A Gaussian Filter-Based 3D Registration Method for Series Section Electron Microscopy

Zhenbang Zhang[1,2†], **Hongjia Li**[3†], **Zhiqiang Xu**[4], **Wenjia Meng**[1*], **Renmin Han**[2*]
[1]School of Software, Shandong University
[2]Research Center for Mathematics and Interdisciplinary Sciences, Shandong University
[3]School of Medical Technology, Beijing Institute of Technology
[4]Mohamed bin Zayed University of Artificial Intelligence
{zhangzhenbang2021,lihongjia96}@gmail.com, zhiqiang.xu@mbzuai.ac.ae, {wjmeng, hanrenmin}@sdu.edu.cn

Abstract

Series Section Electron Microscopy (ssEM) is a crucial technique for visualizing three-dimensional (3D) biological structures, which involves collecting electron microscopy images from a series of biological sections along the z-axis and reconstructing the 3D structure. 3D registration is an essential step in ssEM, designed to eliminate axial misalignment and nonlinear distortions introduced during sample sectioning. A significant challenge in 3D registration is eliminating nonlinear distortions while preserving natural deformations. In this paper, we present a new formulation of the 3D registration problem from a frequency domain perspective and propose a Gaussian filtering-based 3D registration method, which defines 3D registration as a superposition problem of high-frequency and low-frequency components. We extend the concept of a one-dimensional Gaussian filter to three-dimensional image stacks and integrate it with optical flow networks to consolidate the deformation field within the receptive field. Extensive experiments demonstrate that our method can successfully decouple nonlinear distortions and natural deformations in the frequency domain, proving superior to existing methods in rapidly and accurately eliminating nonlinear distortions and restoring biological structures, and has the potential to be extended to large datasets.

Introduction

Series Section Electron Microscopy (ssEM) is a widely used high-resolution three-dimensional (3D) visualization technique for biological tissues (Terasaki et al. 2013; Müller et al. 2021; Xu et al. 2021) and cells (Ryan, Lu, and Meinertzhagen 2016; Wilson et al. 2022), which has been adopted across various domains, including life sciences (Morgan and Lichtman 2020; Baena and Terasaki 2019), medicine (de Senneville et al. 2021; Johnson et al. 2022), and clinical diagnostics (Weidner et al. 2009; Shoemark et al. 2020). Specifically, it has been instrumental in connectomics studies elucidating neuronal connections (Zheng et al. 2018) and in reconstructing the 3D ultrastructure of diverse cells and tissues (Yin et al. 2020; Wilson and Babadi 2023). In ssEM, the workflow typically involves slicing biological samples along the z-axis into a series of microscope sections, which are then subjected to staining, digital imaging, 2D stitching, 3D registration, and finally, reconstruction to obtain the 3D structure (Kievits et al. 2022). Notably, 3D registration is the most critical step in this process, which serves to eliminate axial nonlinear distortions introduced during sectioning and restore the faithful 3D structure. However, the complexity of cellular structures, coupled with the inevitable accumulation of errors in the registration of long-sequence large datasets, poses significant challenges. Furthermore, distinguishing between nonlinear distortions and natural deformation adds to the difficulty in solving the problem of 3D registration.

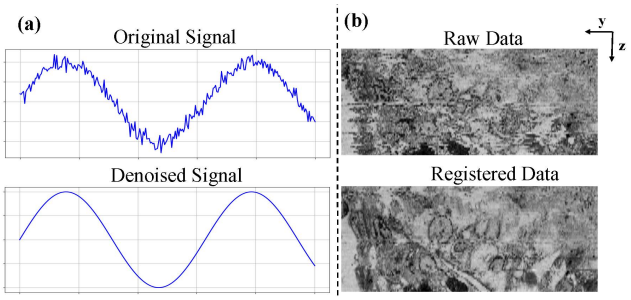

Figure 1: (a) One-dimensional signal with high frequency noise (Original Signal) and Gaussian-filtered signal (Denoised Signal). (b) The side view (y-z plane) of original ssEM data (Raw Data) and registered ssEM data (Registered Data). Our method can effectively remove nonlinear distortions while preserving natural deformations.

