MELBA Journal
Machine Learning
for Biomedical Imaging

Quantifying the Efficacy of Deep Learning-Driven Deformable Registration in Multiplexed-Immunofluorescence Imaging for Nucleus Subtype Classification

Gaurav **Rudravaram** [1], Shunxing Bao [1], Lucas W. Remedios [2], Aravind R. Krishnan [1]. Michael E. Kim [2], Yihao Liu [2], Chenyu Gao [1], Rendong Zhang [2], Bohan Jiang [1], Qi Liu [3,4], Ken S Lau [3,5,6,7], Joseph T. Roland [6], Mary K. Washington [8], Lori A. Coburn [5,9,10,11], Keith T. Wilson [5,9,10,11,12], Yuankai Huo [2], Bennett A. Landman [1,2,13]

1 Department of Electrical and Computer Engineering, Vanderbilt University, Nashville, TN, USA
2 Department of Computer Science, Vanderbilt University, Nashville, TN, USA
3 Center for Quantitative Sciences, Vanderbilt University Medical Center, Nashville, TN, USA
4 Department of Biostatistics, Vanderbilt University Medical Center, Nashville, USA
5 Vanderbilt Center for Mucosal Inflammation and Cancer, Vanderbilt University Medical Center, Nashville, TN, USA
6 Epithelial Biology Center, Vanderbilt University Medical Center, Nashville, TN, USA
7 Department of Cell and Developmental Biology, Vanderbilt University School of Medicine, Nashville, TN, USA
8 Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville, TN, USA
9 Division of Gastroenterology, Hepatology, and Nutrition, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA
10 Program in Cancer Biology, Vanderbilt University School of Medicine, Nashville, TN, USA
11 Veterans Affairs Tennessee Valley Healthcare System, Nashville, TN, USA
12 Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville, TN, USA
13 Department of Biomedical Engineering, Vanderbilt University, Nashville, TN, USA

Abstract

Accurate identification and classification of nucleus subtypes is crucial for cell tracking and uncovering patterns across cell types, such as local cell neighborhoods. Multiplexed immunofluorescence (MxIF) imaging is a process that involves staining, imaging, and then bleaching the same tissue multiple times. Repeating MxIF staining with different marker combinations enables subclassification of cells. However, repeated cycles of staining and bleaching can cause deformation, movement, and tissue loss, resulting in misalignment of markers at the nucleus level. This misalignment can lead to the exclusion of a significant number of cells during downstream analysis. We propose that applying a post hoc deep learning-based deformable registration technique (VoxelMorph) on the respective 4′,6-diamidino-2-phenylindole (DAPI) image for each round of staining can reduce the number of nuclei that are excluded due to spatial misalignment across successive staining rounds. By applying the registration transformations from different DAPI rounds to their corresponding stains, we achieve stain registration at pixel-level. To tackle the challenge of large image sizes, we propose a patch-based training and inference strategy. By analyzing residual displacement from bidirectional registrations, we are able to mask out areas in the tissue with high residual displacement to indicate image regions that should not be included for downstream analyses. For validation, we used a deterministic decision tree, based on biological domain knowledge, to classify MxIF nuclei into either one of 13 different classes or an undefined class. Our proposed registration approach effectively reduced the number of undefined nuclei, and we observed a 17.6% increase in the number of successfully classified nuclei compared to a baseline rigid registration. Our code is available at `https://github.com/MASILab/MxIF_Registration`.

Keywords

Machine Learning, Image Registration, Cell Type Classification

Article informations

`https://doi.org/https://doi.org/10.59275/j.melba.2026-912a`

Volume 2026, Received: 2026-02-27, Published 2026-04-05

Corresponding author:Gaurav Rudravaram gaurav.rudravaram@vanderbilt.edu

1. Introduction

Multiplexed Immunofluorescence (MxIF) is a pivotal imaging technique designed for the simultaneous detection of multiple biomarkers by leveraging a cyclic staining process where fluorescently labeled antibodies are applied, imaged, and subsequently quenched, enabling successive rounds of staining and imaging within the same sample (Lin et al., 2018). By allowing for the visualization of a broad array of proteins, MxIF provides a multifaceted view of biological interactions within tissues, enhancing our understanding of cellular and molecular dynamics critical for investigating pathological states and therapeutic interventions(Goltsev et al., 2018). The layered complexity made possible by MxIF is particularly invaluable in fields like oncology and immunology, where understanding spatial relationships and protein co-localization is essential(Schubert et al., 2006; Angelo et al., 2014). Each marker and 4′,6-diamidino-2-phenylindole (DAPI) image acquired during the MxIF imaging process is itself a whole slide image (WSI), capturing the entire tissue section stained for a specific target. These individual WSIs provide detailed information about the expression of specific proteins or the location of nuclei. Through the combination of these multiple WSIs, each corresponding to a different marker or nuclear stain, MxIF reveals a comprehensive picture of the tissue's biological landscape.

Multiplexed immunofluorescence experiments typically involve sequential rounds of staining and imaging the same tissue section, necessitating image alignment across cycles. In practice, most studies still rely on rigid or affine transformations to register these multi-round images into one composite dataset(Ghahremani et al., 2023). A common strategy is to include a repeat nuclear stain (e.g. DAPI) in every round and use it as the reference channel for alignment(Vizcarra et al., 2022). This approach assumes that only global shifts or linear scaling occur between rounds. For example, the cyclic immunofluorescence method (t-CyCIF) aligns each imaging cycle by applying a rigid-body transform based on the DAPI channel, effectively stacking 16 rounds of staining with only slide wise rotation/translation adjustments(Lin et al., 2018). Likewise, other multiplex protocols explicitly instruct aligning images via a linear transform using the recurrent nuclear or structural marker present in all cycles(Muñoz-Castro et al., 2022).

Numerous recent studies underscore the dominance of rigid/affine registration in MxIF. For instance, Lin et al. (2018) and Muñoz-Castro et al. (2022) employed nuclear markers from each cycle to compute rigid-body transformations, co-registering all cycles into a hyper-stack. Similarly, in a recent multiplex IHC versus IF comparison, an affine registration—allowing translation, rotation, and scaling—was used to accurately align sequential scans and merge all markers(Ghahremani et al., 2023). Muhlich et al. (2022) performed registration of multiplexed WSI by dividing images into overlapping patches and registering each individual patch. Their registration used a phase-correlation algorithm proposed by Guizar-Sicairos et al. (2008), however this method only accounted for translation and is not a complete rigid registration. Gerdes et al. (2013) rigidly aligned all rounds together using the nuclei in the different DAPI images as fiducial points. Similarly, Lin et al. (2015, 2018) performed their registration on cyclic immunofluorescence images using Hoechst as the nuclear stain and applying the transformation on other markers. Open-source tools and pipelines further reinforce this trend: ASHLAR (Alignment by Harmonization of Layers)(Muhlich et al., 2022) selects a representative nuclear channel (e.g., DAPI) to perform adjacent tile and inter-cycle alignment via affine registrations, and the MCMICRO pipeline(Schapiro et al., 2022) employs ASHLAR as its default module for generating stitched mosaic images.

Despite the robustness and simplicity of rigid and affine registrations, these methods inherently assume that only global shifts occur and are unable to correct nonlinear tissue deformations. Factors such as tissue swelling, shrinking, or slight stretching between cycles can introduce local misalignments that a single global transform cannot fully resolve. Each staining cycle has the potential to slightly alter tissue morphology, resulting in nuclei misalignment, tissue deformation, and missing tissue (Figure 1), which cannot be accounted for by rigid registration alone. The accumulation of such artifacts across multiple staining cycles can significantly degrade image quality. In previous works, Bao et al. (2022) have introduced a biology-informed Q-score that assesses the quality of MxIF images, while Jiang et al. (2023) have used the consistency of 4′,6-diamidino-2-phenylindole (DAPI) staining as a benchmark to evaluate the alignment and integrity of tissue sections.

