Registration of Multimodal Microscopy Images using CoMIR – Learned Structural Image Representations

UPPSALA UNIVERSITET

Elisabeth Wetzer[1], Nicolas Pielawski[1], Johan Öfverstedt[1], Jiahao Lu[2], Carolina Wählby[1], Joakim Lindblad[1], Natasa Sladoje[1]

[1]Uppsala University, Sweden; [2]University of Copenhagen, Denmark

Introduction

- Combining information of multiple modalities for one specimen can shed light on properties not detectable by only one modality as they can provide complementary signals.
- Multimodal Registration can be extremely challenging if the appearance or signal expression density differs greatly between the modalities, as is the case for brightfield microscopy (BF) and second harmonic generation (SHG).
- We have developed contrastive learning based on InfoNCE [2] to learn representations from different modalities, called CoMIRs [1], which are visually similar.
- These image-like, dense representations can be succesfully registered by monomodal rigid registration methods, e.g. α-AMD (intensity-based, [3]) or using SIFT (feature-based, [4]).
- No data-specific information is incorporated in the learning, i.e. the method is modality independent and can be applied to other imaging modalities than BF and SHG.
- Very little aligned training data is required, for modalities which share sufficient structural similarities, the required aligned training data can be as little as one image pair.

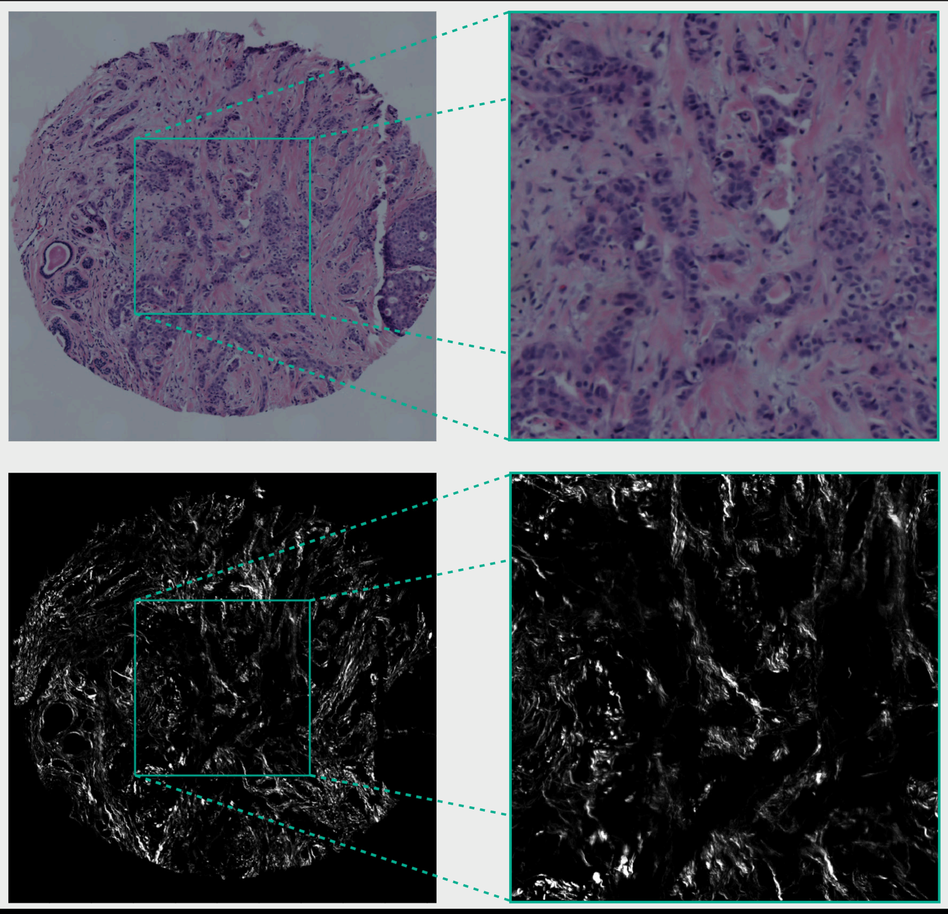

Fig. 1: Patches which are to be registered (available at [5]), cut from TMA cores captured by BF and SHG [6]. In our performance evaluation random rotations by ± 30° and random translations by ±100px were applied to the patches of 834x834px in size.

