Supplementary material for "ANHIR: Automatic Non-rigid Histological Image Registration Challenge"

Jiří Borovec, Jan Kybic* *Senior Member*, Ignacio Arganda-Carreras, Dmitry V. Sorokin, Maria Gloria Bueno Garcia, Alexander V. Khvostikov, Spyridon Bakas, Eric I-Chao Chang, Stefan Heldmann, Kimmo Kartasalo, Leena Latonen, Johannes Lotz, Michelle Noga, Sarthak Pati, Kumaradevan Punithakumar *Senior Member*, Pekka Ruusuvuori, Andrzej Skalski *Senior Member*, Nazanin Tahmasebi, Masi Valkonen, Ludovic Venet, Yizhe Wang, Nick Weiss, Marek Wodzinski, Yu Xiang, Yan Xu, Yan Yan, Paul Yushkevic, Shengyu Zhao, and Arrate Muñoz-Barrutia *Senior Member*

S.I. Dataset descriptions

Lung lesions: Unstained adjacent $3\,\mu$m formalin-fixed paraffin-embedded sections were cut from the blocks and stained with Hematoxylin and Eosin (H&E) or by immunohistochemistry with a specific antibody for platelet endothelial cell adhesion molecule (PECAM-1, also known as CD31), prosurfactant protein C (proSPC), clara cell 10 protein (CC10) or antigen ki-67 (Ki67). Images of three mice lung lesions (adenoma or adenocarcinoma) were acquired with a Zeiss Axio Imager M1 microscope (Carl Zeiss, Jena, Germany) equipped with a dry Plan Apochromat objective (numerical aperture NA=0.95, magnification $40\times$, pixel size $0.174\,\mu$m/pixel). See also Fig. 2.

Lung lobes: Images of four whole mouse lung lobes, corresponding to the same set of histological samples as in the lung lesions dataset. They were also acquired with a Zeiss Axio Imager M1 microscope (Carl Zeiss, Jena, Germany) equipped with a dry EC Plan-Neofluar objective (NA=0.30, magnification $10\times$, pixel size $1.274\,\mu m/pixel$).

Mammary glands: The sections are cuts from two mouse mammary glands blocks stained with H&E in even sections and in odd sections alternatively with an antibody against the estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2). The images were acquired with the same microscope and set of acquisition parameters as the mouse lung lobes, the pixel size was $2.294\,\mu$m/pixel).

COAD: The COlon ADenocarcinoma (COAD) set assembles series of histological sections from colon cancer samples, scanned with a 3DHistech Pannoramic MIDI II scanner at $10\times$ magnification, for a resolution of $0.468\,\mu$m/pixel with a white-balance set to auto. Each series consists of one H&E histopathology section (first cut) followed by a variable number (4–7) of immunohistopathology sections stained with hematoxylin and DAB, with antibodies binding to proteins expressed by various immune cells (T-cells and macrophages). The information about antibodies used was not disclosed on a per-image basis by the owner of the data.