3D registration typically is cast as an optimization problem over slices, with numerous energy-based methods developed (Arganda-Carreras et al. 2006; Saalfeld et al. 2012). However, these traditional techniques often suffer from complex parameter tuning, leading to suboptimal performance. Deep learning methods (Yoo et al. 2017; Zhou et al. 2019; Liu et al. 2023) have shown significant progress in registration speed. Nevertheless, most of these methods often struggle to handle natural deformations between slices, potentially leading to damage to axial structures.

In this paper, we examine the 3D registration problem

*Corresponding authors; †Joint first authors. This work is done during Zhang's internship in RCMIS.
Copyright © 2025, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

from a frequency domain perspective and introduce a Gaussian filter-based 3D registration method designed to eliminate axial nonlinear distortions while preserving natural deformations. As illustrated in Figure 1, the natural deformations, associated with low-frequency signals, change gradually, whereas nonlinear distortions, associated with high-frequency signals, occur abruptly and subtly. Inspired by (Gaffling et al. 2014), we conceptualize the 3D registration problem as a superposition of high-frequency and low-frequency components. We propose applying a one-dimensional (1D) Gaussian filter to a 3D image stack, while integrating the deformation field using neural networks. Specifically, we extend the 1D Gaussian filter to a 3D image stack, leveraging its robust smoothing capabilities to effectively remove nonlinear distortions while retaining natural deformations. We further use a cascaded pyramid optical flow estimation network to estimate deformation fields between adjacent slices from coarse to fine. The estimated deformation fields are integrated by a Gaussian filter to achieve precise distortion elimination. Furthermore, we introduce a weighted loss function over multiple adjacent slices to maintain the axial continuity. We conduct extensive experiments on synthetic and real datasets, which demonstrate that our method outperforms the state-of-the-art in terms of rapidly and accurately eliminating nonlinear distortions and restoring biological structures, and has the potential to be extended to large datasets. Our contributions are summarized as follows:

- We examine the 3D registration problem from a frequency domain perspective and propose a Gaussian filter-based method that can eliminate nonlinear distortions while preserving natural deformations.
- We extend the concept of the 1D Gaussian filter to 3D image stacks, which integrates deformation fields among adjacent slices, for effectively smoothing the structure and precisely eliminating nonlinear distortions.
- We conducted comprehensive experiments on multiple synthetic and real datasets, demonstrating that our method exhibits outstanding robustness in handling complex deformations and long-sequence error accumulation. It surpasses existing methods in terms of registration accuracy and speed and has the potential to be extended to large datasets.

Related Works

3D Registration of ssEM Images Stack. Several traditional software packages have been developed for the 3D registration of ssEM datasets, including AlignTK (Lee et al. 2016), IMOD (Kremer, Mastronarde, and McIntosh 1996), StackReg (Thevenaz, Ruttimann, and Unser 1998) and TrakEM2 (Cardona et al. 2012). TrakEM2 is one of the most popular tools, which achieves 3D registration by iteratively optimizing a spring-connected particle system in which each section is represented as a triangular spring mesh. However, TrakEM2 faces a series of issues, such as high computational cost and limited registration accuracy. Furthermore, Gaffling et al. proposes a Gauss-Seidel iteration-based method aimed at eliminating nonlinear distortions and recovering the correct 3D structure (Gaffling et al. 2014). However, this method does not provide a specific mathematical definition of the 3D registration problem and faces the risk of natural deformation diminishment due to excessive iterations. Recently, deep learning methods (Yoo et al. 2017; Zhou et al. 2019; Liu et al. 2023) have emerged as promising alternatives for 3D registration. For instance, Xin et al. proposed a structured regression approach to reduce error accumulation, but their method faces challenges in extending to long sequences (Xin et al. 2023). Another method, SEAMLeSS (Popovych et al. 2024) can be extended to large-scale datasets through parallel computing. Nevertheless, SEAMLeSS does not specifically address the issue of distinguishing between nonlinear distortions and natural deformations.

Medical Image Registration. Medical image registration serves as the most analogous task to 3D registration and has witnessed numerous exciting advancements (Zhou et al. 2023; Pielawski et al. 2020; Chen et al. 2019; Zhao et al. 2019; Haskins, Kruger, and Yan 2020; Hering et al. 2022). It can be categorized into two types based on parameter transformation models. The first type encompasses dense transformation models (Balakrishnan et al. 2018; Dalca et al. 2018; Krebs et al. 2019; Dalca et al. 2019), which directly estimate the transformation for each pixel to acquire a dense deformation field. The second type involves interpolation-based transformation models (Jud et al. 2016; Vishnevskiy et al. 2016; Sandkühler et al. 2019; Shen, Vialard, and Niethammer 2019), utilizing a set of fixed functions (e.g., B-splines) over a grid in the image domain to approximate the image transformation. While both types have been extensively studied and yielded significant results, they primarily focus on image pairs. In contrast, 3D registration is more complex, requiring consideration not only of the nonlinear distortions but also the axial continuity and biologically natural deformations.

