Fast Computation of Mutual Information with Application to Global Multimodal Image Alignment of Micrographs

Johan Öfverstedt, Joakim Lindblad, Nataša Sladoje

MIDA Group, Department of Information Technology, Uppsala University, Sweden
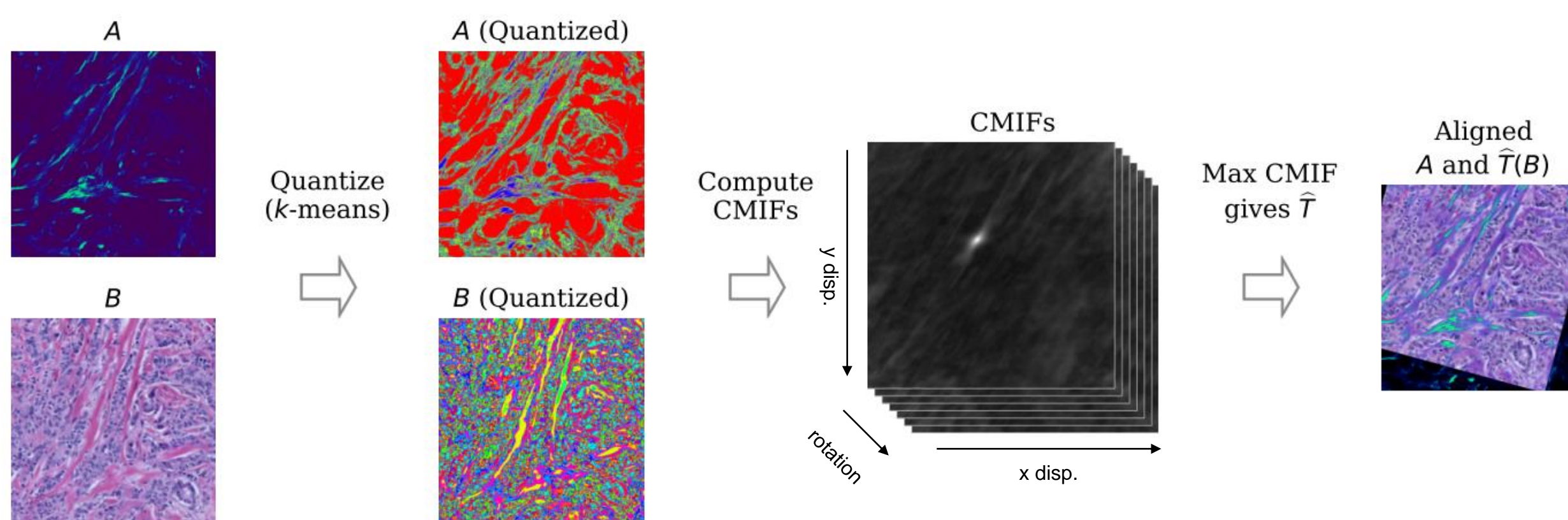
A rigid multimodal image alignment pipeline…

Fig 1: A pipeline for global multimodal rigid alignment of micrographs. The pixel values are clustered with k-means (k=16), and cross mutual-information (CMIF) maps are produced, i.e. mutual information computed for all possible overlaps of image A and B (one CMIF map per considered rotation angle). Finally the optimal transformation is found by locating the global maximum of this stack of CMIF maps.
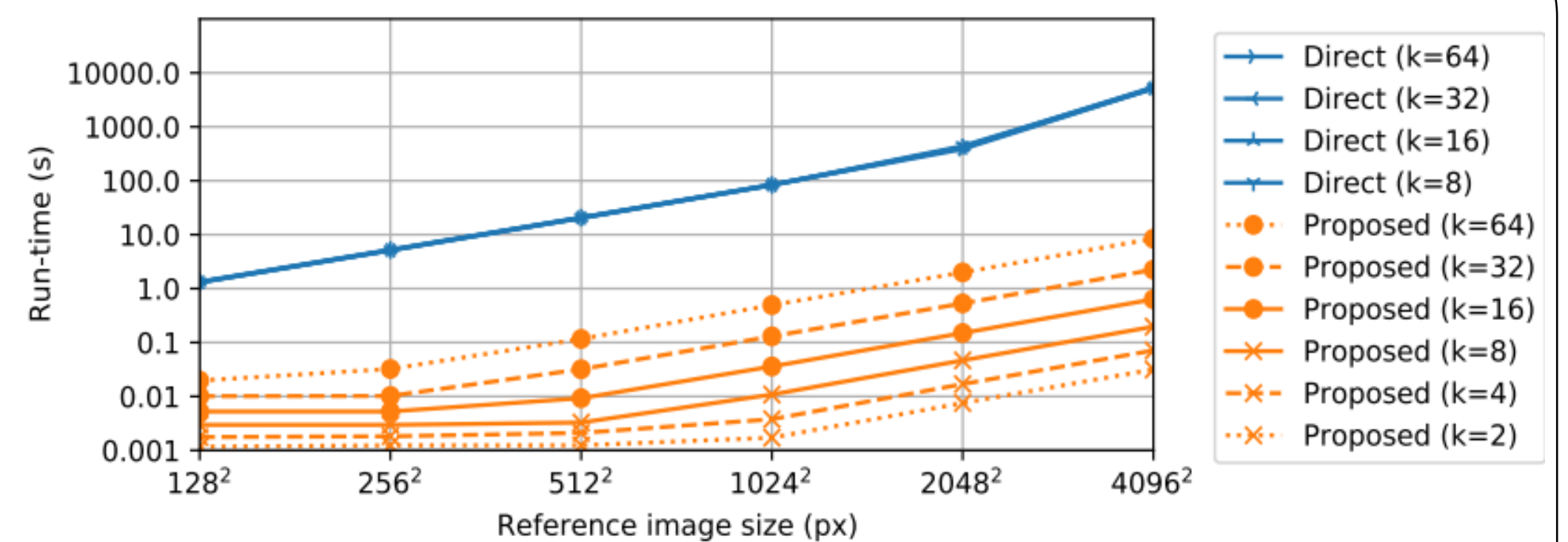
… based on a very fast algorithm.

