Journal of Pathology Informatics 14 (2023) 100311

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

Deep learning based registration of serial whole-slide histopathology images in different stains

Mousumi Roy [a,*], Fusheng Wang [a,b,*], George Teodoro [c], Shristi Bhattarai [d], Mahak Bhargava [d], T. Subbanna Rekha [e], Ritu Aneja [d], Jun Kong [f,g,*]

[a] Department of Computer Science, Stony Brook University, NY 11794, USA
[b] Department of Biomedical Informatics, Stony Brook University, NY 11794, USA
[c] Department of Computer Science, Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil
[d] Department of Clinical and Diagnostic Sciences, School of Health Profession, University of Alabama at Birmingham, Birmingham, AL 35233, USA
[e] Department of Pathology, JSS Medical College, JSS Academy of Higher Education and Research, Mysuru, Karnataka 570009, India
[f] Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303, USA
[g] Department of Computer Science and Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA

ARTICLE INFO

Keywords::
Image registration
Whole-slide image
Multiple stains
Image translation
Deformable vector field

ABSTRACT

For routine pathology diagnosis and imaging-based biomedical research, Whole-slide image (WSI) analyses have been largely limited to a 2D tissue image space. For a more definitive tissue representation to support fine-resolution spatial and integrative analyses, it is critical to extend such tissue-based investigations to a 3D tissue space with spatially aligned serial tissue WSIs in different stains, such as Hematoxylin and Eosin (H&E) and Immunohistochemistry (IHC) biomarkers. However, such WSI registration is technically challenged by the overwhelming image scale, the complex histology structure change, and the significant difference in tissue appearances in different stains. The goal of this study is to register serial sections from multi-stain histopathology whole-slide image blocks. We propose a novel translation-based deep learning registration network CGNReg that spatially aligns serial WSIs stained in H&E and by IHC biomarkers without prior deformation information for the model training. First, synthetic IHC images are produced from H&E slides through a robust image synthesis algorithm. Next, the synthetic and the real IHC images are registered through a Fully Convolutional Network with multi-scaled deformable vector fields and a joint loss optimization. We perform the registration at the full image resolution, retaining the tissue details in the results. Evaluated with a dataset of 76 breast cancer patients with 1 H&E and 2 IHC serial WSIs for each patient, CGNReg presents promising performance as compared with multiple state-of-the-art systems in our evaluation. Our results suggest that CGNReg can produce promising registration results with serial WSIs in different stains, enabling integrative 3D tissue-based biomedical investigations.

Introduction

Histopathology whole-slide images (WSIs) of tissue sections provide high resolution tissue details critical for disease diagnosis and study. Such high resolution WSIs have been largely analyzed in a two-dimensional (2D) tissue imaging space by far.[1–4] As each 2D WSI can only capture information from the tissue cutting plane, it is inevitably subject to information loss and sampling bias problems. Therefore, it is critical to extend such analyses to a three-dimensional (3D) tissue space, especially for those studies requiring a definitive tissue characterization.[5] In the meanwhile, such 3D tissue-based methods are urged by the rapidly increasing analysis and clinical demand on spatial integration of multi-stained serial WSIs. In particular, there is a strong demand to integrate Hematoxylin and Eosin (H&E) WSIs capturing histology tissue phenotype information with serial Immunohistochemistry (IHC) WSIs visualizing the underlying disease molecular underpinnings for a comprehensive understanding of tumor microenvironment.

In the era of precision medicine, biomedical and cancer studies have been strongly supported by high-resolution, and multi-stained tissue-based microscopic pathology image analytics, as it provides a new avenue to describe cellular, molecular and tissue-level interactions in an integrative manner.[7,8] This emerging research field extracts key morphological, spatial, and molecular information from high-resolution pathology images to enable a more precise understanding of biomarker interactions, a prerequisite for optimized clinical decisions and targeted therapy development.

* Corresponding authors.
E-mail addresses: mnroy@cs.stonybrook.edu (M. Roy), fusheng.wang@stonybrook.edu (F. Wang), jkong@gsu.edu (J. Kong).

http://dx.doi.org/10.1016/j.jpi.2023.100311
Received 31 October 2022; Received in revised form 11 April 2023; Accepted 12 April 2023
Available online 23 April 2023

Due to convenience, cost, and preparation technique availability, immunohistochemistry (IHC) has been a widely used to visualize disease-related molecular biomarkers without loss of spatial tissue reference in most pathologic review processes. For a comprehensive characterization and understanding of tissue phenotypic and molecular signatures, adjacent tissue slides in different stains are required to be jointly studied. Thus, there is a crucial need to spatially map molecular signatures from IHC to the adjacent H&E tissue slides within the same histology space, enabling multimodal microscopy integrative analysis for better clinical prognostic and prediction power.[9,10] In breast cancer, for example, Ki67 biomarker from IHC is widely used as a predictor for the recurrence-free survival rate and pathological complete response to the neoadjuvant chemotherapy.[11,12] In current clinical practice, the intensity of proliferative Ki67 positive biomarkers is measured in tumor regions that can be readily recognized by H&E slides. This leads to the need to combine biomarker information from IHC with tumor information from serial H&E slides. Another commonly IHC biomarker Phosphohistone H3 (PHH3) is used to measure the mitotic activity informative for the breast tumor grading by the Nottingham Grading System[13] and neoadjuvant chemotherapy response prediction in breast cancer.[14] As the mitosis is a sub-phase of the entire proliferation cycle,[14,15] an integrated use of mitosis and proliferation information can more precisely characterize the tumor proliferation momentum. The development of such new prognostic biomarkers thus requires to investigate tumor tissues, proliferation, and mitosis biomarkers in a single tissue space that can better represent proliferation and the mitotic activity intensity in tumors. Such phenotype–genotype integrative analyses for new predictive biomarkers for clinical decision-making make it an urgent call to develop registration tools that support spatial alignment of pathology images of serial tissue sections in different stains.

However, WSI registration is technically challenging. Existing methods only achieve a limited success due to the overwhelmingly large image scale, complex histology structure change across adjacent slides, and significant tissue appearance difference in different staining.[16] Although a successful 3D histology image reconstruction method has been developed, it requires ground-truth images (i.e. block-face images) and image stacks of relatively limited size (e.g. 1008×756) for each slide.[6]

Recently, the ANHIR challenge was organized to systematically compare the performances of image registration algorithms for microscopy histology images.[16] Most methods described in the ANHIR challenge are based on a rigid registration followed by a non-rigid registration. Rigid registration is determined by RANSAC from feature points, whereas non-rigid registration is performed by local affine transformation, demons algorithms, or interpolations. For example, a 2-stage image registration method is proposed where the registration is formulated as an optimization problem with an objective function defined by Normalized Gradient Fields (NGF)-based distance measure.[17] In another study, a fine-tuning method is proposed based on integrated landmark evaluation by texture and spatial proximity measures.[18] In a study on multi-stained histology image registration, a method is developed by the feature-based affine registration, rotation alignment followed by non-rigid alignment with Demons algorithm.[19] Moreover, a modified SIFT method is applied to image registration with color interpolation.[20] WSIReg is another method for WSI registration based on elastix.[21] However, all these methods require feature matching that highly depends on feature engineering and parameter tuning.

With the emergence of deep learning methods, it is possible to leverage the image-to-image translation for spatial alignment of images in different appearances.[22–24] Translation-based approaches use Generative Adversarial Network (GAN) to translate images from one modality (e.g. H&E) to another (e.g. IHC), simplifying the image registration task. Although much simplified, such an analysis still presents 2 significant challenges. First, H&E and IHC WSIs are unpaired data as each pathology tissue slide is stained only once in most clinical practice. Second, it is time-consuming and financially costly to have accurate landmark pair annotations from serial WSIs for registration. Recently, CycleGAN with the cycle consistent loss has been developed to learn an image-to-image mapping between two domains from unpaired data.[25] The cycle consistency loss for the adversarial training process forces the generator to find an accurate mapping between 2 different domains with unpaired data. By this approach, synthetic slice-wise computed tomography data has been produced from magnetic resonance images.[26] In another study on thoracic and abdominal organs, a mono-modal image registration with CycleGAN presents a comparable performance to the multi-modal deformable registration with paired image data.[22] Furthermore, a deformation field is used for MRI-CT registration in a dual-stream fashion with CycleGAN.[27] In this work, the MIND-based loss term added to CycleGAN loss describes the local image structure. As such a loss term is computed with grayscale images, and such a method cannot be applied to multi-stained pathology image data directly. In another study, a CycleGAN-based image generation method generates IHC histology images from H&E images without any annotation[28] where the class-related information is added as an additional input patch channel. Therefore, this method is not directly applicable to our study with multi-stained pathology image data.

In this paper, we present a new translation-based deep learning registration approach (CGNReg) for serial WSIs in different stains. It consists of an image translation and an image registration module. The image translation module produces synthetic IHC slides (i.e. synIHC images) from real H&E slides through a robust image synthesis algorithm. With Fully Convolutional Network (FCN) model[29] as a building block, the synIHC and the real IHC image pairs are registered through a multi-scale-based FCN registration model. The major contributions of this paper are summarized in multiple folds:

- We develop a modified CycleGAN method to generate synthetic IHC pathology images (i.e. synIHC) from unpaired serial H&E pathology slides. To enhance the image stain translation ability, we propose to adopt a perceptual loss in the CycleGAN loss function, resulting in a better image mapping from H&E to synthetic IHC images. Such an image translation enables a better registration between synIHC and real IHC images.
- We extend the original FCN model to a multi-scaled architecture for registration. Our proposed multi-scale FCN uses a coarse-to-fine multi-scale deformable image registration strategy that combines the Displacement Vector Fields (DVFs) at multiple resolutions for better image alignment.
- CGNReg is an unsupervised registration approach and can be efficiently trained without ground-truth image deformation information.
- Instead of resizing images to a lower resolution,[16] we recover the WSI registration results with patch-based image registration results at the highest image resolution. Therefore, high resolution tissue details from WSIs are retained.

Methods

We develop a deep learning-based model CGNReg to register serial IHC to H&E histopathology images for molecular biomarker and pathology hallmark integration. It is an end-to-end deep learning process in 2 stages. First, we develop an image translation module with a modified CycleGAN to translate real H&E references to synthetic IHC image patches (i.e. synIHC). Next, we develop a multi-scale FCN in the image registration module to estimate the spatial mapping from the real IHC to the synthetically produced synIHC image patches. Then, the real IHC image is transformed to the reference H&E image space via Spatial Transformation Network (STN).[30] Finally, individual registered image patches are spatially assembled to recover registered WSI blocks. We present the overall schema of CGNReg in Fig. 1.