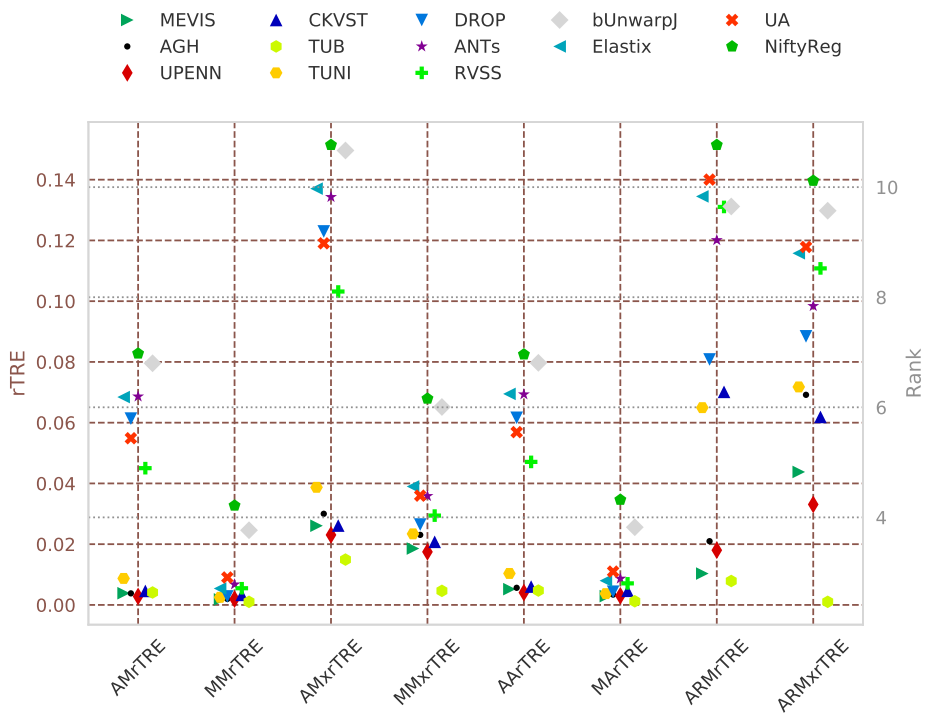
Mouse kidney: The set consists of resected healthy mouse kidneys which show high similarity to human kidneys. We used nine consecutive whole slide images having similar tissue structures. Whole slides were digitized with a NanoZoomer 2.0HT scanner (Hamamatsu) and a $20\times$ objective lens. The images were each roughly of $37\text{k}\times30\text{k}$ pixel size. Each image was dyed with one of the three stains — periodic acid-Schiff (PAS), smooth muscle actin (SMA) or CD31, such that every alternate slide is a PAS image.

Gastric: Surgical material from patients with a histologically verified diagnosis (gastric adenocarcinoma) were used for routine staining with Hematoxylin and Eosin (H&E) or for immunophenotyping. IHC-staining for latent membrane protein 1 (LMP-1) was used for Epstein-Barr virus (EBV) identification. The study of the cellular composition of the tumour tissue infiltrate was performed by immunohistochemical staining on the markers CD4, CD8, CD68 and CD1a. Deparaffinization and antigen recovery was performed by using Thermo Dewax and HIER Bufer L, a pH 6 buffer. The preparations were acquired with a Leica DM LB2 microscope.

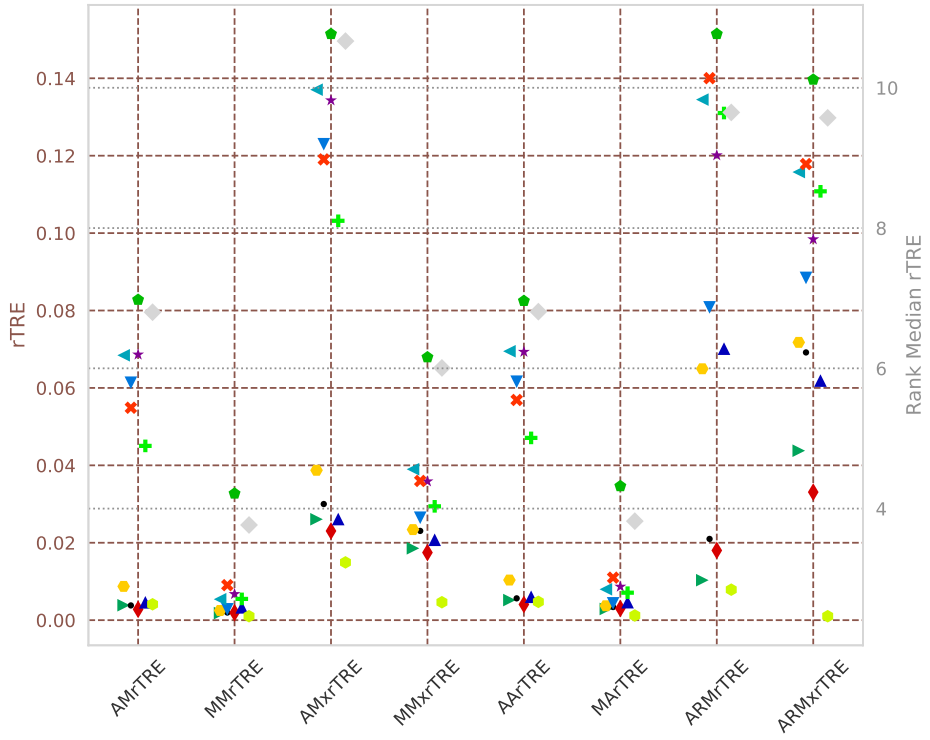
Human breast: Unstained adjacent $3\,\mu$m formalin-fixed paraffin-embedded sections were cut from the blocks, stained with Hematoxylin and Eosin (H&E), and with immunohistochemistry (IHC) with an antibody against ER, PR, and HER2, and imaged with Leica Biosystems Aperio AT2.