Method

Problem Formulation

In the problem of 3D registration, consider an image stack $\{I_i\}_{i=0}^{N}$, which is distorted by nonlinear deformation caused by sample section or optical effect. The objective of 3D registration is to find a series of deformation fields $\{\Phi_i\}_{i=0}^{N}$, where deformation field Φ_i registers image I_i, in order to eliminate distortion and reconstruct a faithful 3D structure from the registered image stack $\{\hat{I}_i\}_{i=0}^{N}$. The primary challenge lies in finding a deformation field Φ_i that not only effectively eliminates nonlinear distortions but also preserves natural deformations. In this paper, we further refine the mathematical definition of 3D registration. Specifically, for each slice I_i and its neighboring slices $\mathcal{N}(I_i)$, we aim to identify a deformation field that eradicates nonlinear distortions while maintaining natural deformations. Motivated by this reconception, the following objective function is specifically designed:

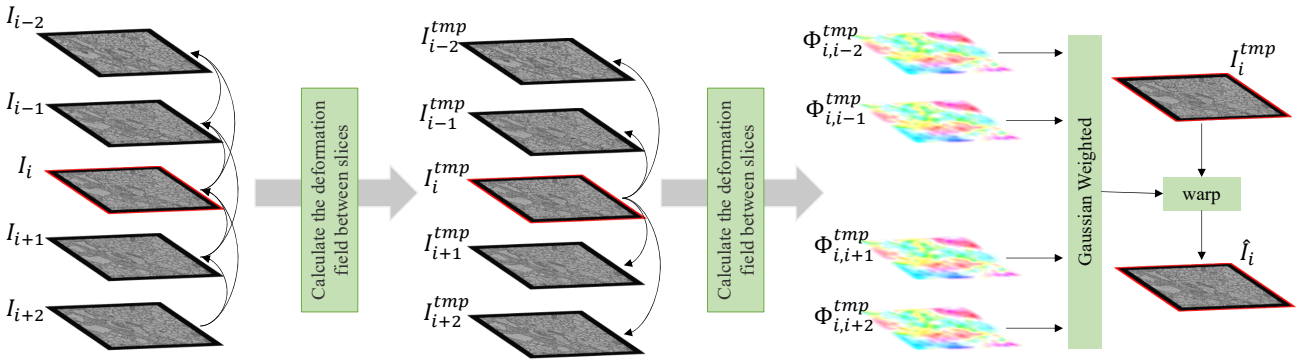

Figure 2: Overview of our method. It predicts the deformation field between slices using dense optical flow estimation network and performs 3D registration through the Gaussian filter algorithm.

Algorithm 1: Algorithm of our method

Input: $stack$: Image stack
Parameter: r: Gaussian kernel radius, L: Maximum registration distance, T: Number of iterations (default 2)
Output: $stack$: Registered image stack

while $T \neq 0$ **do**
  for I_i in $stack$ **do**
    Compute $\{w_k^{tmp}\}_{k=-r+1}^{r}$ and $\{I_k^{tmp}\}_{k=-r+1}^{r}$ using Eq.(8)
    $\{\bar{w}_k^{tmp}\}_{k=-r+1}^{r}$ = norm($\{w_k^{tmp}\}_{k=-r+1}^{r}$)
    Compute $\hat{I}_i$ using Eq.(9)
    $stack \leftarrow \hat{I}_i$
  end for
  $stack \leftarrow$ reverse($stack$)
  $T \leftarrow T - 1$
end while

$$\Phi_i = \arg\min_{\Phi} \sum_{\forall I_j \in \mathcal{N}(I_i)} F(I_i, I_j, \Phi_i) + \lambda ||\nabla \Phi_i||_2^2, \quad (1)$$

where I_i denotes the slice to be registered, $\mathcal{N}(I_i)$ represents the neighboring slices of I_i, function F estimates the difference between the I_i deformed by Φ and the neighboring slice I_j, and $\lambda||\nabla\Phi||_2^2$ acts as a regularization term.