Fig 2: Our algorithm (implemented in PyTorch) for computing MI maps in the frequency domain is orders of magnitude faster than a baseline (direct) algorithm (both running on a GeForce GTX 2080 GPU) on images of practically relevant size, and with suitable choice of the number of discrete values (k). *The algorithm works for both 2D and 3D data.*

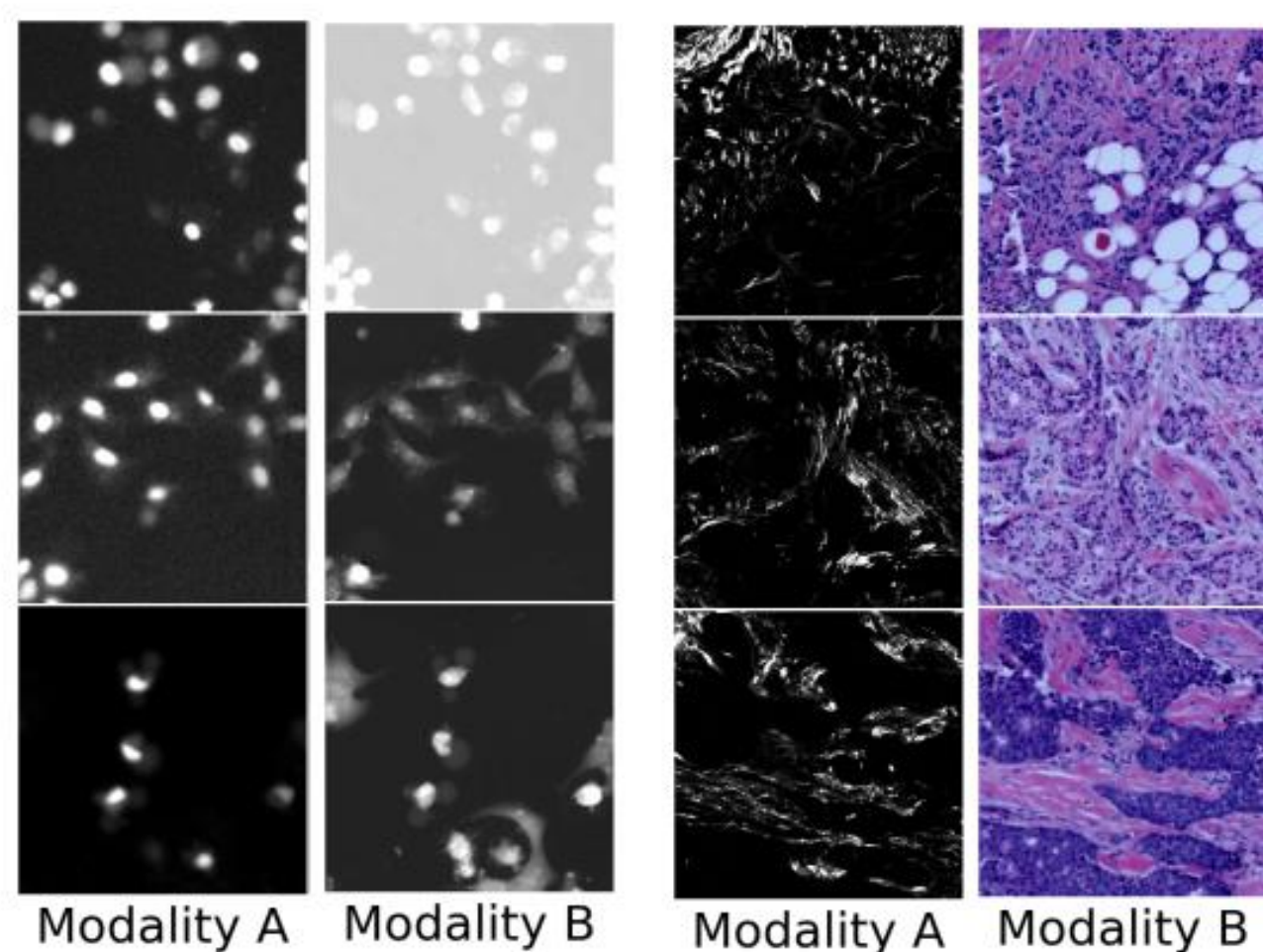## Validated on two pairs of microscopy modalities…