Contrastive Learning

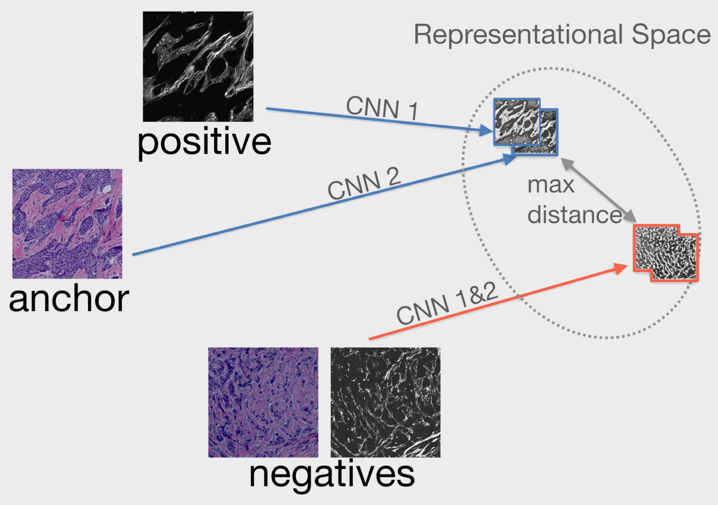

Fig. 2: Through contrastive learning our method produces abstract representations which are very similar for all input modalities.

- A randomly cropped patch in one modality serves as an anchor. Its corresponding patch in the other modality acts as a positive. Any other patch of any modality serves as a negative.
- Two CNNs, sharing no weights, only connected by the loss function, learn dense representations by maximizing the distance between the anchor and the negatives, as well as minimizing the distance between the anchor and the positive.

CoMIRs

- We require certain properties of the representations, such as rotational equivariance and similar intensities, which can be realized through the loss function without any additional hyperparameters.
- The appearance of CoMIRs depends on the choice of similarity function; MSE yieleded the best results.
- The number of channels for the CoMIRs can be chosen; single channel CoMIRs expedite registration.