Framework of Proposed Method

We extend the application of the 1D Gaussian filter to 3D image stacks to leverage its powerful smoothing capabilities (Roberts and Mullis 1987), thereby eliminating nonlinear distortions and restoring a smoother 3D structure. The overview of the algorithm is presented in Algorithm 1. A 1D Gaussian filter is modeled along the axial direction of the image stack $\{I_i\}_{i=0}^{N}$. At each window of the Gaussian filter, an optical flow estimation network is employed to estimate the deformation field between slices, which is then combined by Gaussian filtering. Figure 2 illustrates the filtering operation at the I-th window. At each filtering operation, the deformation field Φ_i is outputted to deform slice I_i and we sequentially apply this process until the last slice. In the next section, we first demonstrate our method on a simple 1D model and subsequently extend it to 3D image stacks.

Gaussian Filter in Simple 1D Model

Given a 1D signal f of length N, and a Gaussian kernel w with the radius r, the filtering operation at point i is expressed as the following:

$$\hat{f}_i = \sum_{k=-r}^{r} w_k f_{i+k}, \quad (2)$$

where w_k denotes the weight of the k-th element of the Gaussian kernel, r is the radius of the Gaussian kernel, f_i represents the i-th value of the signal, and $\hat{f}_i$ represents the filtered value at position i.

Assuming the receptive field radius of $r = 1$, the filtering operation for f_i is expressed as:

$$\hat{f}_i = w_{-1} f_{i-1} + w_0 f_i + w_1 f_{i+1}. \quad (3)$$

Let us define $d_{i,j} = f_j - f_i$ for any indices i and j, then Eq.(3) can be rewritten as:

$$\hat{f}_i = w_{-1} d_{i,i-1} + w_0 d_{i+1,i} + w_{-1} f_i + (w_0 + w_1) f_{i+1}. \quad (4)$$

Gaussian Filter in 3D Image Stacks

Let's consider the application of a similar 1D Gaussian filter to a 3D image stack, where each point f_i in a 1D signal is analogous to an image I_i in the image stack. Here, $d_{i,j}$ represents the deformation field $\Phi_{i,j}$ that register slice I_i to slice I_j. Therefore, the Gaussian filter operation for slice I_i can be expressed as:

$$\hat{I}_i = w_{-1} I_i \circ w_{-1} \Phi_{i,i-1} + (w_0 + w_1) I_{i+1} \circ w_0 \Phi_{i+1,i}, \quad (5)$$

where we assume the radius of the Gaussian kernel to be 1, I_i represents the i-th slice, $\Phi_{i,j}$ represents the deformation field from slice I_i to slice I_j and $\circ$ denotes the register operation.

However, the limitation in Eq.(5) is evident, as each slice is registered only with its immediately preceding slice,

thereby ignoring potentially valuable displacement infor tion from slices with larger distances. As demonstratec Figure 3(a), the slice I_{i+1} is only registered with I_i but I_{i-1}. To expand the registration to include more compreh sive interactions among multiple adjacent slices, we prop the following adjustment:

$$\hat{I}_i = (w_{-1} + w_0)I_i \circ w_{-1}\Phi_{i,i-1} + (w_{-1} + w_0 + w_1)I_{i+1} \circ (w_{-1}\Phi_{i+1,i-1} + w_0\Phi_{i+1,i})$$

where we register each slice with its multiple neighb allowing the registration process to consider a larger a range of natural deformations, as shown in Figure 3(b).

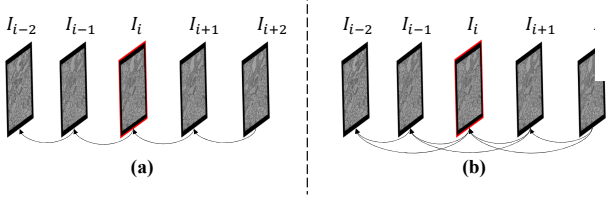

Figure 3: (a) Each slice is registered only with the immediately preceding slice. (b) Each slice is registered with multiple neighboring slices, which considers a larger axial range of natural deformations.

