Distortion-Free Registration of Whole-Slide Images via Tile-Based Refinement

Sungyun Kim[1,*] **Sehoon Kim**[2,*] **Donggyun Kang**[3]
Jaeeun Song[3] **Kwangsoo Kim**[3,†] **Eunna Kim**[2,†]

[1]Interdisciplinary Program of Medical Informatics, Seoul National University
[2]Department of Pathology, Seoul National University Hospital
[3]Department of Transdisciplinary Medicine, Seoul National University Hospital

{sungyun1997, qet8766, kwangsookim, hk1997}@snu.ac.kr,
dkang335@gatech.edu,
0110gengen@naver.com

Abstract

Registering Hematoxylin and Eosin (H&E)-stained tissue images with Immunohistochemistry (IHC)-stained counterparts is a critical step for accurate pathological interpretation. Existing registration methods in digital pathology have primarily focused on aligning corresponding cells at the same pixel coordinates. While this approach achieves low quantitative registration error, nonlinear transformations often distort cellular morphology and tissue architecture, limiting the applicability of such methods in clinical practice. To address this issue, we propose a tile-based local registration method that minimizes image distortion. Our approach relies solely on translation, rotation, and uniform scaling, thereby preserving structural integrity during registration. Experimental results on the ACROBAT 2023 challenge datasets demonstrate that the proposed method achieves lower registration errors compared to conventional rigid registration algorithms, with a 34.99% reduction in error.

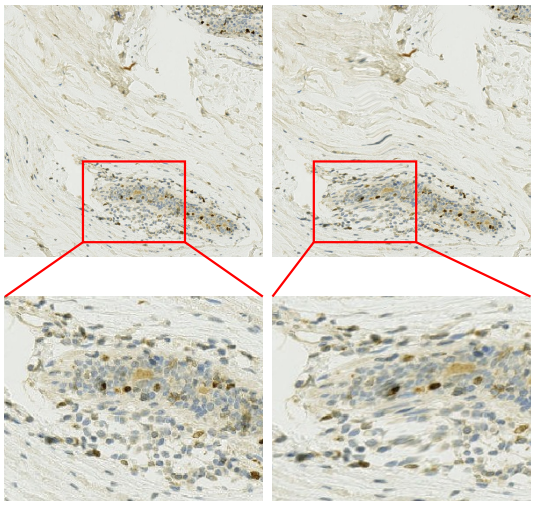
1 Introduction

Hematoxylin and Eosin (H&E)-stained tissue images provide structural information such as nuclei and cytoplasm, whereas Immunohistochemistry (IHC)-stained images reveal biomarker or protein expression patterns targeted by specific antibodies. Since these two staining techniques provide complementary pathological information, accurate registration between H&E and IHC images is essential for reliable interpretation. Local feature matching algorithms such as SuperGlue[7] and LightGlue[5] have recently been introduced for this purpose. These methods establish correspondences between keypoints across images by leveraging learned feature descriptors and attention-based matching.

Image registration can be broadly classified into rigid and non-rigid approaches depending on the type of transformation and the criteria applied to target images. Deep learning-based studies have predominantly focused on non-rigid registration, which estimates deformation fields to align corresponding cells at the same coordinates[13, 12, 4, 6]. While such approaches are advantageous for downstream analyses requiring cell-level correspondence, they often distort the original cellular morphology—particularly under substantial inter-modality appearance differences. As illustrated

[*]Equal contribution.
[†]Corresponding author.

39th Conference on Neural Information Processing Systems (NeurIPS 2025) Workshop: The 3rd Workshop on Imageomics: Discovering Biological Knowledge from Images Using AI.



(a) original (b) non-rigid warped

Figure 1: Distortion of cellular morphology introduced by non-rigid registration under substantial inter-modality differences.

in Figure 1, non-rigid registration can introduce unnatural deformation of cellular structures, which may alter pathological characteristics such as tumor size, location, and growth patterns, thereby compromising diagnostic accuracy in clinical practice.