Unpaired image translation

Although serial slides in different stains look similar at the global tissue level (Supplemental Figure S1), they are unpaired at the pixel level. Although CycleGAN can be applied to unpaired data,[25,28,31] we propose a modified CycleGAN for an enhanced H&E-IHC image translation. Illustrated in Fig. 2, the modified CycleGAN consists of an encoder and a

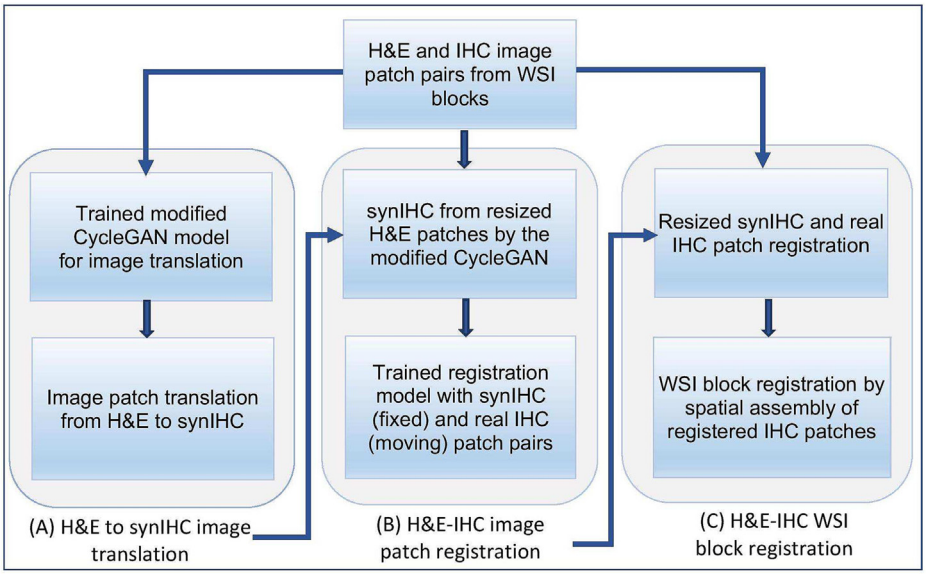

Fig. 1. The overall schema of the developed CGNReg method. (A) H&E images serve as reference images and they are translated to synthetic IHC image patches (i.e. synIHC). (B) Our image registration module estimates the spatial mapping from the real IHC patches to the synthetically produced synIHC image patches, and transforms the real IHC image patch to the reference H&E image space. (C) Individual registered image patches are spatially assembled to recover registered WSI blocks.

decoder. Both have the same network structure that includes a generator and a discriminator. The generator translates an image between stain domains and the discriminator assesses the generated image quality. The modified CycleGAN model consists of 2 generators G_{HE} and G_{IHC}. In the forward translation process, the generator G_{HE} translates a real IHC to a synthetic H&E image (i.e. synHE), while G_{IHC} translates a synHE to a synthetic IHC image (i.e. synIHC) denoted by the red arrows. Similarly, the reverse translation goes from H&E to synIHC and then from synIHC to synHE

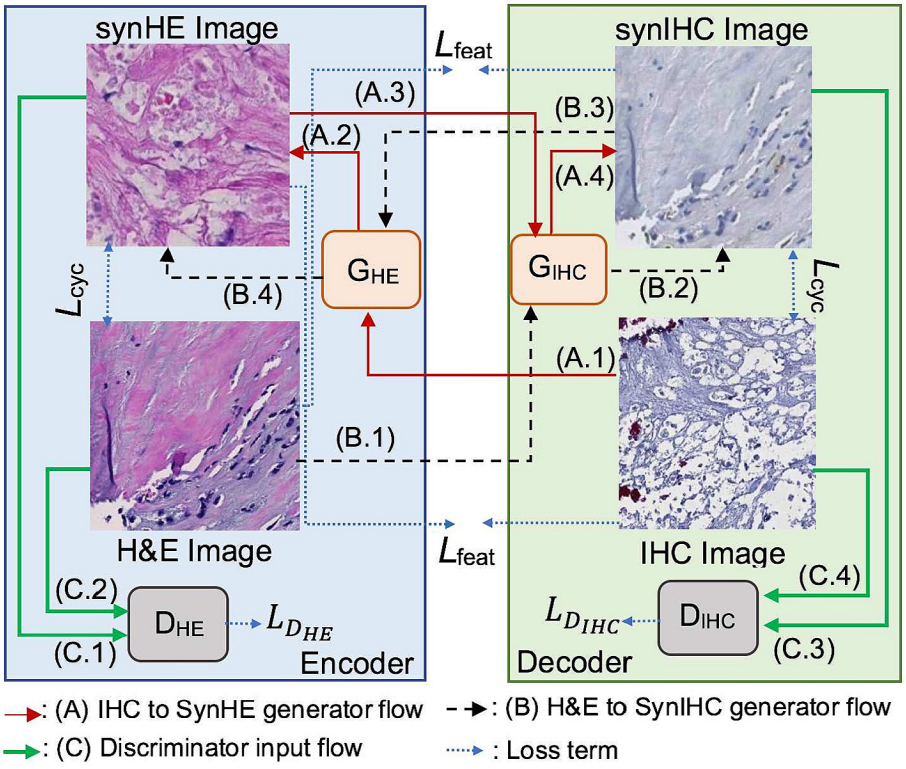

Fig. 2. The overall schema of the modified CycleGAN with the forward and backward translation information flow. (A) The forward generator translation information flow (red solid line). (A.1–A.2) The generator G_{HE} translates a real IHC to a synthetic H&E (synHE) image. (A.3–A.4) Next, the generator G_{IHC} translates the synHE to a synthetic IHC (synIHC) image. (B) The backward translation information flow (black dashed line). (B.1–B.2) The generator G_{IHC} translates a real H&E to a synIHC image. (B.3–B.4) Next, the generator G_{HE} translates the synIHC to a synHE image. (C) The discriminator information flow (green line). (C.1–C.2) The discriminator D_{HE} distinguishes between the translated (i.e. synHE) and real H&E image. (C.3–C.4) The discriminator D_{IHC} distinguishes between the translated (i.e. synIHC) and real IHC image. Blue dashed lines indicate the source of information for individual loss term computation, including the adversarial loss, the cycle-consistent loss, and the feature reconstruction loss associated with H&E or IHC images.

(i.e. the black arrows). Each generator module consists of 2-dimensional fully convolutional networks with 9 residual blocks and 2 fractionally strided convolution layers.[34,32] Additionally, the model has 2 discriminators D_{HE} and D_{IHC} for distinguishing between translated (i.e. synHE) and real H&E images, and between translated (i.e. synIHC) and real IHC images, respectively. Each discriminator has a fully convolutional architecture to predict if overlapping image patches of size 70 × 70 by pixels are real or synthetic.[33] The leaky ReLU activation function with a factor 0.2 is used. All data are normalized with the instance normalization.

The CycleGAN training loss includes the adversarial loss (i.e. $L_{D_{HE}}$ and $L_{D_{IHC}}$) from two discriminators and the cycle-consistent loss (L_{cyc}).[25] Although the cycle-consistent loss is designed to prohibit the generators from generating images not related to the inputs, this loss by itself is not sufficient to enforce either feature or structural similarity between translated and real images. To address this problem, we adopt the VGG-16 based perceptual loss function as an additional constraint in the CycleGAN loss function. Such a perceptual loss addition helps regularize the tissue content and the stain style discrepancies, as it can measure high-level perceptual and semantic differences between each image pair.[34] Such a perceptual loss is produced by a deep convolutional neural network denoted as ϕ, specifically a 16-layer VGG network[35] pre-trained on the ImageNet dataset.[36] The network ϕ forces the transformed network output image $\hat{x}$ to have similar feature representations as the input real image x. The feature reconstruction loss $\mathcal{L}_{feat}$ can be written as:

$$\mathcal{L}_{feat} = l^{\phi_j}_{feat}(\hat{x}, x) = \frac{1}{C_j H_j W_j} \|\phi_j(\hat{x}) - \phi_j(x)\|_2^2 \quad (1)$$

where $\phi_j(x)$ is the activation map of size $C_j \times H_j \times W_j$ at the j-th convolutional layer of network ϕ when processing image x. Although the image transformation network trained by the feature reconstruction loss encourages the output image $\hat{x}$ to be perceptually similar to the target image x, it does not force them to match each other exactly. The total loss $\mathcal{L}$ of our modified CycleGAN can be written as:

$$\mathcal{L} = \mathcal{L}_{D_{HE}} + \mathcal{L}_{D_{IHC}} + \lambda_{cyc}\mathcal{L}_{cyc} + \lambda_{feat}\mathcal{L}_{feat} \quad (2)$$

where $\mathcal{L}_{D_{HE}}$, $\mathcal{L}_{D_{IHC}}$ and $\mathcal{L}_{cyc}$ are defined in prior studies.[25] λ_{cyc} and λ_{feat} are weights for loss term $\mathcal{L}_{cyc}$ and $\mathcal{L}_{feat}$, respectively.

Multi-scaled image patch registration

Numerous state-of-the-art histopathology image registration methods are either based on SIFT-feature matching or MIND-based Demons algorithms. Some learning-based methods (e.g. TUB from ANHIR challenge[16]) have been proposed with limited performance. In this study, we leverage FCN-based deep learning model for histopathology image registration due to its known promising performance.[29] As multiple prior studies on flow estimation have shown the effectiveness of multi-scale strategy,[37,38] we use the FCN and the multi-scale strategy as building blocks for the proposed registration method. Specifically, we create a multi-scale FCN model with multi-scale Deformable Vector Fields (DVFs) that enable a coarse-to-fine multi-scale deformable image registration. For a better deformation estimation, 3 DVF estimation models at 3 different image resolutions are used to define the multi-scale spatial mappings. Our registration pipeline is demonstrated in Fig. 3.

For image registration, a moving I_M and a fixed I_F image are provided to our multi-scale FCN model as inputs. For model training, a pair of real moving and fixed images are concatenated and provided to the first DVF estimation model to estimate the DVF V_1 at scale-1. V_1 has 3 different components (i.e. V_{11}, V_{12} and V_{13}) that are produced by 3 different layers of the FCN model. The resulting warped moving images are $V_{11} \circ I_M$, $V_{12} \circ I_M$, and $V_{13} \circ I_M$, respectively, where $\circ$ is the operator applying a DVF to an image. Next, the fixed I_F and moving I_M image pairs are concatenated and down-sampled by a factor of 4. The resulting image pairs are provided to the second DVF estimation model for DVF estimation at scale-2. The resulting DVF is up-sampled to match the original input image size and denoted as V_2. Similarly, V_2 is applied to the input moving image I_M to generate the warped moving image $V_2 \circ I_M$. In the next step, the warped moving image $V_2 \circ I_M$ at scale-2 and the input fixed image I_F are concatenated and down-sampled by a factor of 2. The resulting image pairs are provided to the third DVF estimation model for the residual DVF estimation. The resulting DVF is up-sampled to the original input image size and denoted as V_3 at scale-3.

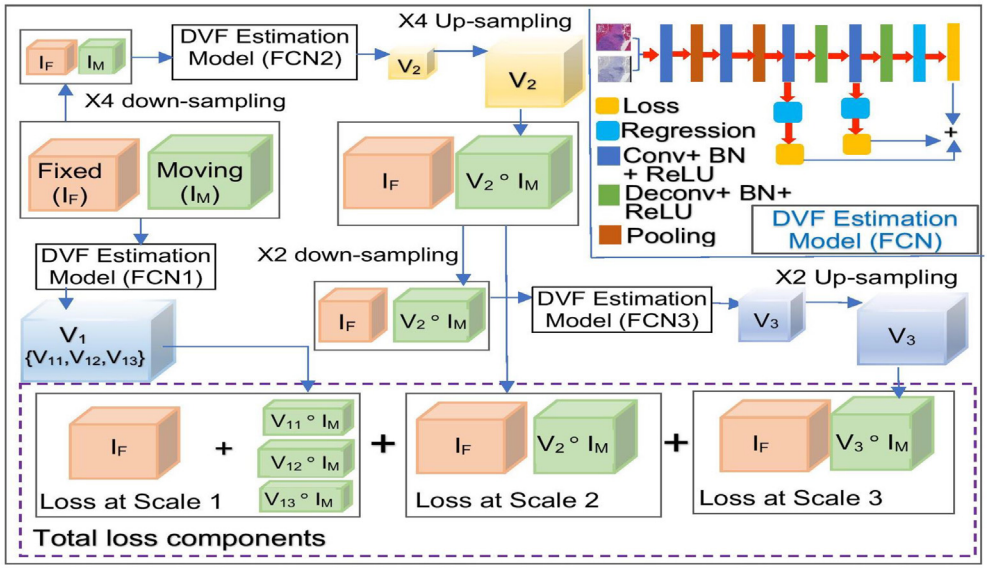

Fig. 3. The overall architecture of the developed multi-scale FCN model for registration is presented with detailed illustrations of FCN layers for the multi-scale DVF estimation and multiple loss function components in the total loss. Paired moving IM and fixed IF images are concatenated and provided to three DVF estimation models for estimated DVFs (i.e. V_1, V_2, and V_3) at three scales. Three DVF estimation models are simultaneously trained for the joint loss minimization (in Eq. 3). The DVF architecture is presented with different network components color-coded, including the loss term computation (yellow), the regression analysis (light blue), layers of convolution, Batch Normalization (BN) and Rectified Linear Unit (ReLU) (blue), layers of deconvolution, BN and ReLU (green), and the pooling layer (brown).