According to Eq.(6), it can be observed that $\hat{I}_i$ incorporates weighted contributions from both the deformed I_i and the deformed I_{i+1}. This unfortunately introduces image information from I_{i+1}, which is not desirable. To retain the true information only from I_i, we made adjustment to Eq.(6). Specifically, we deform I_i using a weighted deformation field calculated among multiple adjacent slices, rather than directly using information from adjacent slices. The deformation fields for I_i with adjacent slices are calculated based on the registered adjacent slices. Therefore, the calculation for the deformed $\hat{I}_i$ is as follows:

$$\begin{cases} I_i^{tmp} = I_i \circ w_{-1}\Phi_{i,i-1}, \\ I_{i+1}^{tmp} = I_{i+1} \circ (w_{-1}\Phi_{i+1,i-1} + w_0\Phi_{i+1,i}), \\ \bar{w}_0^{tmp} = \frac{(w_{-1}+w_0)}{(w_{-1}+w_0)+(w_{-1}+w_0+w_1)}, \\ \bar{w}_1^{tmp} = \frac{(w_{-1}+w_0+w_1))}{(w_{-1}+w_0)+(w_{-1}+w_0+w_1)}, \\ \hat{I}_i = I_i \circ \frac{\Phi_{i,i-1}+w_0^{tmp}\Phi_{i,i}^{tmp}+w_1^{tmp}\Phi_{i,i+1}^{tmp}}{2}, \end{cases} \quad (7)$$

where we denote the deformed $\{I_i, I_{i+1}\}$ as intermediate images $\{I_{i+k}^{tmp}\}_{k=0}^{1}$, and their corresponding normalized weights as $\{\bar{w}_k^{tmp}\}_{k=0}^{1}$. Then,we register I_i to $\{I_{i+k}^{tmp}\}_{k=0}^{1}$ to obtain $\{\Phi_{i,i+k}^{tmp}\}_{k=0}^{1}$. After being weighted by $\{\bar{w}_k^{tmp}\}_{k=0}^{1}$, $\{\Phi_{i,i+k}^{tmp}\}_{k=0}^{1}$ are used for registration to obtain $\hat{I}_i$.

However, this raises another issue: if the kernel radius r is large, the morphological image differences within the same receptive field $[-r, r]$ will be significant, which can affect the registration effectiveness. Therefore, we set the maximum registration distance for each slice to L ($L \leq r$). Generally, assuming the radius is r and the maximum registration distance is L, we have:

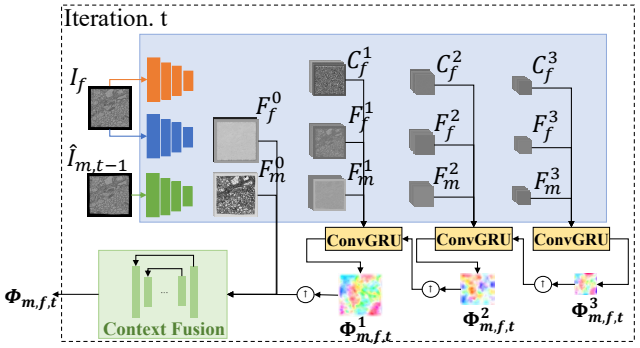

Figure 4: The framework of optical flow estimation network for predicting the deformation field between slices.

$$\begin{cases} w_k^{tmp} = w_k + \sum_{l=\max(-r,k-L)}^{k-1} w_l, \\ I_{i+k}^{tmp} = I_{i+k} \circ \sum_{l=\max(-r,k-L)}^{k-1} (w_l\Phi_{i+k,i+l}), \end{cases} \quad (8)$$

where w_k^{tmp} represents the k-th unnormalized weight in $\{w_k^{tmp}\}_{k=-r+1}^{r}$, and I_{i+k}^{tmp} represents the k-th intermediate image in $\{I_k^{tmp}\}_{k=-r+1}^{r}$. Finally, we obtain the final registered image $\hat{I}_i$ by the following:

$$\hat{I}_i = I_i \circ \frac{\Phi_{i,i-1} + \sum_{k=-r+1}^{r} \bar{w}_k^{tmp}\Phi_{i,i+k}^{tmp}}{2}. \quad (9)$$

Optical Flow Estimation Network

High-precision deformation fields contribute to achieving accurate 3D registration. We use a cascaded optical flow estimation network to estimate the deformation field between slices. We use a cascaded optical flow estimation network to estimate the deformation field between slices using a coarse-to-fine strategy. As demonstrated in Figure 4, we break down the generation of the deformation field $\Phi_{m,f}$ into T iterations. For each iteration, such as the t-th iteration, the input to the network are the fixed image I_f and warped image $\hat{I}_{m,t-1}$ warped by the deformation field $\Phi_{m,t-1}$ from the previous iteration.