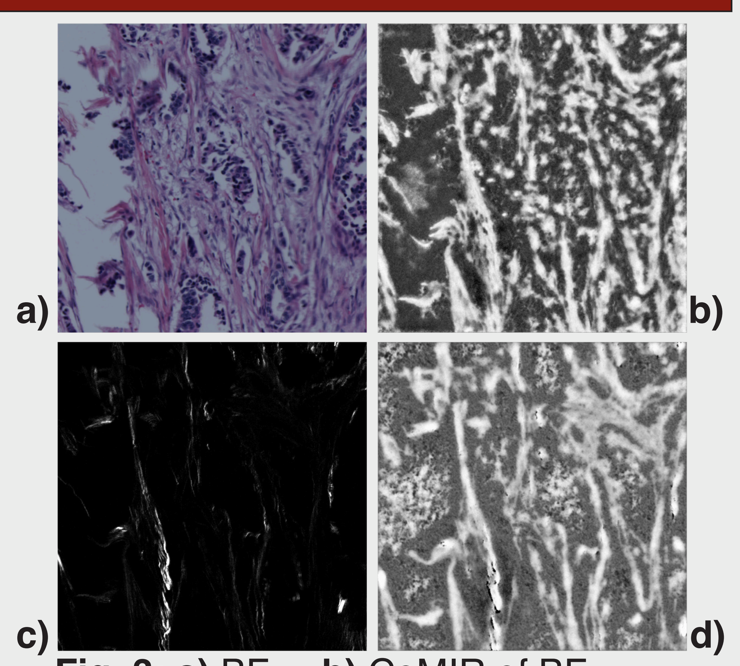

Fig. 3: a) BF **b)** CoMIR of BF **c)** SHG **d)** CoMIR of SHG

Registration of CoMIRs

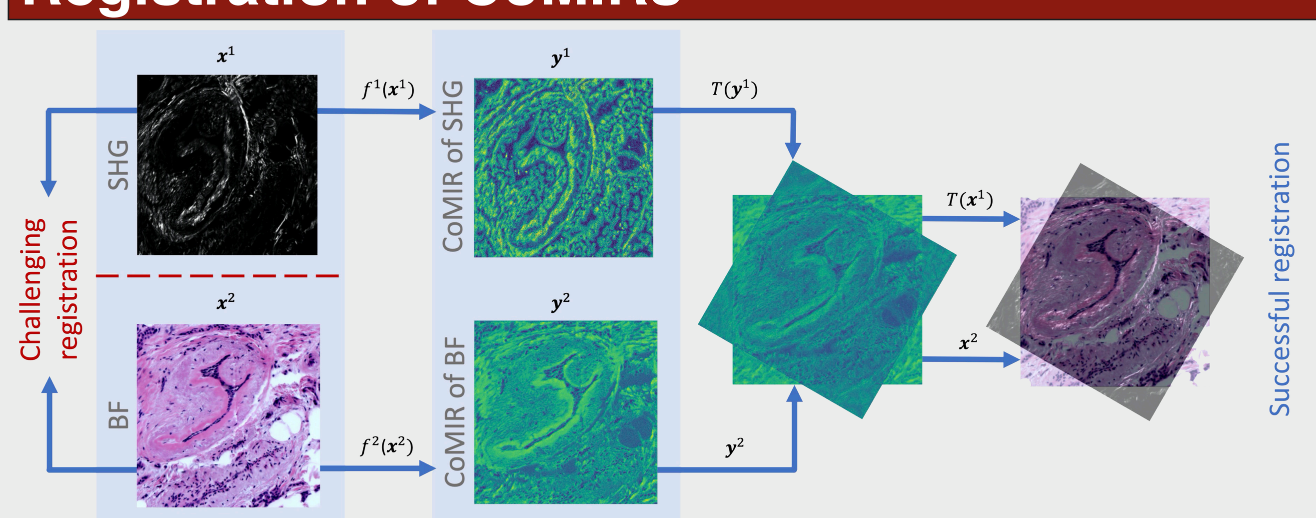

Fig. 4: The transformation found for the CoMIRs is applied to the original modalities to achieve multimodal registration.

- CoMIRs can be registered by common monomodal methods based on their intensities (e.g. α-AMD [3]) or by feature-based methods (e.g. using SIFT [4]).
- A transformation found for CoMIRs can be applied to orginal modalities and solve the multimodal registration.

Competing Methods

- CurveAlign: registers BF and SHG, using modality specific information and mutual information (MI) [6].
- GAN-based Image Translation methods: pix2pix, CycleGAN, DRIT. The resulting “fake” modalities do not qualify for intensity- or feature based monomodal registration methods.

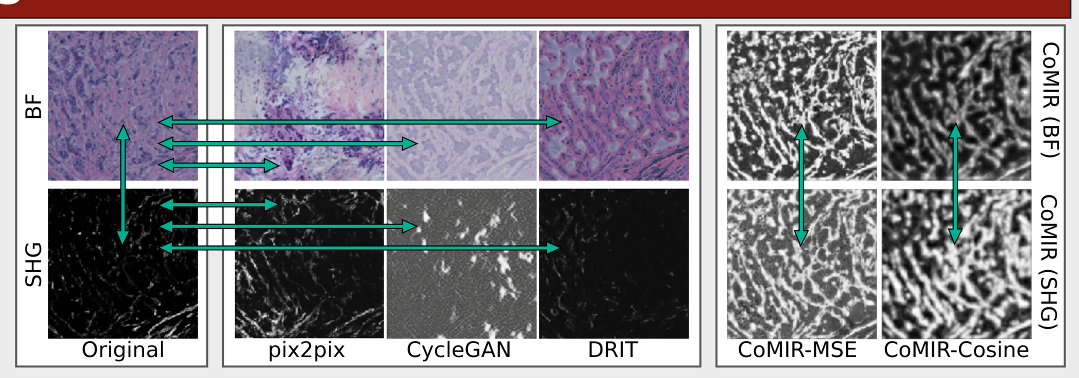

Fig. 5: Image translation methods to transform BF and SHG into one common modality. Arrows indicate resulting pairs for registration.

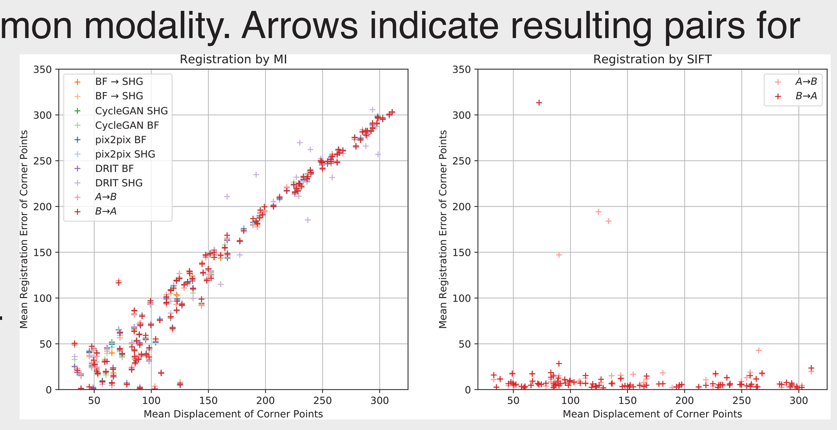

Fig. 6: MI fails to detect global extrema if the initial displacement is too large.