Finally, V_3 is used to deform the moving image I_M for the warped image $V_3 \circ I_M$. We train all 3 DVF estimation models simultaneously to minimize the joint loss at multiple scale levels, achieving an optimal end-to-end performance. Our formulated total loss function can be represented as follows:

$$\mathcal{L} = \sum_{i=1}^{3} \sigma_{1i}\{\mathcal{L}_{sim}(I_F, V_{1i} \circ I_M) + R(V_{1i})\} + \sigma_2 \mathcal{L}_{sim}(I_F, V_2 \circ I_M) + \sigma_3 \mathcal{L}_{sim}(I_F, V_3 \circ I_M) \quad (3)$$

where $\mathcal{L}_{sim}$ is the similarity loss measured by the Negative Normalized Cross-Correlation (NCC),[39] penalizing the differences in appearance between the fixed and moving images. Note σ_{11}, σ_{12} and σ_{13} are weights of the similarity loss metrics at scale level 1, while σ_2 and σ_3 are weights at scale levels 2 and 3, respectively. $R(V)$ is a total variation based regularizer that makes the transformation spatially smooth and physically plausible.[40]

After the weight initialization, all weights are updated by the joint training of 3 DVF estimation models in an end-to-end manner for the harmonic minimization of the composite loss. With displacement vectors between the fixed and moving image pairs, we use Spatial Transformer Network (STN)[30] to deform the moving image.[29] To make the resulting registered images retain more tissue details, we utilize the Enhanced SRGAN (ESRGAN) model[41] in the post-processing step. We demonstrate the registered images by our proposed multi-scale FCN model with and without the post-processing step in Supplemental Figure 2. Note that such a post-processing step improves the registration performance as suggested by Table 2.

WSI block registration

Due to the limited GPU memory size, deep learning methods cannot process giga-pixel WSIs at the full histopathology image resolution all at a time. Therefore, each WSI is partitioned into image blocks of size 8000 × 8000 pixels for tissue pre-alignment. Each H&E block is next translated to a synIHC block by the developed modified CycleGAN. Real IHC and synIHC image blocks are next divided into image patches of size 1024 × 1024 to retain sufficient tissue information for registration. The resulting synIHC and real IHC patch pairs are further resized to 256 × 256 pixels for the deep learning model training and prediction. After registration, the registered real IHC image patches are resized back to 1024 × 1024 size and spatially assembled for WSI block registration.[42]

Experimental result

Dataset and implementation

This study cohort consists of 76 Neoadjuvant Chemotherapy (NAC)-treated Triple Negative Breast Cancer (TNBC) patients from Emory Decatur

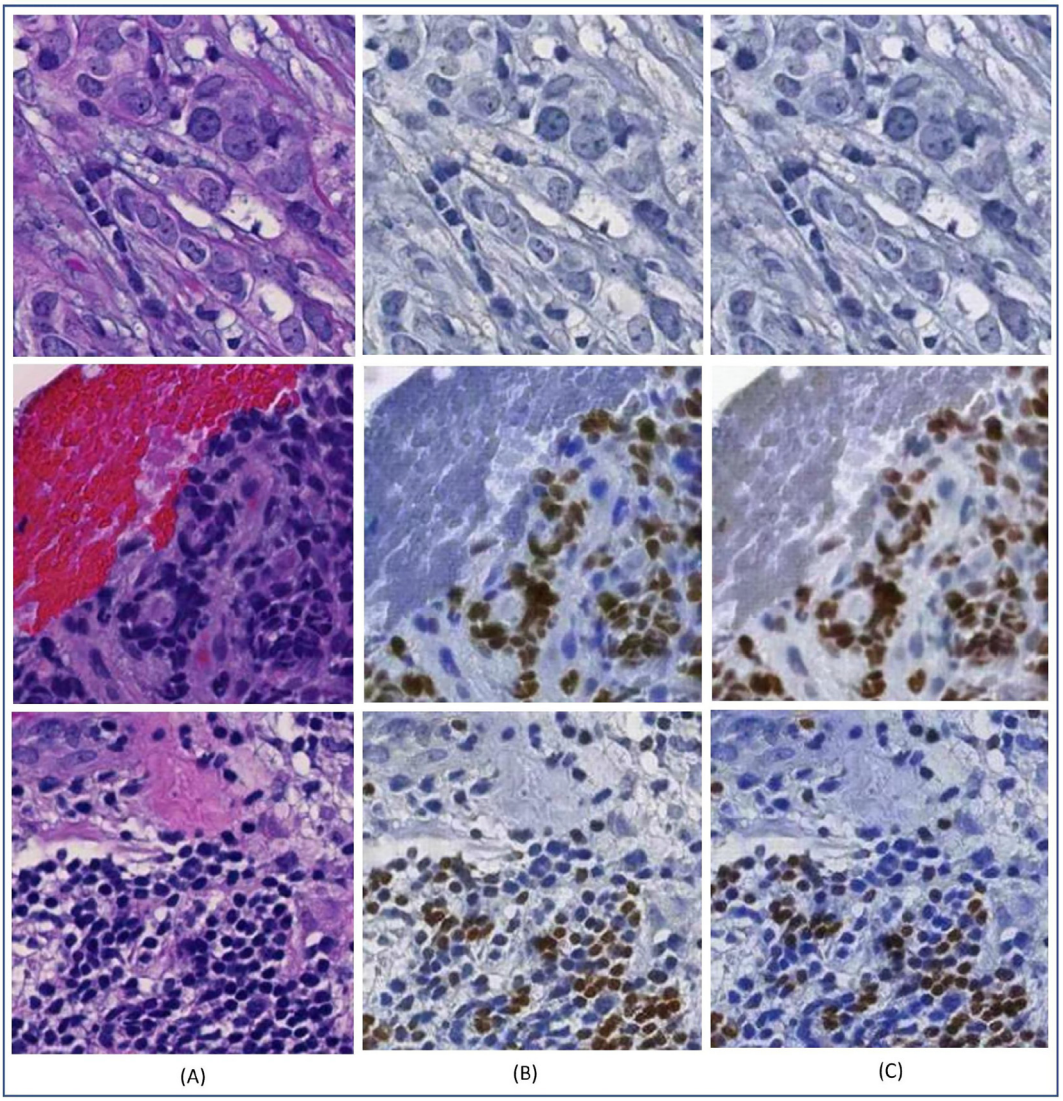

Fig. 4. Representative image patch translation results. (A) Real H&E image patch; (B) synKi67 patch by CycleGAN; and (C) synKi67 patch by our modified CycleGAN.

Hospital. Formalin-fixed paraffin-embedded serial biopsy sections from each patient are H&E and immunohistochemically stained with Ki67 as a biomarker for cell proliferation and Phosphohistone H3 (PHH3) for the mitotic activity. There are about 5 μm between each pair of adjacent slides. In total, we assess our model on 228 WSIs, with 3 serial WSIs from each patient. After the image pre-alignment by the global affine spatial transformation at a low image resolution, the resulting transformation is mapped to the full image resolution level. The pre-aligned tissue regions at the full image resolution level are next partitioned into 1023 WSI blocks of size 8000 × 8000 pixels by each stain. The pre-aligned WSI blocks are further partitioned into non-overlapping image patches of size 1024 × 1024 pixels and resized to 256 × 256 to make the image size appropriate for deep learning models. Patches containing more than 30% background pixels are excluded from further analyses. This process results in 60 000 image patches that are randomly divided into training, validation and testing cohorts by patients at the ratio of 80:10:10.

Our model is first tested with serial slides in H&E and of Ki67 IHC biomarker, followed by additional validation with serial slides in H&E and of PHH3 IHC biomarker. We compare our model with multiple state-of-the-art methods using both 'real' and 'synthetic' datasets. The 'real' dataset includes pairs of H&E and real IHC images, while the 'synthetic' dataset consists of real and synthetic (i.e. synIHC) IHC image pairs with synKi67 for testing and synPHH3 for validation, respectively. Note the dataset with synthetic IHC images by the CycleGAN is labeled 'syn-1', whereas that with synthetic IHC images by our modified CycleGAN model is labeled 'syn-2'. The developed CGNReg is implemented with the open-source deep learning library Tensorflow,[43] while experiments are executed on GPUs (i.e. Tesla K80 and V100) with CUDA 9.1. Adam optimization algorithm[44] with learning rate 0.0001 is used to train both image translation and image registration model. The modified CycleGAN is trained for up to 2 00 000 iterations. Loss weights λ_{cyc} and λ_{feat} are set to 1. All other parameter settings are suggested by the original CycleGAN work.[25] The values of registration loss weights σ_{11}, σ_{12}, σ_{13}, σ_2 and σ_3 are empirically set to 0.9, 0.6, 0.3, 0.05 and 0.05, respectively.

Evaluation of image translation

Both CycleGAN and our modified CycleGAN for image translation are evaluated and compared at the patch and WSI block level. Representative image patches and WSI blocks after translation are demonstrated in Fig. 4, Fig. 5, Supplementary Figure S3, and Supplementary Figure S4, respectively. In addition to the qualitative assessments, we quantitatively evaluate the translated IHC image quality by root mean square error (RMSE), structural similarity index measure (SSIM), and peak signal-to-noise ratio (PSNR) with the maximum value range 40 dB.[45,46,27] In our analysis pipeline, the trained generator G_{IHC} and G_{HE} are used to translate a H&E to a synIHC image, and then back from the synIHC to the synHE (i.e. black arrows in Fig. 2). The similarity between the real H&E and resulting synHE image is quantitatively evaluated and presented in Table 1. Both forward and reversed translation performances by the original and our modified CycleGAN model are presented. In our study, the image translation between H&E and Ki67 is for testing, while the H&E-PHH3 translation is for validation purpose. Suggested by experimental results, our proposed modified CycleGAN presents a consistent superior performance to the original CycleGAN by all evaluation metrics in both forward and reversed translation directions. Additionally, the translated synIHC image patches are spatially combined to generate the translated synIHC WSI blocks as

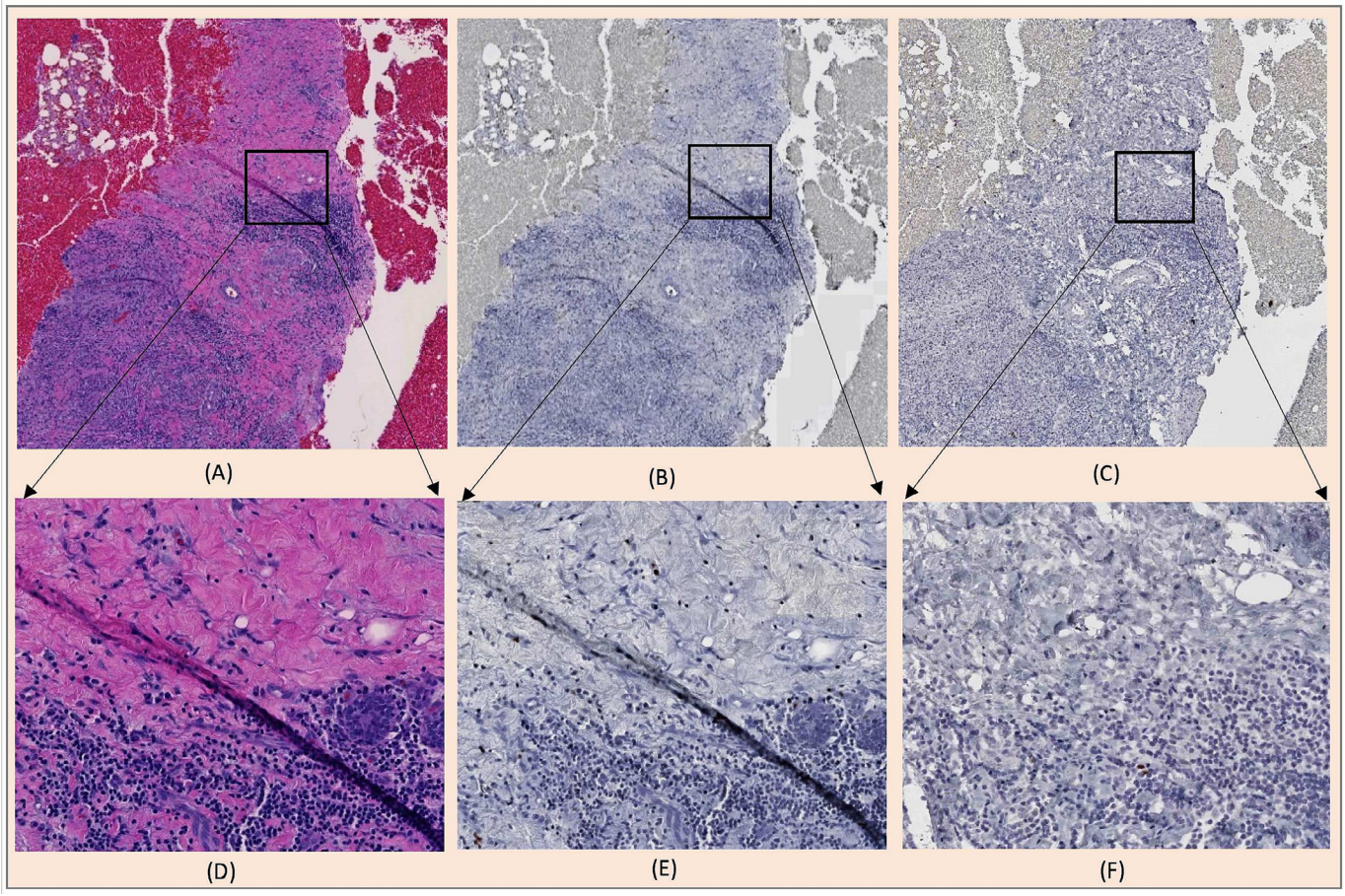

Fig. 5. Representative WSI block translation results. (A) A representative real H&E WSI block; (B) the synPHH3 WSI blocks by our modified CycleGAN; (C) the real PHH3 WSI block; Close-up views of (A–C) are presented in (D–F), respectively.