To address this issue, VALIS[4] allows micro rigid registration on high-resolution tiles. However, this approach incurs excessive computational cost while offering only limited improvements in accuracy. In this study, we propose a tiling-based local registration method that overcomes these limitations. The proposed method significantly improves registration accuracy while maintaining computational efficiency and ensures reliable alignment across all regions of high-resolution images used in clinical diagnosis.

The main contributions of this work are as follows:

- We propose a rigid registration method based on local feature matching, which is practically applicable in clinical settings.
- Without additional training, we adapt and refine existing local feature matching algorithms to the characteristics of digital pathology images.
- By employing tile-based local registration, our method consistently provides optimal alignment across diverse regions, with benchmark experiments demonstrating high performance using rigid registration alone.

2 Related Works

Local Feature Matching Local feature matching algorithms can be categorized into dense feature matching[8, 2], which directly compares pixel intensities across the entire image, and sparse feature matching[7, 5], which extracts keypoints[1, 9] from images and utilizes only those corresponding features. Given the extremely large size of pathology images, sparse feature matching methods are generally preferred for their computational efficiency. In this context, LightGlue[5] has emerged as a state-of-the-art sparse feature matcher, offering high efficiency in memory and computation while maintaining strong accuracy and adaptability to task difficulty.

WSIs Registration VALIS[4] is a widely used open-source software library for pathology image registration. It provides preprocessing modules tailored to pathology images, along with both rigid and non-rigid registration pipelines. While the non-rigid option can achieve higher alignment accuracy, it often introduces structural distortions that may limit clinical applicability. In addition, the optional

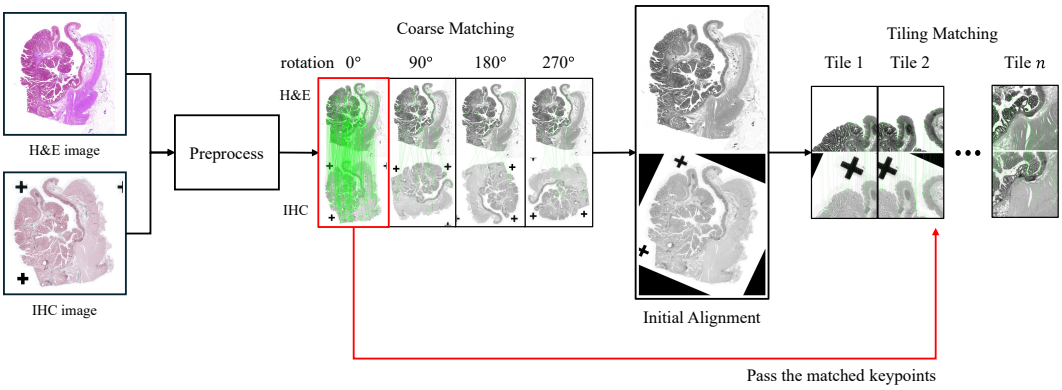


Figure 2: Overall architecture of the proposed pipeline.

micro-registration mode requires substantial computational resources, and the accuracy of the rigid registration remains relatively limited compared to non-rigid approaches.

3 Method

3.1 Overview

The proposed method consists of three stages: (i) preprocessing, (ii) coarse matching, and (iii) tiling matching(i.e., our tile-based local registration stage). In the first stage, preprocessing addresses the limitations of local feature matching algorithms[5, 7, 8, 2] that were originally trained on natural images. In the second stage, coarse matching is performed to obtain an initial alignment between H&E and IHC images. Finally, tiling matching refines the alignment at the local level to improve registration accuracy. An overview of the entire pipeline is shown in Figure 2.

3.2 Preprocessing

To address the computational burden caused by the large number of pixels in whole slide images (WSIs), each image is downscaled while maintaining the aspect ratio so that the longer side is 1024 pixels, and then converted to grayscale to reduce variations in keypoint extraction caused by staining differences. Since the downscaling ratio may differ between H&E and IHC images, the original resolution is later restored through the transformation matrix.

