UPPSALA UNIVERSITET

Jiahao Lu
jiahao.lu.2199@student.uu.se

Johan Öfverstedt
johan.ofverstedt@it.uu.se

Joakim Lindblad
joakim@cb.uu.se

Nataša Sladoje
natasa.sladoje@it.uu.se

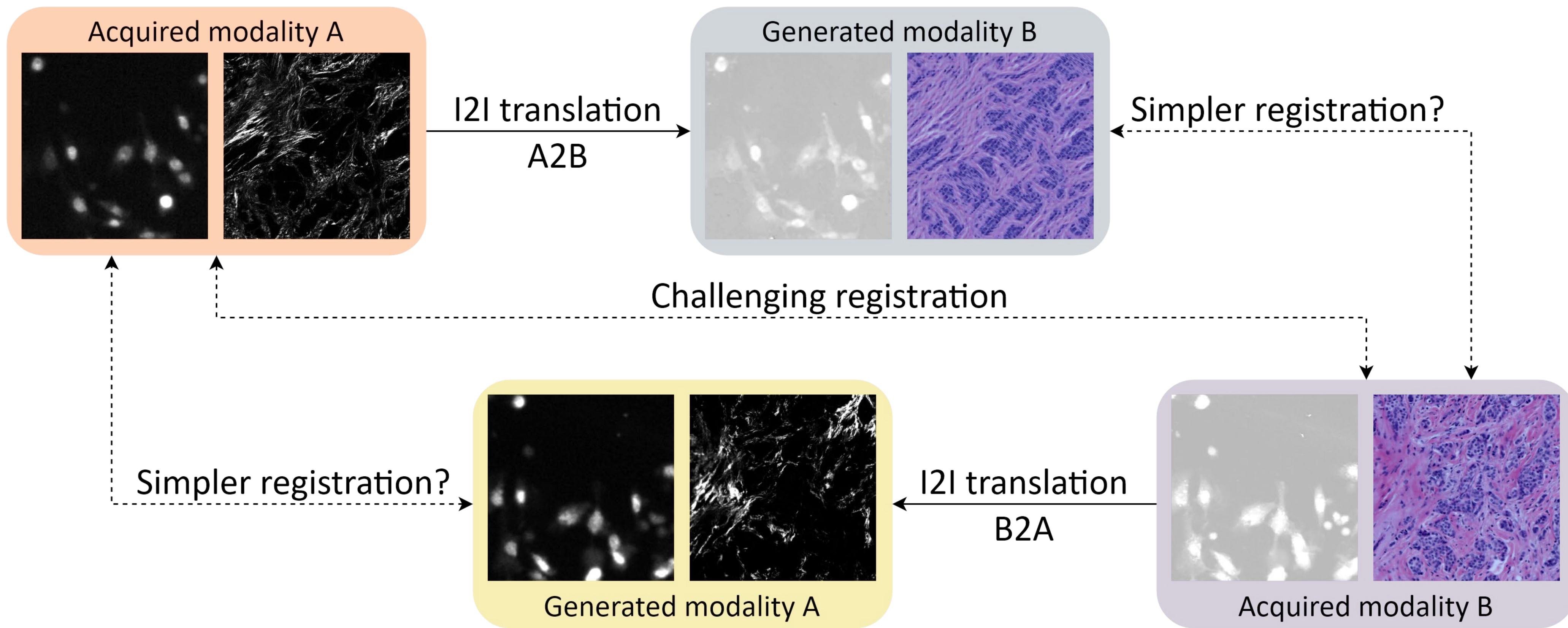
github.com/MIDA-group/MultiRegEval

MIDA Group

Department of Information Technology, Uppsala University

Image-to-Image Translation in Multimodal Image Registration: How Well Does It Work?

Overview



Experiments

- 4 GAN-based methods + 1 contrastive representation learning method
  - pix2pix[1] (supervised, strong baseline)
  - CycleGAN[2] (unsupervised, widely applied in biomedical field)
  - DRIT++[3] (unsupervised, explicitly extract shared information)
  - StyleGAN-v2[4] (unsupervised, injects domain-specific style into a given input)
  - CoMIR[5] (supervised, maps modalities to a "middle ground")
- 2 representative monomodal registration methods: SIFT, α-AMD
- 2 baselines: MI maximisation, CurveAlign[6]
- 3 multimodal datasets of increasing difficulty: aerial (NIR, RGB), cytological (Fluorescence, QPI), histological[7] (SHG, BF)

