.558]	0.482 ± 0.091	0.003 [0.000、0.013]	0.013 ± 0.019	0.474 [0.415, 0.534]	0.475 ± 0.084
10。88 μm	0.467 [0.364, 0.524]	0.450 ± 0.090	0.007 [0.003, 0.019]	0.018 ± 0.024	0.461 [0.397，0.534]	0.452 ± 0.091

这是由下采样—上采样流程引入的。统计结果为 $\chi^2 = 104.61$，$p = 4.86 \times 10^{-19}$(Friedman) 和 $L = 3947$，单侧 $p = 1.0 \times 10^{-4}$(Page 趋势检验)。

跨纸草纸泛化。为探究所学习的形态学线索是否能够跨手稿迁移，我们在原生采样（0.34μm）下进行了留一纸草纸评估。每次选取一份纸莎草纸，将另外两份纸莎草纸的全部样本用于训练 nnU-Net，并在留出的纸莎草纸上进行评估。汇总三个数据划分中的所有留出样本（$n = 14$）后，平均 Dice系数为 0.691。然而，各纸莎草纸的结果差异显著：留出 PHerc. 248（中位数 Dice系数为 0.757）或 PHerc. 250（中位数 Dice系数为 0.817）时，性能仍相对较高；而留出 PHerc. 500P2（中位数 Dice系数为 0.475）时，性能则明显下降，表明该数据集上的跨手稿迁移具有异质性（表2）。

表2。原生横向采样（0.34μm）下的留一纸草文献泛化。对于每份留出的纸草文献，使用另外两份纸草文献的全部样本训练 nnU-Net，并在留出纸草文献的全部样本上进行评估。数值为留出纸草文献样本的中位数 [第一四分位数、Q3] 和均值 ± 标准差。汇总所有留出样本（$n = 14$）后，平均 Dice系数为 0.691。

留出纸莎草纸	n	中位数 [第一四分位数，Q3]	均值 ±标准差
PHerc. 248	5	0.757 [0.733，0.812]	0.777 ± 0.059
PHerc. 250	5	0.817 [0.755，0.822]	0.789 ± 0.053
PHerc. 500P2	4	0.475 [0.454，0.482]	0.461 ± 0.035

讨论

我们的结果表明，高分辨率表面形态包含足够的信息，能够准确分割炭化纸草纸上的碳墨。此外，在我们的实验设置中，横向采样降低时，性能会呈现系统性下降。表1 中一个值得注意的现象是：在原生 0.34 μm 采样条件下训练、并在经过模糊处理且上采样至 0.68 μm 的输入上评估的模型，其 Dice系数高于在匹配的粗分辨率下原生训练的模型。这种反直觉的现象与计算机视觉文献中的报告相呼应：当图像在输入端或网络早期进行统一下采样时，小特征可能会在模型学会检测它们之前消失[26,27]。在高分辨率数据上训练，能够使网络学习全尺度的形态学线索，这些线索在适度模糊后仍可加以利用。相比之下，在粗分辨率输入上原生训练，会迫使编码器处理已经经过低通滤波的信号，从而导致分割效果较差。然而，当有效像素尺寸超过 ∼1μm 时，两种策略的效果都会崩溃，这与小目标检测理论的观点一致，即下采样会造成信息损失[28]。这些考量凸显了在整个处理流程中保留精细结构的必要性，并提醒我们避免过度下采样。随着像素尺寸增大，分割性能系统性下降，这表明判别性形态信号集中在精细空间尺度上，并会随着分辨率变粗而逐渐丢失。在地形图中，这些精细尺度的振荡对应于表面粗糙度；然而，仅通过滤波提取粗糙度，并不能可靠地区分带墨迹的字母与基底。这表明卷积模型利用的是一种复合信号：表面粗糙度，以及墨水沉积造成并编码于形状中的压力诱导变形

模式，而经典轮廓分析算法并不能轻易捕捉这些模式。

我们还进行了补充实验，将垂直高度值离散为与横向像素同样大的区间，以模拟各向同性体素采样，类似于体积X射线断层成像可能实现的效果。表1 中的 z 轴分箱列显示，在我们的数据集中，这种 z 轴量化仅使 Dice系数发生小幅变化（不同分辨率下的最大中位数偏移为 0.026），且总体上仍保持采样越粗、性能越低的趋势。这些观察结果表明，有效各向同性分辨率达到或优于 1μm 的体积扫描，可能足以用于基于形态的墨迹检测。