3.3 Coarse Matching

Coarse matching establishes an initial global alignment. From the preprocessed images, keypoints and descriptors are extracted using SuperPoint[1], and corresponding keypoints are identified with LightGlue[5]. The random sample consensus (RANSAC) algorithm[3] is applied to estimate inliers (correspondences that consistently support a single geometric transformation). Since local feature matching algorithms are typically trained on natural images, they are generally robust only to rotations within approximately 30°. When the IHC image undergoes larger rotations (e.g., 90° or more), the number of valid correspondences decreases significantly. To compensate, the IHC image is rotated by 0°, 90°, 180°, and 270°. For each rotation, matching is performed, and the case with the highest inlier count is selected. Finally, a similarity transformation matrix (translation, rotation, uniform scaling) is estimated and applied to the IHC image.

3.4 Tiling Matching

To address the local misalignment in high resolution after coarse matching, tile-based local registration is introduced. The downscaled image is divided into square tiles with 50% overlap. For each tile, a similarity transformation matrix is estimated using the keypoints extracted in the coarse stage. Tiles containing fewer than three corresponding keypoints are excluded, as matrix estimation is not feasible

Table 1: Comparison of TRE (μm) on the ACROBAT 2023 validation sets and HyReCo. Registration performance is reported as the median of 90th percentile landmark errors (medianp90s), together with twice the standard deviation (twostdp90s) to indicate variability across cases.

Method	ACROBAT	HyReCo
Baseline (SuperPoint + LightGlue)	798.3 $\pm$ 19,610.9	45.7 $\pm$ 73.3
VALIS (rigid only)	425.8 $\pm$ 4,853.4	78.1 $\pm$ 84.7
VALIS (non-rigid)	176.7 $\pm$ 6,986.9	14.9 $\pm$ 55.4
Coarse Matching (ours)	464.5 $\pm$ 5,315.8	45.7 $\pm$ 73.3
Tiling Matching (ours)	**276.8** $\pm$ **1,287.9**	**38.7** $\pm$ **62.7**

Table 2: Comparison of Runtime (sec) on the ACROBAT 2023 validation sets and HyReCo. Registration performance is reported as the mean inference time in seconds (sec), together with the standard deviation to indicate variability across cases.

Method	ACROBAT	HyReCo
Baseline (SuperPoint + LightGlue)	0.052 $\pm$ 0.049	0.114 $\pm$ 0.496
VALIS (rigid only)	2.486 $\pm$ 0.645	2.947 $\pm$ 0.698
VALIS (non-rigid)	39.007 $\pm$ 17.545	11.068 $\pm$ 0.776
Coarse Matching (ours)	0.174 $\pm$ 0.046	0.167 $\pm$ 0.206
Tiling Matching (ours)	**0.215** $\pm$ **0.230**	**0.317** $\pm$ **0.642**

in such cases. This procedure enables accurate local refinement for high-resolution images, without requiring additional feature extraction.

4 Experiments

We evaluated the proposed method on the ACROBAT[11] (Automatic Registration of Breast Cancer Tissue) 2023 challenge dataset and the HyReCo[10] (Hybrid Re-stained and Consecutive) dataset. The ACROBAT 2023 dataset contains H&E–IHC image pairs with expert-annotated landmarks. Since the challenge has officially concluded, evaluations were conducted only on the 100 validation pairs available. The HyReCo dataset further evaluates the method using 9 slides, each containing 5 consecutively stained sections, thus providing multi-stain registration scenarios. Registration accuracy was assessed using Target Registration Error (TRE), defined as the Euclidean distance between annotated landmarks and their corresponding coordinates after registration.