Results

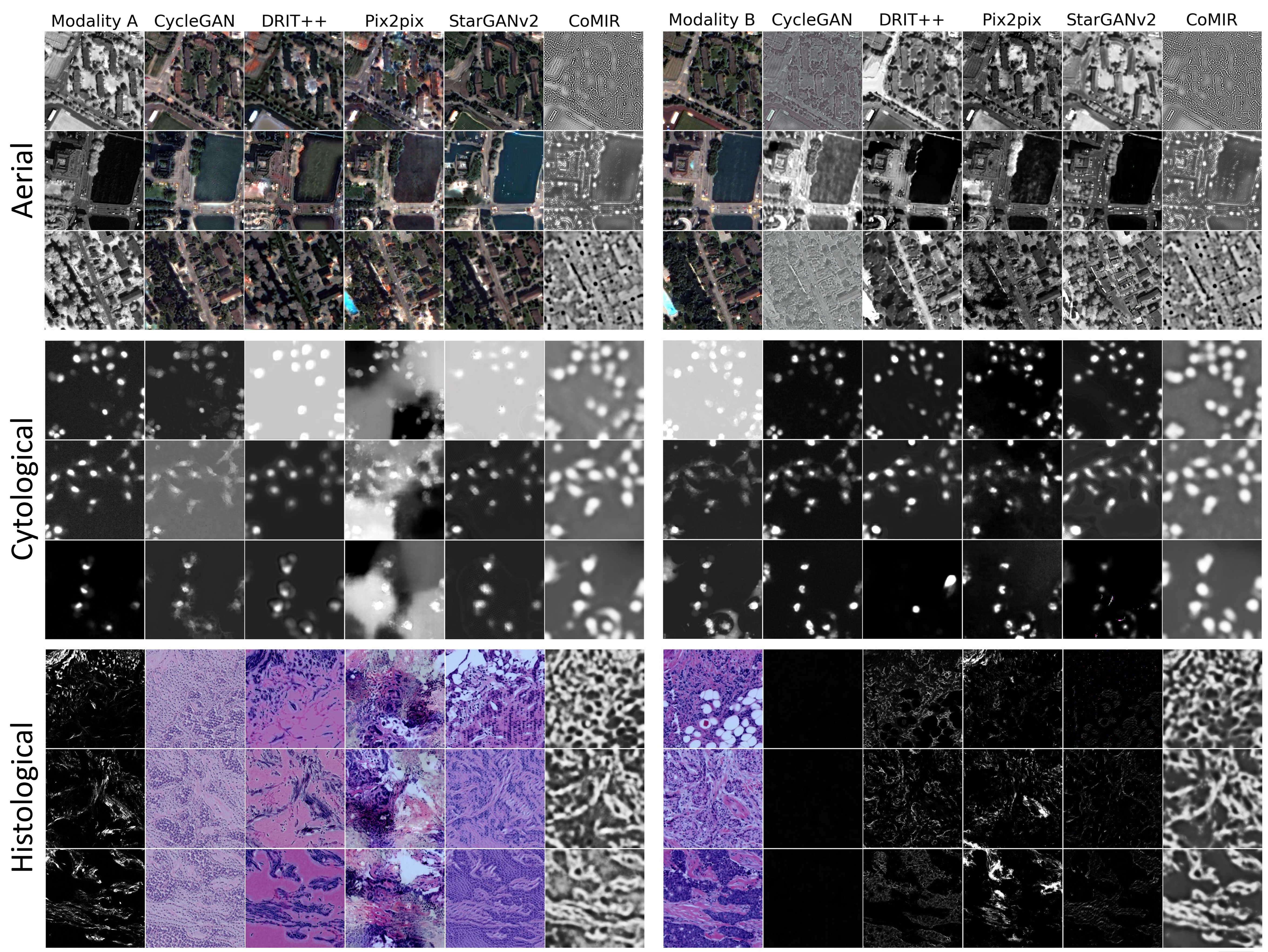


Table: Overall success rate (success: relative registration error $\delta < 2\%$)

Dataset	Aerial Data		Cytological Data		Histological Data	
Method	α-AMD	SIFT	α-AMD	SIFT	α-AMD	SIFT
cyc_A	4.9±2.1	66.4±18.8	**71.1±5.8**	24.4±6.2	0	0
cyc_B	65.0±8.4	83.2±3.1	19.2±2.8	17.6±2.5	13.8	0
drit_A	34.8±5.4	38.0±7.9	61.6±16.2	21.6±3.6	1.7	0
drit_B	18.1±3.1	35.4±3.5	21.0±9.0	4.6±1.3	4.7	0
p2p_A	80.2±3.9	**98.3±0.5**	57.9±7.4	8.6±1.2	**28.4**	0
p2p_B	61.5±4.7	85.0±5.0	0.1±0.1	3.8±2.0	0.4	0
star_A	64.0±7.5	6.5±2.7	57.4±13.0	10.9±2.2	2.6	0
star_B	41.1±3.6	5.9±0.5	17.8±4.9	5.8±0.6	19.6	0
comir	91.8±7.7	**100.0±0.0**	68.0±14.0	72.5±7.1	**81.3**	59.3
B2A	12.8±3.5	72.5±4.8	21.9±10.5	20.8±2.0	0	0
MI_B2A	69.1±3.7		**89.9±3.0**		47.8	
CA_B2A					3.7	

Conclusion

- Popular I2I translation methods show high instability and data dependence, especially when modalities differ considerably
- I2I translation quality (measured by FID[8]) shows to be a reasonably reliable predictor of the success of subsequent monomodal registration
- The supervised representation-learning approach exhibits overall best performance
- An open-source quantitative evaluation framework for multimodal biomedical registration, including all method implementations, evaluation code, and all datasets

References:
[1] P. Isola et al., "Image-To-Image Translation With Conditional Adversarial Networks," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1125–1134.
[2] J.-Y. Zhu et al., "Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks," in 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 2242–2251.
[3] H.-Y. Lee et al., "DRIT++: Diverse Image-to-Image Translation via Disentangled Representations," Int J Comput Vis, Feb. 2020.
[4] Y. Choi et al., "StarGAN v2: Diverse Image Synthesis for Multiple Domains," in 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 8185–8194.
[5] N. Pielawski et al., "CoMIR: Contrastive multimodal image representation for registration," in Advances in neural information processing systems, 2020, vol. 33, pp. 18433–18444.
[6] A. Keikhosravi et al., "Intensity-based registration of bright-field and second-harmonic generation images of histopathology tissue sections," Biomed. Opt. Express, vol. 11, no. 1, p. 160, Jan. 2020.
[7] A. Keikhosravi et al., "Non-disruptive collagen characterization in clinical histopathology using cross-modality image synthesis," Communications Biology, vol. 3, no. 1, pp. 1–12, Jul. 2020.
[8] M. Heusel et al., "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," in Advances in Neural Information Processing Systems, 2017, vol. 30, pp. 6626–6637.