Table 1
Quantitative comparison of image translation performance between the original and our modified CycleGAN model on the testing and validation data. RMSE, SSIM and PSNR stand for root mean square error, structural similarity index and peak signal-to-noise ratio, respectively. The notation "→" represents the image translation direction. For each evaluation metric, the best performance value in each image translation direction is bolded.

Testing data				Validation data			
	RMSE	SSIM	PSNR		RMSE	SSIM	PSNR
CycleGAN (H&E → Ki67)	15.200	0.934	29.928	CycleGAN (H&E → PHH3)	17.641	0.905	28.528
Our (H&E → Ki67)	**14.414**	**0.935**	**30.387**	Our (H&E → PHH3)	**17.210**	**0.911**	**28.754**
CycleGAN (Ki67 → H&E)	13.088	0.955	31.215	CycleGAN (PHH3 → H&E)	15.114	0.928	29.956
Our (Ki67 → H&E)	**12.123**	**0.959**	**31.867**	Our (PHH3 → H&E)	**14.430**	**0.935**	**30.282**

illustrated in Fig. 5. Both qualitative and quantitative experiments with the testing and validation data suggest an enhanced image translation performance by our modified CycleGAN.

Evaluation of image patch registration

After the image translation, the resulting synIHC and real IHC WSI blocks are first pre-aligned by a global intensity-guided rigid transformation.[47] Next, CGNReg takes rigidly registered synIHC and real IHC WSI blocks as inputs and fine tunes image alignment by our proposed multi-scale FCN model. To evaluate the registration performance, we apply our method to the 'real', 'syn-1' and 'syn-2' dataset from both the testing data with H&E and Ki67 image pairs and validation data with H&E and PHH3 image pairs. For performance evaluation, our proposed multi-scale FCN model is compared with multiple state-of-the-art registration methods, including deep learning based DirNet,[48] FCN,[29] VoxelMorph Unet,[49] and the conventional image registration method Elastix.[50]

As the H&E and IHC WSI blocks are pre-aligned by the global affine transformation before registration methods for comparison in this study, the registration effect on certain image patches is not visually salient. To manifest the method efficacy, we thus manually deform the moving images by affine and elastic transformations, resulting in deformed moving images with significant deformations. Additionally, diverse shear transformations with the rotation angle range of $[-40, 30]$ and elastic deformations are applied to moving images. With such artificial transformation operations, 20000 synthetically deformed image patches of size 256×256 are derived from 1023 WSI blocks for the training purpose. Typical registration results by our and state-of-the-art deep learning registration models with moving

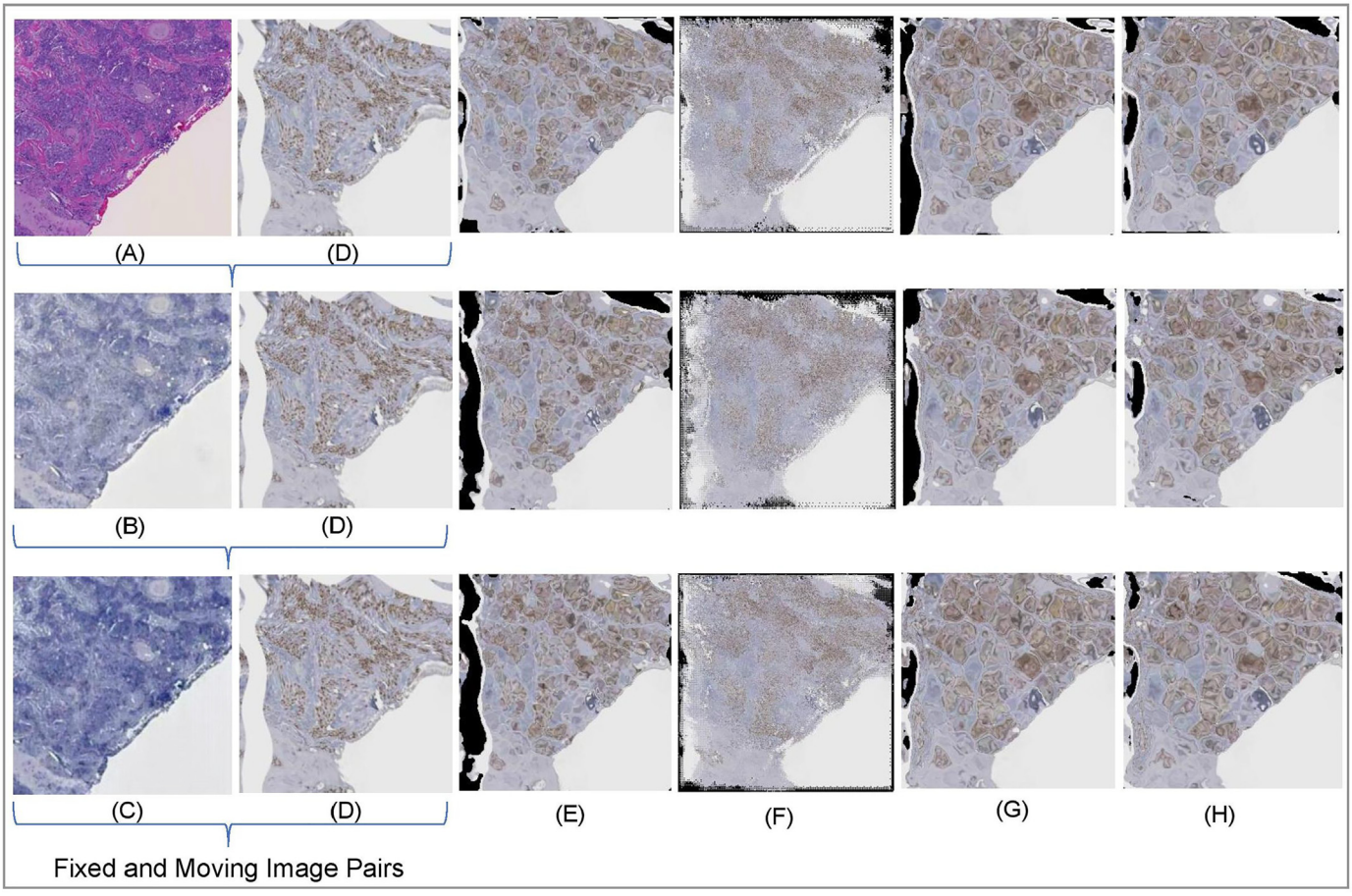

Fig. 6. Patch-based registration performance with the moving image artificially warped by an elastic deformation. (A) Fixed real H&E image. (B)The resulting synIHC image from (A) by CycleGAN. (C) The resulting synIHC image from (A) by our modified CycleGAN. (D) Moving images after manual transformations. Registration results by (E) DirNet, (F) VoxelMorph Unet, (G) FCN, and (H) our proposed multi-scale FCN.

images warped by the elastic deformation and a shear transformation are presented in Fig. 6 and Supplementary Figure S5, respectively. By visual assessments, FCN and our proposed multi-scale FCN model present noticeably better registration results than other deep learning models.

After method evaluations with manually deformed images, the same methods are applied to 'real', 'syn-1' and 'syn-2' from the testing and validation data, respectively. Representative registration results are demonstrated in Fig. 7 and Fig. 8, respectively. By visual comparisons, baseline FCN and our proposed multi-scale FCN models perform better than other methods for comparison when images with our modified CycleGAN from the 'syn-2' translation result set are used. Additionally, we present quantitative performance evaluation results in Table 2 where normalized cross correlation (NCC), SSIM, and normalized mutual information (NMI) are used to report the registration performance. Note our developed multi-scale FCN with the 'syn-2' translation result set from the testing data achieves the best performance (bold) by NCC and the second best (underlined) by SSIM and NMI. For the validation data, our developed multi-scale FCN achieves the third best (italics) with 'syn-2' by NCC, the second best with 'syn-2' by SSIM, and the best value with 'syn-1' by NMI. By contrast, the performance of the conventional registration method Elastix is limited due to its over-deformed image outputs (Supplementary Figure S6). Additionally, all deep learning-based models outperform with 'syn-2' than 'syn-1' or 'real' for the testing dataset, suggesting the efficacy of the enhanced image translation quality by our modified CycleGAN.

Additionally, 2 measures of the predicted DVF are used for deformation quality assessment. First, we compute the Jacobian determinant to evaluate

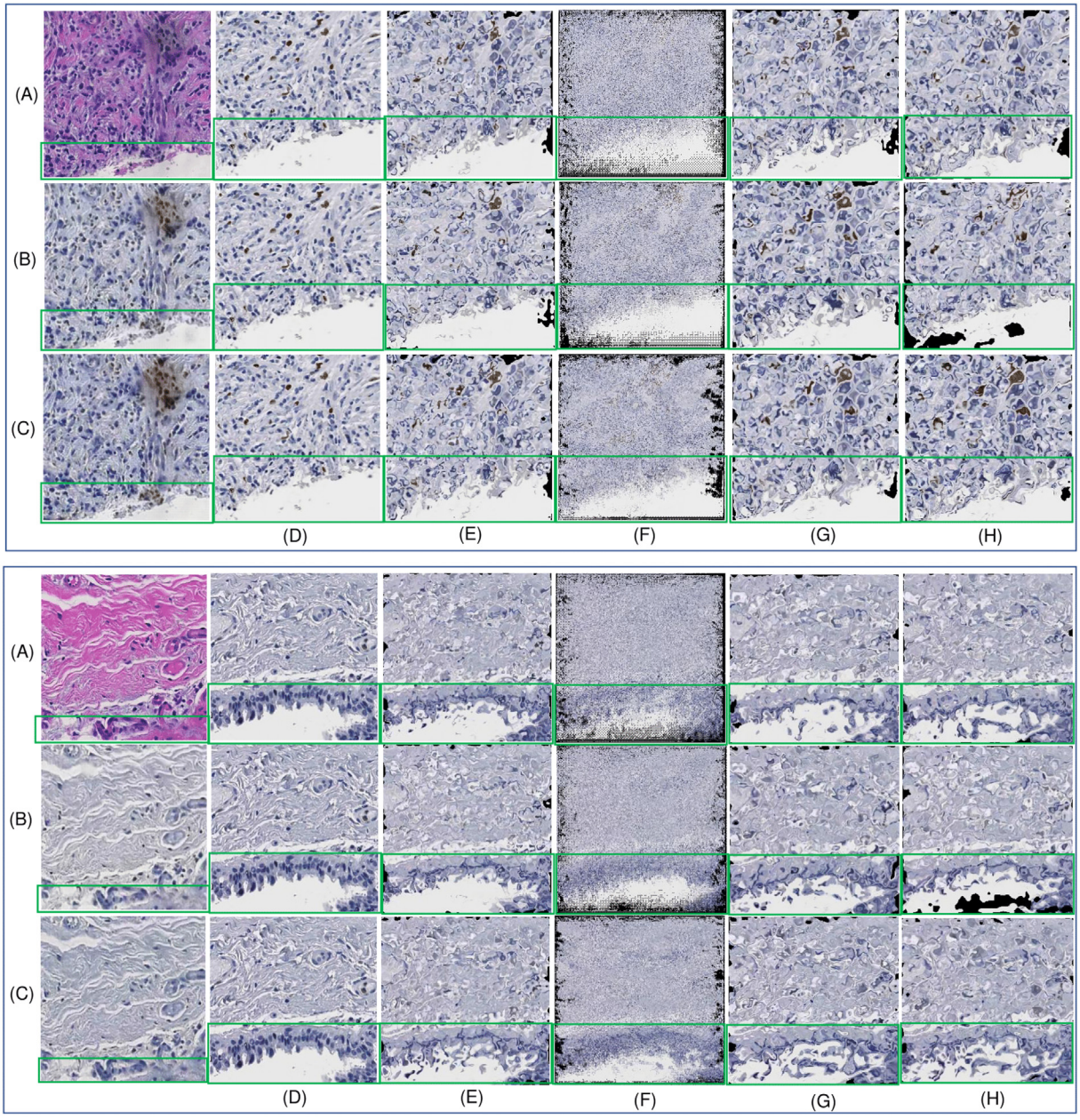

Fig. 7. Typical patch registration performance for the testing data with 2 example patch regions. (A) Fixed real H&E image. (B) The resulting synKi67 image from (A) by CycleGAN. (C) The resulting synKi67 image from (A) by our modified CycleGAN. (D) Real Ki67 moving image. Registration results by (E) DirNet, (F) VoxelMorph Unet, (G) FCN, and (H) our proposed multi-scale FCN. Green boxes are used to highlight the registration results.

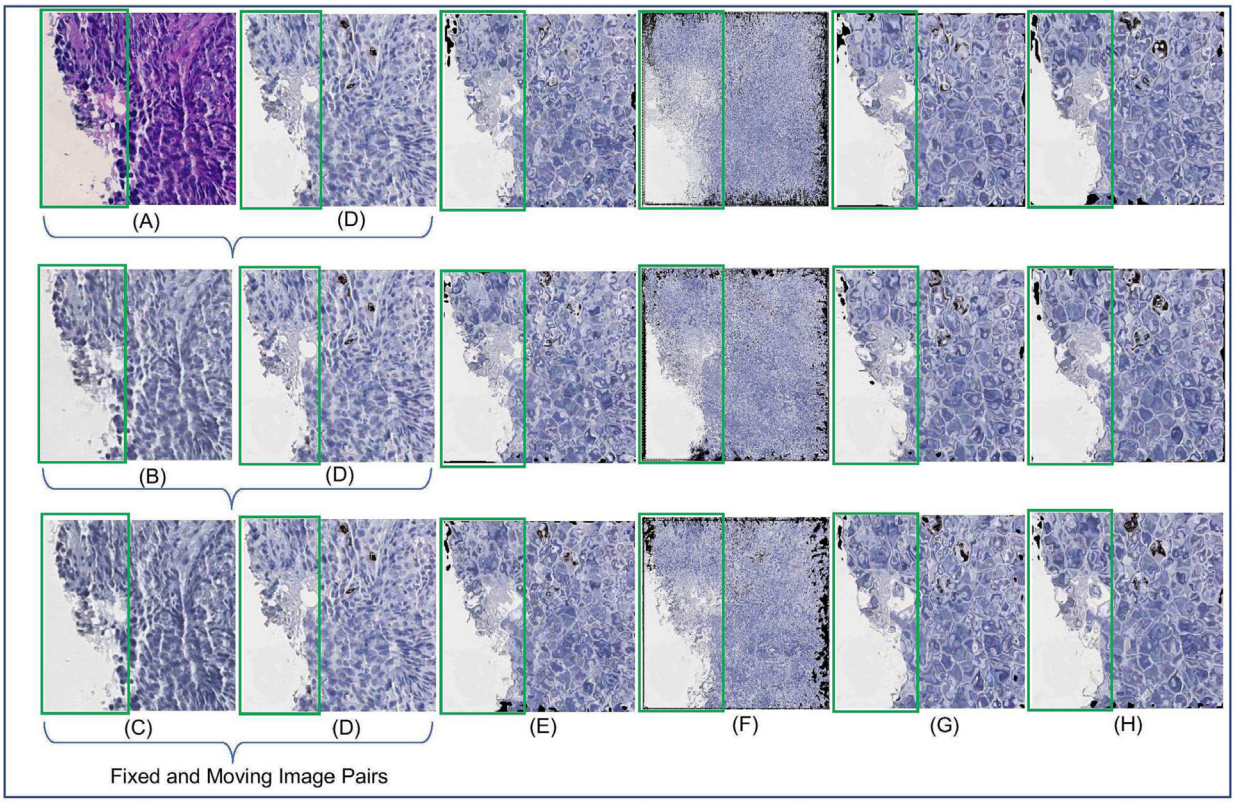

Fig. 8. Typical patch registration performance for the validation data. (A) Fixed real H&E image. (B) Fixed synPHH3 image from (A) by CycleGAN. (C) Fixed synPHH3 image from (A) by our modified CycleGAN. (D) Moving real PHH3 image. Registration results by (E) DirNet, (F) VoxelMorph Unet, (G) FCN, and (H) our proposed multi-scale FCN. Green boxes are used to highlight the registration results.

the invertibility of the DVF transformation.[51] The Jacobian determinant of the DVF (V) at a given point ($\vec{p}$) is defined as:

$$det\left[J\left(\vec{p}+V\left(\vec{p}\right)\right)\right]=\begin{vmatrix}\frac{\partial V_1\left(\vec{p}\right)}{\partial x}+1 & \frac{\partial V_1\left(\vec{p}\right)}{\partial y}\\ \frac{\partial V_2\left(\vec{p}\right)}{\partial x} & \frac{\partial V_2\left(\vec{p}\right)}{\partial y}+1\end{vmatrix} \quad (4)$$

where, $V_1\left(\vec{p}\right)$, $V_2\left(\vec{p}\right)$ are 2 components of the DVF $V\left(\vec{p}\right)$ at pixel $\vec{p}$. Negative Jacobian determinant at a pixel suggests that the local deformation is not invertible, causing the unrealistic folding artifacts within tissues. A Jacobian determinant either way larger than 1, very close to but above 0, or less than 0 indicates a poor deformation field quality. Suggested by literature, 1%–3% of negative Jacobian determinant is generally acceptable for a deformable image registration.[52] In Table 3, we present the percentage of negative Jacobian determinant (i.e. folding%), standard deviation, mean, and median of Jacobian determinant for the testing and validation data. Note that percentage of negative Jacobian is zero in all but VoxelMorph Unet method, suggesting the presence of unrealistic deformations by VoxelMorph Unet. This agrees with our visual assessments.

The second measure is Curl of the DVF that suggests local circular motions within a local tissue region.[52] A rapidly changing Curl in a DVF suggests an unrealistic deformation. In Fig. 9, we overlay the opposite deformation vectors in the associated DVF of a typical image pairs from 'syn-2' result set by DirNet, VoxelMorph Unet, FCN and our proposed multi-scale FCN with the corresponding Curl heatmaps. In this figure, the fixed image in (A) needs to be rotated to match the moving image in (B) denoted by the opposite deformation vectors. Suggested by the deformation vector directions and the Curl magnitudes, the DVF produced by the DirNet model in (C) does not capture such a circulation motion. Additionally, it is noticeable that the DVF from VoxelMorph Unet presents drastically changing vector directions in (D), resulting a high Curl value overall. By contrast, the FCN model (E) and our multi-scale FCN method (F) can produce a smooth DVF as suggested by the quiver plot and the Curl heatmap.

Table 2
Image patch registration performance on the testing (i.e. Ki67) and validation (i.e. PHH3) data. The 'real' dataset includes pairs of H&E and real IHC images. The 'syn-1' dataset includes pairs of real and synIHC images generated by the CycleGAN, while the 'syn-2' dataset includes pairs of real and synIHC images generated by our modified CycleGAN model. NCC, SSIM, and NMI stand for normalized cross correlation, structural similarity index, and normalized mutual information, respectively. For each metric, the best performance is in bold. We highlight our developed method performance by underline and italics when its achieves the second and third best performance, respectively.

Method name	Dataset	Testing data			Validation data		
		NCC	SSIM	NMI	NCC	SSIM	NMI
DirNet	Real	0.2045	0.2307	1.0358	0.2030	0.2413	1.0344
	syn-1	0.2644	0.3320	1.0356	0.2613	0.3161	1.0334
	syn-2	0.2977	**0.3505**	1.0374	0.2553	**0.3443**	1.0324
VoxelMorph UNet	Real	0.0231	0.1922	1.0329	0.0346	0.2078	1.0324
	syn-1	0.0755	0.3164	1.0314	0.0910	0.2828	1.0321
	syn-2	0.1063	0.3384	1.0347	0.1306	0.3243	1.0260
FCN	Real	0.2165	0.2168	1.0373	0.1742	0.2266	1.0362
	syn-1	0.2508	0.3119	1.0363	**0.2688**	0.3050	1.0339
	syn-2	0.3057	0.3356	**1.0392**	0.2683	0.3356	1.0340
Our multi-scale FCN	Real	0.1971	0.2214	1.0370	0.1776	0.2344	1.0360
	syn-1	0.2582	0.3096	1.0370	0.2321	0.2914	**1.0362**
	syn-2	**0.3161**	0.3403	1.0387	*0.2625*	0.3375	1.0343

Table 3
Percentage of folding, standard deviation, mean, and median of the Jacobian of DVFs with testing (i.e. Ki67) and validation (i.e. PHH3) data for image patch registration. The 'real' dataset includes pairs of H&E and real IHC images. The 'syn-1' dataset includes pairs of real and synIHC images generated by the CycleGAN, while the 'syn-2' dataset includes pairs of real and synIHC images generated by our modified CycleGAN model.

Method name	Dataset	Testing data				Validation data			
		Folding%	Std	Mean	Median	Folding%	Std	Mean	Median
DirNet	Real	0.0000	0.0118	0.9999	0.9996	0.0000	0.0131	0.9997	0.9992
	syn-1	0.0000	0.0128	0.9997	0.9994	0.0000	0.0121	0.9997	0.9995
	syn-2	0.0000	0.0118	0.9997	0.9993	0.0000	0.0124	0.9997	0.9992
VoxelMorph Unet	Real	1.4054e-04	0.1506	0.9979	0.9971	7.6896e-04	0.1449	1.0024	1.0006
	syn-1	1.2716e-05	0.1531	0.9999	0.9884	1.2046e-05	0.0963	1.0025	0.9999
	syn-2	2.0077e-06	0.1056	1.0000	0.9947	0.0000	0.0680	0.9973	0.9978
FCN	Real	0.0000	0.0133	0.9999	0.9996	0.0000	0.0156	1.0000	0.9993
	syn-1	0.0000	0.0125	0.9997	0.9991	0.0000	0.0146	0.9995	0.9989
	syn-2	0.0000	0.0133	0.9997	0.9992	0.0000	0.0144	0.9995	0.9985
Our multi-scale FCN	Real	0.0000	0.0131	1.0000	0.9996	0.0000	0.0141	0.9998	0.9990
	syn-1	0.0000	0.0132	0.9998	0.9993	0.0000	0.0126	1.0001	0.9998
	syn-2	0.0000	0.0125	0.9996	0.9990	0.0000	0.0128	0.9996	0.9989

Evaluation of WSI block registration

We further evaluate the image registration with WSI blocks. As each WSI block has a size of 8000 × 8000 pixels, it is partitioned into smaller image patches for the registration process. After individual image patch registrations, they are spatially assembled to generate the registered WSI blocks. In this study, we adopt Dice Similarity Coefficient (DSC),[53] Hausdorff Distance (HD),[54] SSIM, and NCC for WSI block registration accuracy evaluation.[55] For DSC and HD evaluation, 8 region of interest (ROI) pairs are manually annotated from the testing and validation WSI blocks. The complete evaluation process is presented in Supplemental Figure S7 and the quantitative results are shown in Table 4. Our multi-scale FCN model is compared with both state-of-the-art deep learning-based (i.e. DirNet, FCN, and VoxelMorph Unet) and conventional pathology image registration methods (i.e. Elastix, BUnwarpJ,[56] Diffeomorphic Demons,[57] and WSIReg[21]). Of all methods for comparison, our CGNReg consisting of the modified CycleGAN translation and the multi-scale FCN produces the best performance (bold) by NCC for both testing and validation data. Additionally, our method presents the second best performance (underlined) by DSC for the testing and validation data. By the metric of HD, our method achieves the second best and third best (italics) for testing and validation data, respectively. Although no method produces either the best or second best performance by all metrics for both testing and validation data, our multi-scale FCN model with our modified CycleGAN image translation method presents the best, second best or third best performance in most metrics. Results in Table 4 also suggest that the image translation from H&E to synIHC helps to improve registration performance in most cases.

We present the registration results of WSI blocks from the testing and validation data in Fig. 10 and Fig. 11, respectively. Additionally, we superimpose a representative WSI block in H&E with opposite deformation vectors in the associated DVF in Fig. 12. These vectors suggest the direction and magnitude of deformation at a given location in the H&E block to match the IHC block. Both quantitative and qualitative results suggest that our CGNReg exhibits promising registration performance.

Discussion

In this study, we develop a 2-stage CGNReg model for multi-stained serial WSI registration. By both qualitative and quantitative evaluations, CGNReg demonstrates promising registration performance by multiple metrics. Specifically, the registration component of CGNReg follows a coarse-to-fine multi-scale deformable image registration strategy and optimizes the joint loss at multi-scale levels, leading to a more accurate DVF estimation critical for an enhanced image registration. In Table 2, we notice that performance values (e.g. NCC and SSIM) for the patch-based registration are relatively low as some image patches are extracted from poorly matched WSI block pairs after the pre-alignment step (Supplement Figure S8). For patches from well aligned WSI block pairs after pre-alignment step, such performance values are much improved (c.f. Supplement Figure S9). To test the CGNReg efficacy, we synthetically deform moving image patches by different spatial transformations. In such well-controlled experiments, moving images can be well aligned to the fixed images by our method.

Both registration results at the patch and WSI block level suggest the necessity of image translation from one stain to another. Specifically, results in Table 4 suggest that the 'syn-1' and 'syn-2' translation datasets can produce better registration performance than the 'real' dataset in most cases. Additionally, the translated dataset 'syn-2' yields better registration performance than the 'syn-1' dataset by most metrics, suggesting a better DVF estimation enabled by synIHC images from our modified CycleGAN. Meanwhile, our multi-scale FCN with 'syn-2' translated dataset consistently produces the best, second best, or third best performance values for most metrics, suggesting the superiority of our model to other methods for comparison.

Overall, there is a strong need in a large spectrum of disease investigations where integrative analyses with molecular, cellular and tissue-level interactions are in need.[59–65] Among other diseases, one immediate application our developed registration method can enable is to evaluate the predictive value of the combination of tumors, tumor-infiltrating lymphocytes, Ki67, and PHH3 biomarkers for enhanced response to chemotherapy, metastasis, relapse, and clinical survival outcome prediction for TNBCs.[12,11,58] Although we evaluate the efficacy of our registration method by breast cancer serial WSIs in H&E, Ki67, and PHH3 in this paper, it is not necessary to limit to these nuclear markers for the developed registration method validation. Such H&E stain and 2 IHC markers are chosen due to their known predictive values for predicting response to neoadjuvant chemotherapy in breast cancer patients.

As our developed registration method spatially aligns pathology images of serial tissue slides in different stains or biomarkers at the full image resolution, pathologists can review such multiplexed biomarkers within the

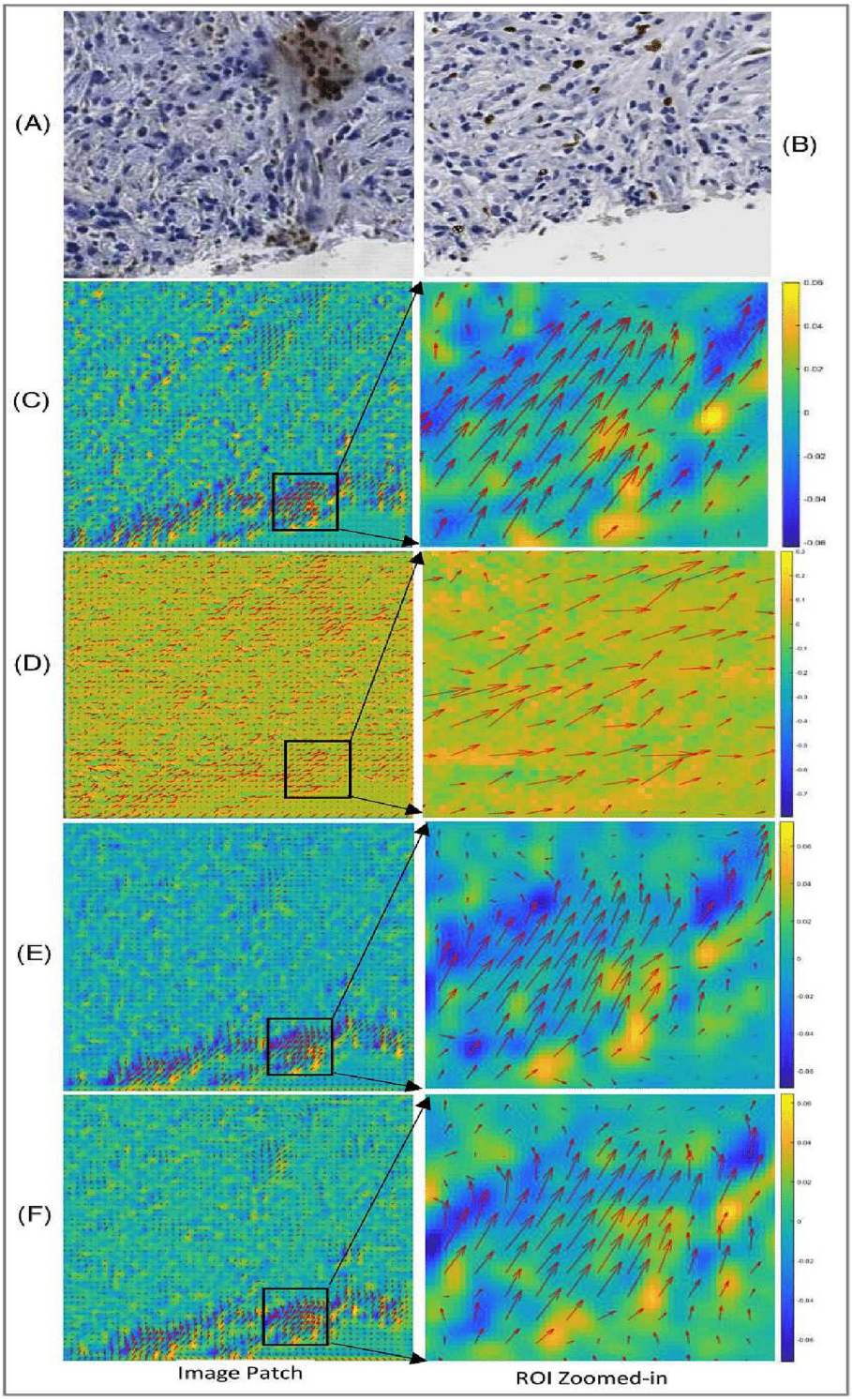

Fig. 9. Typical patch-based DVFs overlaid with Curl heatmaps by different methods. (A) Fixed synKi67 image by our modified CycleGAN. (B) Real Ki67 moving image. Quiver plots of DVFs overlaid with Curl values by (C) DirNet, (D) VoxelMorph Unet, (E) FCN, and (F) our multi-scale FCN.

Table 4
Registration performance of WSI blocks from the testing (i.e. Ki67) and validation (i.e. PHH3) data. The 'real' dataset includes pairs of H&E and real IHC images. The 'syn-1' dataset includes pairs of real and synIHC images generated by the CycleGAN, while the 'syn-2' dataset includes pairs of real and synIHC images generated by our modified CycleGAN model. DSC, HD, SSIM, and NCC stand for dice similarity coefficient, Hausdorff distance, structural similarity index, and normalized cross correlation, respectively. For each metric, the best performance is in bold. We highlight our developed method performance by underline and italics when it achieves the second and third best performance, respectively.

Testing data						Validation data			
Dataset	Method	DSC	HD	SSIM	NCC	DSC	HD	SSIM	NCC
Real	Elastix	0.9423	1240.1000	0.7348	0.6372	0.8881	1214.1727	0.7534	0.5276
	BUnwarpJ	0.9357	3.3348e + 03	0.8056	0.7508	0.9115	634.4545	0.8122	0.7492
	Diffeomorphic Demons	0.9153	367.4675	0.8138	0.7562	0.8880	550.8127	0.8080	0.7336
	WSIReg	0.8797	4154.1401	0.7732	0.5677	0.8510	4447.2181	0.7745	0.4657
	DirNet	0.9516	925.6837	0.7914	0.7983	0.9072	1290.1709	0.8153	0.7447
	FCN	0.9566	606.5160	0.7887	0.7833	0.9088	1302.1997	0.8145	0.7323
	Our multi-scale FCN	0.9454	320.6497	0.8155	0.7905	0.9127	680.1516	0.8173	0.7432
syn-1	BUnwarpJ	0.9578	3.1562e + 03	0.8214	0.7638	0.9078	1.4102e + 03	0.8581	0.7076
	Diffeomorphic Demons	0.9351	329.2142	0.8287	0.7728	0.8884	550.8945	0.8495	0.7378
	WSIReg	0.9439	3159.1797	0.8076	0.6528	0.8481	4533.4090	0.8194	0.4527
	DirNet	0.9510	272.2740	0.8362	0.7958	0.9004	1300.1804	0.8659	0.7241
	FCN	0.9572	272.1319	0.8347	0.7988	**0.9161**	984.5001	**0.8709**	0.7639
	Our Multi-scale FCN	0.9587	1.0520e + 03	0.8316	0.7855	0.9100	1.2442e + 03	0.8690	0.7398
syn-2	BUnwarpJ	0.9171	3.5903e + 03	0.8343	0.6967	0.9073	1.3083e + 03	0.8515	0.7040
	Diffeomorphic Demons	0.9124	368.1703	**0.8465**	0.7641	0.8885	550.8127	0.8488	0.7355
	WSIReg	0.8826	3418.5034	0.8126	0.5816	0.8514	4311.2858	0.8139	0.4571
	DirNet	0.9554	264.6888	0.8369	0.8075	0.9073	**486.4495**	0.8641	0.7576
	FCN	**0.9631**	**244.2928**	0.8345	0.8101	0.9149	1216.2255	0.8637	0.7614
	Our Multi-scale FCN	0.9612	257.2176	0.8347	**0.8175**	0.9152	*724.4312*	0.8654	**0.7683**

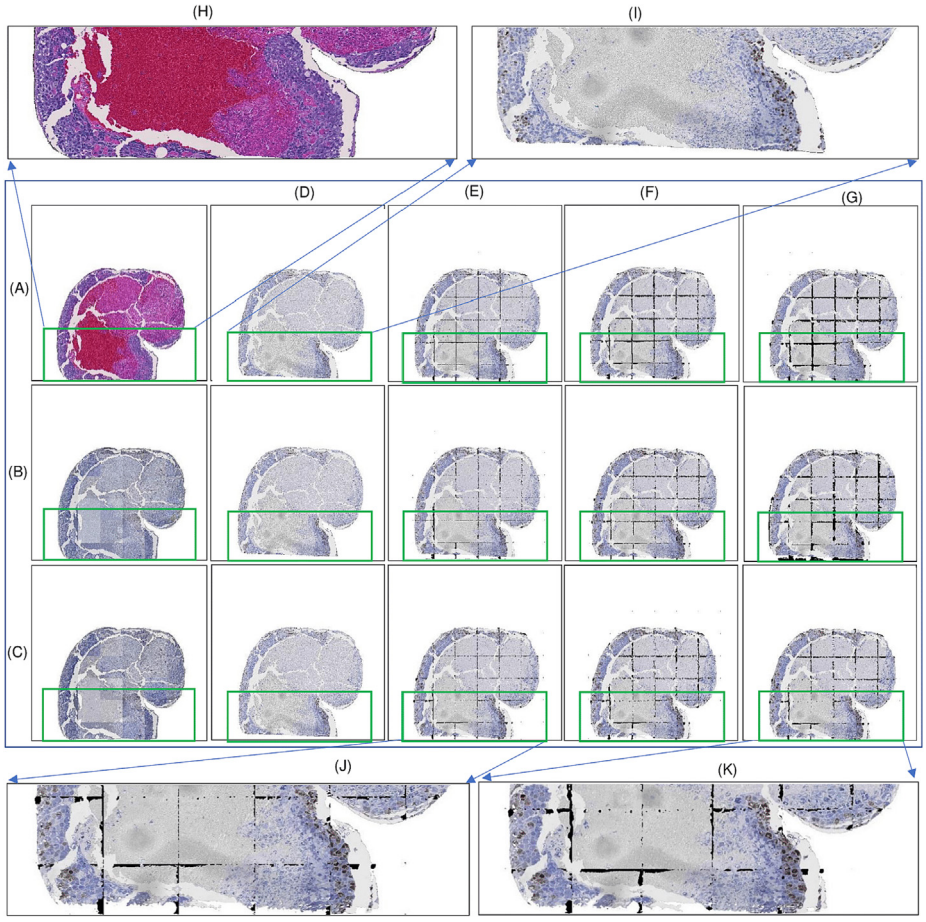

Fig. 10. WSI block registration with H&E and Ki67 IHC testing image data. (A) Fixed real H&E image block. (B) The resulting synKi67 image block translated from (A) by CycleGAN. (C) The resulting synKi67 image block translated from (A) by our modified CycleGAN. (D) Real Ki67 moving image block. Registration results by (E) DirNet, (F) FCN, and (G) our multi-scale FCN. (H) Close-up view of a tissue region in (A). (I) Close-up view of a tissue in (D). (J) Close-up view of a tissue in the registered block by DirNet with the 'syn-2' translated dataset; and (K) close-up view of a tissue in the registered block by CGNReg.

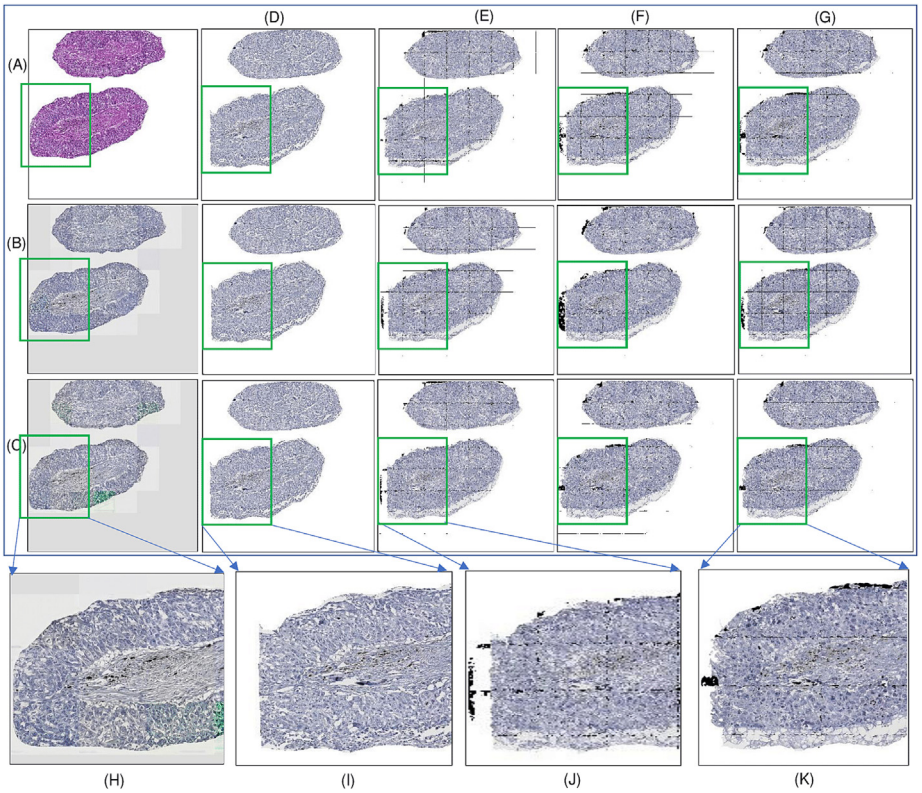

Fig. 11. WSI block registration with H&E and PHH3 IHC validation image data. (A) Fixed real H&E image block. (B) The resulting synPHH3 image block translated from (A) by CycleGAN. (C) The resulting synPHH3 image block translated from (A) by our modified CycleGAN;.(D) Real PHH3 moving image block. Registration results by (E) DirNet, (F) FCN, and (G) our multi-scale FCN. (H) Close-up view of a tissue region in (A). (I) Close-up view of a tissue in (D). (J) Close-up view of a tissue in the registered block by DirNet with the 'syn-2' translated dataset; and (K) close-up view of a tissue in the registered block by CGNReg.

same tissue space simultaneously. Therefore, this upgraded reviewing function overlaying multiple markers with the tissue histology can facilitate human readers to have a more comprehensive and accurate understanding of the phenotype–genotype tissue environment, leading to a better prognosis and treatment decision.

In future research, we plan to improve this work by better learning and integrating spatial transformations from the H&E-IHC and synIHC-IHC pipeline for registration. In our current study, we extract non-overlapping patches from each WSI for deep learning model training and evaluation. We will extract overlapping patches to better control the displacement vectors at the patch borders in future. In this study, our goal is to map a given H&E image to another given IHC stain before registration with 2 images of the same IHC biomarker. As each such mapping from H&E to a given IHC stain is independent, we need to train our modified CycleGAN for each such mapping. In future, such an image translation step can be replaced with an advanced model that supports simultaneous H&E mappings to multiple IHC markers of interest. While this study includes 1023 large image blocks (8000 × 8000) and further 60 000 image patches (1024 × 1024) from 228 WSIs of 76 TNBC patients, we will further validate our developed method by increasing the TNBC patient number and testing on additional disease datasets of serial tissue sections in future study.

Conclusion

To support the common integrative analyses involving histology hallmarks and molecular biomarkers in a large spectrum of tissue based biomedical research, we develop a fully unsupervised translation based deep learning registration network (namely CGNReg) that spatially aligns WSIs of serial tissue sections in H&E stain and IHC biomarkers without prior deformation information for the model training. In the first preparation step, CGNReg translates unpaired H&E to synthetic IHC images through a modified CycleGAN method equipped with a perceptual loss term for an enhanced image stain translation. In the next step, a multi-scale FCN is developed to align the resulting synthetic to the real IHC serial WSIs with a coarse-to-fine multi-scale deformable image registration strategy, preserving histology structures at the full WSI resolution. The efficacy of CGNReg is tested with a dataset of 76 breast cancer patients with 1 H&E and 2 IHC serial WSIs for each patient. After CGNReg is trained without ground-truth image deformation information, it presents competitive performance compared with other state-of-the-art methods by both qualitative and quantitative evaluations. Such promising results at both image patch and WSI level suggest its potential for multi-stained serial WSI registration essential to a large number of tissue-based integrative biomedical research studies.

Author Declaration

There is no conflict of interest to report.

Acknowledgements

This research was partially supported by National Institutes of Health [U01CA242936, R01CA239120], and National Science Foundation [ACI 1443054, IIS 1350885].

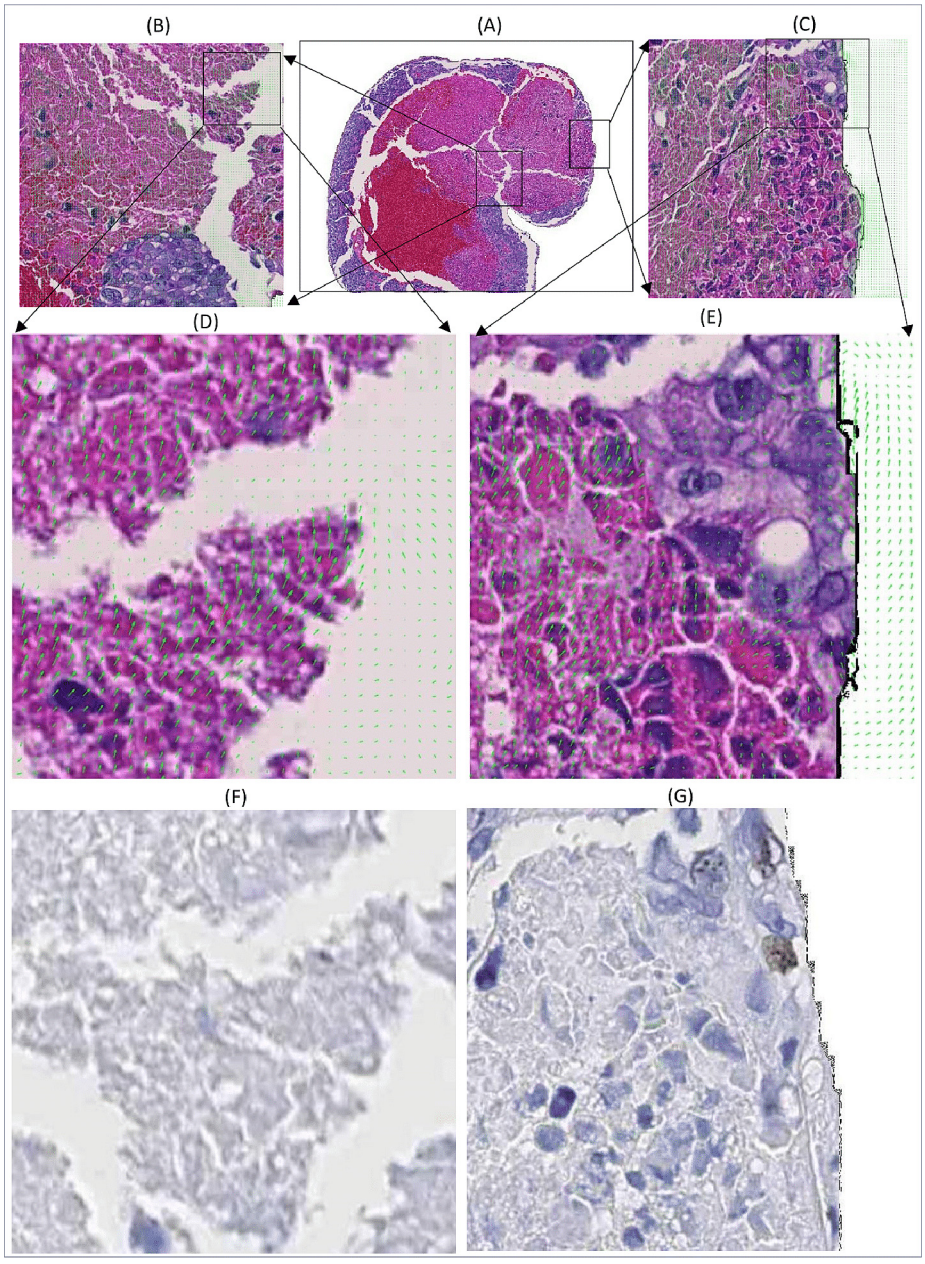

Fig. 12. A typical WSI block overlaid with the negative DVF quiver plot. (A) Fixed real H&E block. (B–C) Close-up views of 2 tissue regions with the negative DVF overlaid. (D–E) Close-up views of (B–C) at a higher resolution. (F–G) Associated regions in the Ki67 IHC moving blocks.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jpi.2023.100311.