2. Related Works

To effectively manage the multifaceted challenges presented by MxIF, it is essential to develop a robust method that not only identifies and excludes areas of significant tissue deformation but also aligns the markers from all staining rounds into a unified space. Deformable registration is especially effective for this task because it uses nonlinear transformations to align images from different staining rounds,

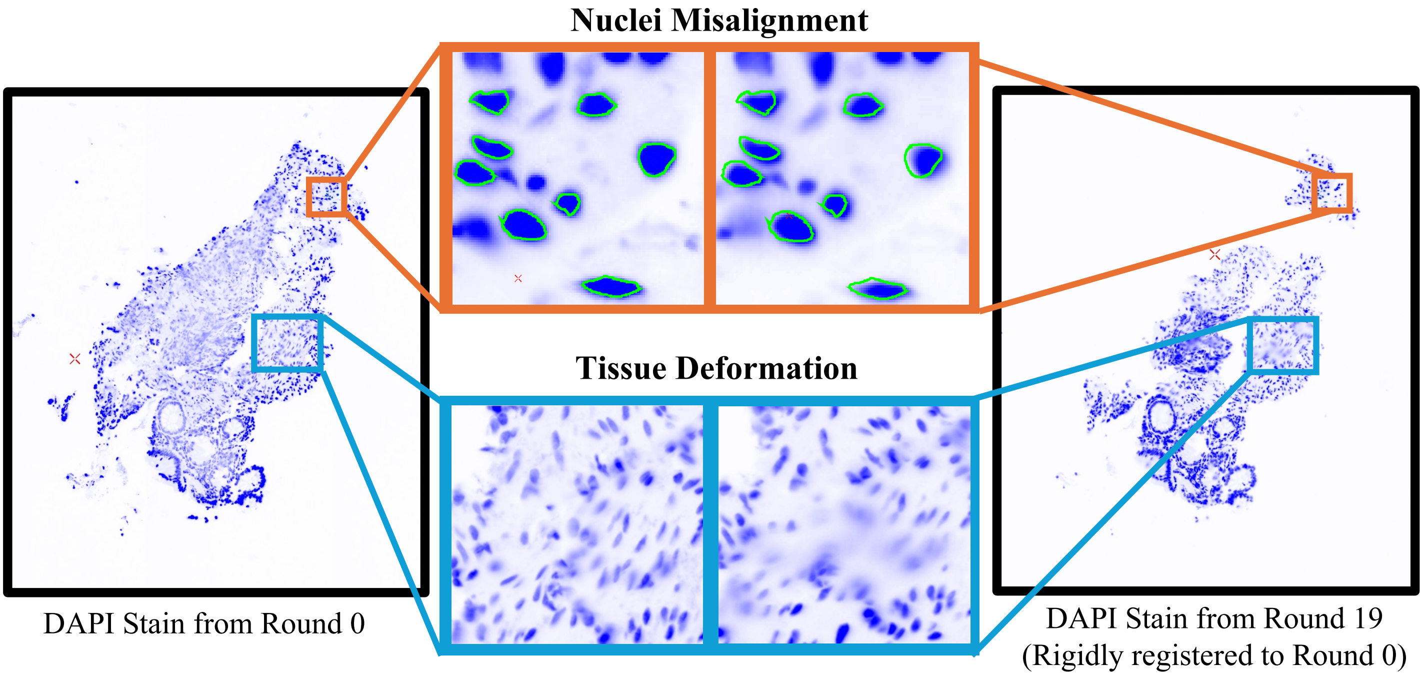

Figure 1: Due to repeated rounds of staining and bleaching, some portions of the tissue are lost or deformed in the later rounds. Moreover, despite the images being rigidly registered, local displacements in nuclei positions persist. **Top panel:** Green markers manually annotated around nuclei in round 0 do not align with the nuclei expression in round 19, demonstrating a misalignment of nuclei positions. **Bottom panel:** In round 0, nuclei markers are clearly expressed with well-defined, intact shapes. By round 19, due to the repeated staining process, these markers spread out, and lose structural integrity, rendering them ineffective for further analysis.

ensuring that the tissue markers are accurately mapped into a unified spatial framework. Deformable registration has been an area of intense focus in medical image processing, evolving from initial methods aimed at optimizing an energy function for pairs of moving and fixed images(Klein et al., 2009), to more advanced deep learning techniques that learns a universal function for estimating deformation fields(Haskins et al., 2020; Fu et al., 2020; Hering et al., 2022). However, the application of deformable registration techniques to WSIs are often limited due to large image size and high computational complexity(Borovec et al., 2018). Traditional registration methods typically commence with affine registration to correct major misalignments and rotations, followed by deformable registration that fine-tunes the alignment by addressing smaller tissue discrepancies. These methods commonly use techniques such as B-spline or feature matching(Gatenbee et al., 2023; Shafique et al., 2021; Wodzinski et al., 2024). Roy et al. (2023) have proposed a novel deep learning-based approach specifically for co-registering paired immunohistochemistry (IHC) and hematoxylin and eosin (H&E) images using a synthesis network for stain conversion, coupled with a registration network that calculates the deformation field necessary for their co-registration. While these advances have predominantly focused on H&E and IHC images, non-rigid registration has had a slow adoption in MxIF due to the considerable challenges and unique necessities required by histology slides, as discussed by Borovec et al. (2020).

In this study, we characterize the impact of deformable registration on MxIF image alignment relative to the standard rigid registration pipeline. Rather than proposing a brand-new registration method, we apply an existing deep learning-based deformable registration approach VoxelMorph framework as described by Balakrishnan et al. (2019) to estimate deformation fields between sequential DAPI images, subsequently propagating these corrections to associated marker channels. We also introduce a strategy to identify and exclude areas of significant tissue deformation by analyzing residual displacement fields. Our evaluation focuses on the downstream effects of improved alignment on cell-type classification and the preservation of essential biological information across imaging rounds, providing a systematic characterization of the benefits and limitations of deformable registration in the context of MxIF.

3. Methods

3.1 Dataset

We conducted our analysis using anonymized samples obtained from Vanderbilt University Medical Center, adhering

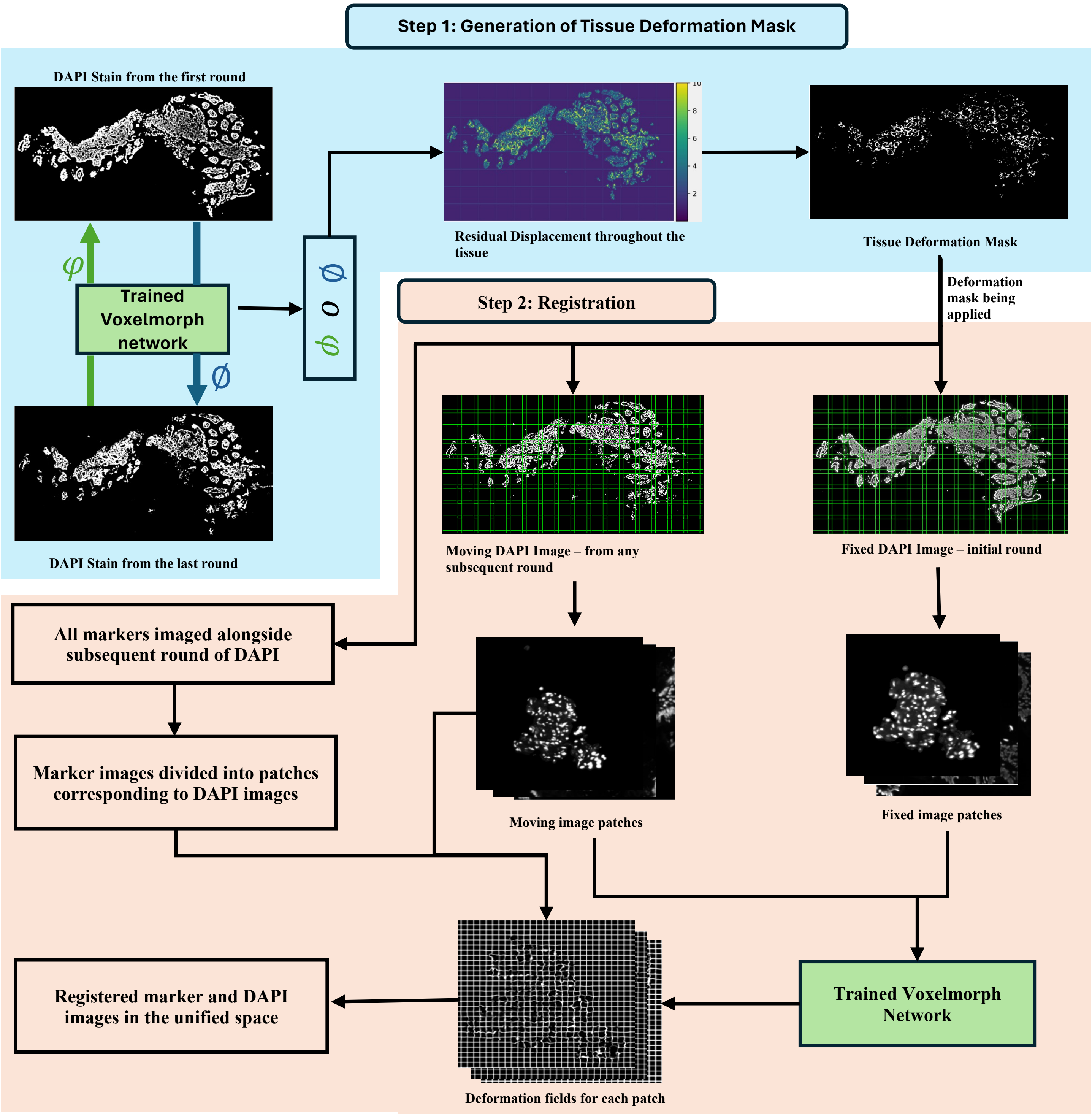

Figure 2: Workflow of the registration process. The initial step involves calculating the composition of bidirectional fields between the first and last rounds of DAPI staining, followed by generating a mask to exclude regions of high residual displacement. In the second step, DAPI images and corresponding markers are segmented into patches. Registration is then calculated between each round of DAPI and the initial round, and the resulting deformations are applied to the marker patches after applying the tissue deformation mask on all images. The second step is repeated for all rounds of staining, aligning all the DAPI images and corresponding markers into a unified spatial framework.

to ethical guidelines approved by the Institutional Review Board (IRB #191738 and #191777). Each sample was formalin-fixed and paraffin-embedded. The dataset featured tissue samples from two intestinal regions: the ascending colon (AC) and the terminal ileum (TI). Fluorescent, immunostained tissue slides were scanned using an Aperio Versa 200 automated slide scanner (Leica Biosystems), with tissue cores imaged at 20× magnification, yielding a spatial resolution of 0.323 μm per pixel. Our study used 53 WSIs captured at a resolution of 0.324 microns per pixel in 20× magnification, derived from 31 subjects. Our imaging protocol included 17 to 19 staining rounds per sample (across from 4 MxIF batches), starting each round with a DAPI stain for nuclear visualization followed by three specific markers. Each marker, along with the DAPI, is imaged in a separate channel. The markers and DAPI are alternately stained and imaged every other round. In the rounds without marker staining, background elements were imaged across different channels to capture autofluorescence, enhancing our ability to effectively remove background during post-processing. The dataset includes images acquired across four different MxIF batches, reflecting realistic inter-batch acquisition variability.

3.2 Preprocessing

Prior to training the neural network, we performed a series of preprocessing steps to ensure optimal image quality and compatibility with our analytical framework. In the preprocessing stage, all images, initially represented as unsigned 8-bit grayscale with pixel intensity values ranging from 0 to 255, underwent normalization to a standardized floating-point scale. This is achieved by dividing each pixel value by 255, effectively rescaling the intensity values to the range between 0 and 1. Following normalization, the images underwent a padding process. The amount of zero-padding applied was calculated to enable the division of images into patches of 1024×1024 pixels, while ensuring an overlapping region of 200 pixels between adjacent patches. This overlap was crucial for maintaining continuity and alignment across the image patches during the registration process, facilitating accurate analysis and comparisons.

3.3 Rigid Registration of MxIF Images as a Baseline for Comparative Analysis

Following the MxIF image processing protocol outlined by McKinley et al.McKinley et al. (2017), we performed a preliminary rigid registration to align all the MxIF images with the first round. Rigid alignment was formulated as a two-dimensional transformation consisting of an in-plane rotation and translation to standardize the orientation and position of each subsequent DAPI image relative to the reference (Round 1). Transformation parameters were estimated using MATLAB's intensity-based image registration function `imregtform` with the transformation type set to rigid, optimizing the mean squared error between the fixed and moving DAPI images. The resulting transformation, represented as a `rigid2d` object, was then applied to both the DAPI images and the corresponding marker images acquired in the same round to rigidly align all imaging rounds.

3.4 Deformable registration network and workflow

We trained our VoxelMorph(Balakrishnan et al., 2019) network using data derived from 10 MxIF images. The WSIs used for training were randomly sampled across three of the four imaging batches. These MxIF WSI were segmented into 1024×1024 patches, with the fixed images being patches from the initial round and the moving images being patches from the last round. We selected only those patches where at least 50% of the area was non-zero, effectively disregarding patches that contained zero intensities throughout, as these areas outside the tissue in WSIs typically show zero intensities and contribute no meaningful information to the network. After this selection, we obtained a total of 1316 patches, which were then split into training and validation sets with an 80:20 ratio. The trained network was subsequently tested on the entire dataset of 53 MxIF images to compare the number of labeled nuclei. The network optimized a single objective function,

$$Loss = - \sum_{p \in \Omega} \left(\frac{\left(\sum_{p_i} (f(p_i) - \hat{f}(p)) (w(p_i) - \hat{w}(p)) \right)^2}{\left(\sum_{p_i} (f(p_i) - \hat{f}(p))^2 \right) \left(\sum_{p_i} (w(p_i) - \hat{w}(p))^2 \right)} \right) + \lambda \sum_{p \in \Omega} \|\nabla u(p)\|^2, \quad (1)$$

where Ω is the 2-dimensional spatial domain, f is the fixed image and w is the warped image defined over Ω, $u(p)$ is the displacement that is applied on the moving image to produce the warped image and λ is the weight associated to the regularization of the displacement field. The first component of the loss function is the local cross-correlation, where $\hat{f}(p) = \frac{1}{n^2} \sum_{p_i} f(p_i)$ and $\hat{w}(p) = \frac{1}{n^2} \sum_{p_i} w(p_i)$ represent the local mean intensities of the fixed and warped image such that p_i iterates over an area of n^2 around pixel p. The second component was the L2 regularization term of the displacement field produced by the network. This term encouraged the network to generate a smooth deformation field. In the original implementation, the weight assigned to the smoothness loss was 0.01. The weight on the smoothness loss was set as 1, and we used a window size of 128 for the local cross-correlation. The model was trained on Nvidia Quadro RTX 5000 for a total of 25 epochs. The optimizer chosen was Adam(Kingma and Ba, 2014) and the learning rate was set to 0.001.

The workflow for registering all markers into a common space was done in two parts (Figure 2). The first part of the process is to remove areas of tissue with significant deformation. For this, we compare the first and last rounds of DAPI-stained images. The network first generates a deformation field (ϕ) mapping the first round of DAPI to the last. Subsequently, a reverse deformation field (φ) is computed by registering the last round back to the first. By composing these deformation fields ($\phi \circ \varphi$), we calculated the L2 norm of the resultant field to measure residual displacement per pixel across the tissue. We then eliminated regions exhibiting a residual displacement exceeding 10 pixels (Supplementary material 2), to ensure that only minimally deformed areas are retained for further analysis. This design choice was motivated by the nature of multiplexed immunofluorescence imaging, in which cell-type identification relies on the combined expression of markers acquired across multiple staining rounds. Consequently, only cells that remain intact and consistently present from the first to the final staining round can be reliably classified. While intermediate rounds can provide information about round-to-round deformation, our objective was to preserve cells suitable for downstream cell-type classification across all markers. Therefore, we used the first and final rounds of DAPI staining to identify and exclude regions affected by cumulative tissue deformation, as deformation artifacts can accumulate over successive staining and imaging rounds.

The next step in the workflow was to register all elements, including stains, background, and autofluorescence images, into one common space. Each DAPI image was subdivided into overlapping patches of size 1024×1024 pixels, with a 200-pixel overlap between adjacent patches. For each patch, the network computed a deformation field that was applied to warp the corresponding image patch into the reference (first-round) space. During reconstruction of the full-resolution registered image, overlapping regions between neighboring warped patches were combined by averaging the spatially averaging the overlapping registered patches(Atkins et al., 2025), which reduces boundary artifacts and ensures smooth transitions across patch boundaries. The deformation fields, obtained from the DAPI registration, were then applied to all other elements captured in conjunction with that round, including other stains and background images. This process was repeated for all subsequent rounds of imaging to align all data across the round. After registering everything into a unified space, autofluorescence removal and background subtraction were performed. We made a design choice to use deep learning based deformable registration for this analysis after evaluating traditional registration methods. Specifically, we found that applying ANTs registration at the WSI scale was computationally extensive. Given the gigapixel nature of MxIF images and the presence of 17–19 rounds of staining and bleaching per image, the total computation time for registering all rounds of a single MxIF image using ANTs ranged from 24 to 48 hours depending on image size. In contrast, since Voxelmorph is a deep learning based model, during inference it takes 2 to 4 min to per MxIF image across all rounds.

3.5 Experimental Design

To standardize the variability observed in DAPI stain intensities across different imaging rounds, we calculated the intensity ratio between the DAPI stain from the initial imaging round and those from all registered rounds (γ_n). This ratio is used to adjust the intensities of the registered DAPI images to match those of the original round. Each registered image is then multiplied by its respective intensity ratio, ensuring that the intensities are approximately equal to the baseline round. After this adjustment, we averaged these normalized DAPI images to create a unified DAPI image that represents all rounds post-registration (D_{avg})

$$D_{avg} = \frac{D_0 + \sum_{n=1}^{N} \gamma_n D_n}{N+1} \tag{2}$$

where N is the total number of subsequent rounds of DAPI staining and D_0 is the DAPI image from the initial round. Using this unified DAPI image, we then employed the DeepCell Mesmer model(Greenwald et al., 2022) to perform instance segmentation of nuclei. This segmentation allowed us to calculate the mean intensities of all markers for each segmented nucleus accurately. Following segmentation, we analyzed the intensities to determine if they were positive or negative for each stain using a semi-automatic marker gating approach called GammaGateR, with the posterior probability threshold for marker expression set at 0.5 as suggested by Xiong et al. (2024).

Post-segmentation, each nucleus was annotated into one of 13 sub-types based on a biology-informed decision tree previously established by our group(Remedios et al., 2024b,a). We used 15 of the 27 markers available in our data for nucleus subtype classification: NaKATPase, PanCK, CgA, Vimentin, SMA, Sox9, OLFM4, Lysozome, CD45, CD20, CD68, CD11B, CD3d, CD8, and CD4. These markers facilitated the classification of nuclei into the following subtypes: enteroendocrine, enterocytes, fibroblasts, stromal(undetermined), myeloid, helper T, cytotoxic T, $CD3^{+}CD4^{-}CD8^{-}$ T cell nuclei, monocytes, macrophages, B cell nuclei, leukocytes, and progenitor. Nuclei that did not conform to any specified categories were labeled as unknown as per Table 1.

4. Results

Table 1: The nuclei in MxIF were classified into 13 distinct subtypes through a sequential deterministic decision tree and the nuclei which did not confirm to the above subtypes were labelled as undefined.

Step	Purpose	Stain Combinations
1	Group Epi^{+} nuclei	NaKATPase^{+} or PanCK^{+} or CgA^{+}
2	Group Stroma^{+} nuclei	Vimentin^{+} or SMA^{+}
3	Exclude nuclei that are both Epi^{+} and Stroma^{+}	Exclude nuclei that are marked as (Epi^{+} and Stroma^{+})
4	Group Immune^{+} nuclei	Group nuclei that are CD45^{+} or CD20^{+} or CD68^{+} or CD11B^{+} or Lysozyme^{+} or CD3d^{+} or CD8^{+} or CD4^{+}
5	Remove immune conflicts for macrophage nuclei across all nuclei	Exclude nuclei where (CD68^{+} and CD3d^{+}), (CD68^{+} and CD20^{+}), (CD68^{+} and CD4^{+}), (CD68^{+} and CD8^{+}), or (CD68^{+} and CD11B^{+})
6	Remove immune conflicts for monocyte nuclei across all nuclei	Exclude nuclei where (CD11B^{+} and CD3d^{+}), (CD11B^{+} and CD20^{+}), (CD11B^{+} and CD4^{+}), (CD11B^{+} and CD8^{+}), or (CD11B^{+} and CD68^{+})
7	Remove immune conflicts for B cell nuclei across all nuclei	Exclude nuclei where (CD20^{+} and CD3d^{+}), (CD20^{+} and CD4^{+}), or (CD20^{+} and CD8^{+})
8	Remove conflicts for helper T and cytotoxic T nuclei across all nuclei	Exclude nuclei where (CD3d^{-} and CD45^{-} and CD4^{+}), (CD3d^{-} and CD45^{-} and CD8^{+}), or (CD4^{+} and CD8^{+})
9	Group Progenitor^{+} nuclei	Sox9^{+} or OLFM4^{+}
10	Exclude nuclei that are not in either the epithelium or stroma	Exclude nuclei that are (Epi^{-} and Stroma^{-})
11	Remove conflicts for endocrine nuclei across all nuclei	Exclude nuclei where (CgA^{+} and Immune^{+}), (CgA^{+} and SMA^{+}), (CgA^{+} and Progenitor^{+})
12	Remove conflicts for fibroblasts across all nuclei	Exclude nuclei that are (SMA^{+} and Immune^{+})
13	Remove conflicts for progenitors across all nuclei	Exclude nuclei where (Immune^{+} and Progenitor^{+})
14	Remove nuclei that are negative for all the large groupings	Exclude nuclei where (Epi^{-} and Stroma^{-} and Progenitor^{-} and Immune^{-})
15	Remove any immune nuclei from Epi^{+} group	Exclude Epi^{+} nuclei where (Epi^{+} and Immune^{+})
16	Final annotation for endocrine cells	Group Epi^{+} nuclei where (CgA^{+} and Progenitor^{-})
17	Final annotation for enterocytes	Group Epi^{+} nuclei where (CgA^{-} and Progenitor^{-})
18	Group stromal/fibroblasts	Group nuclei that are Stroma^{+} and Immune^{-}
19	Final annotation for fibroblasts	Group stromal/fibroblast nuclei where (SMA^{+} and Progenitor^{-})
20	Final annotation for stromal(undefined) nuclei	Group stromal/fibroblast nuclei where (SMA^{-} and Progenitor^{-})
21	Final annotation for myeloid nuclei	Group Immune^{+} nuclei that are ((Lysozyme^{+} and CD68^{-} and CD11B^{-} and Progenitor^{-} and CD20^{-}) and CD3d^{-} and CD8^{-} and CD4^{-}))
22	Final annotation for helper T nuclei	Group Immune^{+} nuclei where (CD4^{+} and Progenitor^{-})
23	Final annotation for cytotoxic T nuclei	Group Immune^{+} nuclei where (CD8^{+} and Progenitor^{-})
24	Final annotation for CD3^{+}CD4^{-}CD8^{-} T cell nuclei	Group Immune^{+} nuclei where (CD3d^{+} and CD4^{-} and CD8^{-})
25	Final annotation for monocyte nuclei	Group Immune^{+} nuclei where (CD11b^{+} and CD3d^{-} and Progenitor^{-} and CD4^{-} and CD8^{-})
26	Final annotation for macrophage nuclei	Group Immune^{+} nuclei where (CD68^{+} and CD3d^{-} and Progenitor^{-} and CD4^{-} and CD8^{-})
27	Final annotation for B cell nuclei	Group Immune^{+} nuclei where (CD20^{+} and CD68^{-} and CD3d^{-} and Progenitor^{-} and CD4^{-} and CD8^{-})
28	Final annotation for leukocyte nuclei	Group Immune^{+} nuclei where (CD45^{+} and CD20^{-} and CD68^{-} and CD3d^{-} and Progenitor^{-} and CD4^{-} and CD8^{-} and CD11B^{-} and Lysozyme^{-})
29	Final annotation for progenitor nuclei	Group all nuclei that are Progenitor^{+}

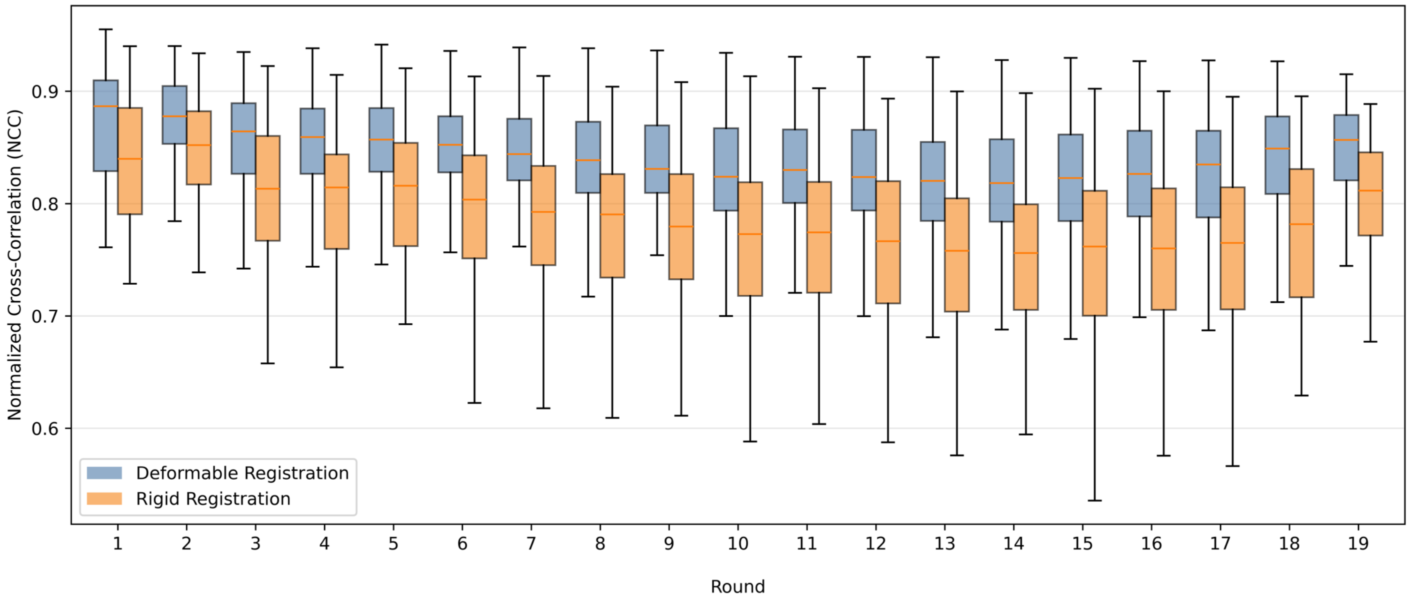

Figure 3: Image similarity quantified using normalized cross-correlation (NCC) between the first-round DAPI image and all subsequent imaging rounds. As repeated staining and imaging introduce cumulative tissue deformation, rigid registration exhibits a progressive decline in similarity, with the lower quartile (25th percentile) dropping to approximately 0.7 in later rounds. In contrast, deformable registration consistently preserves higher similarity across rounds, with the lower quartile remaining near 0.8 even at later stages. Deformable registration achieved significantly higher NCC than rigid registration across rounds ($p < 0.01$).

4.1 Impact of Rigid vs. Deformable Registration on DAPI Image Similarity

We quantified image similarity using NCC between the first-round DAPI image and all subsequent rounds across 43 MxIF images in our dataset, excluding the 10 MxIF images whose patches were used to train the deformable registration network. Deformable registration consistently produced higher NCC values than rigid registration across all imaging rounds (Figure 3), indicating improved alignment between successive DAPI stains. This improvement was statistically significant across rounds ($p < 0.01$, Wilcoxon signed-rank test). Notably, while rigid registration exhibited a pronounced decline in similarity in later rounds due to cumulative tissue deformation, deformable registration maintained higher similarity throughout repeated staining and imaging cycles.

4.2 Evaluation of Rigid vs. Deformable Registration on Cell Annotation

We applied both rigid and deformable registration through our proposed pipeline to our dataset of 53 WSI. The performance of each method was evaluated by comparing the total number of cells annotated, as shown in Figure 4. Visual comparisons revealed that deformable registration typically enhanced the labelled nuclei counts for the majority of the slides. Additionally, the tissue deformation mask excluded $6.65 \pm 4.46\%$ of the WSI area, depending on the extent of deformation and tissue loss observed in the tissues.

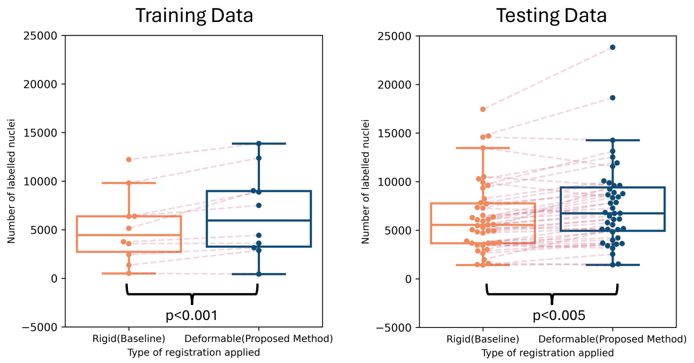

Figure 4: Comparison of number of classified nuclei using rigid registration and deformable registration via the proposed deformable registration pipeline for all 53 MxIF images. We observe an increase in the number of classified nuclei for most images when deformable registration is used, as highlighted by the correspondence lines. A Wilcoxon signed rank test yields a statistically significant difference ($p < 0.01$ on testing data and $p < 0.005$ on the training data) between the two methods.

The variations in increases and decreases in nuclei subtype classification primarily stemmed from nuclei transitioning between undefined and defined states as a result of

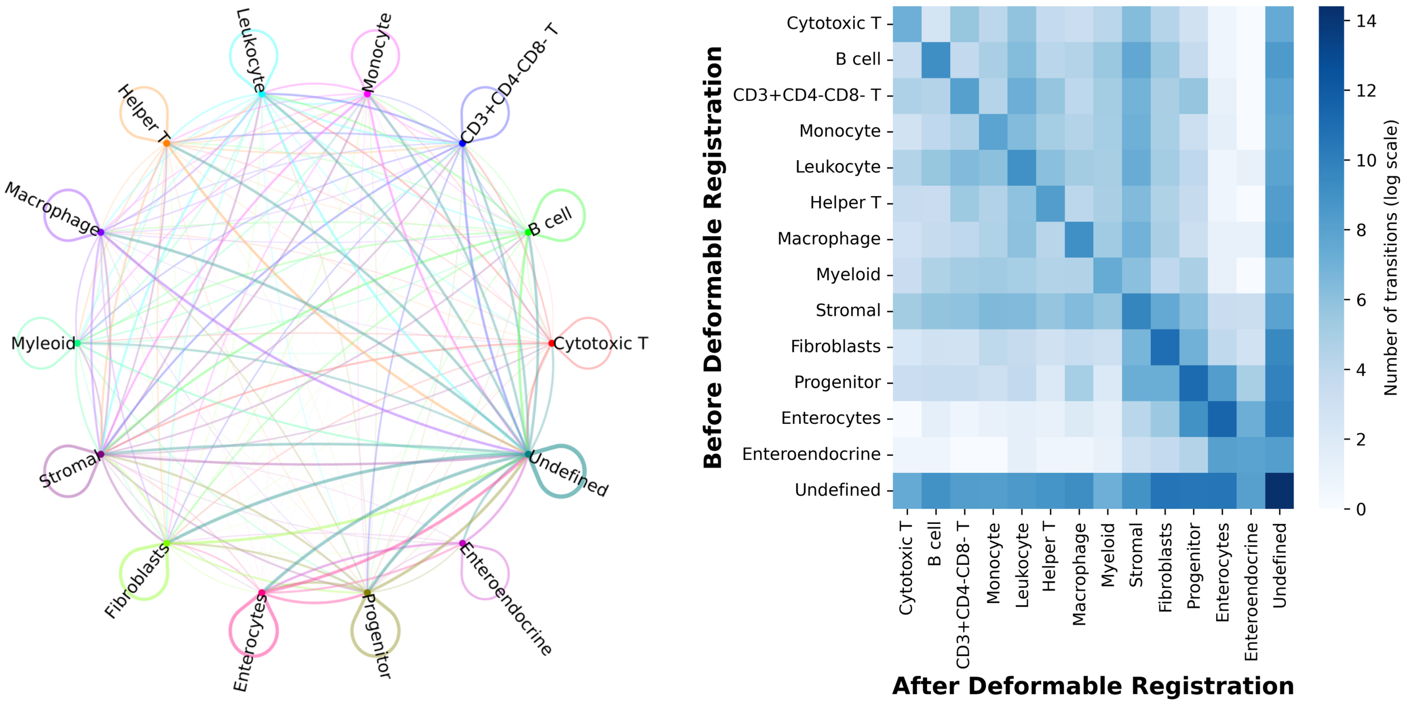

Figure 5: We find that deformable registration results in nuclei transitions between subtypes. **(Left)** A directed graph displaying cell interactions, with edges weighted by the square root of the number of transitions and scaled down by a factor of 50. The edges are colored based on the subtype classification from rigid registration, and the transparency of each edge is adjusted according to the edge weights. We observe that transitions out from the "Undefined" state are more heavily weighted than other transitions. **(Right)** A matrix visualizing the dynamics of cell type changes post-deformable registration. Values are shown using natural log scaling for enhanced clarity. The main diagonal entries denote cells retaining their classification pre- and post- deformable registration. The last row indicates the number of cells transitioning from an undefined state before deformable registration to defined classes after, highlighting the gains for each cell type. Conversely, the last column shows the cells that shifted from specific classes to an undefined state, reflecting losses per class. Other entries detail the transitions between different cell types. As the last row and column are relatively darker than corresponding rows and columns, most transitions happen with respect to the "Undefined" state.

applying deformable registration. Additionally, changes in subtype classification post-registration also impacted these results. To quantitatively assess the significance of these registration-induced changes, a Wilcoxon signed rank test was performed on the number of labeled nuclei before and after applying deformable registration. This test revealed a statistically significant increase in labeled nuclei with a $p < 0.01$, emphasizing the effectiveness of the deformable registration technique. Shifts in marker position due to deformable registration can change the class of a nuclei from its initial designation with rigid registration. Figure 5 provides an illustration of the transitions that happened between nuclei subtypes due to the application of deformable registration. A few instances of this effect are shown in Figure 6, where we pick some nuclei at random and show how the classification of the nuclei changed due to the application of deformable registration. For visualization purposes, the intensity of the marker signal in these figures has been increased by a factor of 20 and clipped within the range $[0, 255]$. Additional images from multiple MxIF images showing the transitions for all classes are provided in the supplementary figures.

Within the immune cell groupings, we observed increases in annotation counts for B cell nuclei (2.8%), monocytes (12.3%), leukocytes (18.1%), helper T cell nuclei (25.7%), macrophages (24.5%), and myeloid cell nuclei (8.6%). Conversely, there was a decrease of 21% in cytotoxic T cell nuclei and 2.29% in CD3^{+}CD4^{-}CD8^{-} T cell nuclei. In the stromal grouping, fibroblasts showed a 40% increase, and the stromal (undefined) class saw a 37.6% increase. Among epithelial cells, enterocytes increased by 6.8%, whereas enteroendocrine nuclei decreased by 24%. The progenitor group exhibited a 19.4% increase in the number of nuclei.

4.3 Sensitivity Analysis

We investigated the impact of varying the smoothness hyperparameter weight on the performance of our deformable

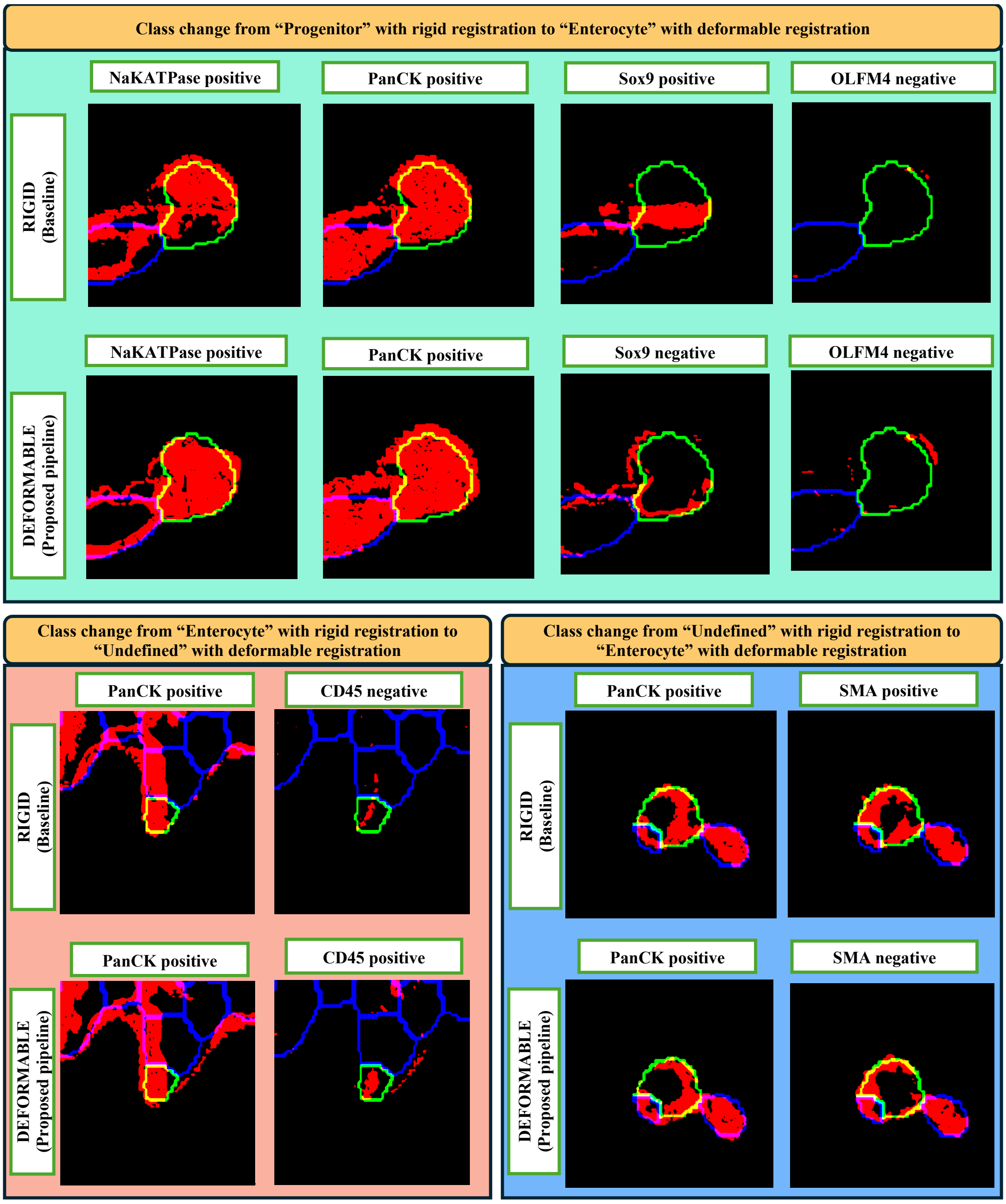

Figure 6: The type of registration applied has a profound effect on nucleus sub-type classification. The targeted nucleus is highlighted in green, surrounding nuclei in blue, and marker expressions in red.**(Top)** Under rigid registration, the nucleus was initially labeled as Progenitor being positive for NaKATPase, PanCK and Sox9. After alignment using deformable registration, the nucleus was negative for Sox9, changing its classification to Enterocyte.**(Bottom Left)** Here, rigid registration identified the nucleus positively with only PanCK, classifying it as an Enterocyte. However, deformable registration caused the marker for CD45 to become more pronounced, creating an epithelial and immune conflict and leading to an undefined label.**(Bottom Right)** Initially, with rigid registration, the nucleus tested positive for PanCK and SMA, classified as Undefined due to epithelial and stromal conflict. Deformable registration subsequently aligned the marker from SMA more closely around the nucleus causing it to be negative for SMA, thus classifying it as Enterocyte.

registration pipeline, using the number of nuclei labels gained as a measure of registration effectiveness. This sensitivity analysis involved training models with different smoothness weights, applying each model to perform registration, and quantifying the resulting gain in classified nuclei. We observed that the number of nuclei labels gained peaked at a smoothness weight of 0.5, with reduced gains observed for both lower and higher values (Figure 7). However, as the optimal smoothness weight can vary depending on dataset quality, tissue characteristics, and imaging conditions—consistent with the behavior of hyperparameters in deep learning models—we conservatively selected a smoothness weight of 1 for all experiments reported in this study. This choice lies toward the edge of the optimal region and was intentionally made to demonstrate that the observed improvements in registration quality and downstream nucleus subtype classification are not overly dependent on a finely tuned hyperparameter value.

5. Discussion and Conclusion

Our study underscores the significance of registration on the classification and interpretation of MxIF data. Our hypothesis that improved alignment of markers with nuclei would enhance the proportion of nuclei classified was supported through both quantitative and qualitative analyses. The deployment of deformable registration markedly increased number of labelled nuclei. Additionally, our approach aids in identifying areas of tissue loss and deformation, allowing for their exclusion from the analysis.

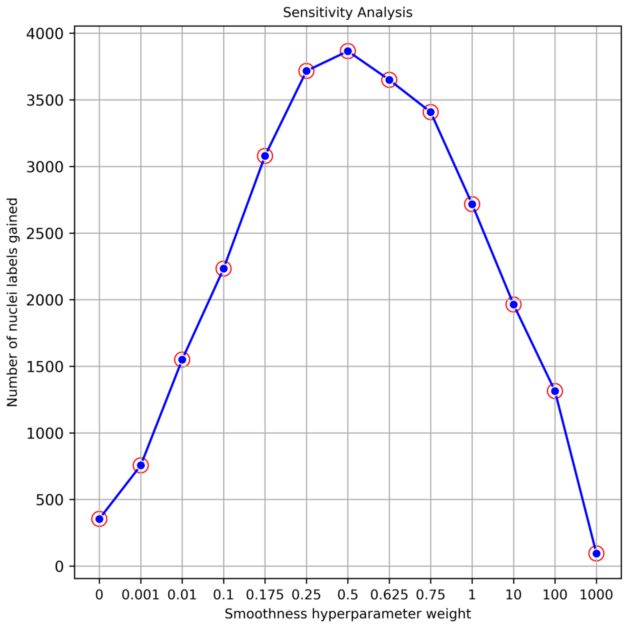

Figure 7: Variation in the weight of the smoothness hyperparameter effects the number of nuclei gained. This impact is illustrated on one WSI in the figure. The number of labels is maximized is maximized at 0.5 and falls off beyond 1 and below 0.1. Lower weights lead to non-realistic deformation fields, while higher weights overly constrain the vector fields, resulting in suboptimal registration.

In our manuscript, we aim to show the impact of deformable registration on the quantification of cells in MxIF images. One of the key benefits of using VoxelMorph, which is a deep learning-based method, over numerical optimization is that during inference, the method is faster than numerical optimization. Considering that each MxIF image in our dataset consists of around 17–19 rounds of imaging, and owing to the size of WSI images, numerical optimization techniques would require significantly more computational time highlighting the efficiency and practicality of deep learning-based registration. We used the VoxelMorph framework to quantify how effectively deep learning based deformable registration helps in classifying the nuclei in MxIF. Our approach involved training the model on a subset of 10 images, then applying the model to both the training subset and the remaining images to evaluate its overall performance. This inclusion of the training data in our results aligns with principles from classical registration methods, where the algorithm typically minimizes a metric between the moving and fixed images directly. This approach is supported by the concept of instance-specific optimization(Balakrishnan et al., 2019), which suggests that optimizing on a per-instance basis can enhance performance. Our results indicated that once trained, the model could be applied to new images from the same dataset without the need for re-training, suggesting a reasonable level of generalizability within data acquired from the same site.

Interestingly, the transition matrix shows rare reclassifications between the stromal and epithelial groups, reinforcing the delineation between them. Furthermore, few transitions were noted between the immune and epithelial groups, suggesting that the application of deformable registration maintains the distinct groupings, with reclassifications predominantly occurring within nuclei subtypes that belong to the same group.

In our sensitivity analysis, we observed that when we reduce or increase the weight of the smoothness hyperparameter by a large factor, the gains in the nuclei labels became very low. This could be because reducing the weight of the smoothness parameter leads the network to prioritize image matching, which can generate non-realistic deformation fields. These deformation fields, when applied to the marker images, cause the images to spread out sporadically instead of moving coherently with the nucleus signal. Conversely, increasing the parameter too much results in applying excessive constraint on the vector fields, leading the network to learn suboptimal registration that almost approaches rigid registration . While it is relatively robust in reasonable ranges, the optimal smoothness weight might be dataset- and image-quality dependent, as variations in tissue integrity, staining consistency, and imaging noise can influence the balance between deformation flexibility and regularization. Therefore, we recommend performing a small hyperparameter sweep to identify a smoothness weight that yields satisfactory performance for a given dataset.

Our findings are consistent with prior studies in multiplexed immunofluorescence and digital pathology that rely predominantly on rigid or affine registration for marker alignment across staining rounds. Several existing MxIF workflows repeatedly acquire a nuclear or structural marker and apply rigid registration to align successive imaging cycles into a common coordinate system(Lin et al., 2018; Muñoz-Castro et al., 2022; Liu and Ding, 2025). While these approaches are effective for correcting global shifts and rotations, they do not explicitly address the local, nonlinear tissue deformations that can accumulate over repeated staining, bleaching, and imaging cycles(Bao et al., 2022). More broadly, prior work in histological image registration has shown that deformable registration can improve local alignment and downstream quantitative analyses in the presence of tissue distortion, albeit often at substantial computational cost when using numerical optimization–based methods(Elhaminia et al., 2025). Our results extend these observations to the MxIF setting by demonstrating that deep learning–based deformable registration can improve marker–nucleus alignment and nucleus classification at whole-slide scale, while remaining computationally feasible for datasets involving many imaging rounds.

Acknowledgments

This publication is part of the Gut Cell Atlas Crohn's Disease Consortium, funded by The Leona M. and Harry B. Helmsley Charitable Trust and supported by a grant from Helmsley to Vanderbilt University, with specific funding from grants G-1903-03793 and G-2103-0512. Additionally, we acknowledge support from the Integrated Training in Engineering and Diabetes program, grant number T32 DK101003, the VUMC Digestive Disease Research Center supported by NIH grant P30DK058404, NIH grant R01DK103831, NIH grant R01DK135597, NIH grant R01DK128200, the Department of Veterans Affairs I01CX002662 and I01CX002473, NSF CAREER award 1452485, and NSF grant 2040462. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, TN. We also acknowledge support from the Vanderbilt Institute for Clinical and Translational Research (VICTR), funded by the National Center for Advancing Translational Sciences (NCATS) Clinical Translational Science Award (CTSA) Program, Award Number 5UL1TR002243-03. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. We extend our sincere thanks to all funding bodies and institutions that have made this research possible.

We utilized generative artificial intelligence (AI) to assist in creating and refining code segments based on specified tasks, including debugging, editing, and autocompleting code. Additionally, generative AI tools were used to help structure sentences and perform grammar checks. The conceptualization and ideation, as well as all instructions given to the AI, were exclusively the result of the authors' creative and intellectual input. We take full responsibility for reviewing and ensuring the accuracy and quality of all AI-generated content included in this work.

Ethical Standards

The work follows appropriate ethical standards in conducting research and writing the manuscript, following all applicable laws and regulations regarding treatment of animals or human subjects.

Conflicts of Interest

We declare we don't have conflicts of interest

Data availability

The data used in this study were collected under Vanderbilt University Medical Center Institutional Review Board protocols (IRB #191738 and #191777) that preclude public sharing. De-identified data may be made available upon reasonable request to the corresponding authors, subject to institutional approval and a data use agreement.

References

Michael Angelo, Sean C Bendall, Rachel Finck, Matthew B Hale, Chuck Hitzman, Alexander D Borowsky, Richard M Levenson, John B Lowe, Scot D Liu, Shuchun Zhao, et al. Multiplexed ion beam imaging of human breast tumors. *Nature medicine*, 20(4):436–442, 2014.

Connor Atkins, Gary KL Tam, Michael Edwards, Muhammad Aslam, and Jiaxiang Zhang. Multi-scale wsi analysis: A cascade framework for efficient breast cancer metastasis detection. In *Annual Conference on Medical Image Understanding and Analysis*, pages 87–101. Springer, 2025.

Guha Balakrishnan, Amy Zhao, Mert R Sabuncu, John Guttag, and Adrian V Dalca. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*, 38(8):1788–1800, 2019.

Shunxing Bao, Jia Li, Can Cui, Yucheng Tang, Ruining Deng, Lucas W Remedios, Ho Hin Lee, Sophie Chiron, Nathan Heath Patterson, Ken S Lau, et al. Mxif q-score: Biology-informed quality assurance for multiplexed immunofluorescence imaging. In *International Workshop on Medical Optical Imaging and Virtual Microscopy Image Analysis*, pages 42–52. Springer, 2022.

Jiří Borovec, Arrate Munoz-Barrutia, and Jan Kybic. Benchmarking of image registration methods for differently stained histological slides. In *2018 25th IEEE International Conference on Image Processing (ICIP)*, pages 3368–3372. IEEE, 2018.

Jiří Borovec, Jan Kybic, Ignacio Arganda-Carreras, Dmitry V Sorokin, Gloria Bueno, Alexander V Khvostikov, Spyridon Bakas, Eric I-Chao Chang, Stefan Heldmann, Kimmo Kartasalo, et al. Anhir: automatic non-rigid histological image registration challenge. *IEEE transactions on medical imaging*, 39(10):3042–3052, 2020.