Fig 3: The method is evaluated on aligned image pairs* of fluorescence / quantitative phase imaging (cytological) and second-harmonic generation / bright-field (histological). The image from Modality B is synthetically (spatially) rotated and displaced. The task is to recover the position of landmarks corresponding to the corners of the reference image (Modality A) by aligning the image pair.

*The cytological dataset consists of 5040 image pairs and the histological dataset consists of 536 image pairs.

… exhibiting excellent performance.

Dataset	Cytological Data		Histological Data	
Method	α-AMD	SIFT	α-AMD	SIFT
I2I	71.1 ± 5.8	24.4 ± 6.2	28.4	0
CoMIR	68.0 ± 14.0	72.5 ± 7.1	81.3	59.3
Local MI	89.9 ± 3.0		47.8	
Proposed	**99.4 ± 0.4**		**91.2**	

Table 1: Quantitative evaluation of our image alignment pipeline on synthetically (spatially) transformed images, evaluated in terms of success-rate (success criterion: relative error* < 2%). Our method outperforms gradient-based MI maximization and several recent Deep Learning-based methods. [1, 2]

*The relative error is computed as the mean Euclidean displacement of recovered landmarks (the 4 corners of the reference image) divided by the side-length of the image.
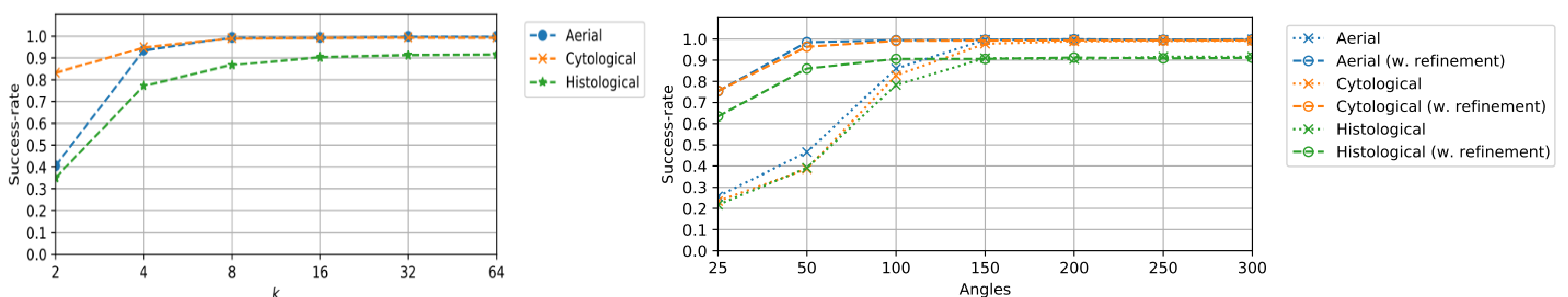
Easy to use.

Fig 4: The method is robust (in terms of success-rate) to the choice of the parameter k and the number of rotation angles explored in the grid search (especially when including a refinement step of random search around the best solution found with grid search). No training (data) is required.

Open source.

github.com/MIDA-group/globalign/

Acknowledgements.

[1] "Is Image-to-Image Translation the Panacea for Multimodal Image Registration: A Comparative Study", J. Lu, J. Öfverstedt, J. Lindblad, N. Sladoje, arXiv:2103.16262 [eess.IV], [cs.CV].
[2] "CoMIR: Contrastive Multimodal Image Representation for Registration", N. Pielawski, E. Wetzer, J. Öfverstedt, J. Lu, C. Wählby, J. Lindblad, N. Sladoje, NeurIPS 2020.

Contact us at: johan.ofverstedt@it.uu.se, joakim.lindblad@it.uu.se, natasa.sladoje@it.uu.se