Results and Conclusions

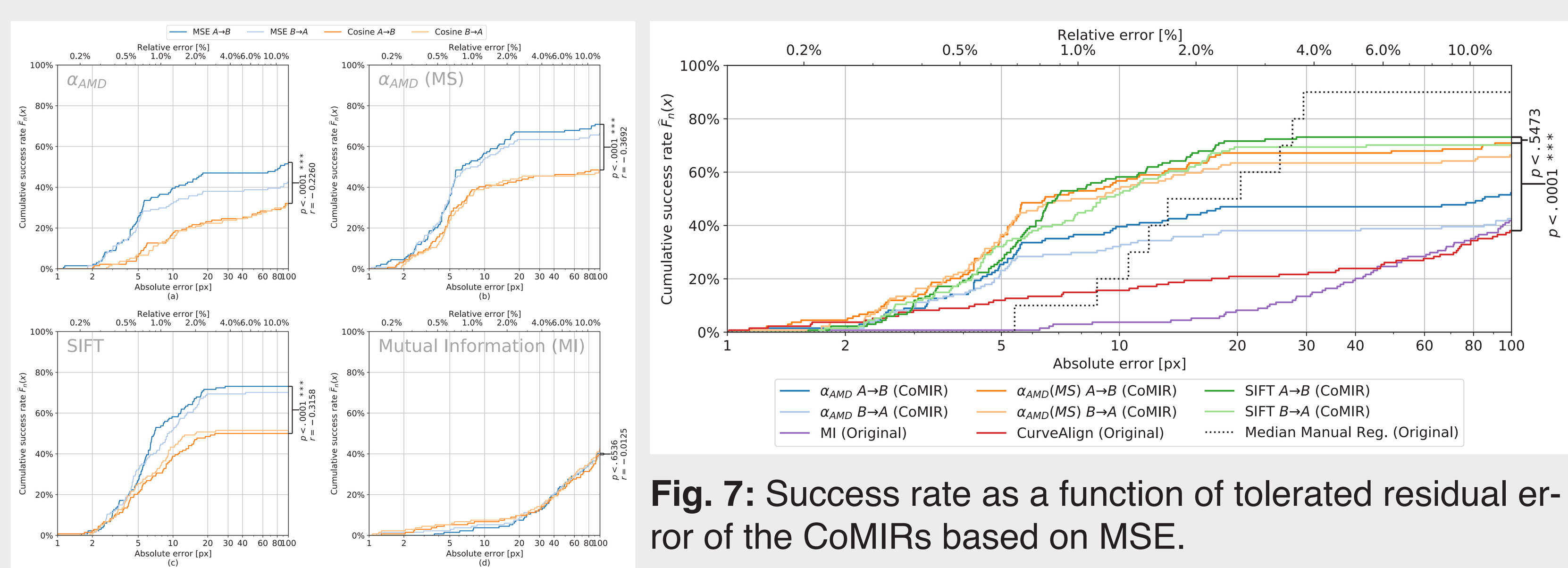

Fig. 7: Success rate as a function of tolerated residual error of the CoMIRs based on MSE.

- CoMIRs extract shared content in mutlimodal images and enable multimodal registration by reducing the problem to a monomodal one.
- CoMIRs combined with monomodal intensity- and feature-based registration methods significantly outperform multimodal registration by MI as well as a state-of-the-art data-specific approach. Best results were obtained using SIFT.
- CoMIRs contain more valuable information than GAN generated images obtained by image-to-image translation from one modality to the other.
- Feature-based registration does not depend on the initial displacement of the images as is the case for MI-based approaches.

References and Code

Code available at https://github.com/MIDA-group/CoMIR

[1] Pielawski, Wetzer et al.: CoMIR: Contrastive Multimodal Image Representation for Registration. NeurIPS, 2020

[2] Hjelm et al.: Learning deep representations by mutual information estimation and maximization. ICLR 2019

[3] Öfverstedt et al.: Fast and robust symmetric image registration based on distances combining intensity and spatial information. Trans. on Img Proc, 2019

[4] Lowe et al.: Object recognition from local scale-invariant features. ICCV 1999

[5] Eliceiri et al.: Multimodal Biomedical Dataset for Evaluating Registration Methods (patches from TMA Cores). 10.5281/zenodo.3874362, 2020

[6] Keikhosravi et al.: Intensity-based registration of bright-field and secondharmonic generation images of histopathology tissue sections. Biomed. Opt. Express, 2020

ELISABETH.WETZER @IT.UU.SE
NICOLAS.PIELAWSKI @IT.UU.SE
JOHAN.OFVERSTEDT @IT.UU.SE
JIAHAO.LU @DI.KU.DK
CAROLINA.WAHLBY @IT.UU.SE
JOAKIM.LINDBLAD @IT.UU.SE
NATASA.SLADOJE @IT.UU.SE

erc
Centre for Image Analysis, Uppsala University
VINNOVA Sweden's Innovation Agency
WASP WALLENBERG AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM