Registration of Spatial Transcriptomics Images to an Atlas using Implicit Neural Representations

Meghane Decroocq[1] MEGHANE.DECROOCQ@RIKEN.JP
[1] *Brain Image Analysis Unit, Center for Brain Science, RIKEN, Japan*
Kohei Onishi [2] KOHEI.ONISHI@RIKEN.JP
[2] *Laboratory for Molecular Mechanisms of Brain Development, Center for Brain Science, RIKEN, Japan*
Tomomi Shimogori [2] TOMOMI.SHIMOGORI@RIKEN.JP
Henrik Skibbe[1,3] HENRIK.SKIBBE@RIKEN.JP
[3] *Department of Informatics, Faculty of Informatics, Matsuyama University, Japan*

Editors: Under Review for MIDL 2026

Abstract

Spatial transcriptomics enables the visualization of gene expression in ex vivo samples, offering unique insights into the brain structure and function. However, downstream analysis requires to match gene expression patterns to specific brain regions by aligning highly deformed 2d brain slices with an atlas. In this work, we propose a registration method using Implicit Neural Representations (INR) specifically designed to address the challenges presented by spatial transcriptomics data.

Keywords: Implicit Neural Representations, Registration, Spatial Transcriptomics

1. Introduction

Spatial transcriptomics allows to visualize the expression of genes in spatial context. In particular, the MERFISH technology, by enabling to image hundreds of genes simultaneously, have contributed greatly to advance our understanding of the brain of small mammals like mice (Du et al., 2023). However, the imaging process requires to slice the brain ex vivo into 2d sections, often resulting in tissue tearing and large deformations. In this context, the registration to a reference atlas becomes a crucial step in the post-processing pipeline. Gene expression occurs in specific regions of the brain, highlighting specific structures. This provides rich information that could support the registration, but manually processing the different gene maps is burdensome.

Although the adoption of spatial transcriptomics is accelerating, only a few methods in the literature address the problem of its registration. (Yao et al., 2023; Tustison et al., 2025) proposed a pipeline using the Advanced Normalization Tools (ANTs) to perform the registration. However, this approach requires the creation of a specific target image by matching genes to brain regions. Meanwhile, recent advances in deep learning, particularly Implicit Neural Representations (INRs), have emerged as promising solutions for registration tasks (Wolterink et al., 2022), and have been successfully applied to the registration of in-situ hybridization images (Byra et al., 2023). In this work, we propose an INR model that automatically extracts the necessary information for registration from the different gene maps, removing the need for prior knowledge of gene expression patterns.

2. Methods

We train a registration INR to output a displacement vector Δc from a set of coordinates c sampled from random patches. The deformation field ϕ, which warps the moving image to the atlas reference space, is defined as $\phi(c) = c + \Delta c$. A key feature of our model is that instead of using a single image, the moving image is a linear combination of n gene expression maps, with learnable parameters $(a_1, ..., a_n)$ optimized during training. It allows the model to reconstruct a moving image that resembles the target atlas, thereby facilitating the registration. We train the INR using three loss functions. L_{sim} is a similarity loss between the fixed image and the moved image. It is composed of a Normalized Cross Correlation (NCC) term and a local NCC term calculated on sliding patches, similarly to (Byra et al., 2023). We regularize the deformation field ϕ by a loss L_{reg} based on the determinant of its Jacobian J_ϕ. Finally, we use the symmetry between brain hemispheres as a prior to guide the model: we compute the L_{sym} loss as the NCC between the moved image patch and its symmetrical counterpart. The total loss is written as follows: $L = \alpha L_{sim} + \beta L_{reg} + \gamma L_{sym}$. The loss equations are given in Appendix A.

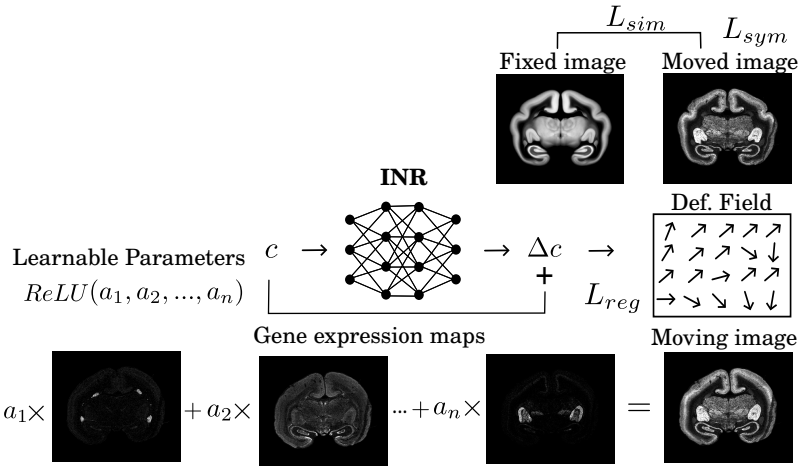


Figure 1: Overview of the proposed INR registration model.

3. Evaluation

Dataset. We tested our model on three 2d slices of the marmoset brain acquired at our institute, including one whole-brain slice of the neonate marmoset, and two half-brain slices of the adult marmoset. The data include a DAPI image, and a set of gene expression maps obtained by MERFISH (300 - 960 genes). We manually annotated the images by segmenting different brain regions based on the gene maps and DAPI image.

Parameters. The images are pre-aligned to the atlas by affine registration using manual landmarks, and normalized between 0 and 1. We compare our approach to the registration obtained by ANTsPy SyN method on three images: the DAPI, the "best" pan-neuronal gene, and an averaged image of all gene maps and DAPI. Our INR model is trained in a multiscale fashion by sampling coordinates from dilated patches (32×32), across three scales: 4, 2, and 1. We optimized the loss parameter β by grid search for each dataset.

Table 1: Evaluation metrics for different registration methods. The metrics are averaged over 8 labeled regions, and over 5 runs for our model. We masked the right hemisphere for the adult marmoset data.

model	image	$\lvert J_\phi\rvert < 0$	Dice	HD
Neonate Marmoset (1 slice)				
affine	manual landmarks	0.0	0.65	17.93
ANTs SyN	DAPI	0.03	0.5	19.5
ANTs SyN	avg	0.02	0.7	16.74
ANTs SyN	best gene (BMERB1)	0.01	0.72	16.5
our model ($\alpha = 1$, $\beta = 2.5$, $\gamma = 0.5$)	all	0.06	**0.8**	**16.21**
Adult Marmoset (2 slices)				
ANTs SyN	best gene (ATP1A1)	0.0±0.0	0.75±0.04	17.02±0.57
our model ($\alpha = 1$, $\beta = 0.5$, $\gamma = 0$)	all	0.04±0.02	**0.82±0.05**	**13.9±1.49**

Results. In Table 1, we report the Dice and the Hausdorff distance between the warped segmentation and the ground-truth segmentation based on the atlas, as well as the percentage of folding ($|J_\phi| < 0$) in the deformation field. ANTs performs the best when a pan-neuronal gene is used for the registration. Our model shows an improvement over ANTs registration in terms of segmentation accuracy, while maintaining a low percentage of folding in the deformation field. We provide a visualization of the registration results in Figure 2. Furthermore, the proposed model allows us to identify the genes that are essential to the registration by analyzing the weights attributed to the different gene maps, as we illustrate in Appendix B.

4. Conclusion

We proposed a method to register spatial transcriptomics images to an atlas using INRs, showing promising results on marmoset brain data. Nevertheless, our model remains to be tested on a larger dataset to draw reliable conclusion on its performance. Besides, the current 2d version of the model requires to manually select the corresponding coronal slice from the atlas. In future work, we will investigate direct 2d to 3d registration.

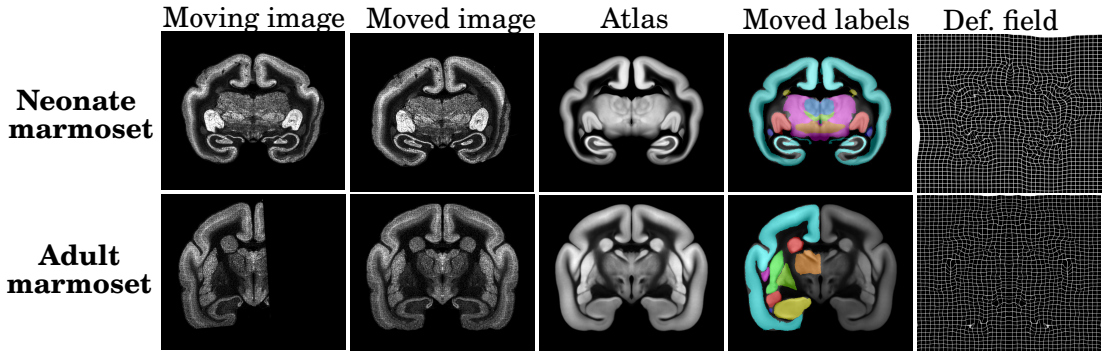


Figure 2: Visualization of the proposed model. For the adult slices, the moved image and deformation field were obtained by mirroring the left part of the output.

Acknowledgments

This work was founded by the Multidisciplinary Frontier Brain and Neuroscience Discoveries (Brain/MINDS 2.0) from the Japan Agency for Medical Research and Development AMED JP23wm0625001.

References

Michal Byra, Charissa Poon, Tomomi Shimogori, and Henrik Skibbe. Implicit neural representations for joint decomposition and registration of gene expression images in the marmoset brain. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 645–654. Springer, 2023.

Jun Du, Yu-Chen Yang, Zhi-Jie An, Ming-Hui Zhang, Xue-Hang Fu, Zou-Fang Huang, Ye Yuan, and Jian Hou. Advances in spatial transcriptomics and related data analysis strategies. *Journal of translational medicine*, 21(1):330, 2023.

Nicholas J Tustison, Min Chen, Fae N Kronman, Jeffrey T Duda, Clare Gamlin, Mia G Tustison, Michael Kunst, Rachel Dalley, Staci Sorenson, Quanxin Wang, et al. The antsx ecosystem for mapping the mouse brain. *Nature communications*, 2025.

Jelmer M Wolterink, Jesse C Zwienenberg, and Christoph Brune. Implicit neural representations for deformable image registration. In *International Conference on medical imaging with deep learning*, pages 1349–1359. PMLR, 2022.

Zizhen Yao, Cindy TJ Van Velthoven, Michael Kunst, Meng Zhang, Delissa McMillen, Changkyu Lee, Won Jung, Jeff Goldy, Aliya Abdelhak, Matthew Aitken, et al. A high-resolution transcriptomic and spatial atlas of cell types in the whole mouse brain. *Nature*, 624(7991):317–332, 2023.

Appendix A. Loss functions

The similarity loss function is written as:

$$L_{sim}(F, P, c) = -NCC(F(c), M(\phi(c))) - LNCC(F(c), M(\phi(c))), \tag{1}$$

where F denotes the fixed image, M the moving image, ϕ the deformation field and c the coordinates sampled from a random patch. NCC is the Normalized Cross Correlation. The Local Normalized Cross Correlation $LNCC$ is calculated as the mean of NCC on sliding patches of size 9×9.

The regularization function L_{reg} is written as follows:

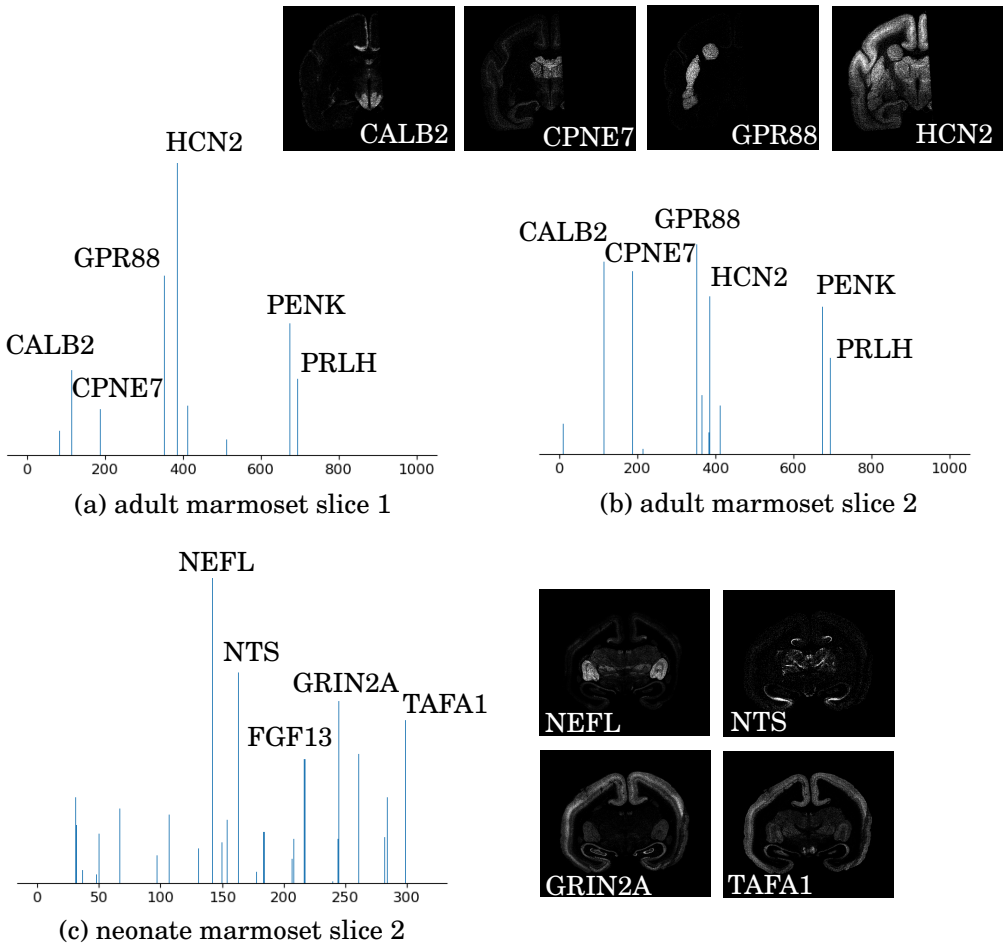
$$L_{reg}(\phi, c) = \frac{1}{n} \sum_{i=1}^{n} (|1 - |J_{\phi(c_i)}|), \tag{2}$$

with J_ϕ the Jacobian of the deformation field ϕ. The symmetrical prior loss L_{sym} is written as :

$$L_{sym}(M, \phi, c, c') = -NCC(M(\phi(c)), M(\phi(c'))), \tag{3}$$

where c' denotes the symmetrical counterpart of the coordinates c.

Appendix B. Weight analysis



(a) adult marmoset slice 1

(b) adult marmoset slice 2

(c) neonate marmoset slice 2

Figure 3: Bar plot illustrating the weights assigned to each gene map by the model to generate the moving image. Weights are shown for two adult marmoset slices (a) and (b), as well as a neonate slice (c). The model relies on a small number of genes (10) to perform the registration. This limited selection makes it easier to include the genes in a MERFISH experiment and can help reduce memory requirements by reducing the number of channels used during the training of the INR. The genes used for the adult marmoset are similar for both slices, showing the robustness of the gene selection process.