Human kidney: Unstained adjacent $3\,\mu$m formalin-fixed paraffin-embedded sections were cut from the glomerulopathies blocks, stained with Hematoxylin and Eosin (H&E) and PAS, Masson and Methenamine, and imaged with Leica Biosystems Aperio AT2.

See Fig. S4 for examples of the appearance differences due to staining between pairs of images from the same sets to be registered. Fig. S3 shows examples of the differences of the local structure.



(a)



(b)

Fig. S1. Relative performance on all data (a) and on test data (b) of all methods measured by AMrTRE (average median rTRE), MMrTRE (median of the median rTRE), AMxrTRE (average maximum rTRE), MMxTRE (median of the maxium rTRE), AArTRE (average of the average rTRE), MArTRE (median of the average rTRE), ARMrTRE (average rank of the median rTRE), and ARMxrTRE (average rank of the maximum rTRE), all based on aggregating rTRE.

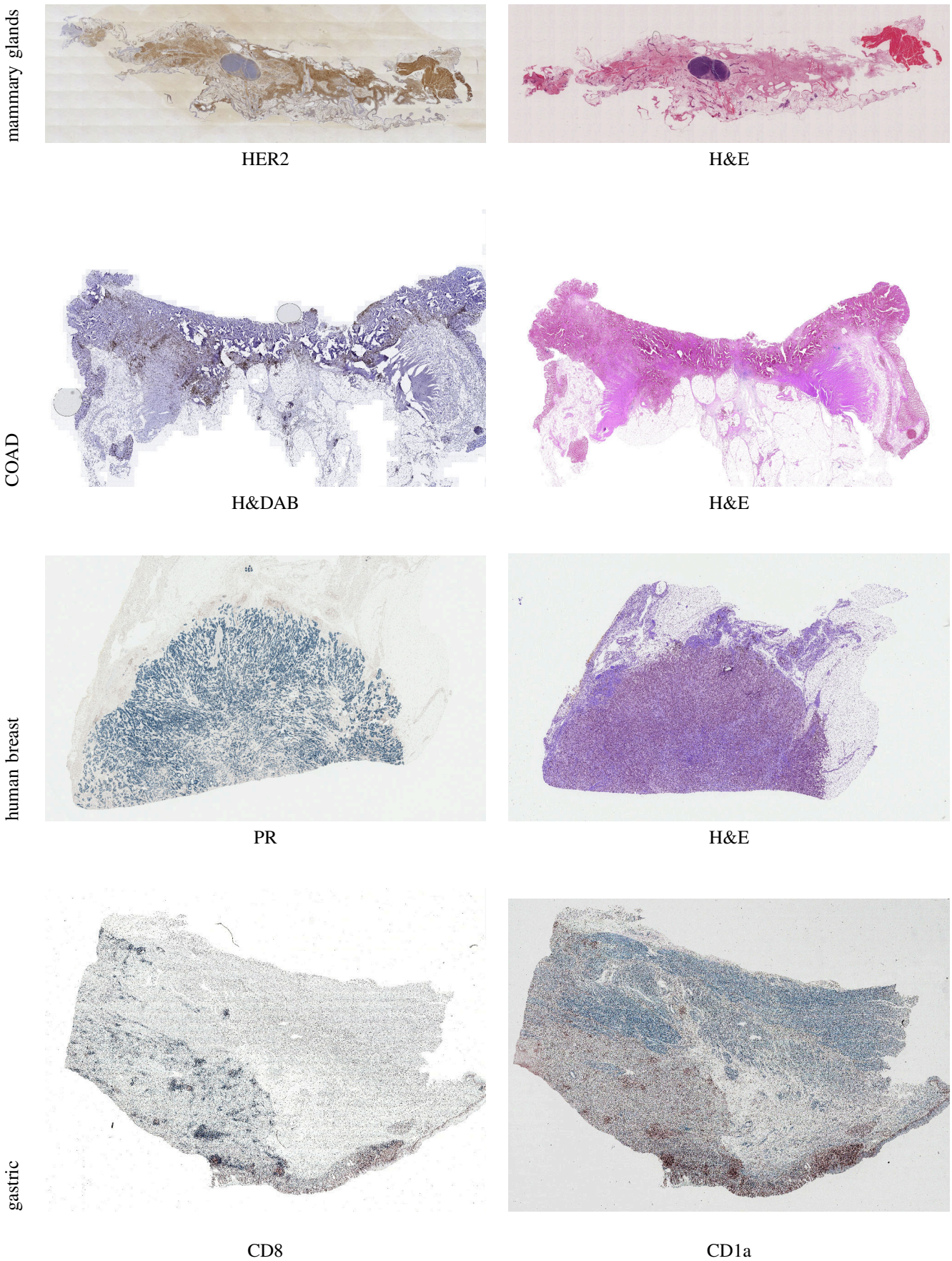
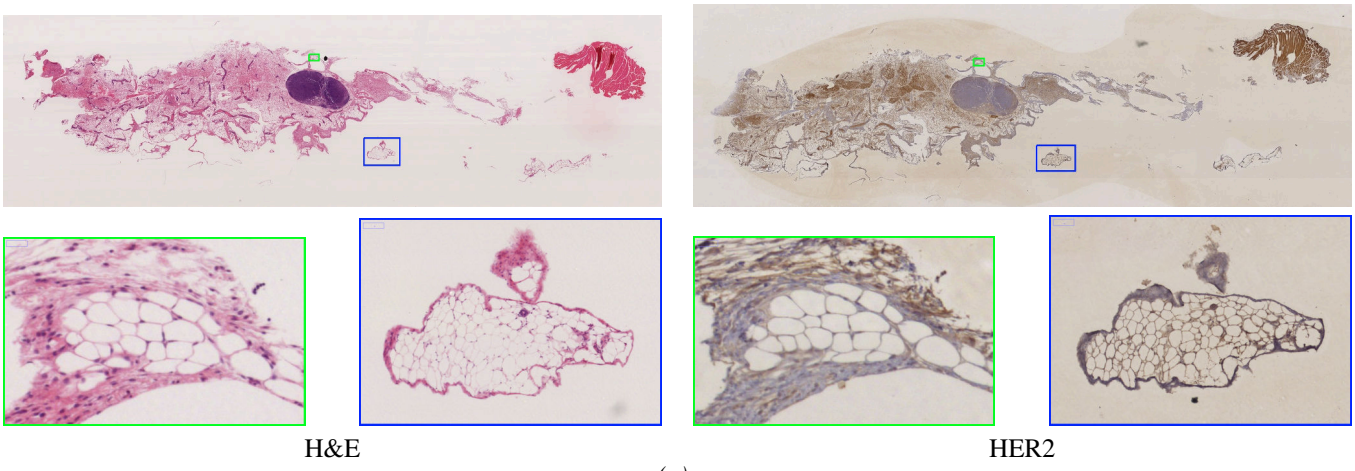
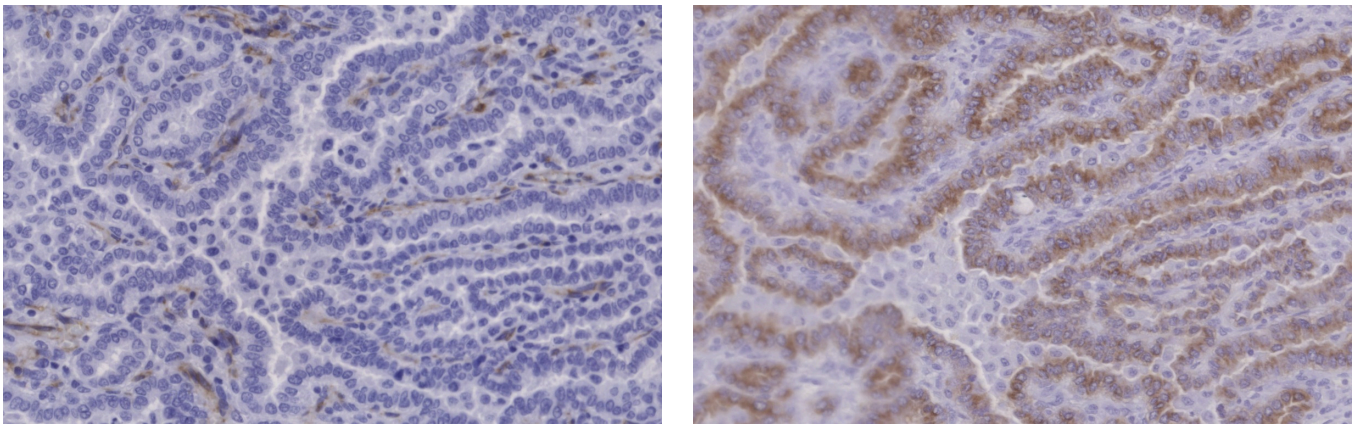


Fig. S2. Examples showing the differences between images to be registered. Each row shows two images from the same dataset with different stains. Some images were clipped and color-enhanced for visualization purposes.



(a)



(b)

Fig. S3. *(a)* Two differently stained images of the mammary gland. The contents of the blue and green rectangles in the first row are shown magnified in the second row. *(b)* Differences in the local structure of two differently stained images of the lung tissue at the same locations from the same set of images.

TABLE SI
GIVEN A SET OF FIVE IMAGES, THE LANDMARKS FOR EVERY THIRD IMAGE (IN BOLD) ARE WITHHELD AND THE REST IS PROVIDED TO PARTICIPANTS. THIS LEADS TO 3 IMAGE PAIRS FOR TRAINING (MARKED BY △) AND 6 IMAGE PAIRS FOR TESTING (MARKED BY ⋆). INVERTED PAIRS ARE NOT COUNTED.

		target image				
		1	2	3	**4**	5
source image	**1**		⋆	⋆		⋆
	2			△		△
	3					△
	4		⋆	⋆		⋆
	5					

S.II. LANDMARK ANNOTATION

Annotation was performed in ImageJ [S1][1] with the help of simple custom macros and scripts, which we provide[2]. The correctness of the landmarks was checked visually in high magnification.

Within each set, landmarks for every third image were withheld by the organizers and the remaining ones were made available to the participants to be used for training. For example, for 5 images in the set, we would withhold landmarks for images 1 and 4, resulting in three image pairs for training $((2,3),(2,5),(3,5))$ and 6 for testing $((1,2),(1,3),(1,5),(4,2),(4,3),(4,5))$, as shown in Table SI.

S.III. WELL KNOWN METHODS

Here we provide pointers to the implementation of existing registration methods which we used for comparison (see Section III of the main article):

bUnwarpJ: https://imagej.net/BUnwarpJ
RVSS: https://imagej.net/Register_Virtual_Stack_Slices
NiftyReg: https://github.com/jonclayden/RNiftyReg
Elastix: http://elastix.isi.uu.nl, based on ITK (https://itk.org/)
ANTs: http://stnava.github.io/ANTs/
DROP: https://www.mrf-registration.net/, https://github.com/biomedia-mira/drop2

[1] https://imagej.net/ImageJ
[2] https://borda.github.io/dataset-histology-landmarks/

TABLE SII

QUANTITATIVE RESULT OF ALL METHODS ON ALL DATA (TRAINING AND TESTING COMBINED), TO BE COMPARED WITH TABLE III OF THE MAIN ARTICLE. THE FIRST AND SECOND ROWS ('AVERAGE', 'MEDIAN' ETC.) CORRESPOND TO THE AGGREGATION METHOD WITHIN EACH IMAGE PAIR AND OVER ALL IMAGE PAIRS, RESPECTIVELY. THE TABLE IS SORTED BY ARMRTRE (IN BOLD). ⋆ = METHODS ADDED BY THE ORGANIZERS.

method	Average rTRE		Median rTRE		Max rTRE		Robustness		Median rTRE	Max rTRE	*Average*
	Average	Median	Average	Median	Average	Median	Average	Median	Average Rank		*time*
	(AArTRE)		(AMrTRE)		(AMxrTRE)		R		(**ARMrTRE**)	(ARMxrTRE)	[min]
initial	*0.1340*	*0.0684*	*0.1354*	*0.0665*	*0.2338*	*0.1157*	-	-	-		-
TUB	0.0047	0.0012	0.0041	0.001	0.0149	0.0046	0.9919	1.0000	2.84	2.47	0.02
MEVIS	0.0052	0.0029	0.0039	0.0018	0.0261	0.0186	0.9845	1.0000	2.98	4.83	0.15
UPENN	0.0041	0.0030	0.0028	0.0019	0.0230	0.0175	0.9888	1.0000	3.40	4.23	1.45
AGH	0.0056	0.0034	0.0038	0.0020	0.0300	0.0231	0.9770	1.0000	3.57	6.23	6.86
TUNI	0.0104	0.0037	0.0087	0.0025	0.0387	0.0234	0.8899	1.0000	5.99	6.37	10.32
CKVST	0.0060	0.0047	0.0046	0.0033	0.0261	0.0208	0.9730	1.0000	6.28	5.83	7.13
DROP⋆	0.0616	0.0043	0.0613	0.0028	0.1230	0.0265	0.8861	0.9907	6.87	7.29	3.41
ANTs⋆	0.0693	0.0087	0.0686	0.0067	0.1343	0.0359	0.8137	0.9718	9.04	7.84	43.09
RVSS⋆	0.0471	0.0071	0.0450	0.0055	0.1032	0.0294	0.7958	0.9875	9.64	8.52	4.72
bUnwarpJ⋆	0.0797	0.0256	0.0796	0.0246	0.1496	0.0652	0.7940	0.9310	9.65	9.57	9.15
Elastix⋆	0.0695	0.0080	0.0684	0.0054	0.1371	0.0390	0.7668	0.9706	9.83	8.80	2.96
UA	0.0569	0.0110	0.0549	0.0090	0.1190	0.0360	0.8076	0.9737	10.14	8.91	1.47
NiftyReg⋆	0.0825	0.0346	0.0828	0.0327	0.1514	0.0679	0.7495	0.8519	10.77	10.12	0.15

TABLE SIII

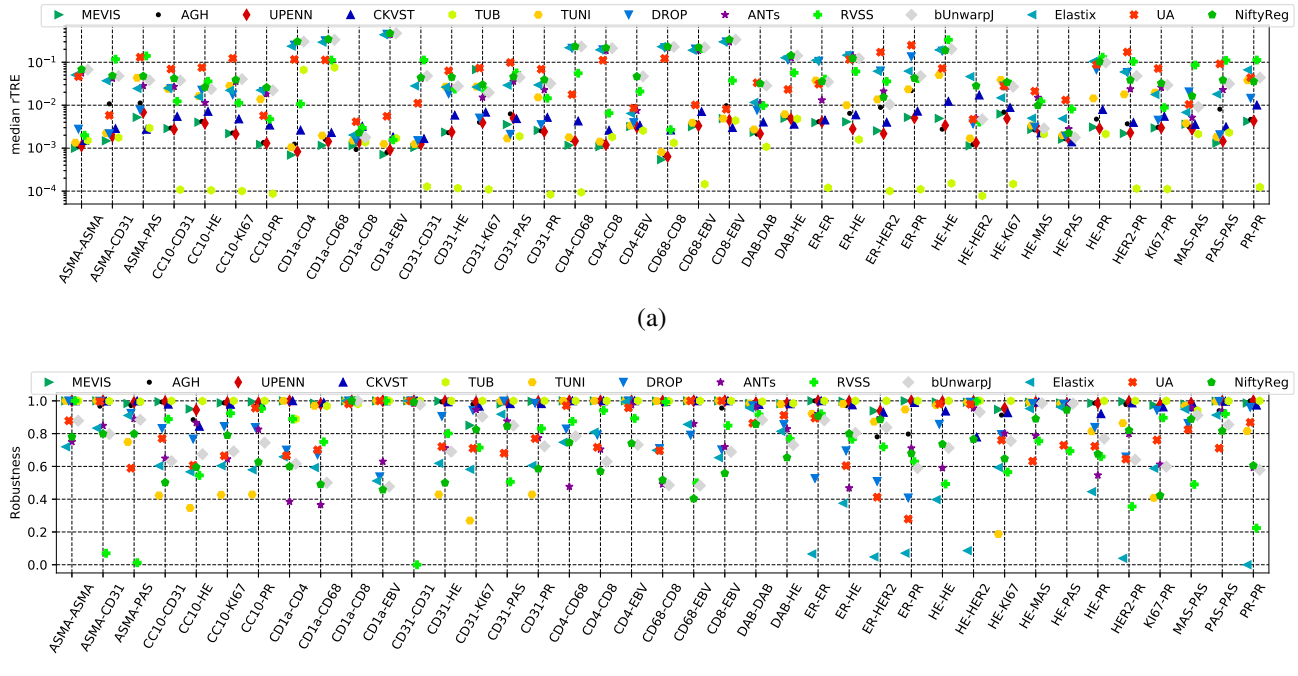
NUMBER OF IMAGE PAIRS BY STAINING FOR THE EVALUATION DATASETS.

reference/moving	ASMA	CC10	CD1a	CD31	CD4	CD68	CD8	DAB	EBV	ER	HE	HER2	KI67	MAS	PAS	PR	*sum*
ASMA	1	-	-	4	-	-	-	-	-	-	-	-	-	-	6	-	11
CC10	-	-	-	1	-	-	-	-	-	-	3	-	6	-	-	3	13
CD1a	-	-	-	-	8	7	1	-	2	-	-	-	-	-	-	-	18
CD31	-	6	-	1	-	-	-	-	-	-	3	-	6	-	2	6	24
CD4	-	-	1	-	-	7	2	-	2	-	-	-	-	-	-	-	12
CD68	-	-	1	-	1	-	1	-	2	-	-	-	-	-	-	-	5
CD8	-	-	-	-	7	7	-	-	2	-	-	-	-	-	-	-	16
DAB	-	-	-	-	-	-	-	171	-	-	5	-	-	-	-	-	176
EBV	-	-	-	-	1	-	1	-	-	-	-	-	-	-	-	-	2
ER	-	-	-	-	-	-	-	-	-	1	12	1	-	-	-	4	18
HE	-	4	-	4	-	-	-	61	-	6	9	3	6	5	5	14	117
HER2	-	-	-	-	-	-	-	-	-	2	2	-	-	-	-	2	6
KI67	-	1	-	1	-	-	-	-	-	-	1	-	-	-	-	3	6
MAS	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	5
PAS	2	-	-	8	-	-	-	-	-	-	-	-	-	5	14	-	29
PR	-	4	-	1	-	-	-	-	-	1	11	1	4	-	-	1	23
sum	3	15	2	20	17	21	5	232	8	10	46	5	22	10	32	33	481