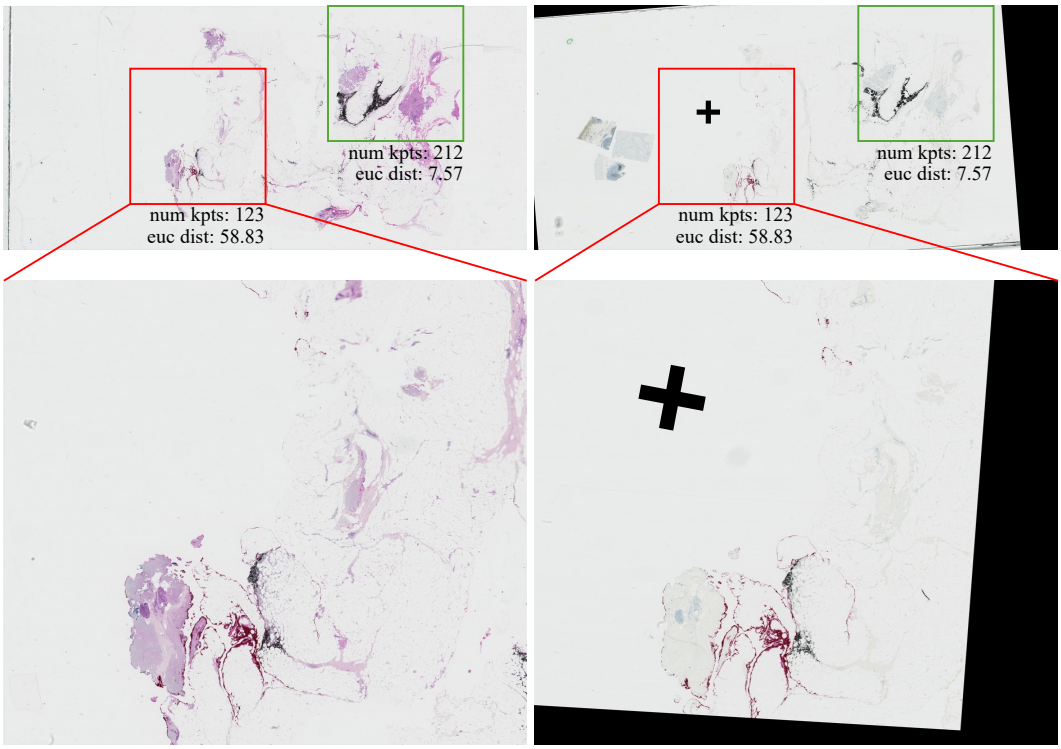
For comparison, we included the following baselines: Baseline (direct application of SuperPoint[1] and LightGlue[5]), VALIS[4] library (rigid and non-rigid), Coarse Matching (ours), and Tiling Matching (ours). In our implementation, the maximum number of keypoints for SuperPoint[1] was set to 1,024, and we used the publicly released pretrained models for both SuperPoint[1] and LightGlue[5].AMD EPYC 7763 CPU, an NVIDIA RTX A6000 GPU, and 512 GB of RAM. Nevertheless, the pipeline was designed for general computing environments and can be executed efficiently even with limited computational resources.

5 Results

Compared with the baseline, both coarse and tiling matching progressively reduced registration errors. The proposed tiling strategy achieved the most accurate and stable alignment, outperforming rigid VALIS[4] while avoiding the instability observed in non-rigid registration, as summarized in Table 1.

Significantly, the improved alignment accuracy did not come at the expense of speed. As detailed in Table 2, our tiling matching method maintains a fast average runtime, substantially outperforming non-rigid registration techniques.

To further illustrate the effectiveness of the proposed method, Figure 3 shows qualitative examples of registration. In the coarse matching result (top right), the green box highlights a well-aligned region, whereas the red box reveals local misalignment. After applying tiling matching (bottom), the



(a) H&E source image (b) IHC target image

Figure 3: Qualitative examples of registration results. (Top) Coarse matching of an IHC image against an H&E reference. The green box shows a well-aligned region, whereas the red box highlights a misaligned region. (Bottom) Tiling matching corrects the local misalignment. Note: "num kpts" denotes the number of matched keypoints, and "euc dist" refers to the Euclidean distance between corresponding landmarks.

previously misaligned region is correctly aligned, demonstrating the advantage of local refinement. These examples confirm that tile-based local registration is essential for maintaining accuracy in high-resolution pathology images while preserving computational efficiency.

6 Conclusion

In this study, we proposed a distortion-free registration method for H&E- and IHC-stained pathology images to address cellular morphology distortion commonly observed in conventional non-rigid registration. The proposed approach combines global coarse matching with tile-level local refinement, while relying solely on rigid transformations to preserve structural integrity. Evaluations on the ACROBAT 2023 challenge dataset demonstrated that our method achieved significantly lower TRE (34.99% reduction in error) compared with direct feature matching and global rigid alignment. Additionally, applying our method to the HyReCo dataset resulted in a 15.32% reduction in TRE compared to the benchmark. Our approach delivers performance comparable to non-rigid registration without introducing deformation. Both quantitative and qualitative analyses confirmed that the proposed method enables accurate and stable registration of high-resolution pathology images, providing an important foundation for practical use in clinical diagnosis.

Acknowledgments and Disclosure of Funding

This study was supported by the NAVER Digital Bio Innovation Research Fund, funded by NAVER Corporation (Grant No. 3720242040).

References

[1] Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. Superpoint: Self-supervised interest point detection and description. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 224–236, 2018.

[2] Johan Edstedt, Qianqian Sun, Gustav Bökman, Magnus Wadenbäck, and Michael Felsberg. Roma: Robust dense feature matching. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19790–19800, 2024.

[3] Martin A Fischler and Robert C Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.

[4] Cody D Gatenbee, Anne M Baker, S Prabhakaran, O Swinyard, Robert J Slebos, G Mandal, and Alexander R Anderson. Virtual alignment of pathology image series for multi-gigapixel whole slide images. *Nature Communications*, 14(1):4502, 2023.

[5] Philipp Lindenberger, Paul-Edouard Sarlin, and Marc Pollefeys. Lightglue: Local feature matching at light speed. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 17627–17638, 2023.

[6] Christian Marzahl, Fabian Wilm, L Tharun, S Perner, Claus A Bertram, C Kröger, and K Breininger. Robust quad-tree based registration on whole slide images. In *MICCAI Workshop on Computational Pathology*, pages 181–190. PMLR, 2021.

[7] Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. Superglue: Learning feature matching with graph neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4938–4947, 2020.

[8] Jingdong Sun, Zhuliang Shen, Yu Wang, Hujun Bao, and Xiaowei Zhou. Loftr: Detector-free local feature matching with transformers. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 8922–8931, 2021.

[9] Michał Tyszkiewicz, Pascal Fua, and Eduard Trulls. Disk: Learning local features with policy gradient. *Advances in neural information processing systems*, 33:14254–14265, 2020.

[10] Jeroen van der Laak, Johannes Lotz, Nick Weiss, and Stefan Heldmann. Hyreco - hybrid re-stained and consecutive histological serial sections (cc-by-sa 4.0), 2021.

[11] Paul Weitz, Mikko Valkonen, Luis Solorzano, Cameron Carr, Kimmo Kartasalo, Camille Boissin, and Mikko Rantalainen. Acrobat–a multi-stain breast cancer histological whole-slide-image data set from routine diagnostics for computational pathology. *arXiv preprint arXiv:2211.13621*, 2022.

[12] Michal Wodzinski, Nicolò Marini, Manfredo Atzori, and Henning Müller. Regwsi: Whole slide image registration using combined deep feature-and intensity-based methods: Winner of the acrobat 2023 challenge. *Computer Methods and Programs in Biomedicine*, 250:108187, 2024.

[13] Michal Wodzinski and Henning Müller. Deephistreg: Unsupervised deep learning registration framework for differently stained histology samples. *Computer Methods and Programs in Biomedicine*, 198:105799, 2021.

A Additional Experimental Results

Table 3 presents the per-sample TRE(μm) errors on the ACROBAT 2023 validation set (100 cases), and Table 4 shows the per-sample TRE(μm) errors on the 36 cases from the HyReCo dataset.

Table 3: Comparison of registration performance on the ACROBAT 2023 validation set (measured in TRE(μm), lower is better).

Slide ID	Baseline	VALIS (Rigid)	VALIS (Non-rigid)	Coarse Matching (ours)	Tiling Matching (ours)
0	436.75	830.13	75.13	849.83	87.16
1	23699.03	244.76	115.74	618.71	189.42
2	19496.61	238.81	262.97	295.42	544.38
3	22825.41	514.40	109.85	935.95	1246.15
4	346.96	375.63	312.51	342.12	304.71
5	194.68	123.00	112.29	153.57	161.89
6	523.97	137.41	72.40	140.14	85.44
7	1279.43	99.24	62.65	145.68	245.81
8	1217.98	623.96	418.00	682.42	373.03
9	20908.68	2337.04	3181.90	2674.34	3892.67
10	17031.94	496.87	263.94	1407.31	458.37
11	36419.78	663.74	503.67	683.57	582.09
12	370.37	233.51	156.95	253.83	318.94
13	22136.21	121.22	40.24	239.75	138.65
14	1081.70	933.51	176.18	1075.37	271.79
15	590.25	433.11	25.45	298.46	217.85
16	2350.97	1355.11	5065.65	3015.67	2199.12
17	426.35	241.62	60.52	266.61	83.90
18	232.74	122.53	84.60	128.52	184.81
19	14424.06	251.70	82.14	280.04	115.90
20	12139.51	768.54	790.95	854.16	947.08
21	1356.59	631.23	222.56	758.98	360.28
22	21583.35	1221.65	1030.54	1204.81	2112.39
23	95.75	85.12	27.44	82.44	72.86
24	127.00	130.00	58.71	151.35	87.87
25	19962.14	927.45	333.83	1968.96	783.72
26	930.62	1040.11	231.26	1992.83	281.90
27	22852.87	411.04	198.15	1056.14	456.41
28	289.91	202.89	88.83	343.71	173.19
29	1609.73	1145.73	705.33	2525.01	1120.20
30	251.40	209.12	164.94	396.32	170.95
31	394.51	301.61	225.83	545.56	302.17
32	672.05	544.60	72.96	790.35	291.52
33	150.45	106.11	23.61	106.23	53.80
34	221.90	145.90	36.21	164.52	98.30
35	232.05	231.42	39.91	266.64	220.53
36	20589.09	1563.23	194.36	26621.43	740.01
37	15823.47	1501.03	1501.03	984.22	1501.03
38	801.73	238.10	48.14	321.49	140.87
39	163.55	143.33	41.26	152.83	97.75
40	622.65	223.43	160.52	198.07	197.29
41	350.34	353.97	96.21	405.07	176.43
42	1284.18	1147.06	471.58	1454.82	560.29
43	221.89	165.34	41.24	175.51	113.83
44	18120.02	1338.60	921.27	788.24	1063.58
45	577.60	415.42	278.64	644.67	671.00

Continued on next page

Slide ID	Baseline	VALIS (Rigid)	VALIS (Non-rigid)	Coarse Matching (ours)	Tiling Matching (ours)
46	485.84	218.35	211.00	385.16	160.93
47	14286.10	1541.70	792.20	1425.70	2327.85
48	19218.13	778.36	907.10	896.83	636.37
49	794.79	204.09	120.72	258.15	249.06
50	27208.02	1083.12	177.23	1570.53	789.12
51	450.61	314.72	172.68	325.89	216.85
52	31026.47	449.99	92.41	654.65	144.62
53	346.79	262.66	67.74	220.49	199.09
54	334.10	301.45	52.66	412.19	243.48
55	23213.94	291.16	163.59	286.33	265.08
56	1584.62	1857.06	87.10	1558.70	743.48
57	19168.92	208.67	55.13	235.98	148.07
58	570.19	543.56	100.44	732.05	186.29
59	86.72	66.68	42.28	96.88	111.17
60	520.54	334.69	126.51	428.30	179.39
61	16583.92	13474.61	16302.55	321.14	331.15
62	1374.97	15000.16	24234.44	1580.81	1531.37
63	218.65	189.38	92.49	203.18	268.10
64	24056.43	1421.70	851.19	1620.89	1794.26
65	905.74	638.20	203.10	805.61	668.94
66	20357.93	840.09	568.73	861.99	392.19
67	287.56	324.80	136.26	281.82	160.52
68	23188.87	580.86	260.72	1012.14	467.79
69	140.09	74.91	44.47	65.67	161.64
70	363.47	231.87	201.13	255.69	229.80
71	1470.83	918.98	836.91	967.16	728.93
72	719.79	1738.88	324.53	747.53	701.63
73	1319.86	1271.25	274.36	1697.45	462.02
74	726.02	552.72	600.81	561.21	677.80
75	892.65	689.46	551.49	706.35	740.74
76	206.36	123.99	39.79	127.99	168.58
77	127.21	86.10	61.09	103.06	56.06
78	546.08	504.36	413.72	741.89	554.36
79	5724.90	554.38	637.43	445.08	411.80
80	991.56	1224.79	365.64	1044.99	456.06
81	1861.11	15047.35	20141.03	1326.24	1374.39
82	510.47	401.68	180.36	462.41	268.01
83	658.07	903.65	327.54	768.88	647.53
84	1109.99	532.78	216.64	723.23	333.69
85	1389.43	1489.35	1566.27	1524.62	1650.42
86	234.42	235.64	96.59	225.35	121.95
87	23417.06	202.11	82.17	203.04	475.51
88	30740.11	1318.46	1197.98	3316.04	3206.78
89	350.46	429.76	143.06	429.10	370.78
90	1038.29	736.82	531.55	698.30	362.56
91	318.30	291.67	76.34	276.14	191.83
92	200.33	157.18	32.92	146.07	91.09
93	682.22	456.21	107.63	466.58	251.13
94	215.58	193.62	141.93	162.88	159.75
95	21097.17	539.91	201.23	647.90	256.74
96	658.65	421.80	199.30	590.78	246.61
97	410.13	381.00	272.38	389.96	220.95
98	156.43	155.48	57.23	157.05	111.21
99	18992.75	512.19	343.99	451.63	515.01

Table 4: Comparison of registration performance on the HyReCo dataset (measured in TRE(μm), lower is better).

Slide ID	**Stain**	**Baseline**	**VALIS (Rigid)**	**VALIS (Non-rigid)**	**Coarse Matching (ours)**	**Tiling Matching (ours)**
29	CD8	51.93	114.42	14.78	51.93	33.89
29	CD45	59.46	57.67	14.95	59.46	52.87
29	KI67	72.46	104.08	5.99	72.46	38.30
29	PHH3	15.98	31.32	6.47	15.98	17.61
108	CD8	49.33	71.83	16.29	49.33	50.11
108	CD45	133.56	144.40	70.07	133.56	115.58
108	KI67	41.25	66.52	16.19	41.25	30.92
108	PHH3	11.07	17.44	7.37	11.07	16.42
361	CD8	43.33	114.08	10.66	43.33	28.10
361	CD45	45.41	119.11	14.20	45.41	39.12
361	KI67	60.43	129.93	10.50	60.43	30.70
361	PHH3	160.55	178.37	157.99	160.55	176.36
464	CD8	27.42	36.46	13.04	27.42	30.74
464	CD45	148.92	136.75	20.01	148.92	41.54
464	KI67	43.76	60.67	9.83	43.76	26.56
464	PHH3	13.10	19.49	8.81	13.10	22.75
533	CD8	65.27	62.36	21.92	65.27	40.25
533	CD45	64.85	74.38	26.22	64.85	48.46
533	KI67	22.39	30.83	17.14	22.39	31.18
533	PHH3	12.68	16.08	9.67	12.68	15.56
611	CD8	56.80	105.66	13.93	56.80	39.11
611	CD45	48.37	81.71	22.37	48.37	46.17
611	KI67	33.32	63.59	11.91	33.32	26.60
611	PHH3	25.73	12.37	11.19	25.73	31.10
628	CD8	27.22	110.31	12.81	27.22	24.64
628	CD45	106.82	95.94	20.27	106.82	42.47
628	KI67	37.28	42.98	9.45	37.28	29.99
628	PHH3	26.73	71.93	28.53	26.73	29.17
644	CD8	45.96	147.07	24.25	45.96	48.16
644	CD45	90.11	131.23	27.14	90.11	58.14
644	KI67	41.75	43.89	31.63	41.75	49.39
644	PHH3	36.18	96.30	12.30	36.18	45.19
679	CD8	54.83	85.34	30.16	54.83	58.24
679	CD45	60.47	127.68	35.93	60.47	106.97
679	KI67	113.18	104.78	85.80	113.18	96.83
679	PHH3	22.38	45.50	14.45	22.38	32.14