References

1. Bychkov D, Linder N, Turkki R, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. Scient Rep 2018;8(1):3395.
2. Kather JN, Krisam J, Charoentong P, et al. Predicting survival from colorectal cancer histology slides using deep learning: a retrospective multicenter study. PLoS Med 2019;16 (1), e1002730.
3. Kather JN, Pearson AT, Halama N, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med 2019;25(7):1054–1056.
4. Lu MY, Williamson DFK, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. Nat Biomed Eng 2021;5(6):555–570.
5. Long C, Jay S, Cole T, et al. The human body at cellular resolution: The NIH human biomolecular atlas program. Nature 2019;574(7777):187–192.
6. Marco F, Heibel H, Gardiazabal J, Navab N, Groher M. Reconstruction of 3-D histology images by simultaneous deformable registration. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011: 14th International Conference, Toronto, Canada, September 18-22, 2011, Proceedings, Part II 14. Springer Berlin Heidelberg; 2011. p. 582–589.
7. Minoru M, Sasano H, Tamaki K, et al. Prognostic significance of tumor-infiltrating CD8 + and FOXP3 + lymphocytes in residual tumors and alterations in these parameters after neoadjuvant chemotherapy in triple-negative breast cancer: a retrospective multicenter study. Breast Cancer Res 2015;17:1-13.
8. Hae PI, Kong S-Y, Ro JY, et al. Prognostic implications of tumor-infiltrating lymphocytes in association with programmed death ligand 1 expression in early-stage breast cancer. Clin Breast Cancer 2016;16(1):51–58.
9. Liu X, Qiu H, Zhang P, et al. Ki-67 labeling index may be a promising indicator to identify "very high-risk" gastrointestinal stromal tumor: a multicenter retrospective study of 1022 patients. Human Pathol 2018;74:17–24.
10. Boutros PC, Fraser M, Harding NJ, et al. Spatial genomic heterogeneity within localized, multifocal prostate cancer. Nat Genet 2015;47(7):736–745.
11. Wang H, Mao X. Evaluation of the efficacy of neoadjuvant chemotherapy for breast cancer. Drug Design Dev Therapy 2020:2423–2433.
12. Zhang A, Wang X, Fan C, Mao X. The role of Ki67 in evaluating neoadjuvant endocrine therapy of hormone receptor-positive breast cancer. Front Endocrinol 2021;1426.

13. Elston CW, Ellis IO. Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up. Histopathology 1991;19(5):403–410.
14. Sillem M, Timme S, Bronsert P, et al. Anti-phosphohistone H3-positive mitoses are linked to pathological response in neoadjuvantly treated breast cancer. Breast Care 2017;12(4): 244–250.
15. van Steenhoven JEC, Kuijer A, Kornegoor R, et al. Assessment of tumour proliferation by use of the mitotic activity index, and Ki67 and phosphohistone H3 expression, in early-stage luminal breast cancer. Histopathology 2020;77(4):579–587.
16. Borovec J, Kybic J, Arganda-Carreras I, et al. ANHIR: automatic non-rigid histological image registration challenge. IEEE Trans Med Imaging 2020;39(10):3042–3052.
17. Lotz J, Weiss N, van der Laak J. High-resolution image registration of consecutive and restained sections in histopathology. 2021:1–19. arXiv preprint arXiv:abs/2106.13150.
18. Liu P, Wang F, Teodoro G, Kong J. Histopathology image registration by integrated texture and spatial proximity based landmark selection and modification. In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). IEEE; 2021. p. 1827–1830.
19. Wodzinski M, Skalski A. Multistep, automatic and nonrigid image registration method for histology samples acquired using multiple stains. Phys Med Biol 2021;66(2), 025006.
20. Hoque MZ, Keskinarkaus A, Nyberg P, Mattila T, Seppänen T. Whole slide image registration via multi-stained feature matching. Comput Biol Med 2022;144, 105301.
21. https://github.com/NHPatterson/wsireg.
22. Tanner C, Ozdemir F, Profanter R, Vishnevsky V, Konukoglu E, Goksel O. Generative adversarial networks for MR-CT deformable image registration. 2018:1–11. arXiv preprint arXiv:abs/1807.07349.
23. Wei D, Ahmad S, Huo J, et al. Synthesis and inpainting-based mr-ct registration for image-guided thermal ablation of liver tumors. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer; 2019. p. 512–520.
24. Qin C, Shi B, Liao R, Mansi T, Rueckert D, Kamen A. Unsupervised deformable registration for multi-modal images via disentangled representations. International Conference on Information Processing in Medical Imaging. Cham: Springer; 2019. p. 249–261.
25. Zhu J-Y, Park T, Isola P, Efros AA. Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE International Conference on Computer Vision; 2017. p. 2223–2232.
26. Wolterink JM, Dinkla AM, Savenije MHF, Seevinck PR, van den Berg CAT, Isgum I. Deep MR to CT synthesis using unpaired data. International Workshop on Simulation and Synthesis in Medical Imaging. Cham: Springer; 2017. p. 14–23.
27. Xu Z, Luo J, Yan J, et al. Adversarial uni-and multi-modal stream networks for multimodal image registration. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer; 2020. p. 222–232.
28. Xu Z, Li X, Zhu X, Chen L, He Y, Chen Y. Effective Immunohistochemistry Pathology Microscopy Image Generation Using CycleGAN. Front Mol Biosci 2020;7.
29. Li H, Fan Y. Non-rigid image registration using self-supervised fully convolutional networks without training data. In 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018). IEEE; 2018. p. 1075–1078.
30. Jaderberg M, Simonyan K, Zisserman A. Spatial transformer networks. Adv Neural Inform Process Syst 2015;28:2017–2025.
31. Hiasa Y, Otake Y, Takao M, et al. Cross-modality image synthesis from unpaired data using CycleGAN. International Workshop on Simulation and Synthesis in Medical Imaging. Cham: Springer; 2018. p. 31–41.
32. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2016. p. 770–778.
33. Isola P, Zhu J-Y, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017. p. 1125–1134.
34. Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. European Conference on Computer Vision. Cham: Springer; 2016. p. 694–711.
35. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Proceedings of the International Conference on Learning Representations; 2015. p. 1-14. arXiv:1409.1556.
36. Russakovsky O, Deng J, Su H, et al. Imagenet large scale visual recognition challenge. Int J Comput Vision 2015;115(3):211–252.
37. Sokooti H, De Vos B, Berendsen F, Lelieveldt BPF, Isgum I, Staring M. Nonrigid image registration using multi-scale 3D convolutional neural networks. International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer; 2017. p. 232–239.
38. de Vos BD, Berendsen FF, Viergever MA, Sokooti H, Staring M, Isgum I. A deep learning framework for unsupervised affine and deformable image registration. Med Image Anal 2019;52:128–143.
39. Briechle K, Hanebeck UD. Template matching using fast normalized cross correlation. Optical Pattern Recognition XII. International Society for Optics and Photonics; 2001. p. 95-102.
40. Vishnevskiy V, Gass T, Szekely G, Tanner C, Goksel O. Isotropic total variation regularization of displacements in parametric image registration. IEEE Trans Med Imaging 2016;36(2):385–395.
41. Wang X, Yu K, Wu S, et al. Esrgan: enhanced super-resolution generative adversarial networks. Proceedings of the European Conference on Computer Vision (ECCV) Workshops; 2018.
42. https://blogs.mathworks.com/steve/2013/08/07/tiff-bigtiff-and-blockproc/.
43. Abadi M, Barham P, Chen J, et al. Tensorflow: a system for large-scale machine learning. 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16); 2016. p. 265–283.
44. Kingma DP, Ba J. Adam: a method for stochastic optimization. Proceedings of the International Conference on Learning Representations; 2015. p. 1-13. 2018:1-11. arXiv: 1412.6980.
45. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13(4):600–612.
46. Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM. In 2010 20th International Conference on Pattern Recognition. IEEE; 2010. p. 2366–2369.
47. Knaan D, Joskowicz L. Effective intensity-based 2D/3D rigid registration between fluoroscopic X-ray and CT. International Conference on Medical Image Computing and Computer-Assisted Intervention. Berlin, Heidelberg: Springer; 2003. p. 351–358.
48. de Vos BD, Berendsen FF, Viergever MA, Staring M, Isgum I. End-to-end unsupervised deformable image registration with a convolutional neural network. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. Cham: Springer; 2017. p. 204–212.
49. Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. Voxelmorph: a learning framework for deformable medical image registration. IEEE Trans Med Imaging 2019;38(8): 1788–1800.
50. Marstal K, Berendsen F, Staring M, Klein S. SimpleElastix: a user-friendly, multi-lingual library for medical image registration. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops; 2016. p. 134–142.
51. Heath Emily, Collins DL, Keall PJ, Dong L, Seuntjens J. Quantification of accuracy of the automated nonlinear image matching and anatomical labeling (ANIMAL) nonlinear registration algorithm for 4D CT images of lung. Med Phys 2007;34(11):4409–4421.
52. Nie X, Huang K, Deasy J, Rimner A, Li G. Enhanced super-resolution reconstruction of T1w time-resolved 4DMRI in low-contrast tissue using 2-step hybrid deformable image registration. J Appl Clin Med Phys 2020;21(10):25–39.
53. Thada V, Jaglan V. Comparison of jaccard, dice, cosine similarity coefficient to find best fitness value for web retrieved documents using genetic algorithm. Int J Innov Eng Technol 2013;2(4):202–205.
54. Huttenlocher DP, Klanderman GA, Rucklidge WJ. Comparing images using the Hausdorff distance. IEEE Trans Pattern Anal Mach Intel 1993;15(9):850–863.
55. Hoque MZ, Keskinarkaus A, Nyberg P, Mattila T, Seppänen T. Whole slide image registration via multi-stained feature matching. Comput Biol Med 2022;144, 105301.
56. Arganda-Carreras I, Sorzano COS, Kybic J, Ortiz-de-Solorzano C. bunwarpj: consistent and elastic registration in imagej, methods and applications. Second ImageJ User & Developer Conference, vol. 12. ; 2008.
57. Vercauteren T, Pennec X, Perchant A, Ayache N. Diffeomorphic demons: efficient non-parametric image registration. NeuroImage 2009;45(1):S61–S72.
58. Ellis MJ, Suman VJ, Hoog J, et al. Ki67 proliferation index as a tool for chemotherapy decisions during and after neoadjuvant aromatase inhibitor treatment of breast cancer: results from the American College of Surgeons Oncology Group Z1031 Trial (Alliance). J Clin Oncol 2017;35(10):1061.
59. Bhargava R, Madabhushi A. Emerging themes in image informatics and molecular analysis for digital pathology. Annu Rev Biomed Eng 2016;18:387–412.
60. Wang Q, Lou W, Di W, Wu X. Prognostic value of tumor PD-L1 expression combined with CD8+ tumor infiltrating lymphocytes in high grade serous ovarian cancer. Int Immunopharmacol 2017;52:7-14.
61. Rabjerg M. Identification and validation of novel prognostic markers in renal cell carcinoma. PhD diss. Syddansk Universitet. 2016.
62. Liu H, Lin C, Shen Z, et al. Decreased expression of granulocyte-macrophage colony-stimulating factor is associated with adverse clinical outcome in patients with gastric cancer undergoing gastrectomy. Oncol Lett 2017;14(4):4701–4707.
63. Wang P, Zhou Y, Mei Q, Zhao J, Huang L, Fu Q. PPA1 regulates tumor malignant potential and clinical outcome of colon adenocarcinoma through JNK pathways. Oncotarget 2017;8(35):58611.
64. Gomez-Brouchet A, Illac C, Gilhodes J, et al. CD163-positive tumor-associated macrophages and CD8-positive cytotoxic lymphocytes are powerful diagnostic markers for the therapeutic stratification of osteosarcoma patients: an immunohistochemical analysis of the biopsies fromthe French OS2006 phase 3 trial. Oncoimmunology 2017;6(9), e1331193.
65. Ichiki A, Carreras J, Miyaoka M, et al. Clinicopathological analysis of 320 cases of diffuse large B-cell lymphoma using the Hans classifier. J Clin Exp Hematopathol 2017;57(2): 54–63.