Behnaz Elhaminia, Abdullah Alsalemi, Esha Nasir, Mostafa Jahanifar, Ruqayya Awan, Lawrence S Young, Nasir M Rajpoot, Fayyaz Minhas, and Shan E Ahmed Raza. From traditional to deep learning approaches in whole slide image registration: A methodological review. *Journal of Pathology Informatics*, page 100512, 2025.

Yabo Fu, Yang Lei, Tonghe Wang, Walter J Curran, Tian Liu, and Xiaofeng Yang. Deep learning in medical image registration: a review. *Physics in Medicine & Biology*, 65(20):20TR01, 2020.

Chandler D Gatenbee, Ann-Marie Baker, Sandhya Prabhakaran, Ottilie Swinyard, Robbert JC Slebos, Gunjan Mandal, Eoghan Mulholland, Noemi Andor, Andriy Marusyk, Simon Leedham, et al. Virtual alignment of pathology image series for multi-gigapixel whole slide images. *Nature communications*, 14(1):4502, 2023.

Michael J Gerdes, Christopher J Sevinsky, Anup Sood, Sudeshna Adak, Musodiq O Bello, Alexander Bordwell, Ali Can, Alex Corwin, Sean Dinn, Robert J Filkins, et al. Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue. *Proceedings of the National Academy of Sciences*, 110(29):11982–11987, 2013.

Parmida Ghahremani, Joseph Marino, Juan Hernandez-Prera, Janis V de la Iglesia, Robbert JC Slebos, Christine H Chung, and Saad Nadeem. An ai-ready multiplex staining dataset for reproducible and accurate characterization of tumor immune microenvironment. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 704–713. Springer, 2023.

Yury Goltsev, Nikolay Samusik, Julia Kennedy-Darling, Salil Bhate, Matthew Hale, Gustavo Vazquez, Sarah Black, and Garry P Nolan. Deep profiling of mouse splenic architecture with codex multiplexed imaging. *Cell*, 174 (4):968–981, 2018.

Noah F Greenwald, Geneva Miller, Erick Moen, Alex Kong, Adam Kagel, Thomas Dougherty, Christine Camacho Fullaway, Brianna J McIntosh, Ke Xuan Leow, Morgan Sarah Schwartz, et al. Whole-cell segmentation of tissue images with human-level performance using large-scale data annotation and deep learning. *Nature biotechnology*, 40 (4):555–565, 2022.

Manuel Guizar-Sicairos, Samuel T Thurman, and James R Fienup. Efficient subpixel image registration algorithms. *Optics letters*, 33(2):156–158, 2008.

Grant Haskins, Uwe Kruger, and Pingkun Yan. Deep learning in medical image registration: a survey. *Machine Vision and Applications*, 31(1):8, 2020.

Alessa Hering, Lasse Hansen, Tony CW Mok, Albert CS Chung, Hanna Siebert, Stephanie Häger, Annkristin Lange, Sven Kuckertz, Stefan Heldmann, Wei Shao, et al. Learn2reg: comprehensive multi-task medical image registration challenge, dataset and evaluation in the era of deep learning. *IEEE Transactions on Medical Imaging*, 42(3):697–712, 2022.

Jun Jiang, Raymond Moore, Clarissa E Jordan, Ruifeng Guo, Rachel L Maus, Hongfang Liu, Ellen Goode, Svetomir N Markovic, and Chen Wang. Multiplex immunofluorescence image quality checking using dapi channel–referenced evaluation. *Journal of Histochemistry & Cytochemistry*, 71(3):121–130, 2023.

Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Stefan Klein, Marius Staring, Keelin Murphy, Max A Viergever, and Josien PW Pluim. Elastix: a toolbox for intensity-based medical image registration. *IEEE transactions on medical imaging*, 29(1):196–205, 2009.

Jia-Ren Lin, Mohammad Fallahi-Sichani, and Peter K Sorger. Highly multiplexed imaging of single cells using a high-throughput cyclic immunofluorescence method. *Nature communications*, 6(1):8390, 2015.

Jia-Ren Lin, Benjamin Izar, Shu Wang, Clarence Yapp, Shaolin Mei, Parin M Shah, Sandro Santagata, and Peter K Sorger. Highly multiplexed immunofluorescence imaging of human tissues and tumors using t-cycif and conventional optical microscopes. *elife*, 7:e31657, 2018.

Zeyu Liu and Shengwei Ding. Star: A fast and robust rigid registration framework for serial histopathological images. *arXiv preprint arXiv:2509.02952*, 2025.

Eliot T McKinley, Yunxia Sui, Yousef Al-Kofahi, Bryan A Millis, Matthew J Tyska, Joseph T Roland, Alberto Santamaria-Pang, Christina L Ohland, Christian Jobin, Jeffrey L Franklin, et al. Optimized multiplex immunofluorescence single-cell analysis reveals tuft cell heterogeneity. *JCI insight*, 2(11):e93487, 2017.

Jeremy L Muhlich, Yu-An Chen, Clarence Yapp, Douglas Russell, Sandro Santagata, and Peter K Sorger. Stitching and registering highly multiplexed whole-slide images of tissues and tumors using ashlar. *Bioinformatics*, 38(19): 4613–4621, 2022.

Clara Muñoz-Castro, Ayush Noori, Bradley T Hyman, and Alberto Serrano-Pozo. Cyclic multiplex fluorescent immunohistochemistry protocol to phenotype glial cells in formalin-fixed paraffin-embedded human brain sections. In *Signal Transduction Immunohistochemistry: Methods and Protocols*, pages 283–305. Springer, 2022.

Lucas W Remedios, Shunxing Bao, Samuel W Remedios, Ho Hin Lee, Leon Y Cai, Thomas Li, Ruining Deng, Can Cui, Jia Li, Qi Liu, et al. Nucleus subtype classification using inter-modality learning. In *Proceedings of SPIE–the International Society for Optical Engineering*, volume 12933, page 129330F, 2024a.

Lucas W Remedios, Shunxing Bao, Samuel W Remedios, Ho Hin Lee, Leon Y Cai, Thomas Li, Ruining Deng, Nancy R Newlin, Adam M Saunders, Can Cui, et al. Data-driven nucleus subclassification on colon h&e using style-transferred digital pathology. *arXiv preprint arXiv:2407.06116*, 2024b.

Mousumi Roy, Fusheng Wang, George Teodoro, Shristi Bhattarai, Mahak Bhargava, T Subbanna Rekha, Ritu Aneja, and Jun Kong. Deep learning based registration of serial whole-slide histopathology images in different stains. *Journal of Pathology Informatics*, 14:100311, 2023.

Denis Schapiro, Artem Sokolov, Clarence Yapp, Yu-An Chen, Jeremy L Muhlich, Joshua Hess, Allison L Creason, Ajit J Nirmal, Gregory J Baker, Maulik K Nariya, et al. Mcmicro: a scalable, modular image-processing pipeline for multiplexed tissue imaging. *Nature methods*, 19(3): 311–315, 2022.

Walter Schubert, Bernd Bonnekoh, Ansgar J Pommer, Lars Philipsen, Raik Böckelmann, Yanina Malykh, Harald Gollnick, Manuela Friedenberger, Marcus Bode, and Andreas WM Dress. Analyzing proteome topology and function by automated multidimensional fluorescence microscopy. *Nature biotechnology*, 24(10):1270–1278, 2006.

Abubakr Shafique, Morteza Babaie, Mahjabin Sajadi, Adrian Batten, Soma Skdar, and Hamid R Tizhoosh. Automatic multi-stain registration of whole slide images in histopathology. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 3622–3625. IEEE, 2021.

Juan Carlos Vizcarra, Erik A Burlingame, Clemens B Hug, Yury Goltsev, Brian S White, Darren R Tyson, Artem Sokolov, et al. A community-based approach to image analysis of cells, tissues and tumors. *Computerized Medical Imaging and Graphics*, 95:102013, 2022.

Marek Wodzinski, Niccolo Marini, Manfredo Atzori, and Henning Müller. Deeperhistreg: robust whole slide images registration framework. *arXiv preprint arXiv:2404.14434*, 2024.

Jiangmei Xiong, Harsimran Kaur, Cody N Heiser, Eliot T McKinley, Joseph T Roland, Robert J Coffey, Martha J Shrubsole, Julia Wrobel, Siyuan Ma, Ken S Lau, et al. Gammagater: semi-automated marker gating for single-cell multiplexed imaging. *Bioinformatics*, 40(6):btae356, 2024.