本研究受到以下局限性的影响。我们的数据集规模较小（来自三份纸莎草纸的十四个样本中，共有十六个字母），因此统计功效有限；在最粗的采样条件下，方差增大，这一问题尤为明显。尽管我们报告了 0.34 μm 条件下的留一纸草文献泛化结果（表2），但每份纸草文献仅贡献 4–5 个样本；因此，跨纸草文献的估计结果尚属初步，且可能容易受到样本选择和保存状态的影响。真实标注掩膜由配准的明场照片生成；这可能在羽化墨迹边缘引入偏差，并使配准误差传递到评估中。通过分块平均进行合成下采样虽能近似真实仪器的点扩散和调制传递函数，却无法完全复现它们，因此，我们的趋势应被视为定性结果，而非普遍规律。我们仅采用了一种编码器–解码器架构和默认训练方案；其他损失函数、感受野尺寸或多尺度设计可能影响粗采样条件下的性能。尽管存在这些局限性，观察到的像素尺寸增大时性能单调下降的现象，以及高分辨率模型在输入大幅退化时评估所呈现的急剧性能下降，仍为未来实验设计提供了重要指导。最后，我们评估了共聚焦重建缺失是否可能解释分割性能。非有限像素很少（∼1–2%），且与墨迹没有系统性关联；缺失掩膜与墨迹标注的重叠也可以忽略不计（Dice系数 ≈ 0.03）。这表明，简单的缺失值模式不太可能解释观察到的基于形貌的分割结果；然而，与表面光学特性相关的系统性高度重建偏差仍是一个潜在混杂因素，未来研究应对此进行检验。

结论与展望

采用亚微米级横向采样的光学轮廓测量，仅凭表面形态即可准确分割炭化纸草纸上的碳墨。在我们的数据集中，随着横向采样变粗，准确率随之下降：在匹配分辨率下训练和测试的模型性能有所降低，而单个高分辨率模型在上采样的粗分辨率输入上进行评估时，性能下降得更为显著。在我们的数据集中，要实现可靠的分割，横向采样需达到约 1μm 或更精细的水平。具体要求取决于数据的数量和多样性、信噪比、模型架构，以及至关重要的墨迹数量和保存状态（例如，不同书写者的沉积材料存在差异，以及长期暴露的已展开残片上可能发生磨蚀）。

如果对封闭卷轴进行体积成像时，能够在层表面实现相当的有效分辨率，并保留类似的微地形和信噪比，那么在密封卷轴内部进行仅基于形态的检测便有望实现。我们强调的是“有效”分辨率，而非名义像素尺寸，因为点扩散、部分容积效应、散射、重建正则化以及其他伪影都可能降低可分辨的细节。尽管存在这些限制，当前结果通过量化有效表面分辨率如何影响仅基于地形的墨迹可检测性，进一步完善并拓展了基于形态辅助阅读的研究。这些结果也为未来的成像项目提供了依据：在密封卷轴的体积扫描中，应保留纸莎草纸的微地形，并以更高的有效分辨率为目标，从而检测出仍然存在的基于形态的线索。

数据可用性

数据样本、标签、折分组和配置文件将在发表后发布于公共存储库。在此之前，可应同行评审请求向通讯作者索取。

参考文献

1. Sider, D. 赫库兰尼姆帕皮里别墅的图书馆.（盖蒂出版公司，洛杉矶，2005年）。**2.** Longo Auricchio, F., Indelli, G., Leone, G., Del Mastro, G. 《帕皮里别墅：一座古代宅邸及其图书馆》。（Carocci，罗马，2020年）。**3.** Seales, W. B. and Delattre, D. 碳化赫库兰尼姆卷轴的虚拟展开：研究现状（2007—2012年）。《赫库兰尼姆纪事》**43**，191–208（2013年）。**4.** Nicolardi, F. 关于尚未展开的赫库兰尼姆卷轴的一些思考，展望未来的虚拟展开。《雅典与罗马》**18,1-4**，121–130（2024年）。doi:10.7347/AR-2024-p1215. Mocella, V., Brun, E., Ferrero, C. and Delattre, D. 通过X射线相衬成像揭示卷曲的赫库兰尼姆纸莎草文献中的字母。《自然·通讯》**6**，5895（2015年）。doi:10.1038/ncomms6895。

6. Tack, P. 等。通过基于X射线的技术和蒙特卡罗模拟追踪赫库兰尼姆纸草卷上的墨水成分：铅的定量与形态分析。《科学报告》6, 20763（2016年）。doi:10.1038/srep20763。7. Seales, W. B. 等。借助虚拟展开从损毁走向发现：阅读恩盖迪卷轴。《科学进展》2, e1601247（2016年）。doi:10.1126/sciadv.1601247。8. Bukreeva, I. 等。通过X射线相位衬度断层成像虚拟展开并解读赫库兰尼姆纸草文献。《科学报告》6, 27227（2016年）。doi:10.1038/srep27227。9. Nicolardi, F., Parsons, S. 等。从一卷仍未展开的赫库兰尼姆纸草卷中揭示文本（PHerc.Paris. 4）。《纸草学与铭文学杂志》229, 1–13（2024年）。10. Nicolardi, F., Delattre, D. 等。通过虚拟展开揭示PHerc.Paris. 4的最后几栏文字，《赫库兰尼姆纪事》54, 9–27（2024年）11. Henderson, P. 通过微分同胚螺旋拟合虚拟展开赫库兰尼姆纸草文献。载于《IEEE/CVF冬季计算机视觉会议论文集》（WACV）（2026年）。12. Brun, E. 等。揭示赫库兰尼姆纸草文献中的金属墨水。《美国国家科学院院刊》113, 3751–3754（2016年）。doi:10.1073/pnas.1519958113。13. Parker, C. S. 等。从不可见到可读：恢复赫库兰尼姆的墨迹。《公共科学图书馆·综合》14, e0215775（2019年）。doi:10.1371/journal.pone.0215775。14. Parsons, Stephen《坚硬的纸草卷：一种阅读赫库兰尼姆纸草文献的非侵入式方法》。计算机科学论文与学位论文。138。（2023年）doi:10.13023/etd.2023.372 15. Bonnerot, O., Del Mastro, G., Hammerstaedt, J., Mocella, V., Rabin, I. 对一些赫库兰尼姆纸草文献进行X射线荧光墨水分析。《纸草学与铭文学杂志》216, 50–52。16. Booras, S.W., Seely, D.R. 赫库兰尼姆纸草文献的多光谱成像，《赫库兰尼姆纪事》29, 95–100（1999年）。17. Tourni’e, A. 等。通过短波红外高光谱成像，揭示未展开纸莎草纸背面隐藏的古希腊文本。《科学进展》5, eaav8936（2019年）。doi:10.1126/sciadv.aav8936。18. Seales, B., Chapman, C., Nicolardi, F., Parker, C.S. 赫库兰尼姆纸草文献的数字修复。《赫库兰尼姆纪事》52, 201–211（2023年）。19. Ceccarelli, S. 等。赫库兰尼姆纸草文献的脉冲热成像分析。《科学报告》15, 34466（2025年）。doi:10.1038/s41598-025-19911-w。20. Mazzocato, S., Cimino, D. and Daffara, C. 综合运用微观轮廓测量与多光谱成像，对古代手稿进行全场分析。《文化遗产杂志》66, 110–116（2024年）。doi:10.1016/j.culher.2023.11.014。21. Borch, J., Lyne, M. B., Mark, R. E. and Habeger, C. C. 使用光学轮廓测量表征纸张表面。载于《纸张物理测试手册》第2版（编者：Borch, J., Lyne, M. B., Mark, R. E. and Habeger, C. C.），429–450页（CRC Press，博卡拉顿，2002年）。doi:10.1201/9780203910498-25。22. Zienkiewicz, E. 等。光学轮廓测量作为检测压痕文字的方法。Kryminalistyka i S ˛adowa Ekspertyza 66（2020年）。23. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J. and Maier-Hein, K. H. nnU-Net：一种用于基于深度学习的生物医学图像分割的自配置方法。《自然·方法》18, 203–211（2021年）。doi:10.1038/s41592-020-01008-z。24. Zijdenbos AP, Dawant BM, Margolin RA, Palmer AC. 磁共振图像中白质病变的形态计量分析：方法与验证。IEEE Trans Med Imaging. 1994;13(4):716-24. doi: 10.1109/42.363096. PMID: 18218550。25. Telea, A. 一种基于快速行进法的图像修复技术。《图形工具杂志》9(1), 25–36（2004年）。doi:10.1080/10867651.2004.10487596。26. Marin, D., He, Z., Vajda, P., Chatterjee, P., Tsai, S., Yang, F. and Boykov, Y. 高效分割：学习语义边界附近的下采样。载于《IEEE/CVF国际计算机视觉大会论文集》（ICCV），2131–2141页（2019年）。doi:10.1109/ICCV.2019.00222。27. Ben˘cevi’c, M., Qiu, Y., Gali’c, I. and Pizurica, A. 先分割再分割：面向大型生物医学图像的保留上下文的裁剪式分割。《传感器》23, 633（2023年）。doi:10.3390/s23020633。28. Li, L., Gao, S., Wu, F. and An, X. MBAN：用于小目标检测的多分支注意力网络。《PeerJ计算机科学》10, e1965（2024年）。doi:10.7717/peerj-cs.1965。

致谢

我们感谢负责那不勒斯国家图书馆“维托里奥·埃马努埃莱三世”馆藏 PHerc.（赫库兰尼姆纸草文献编号）文献的策展人与修复人员，感谢他们提供访问机会和指导，尤其感谢 Silvia Scipioni 和 Giovanni Bova。本研究得到维苏威挑战赛（https://scrollprize.org）的支持。该非营利组织由捐赠资助，旨在完整解读赫库兰尼姆的全部密封卷轴。我们尤其感谢 Nat Friedman 发起并支持维苏威挑战赛。我们也感谢肯塔基大学 EduceLab 在梅隆基金会资助下开展的前期工作，这些工作为本研究的实现奠定了基础。

作者贡献声明

G.A. 构思了本研究，收集数据，开展分析并起草了稿件。F.N. 挑选并鉴定了纸草文献，并提供纸草学专业知识。P.H. 提供了计算机视觉和机器学习方面的意见。W.B.S. 就最先进水平、相关研究以及结果的呈现提供了意见。所有作者均审阅并批准了稿件。

附加信息

本研究未获得任何资助机构提供的正式研究经费。研究支持来自维苏威挑战赛，这是一个由捐款资助的非营利项目。作者声明不存在利益冲突。