S.IV. DATASET ACKNOWLEDGEMENTS

The lesions, lung-lobes and mammary-gland images were provided by Prof. Carlos Ortiz de Solórzano and Dr. Arrate Munoz Barrutia, Center for Applied Medical Research (CIMA), University of Navarra, Pamplona Spain [S2, S3]. The mice kidney images were provided by Prof. Peter Boor and Dr. Barbara M. Klinkhammer, Institute of Pathology, University Hospital Aachen, RWTH Aachen University [S4]. The colorectal cancer images were provided by Dr. Rudolf Nenutil (Masaryk Memorial Cancer Institute Brno), and Dr. Eva Budinska and Dr. Vlad Popovici (Masaryk University Brno) and were collected under grant nr.16-31966A by Ministry of Health of the Czech Republic. Gastric mucosa and gastric adenocarcinoma tissue images were provided by Prof. Pavel G. Malkov, Dr. Natalya V. Danilova, Dr. Nina A. Oleynikova and Ilya A. Mikhailov, Department of Pathology, Lomonosov Moscow State University [S5]. The kidney and breast cancer whole slide images were provided by Dr. Gloria Bueno and Dr. Oscar Deniz from Grupo VISILAB, Universidad de Castilla-La Mancha (UCLM). The images were obtained and prepared thanks to the AIDPATH European project[3] coordinated by UCLM.

[3]http://aidpath.eu



(a)

(b)

Fig. S4. Quantitative comparison of methods performance — *(a):* median rTRE and *(b):* robustness R) — as a function of the staining combination.

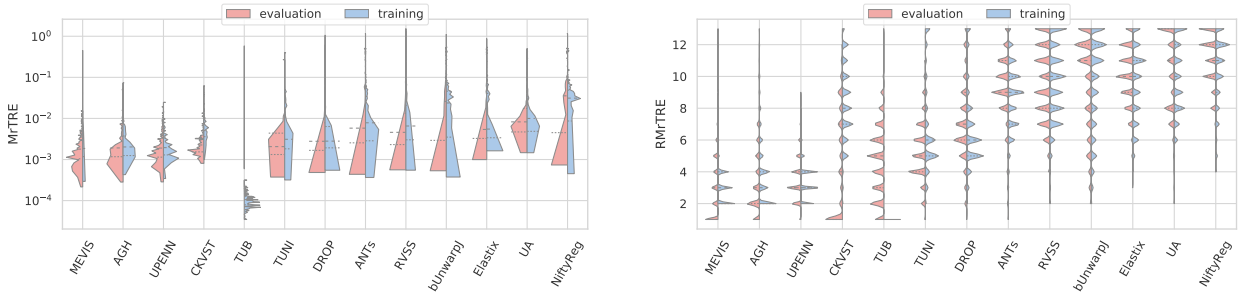


Fig. S5. Comparison of method performances, shown as histogram of the median rTRE and rank median rTRE, between training and testing (evaluation) datasets, to detect overfitting.

References

[S1] C. Schneider, W. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nature Methods*, vol. 9, pp. 671–675, 2012.

[S2] J. Borovec, A. Munoz-Barrutia, and J. Kybic, "Benchmarking of Image Registration Methods for Differently Stained Histological Slides," in *Proc. Int. Conf. Image Process.*, Athens, 2018, pp. 3368–3372.

[S3] R. Fernandez-Gonzalez, A. Jones, E. Garcia-Rodriguez, P. Chen, A. Idica, S. Lockett, M. Barcellos-Hoff, and C. Ortiz de Solórzano, "System for combined three-dimensional morphological and molecular analysis of thick tissue specimens," *Microscopy Research & Techniques*, no. 59, pp. 522–530, 2002.

[S4] L. Gupta, B. M. Klinkhammer, P. Boor, D. Merhof, and M. Gadermayr, "Stain independent segmentation of whole slide images: A case study in renal histology," in *Int. Symp. Biomed. Imag.*, 2018, pp. 1360–1364.

[S5] I. Mikhailov, N. Danilova, and P. Malkov, "The immune microenvironment of various histological types of EBV-associated gastric cancer," presented at European Congress on Pathology, 2018.