Feature Pyramid. Our network employs a 2D convolution encoder to extract feature pyramids. Specifically, the encoder processes the current registered image $\hat{I}_{m,t-1}$ to generate a feature pyramid $\{F_m^s\}_{s=0}^{3}$, and it simultaneously extracts both a feature pyramid $\{F_f^s\}_{s=0}^{3}$ and a contextual feature pyramid $\{C_f^s\}_{s=1}^{3}$ from the preceding image I_f. The dimensions of F_m^0, F_m^1, F_m^2, and F_m^3 are $1/2$, $1/4$, $1/8$, and $1/16$ of I_m, respectively, with channel numbers of 8, 16, 32, and 64. The number of channels in each layer of $\{C_f^s\}_{s=1}^{3}$ and $\{F_f^s\}_{s=0}^{3}$ is kept consistent with that of $\{F_m^s\}_{s=0}^{3}$.

Optical Flow Estimation. Given $\{F_f^s\}_{s=1}^{3}$, $\{F_m^s\}_{s=1}^{3}$, and $\{C_f^s\}_{s=1}^{3}$, we divide them into three sets of features

Datasets\Method		EMReg	EFSR	TrakEM2	SEMLeSS	Ours
Mouse Kidney	MI (GT)	0.65 ± 0.16	1.26 ± 0.15	1.24 ± 0.10	1.07 ± 0.14	**1.54 ± 0.11**
	SSIM (GT)	0.44 ± 0.04	0.63 ± 0.05	0.61 ± 0.03	0.52 ± 0.04	**0.78 ± 0.03**
	NCC (GT)	0.81 ± 0.05	0.97 ± 0.01	0.97 ± 0.00	0.96 ± 0.01	**0.99 ± 0.00**
Mouse Liver	MI (GT)	0.79 ± 0.18	1.36 ± 0.16	1.47 ± 0.08	1.29 ± 0.19	**1.83 ± 0.08**
	SSIM (GT)	0.43 ± 0.03	0.56 ± 0.04	0.61 ± 0.03	0.52 ± 0.06	**0.80 ± 0.02**
	NCC (GT)	0.78 ± 0.05	0.95 ± 0.01	0.97 ± 0.00	0.95 ± 0.01	**0.99 ± 0.00**
Mouse Skin	MI (GT)	0.66 ± 0.18	1.03 ± 0.20	1.08 ± 0.15	0.89 ± 0.18	**1.40 ± 0.16**
	SSIM (GT)	0.50 ± 0.07	0.60 ± 0.07	0.64 ± 0.06	0.51 ± 0.07	**0.82 ± 0.03**
	NCC (GT)	0.87 ± 0.04	0.97 ± 0.01	0.98 ± 0.01	0.96 ± 0.01	**0.99 ± 0.00**

Table 1: Quantitative results of various methods on the Mouse Kidney, Mouse Liver, and Mouse Skin datasets.

based on their scale, denoted as $\{F_f^s, F_m^s, C_f^s\}_{s=1}^3$. These features are then separately input into three individual recurrent generation modules (ConvGRU). Simultaneously, each ConvGRU takes the deformation field $\Phi_{m,f,t}^{s+1}$ as supplementary input, achieving a coarse-to-fine update of the deformation field $\Phi_{m,f,t}^s$, and the update formula is as follows:

$$\begin{cases} z^s = \sigma(Conv_{3\times3}([h_f^s, F_f^s, F_m^s], W_z^s)), \\ r^s = \sigma(Conv_{3\times3}([h_f^s, F_f^s, F_m^s], W_r^s)), \\ \tilde{h}_f^s = tanh(Conv_{3\times3}([r^s \odot h_f^s, F_f^s, F_m^i], W_h^s)), \\ h_f^s = (1 - z^s) \odot h_f^s + z^s \odot \tilde{h}_f^s, \end{cases} \tag{10}$$

where $Conv_{3\times3}$ represents a 3×3 convolutional layer, and W denotes the corresponding learnable kernel parameters.

Context Fusion. In the final stage of the network, we input the deformation field $\Phi_{m,f,t}^1$ and F_f^0, F_m^0 into the Context Fusion Module, thereby introducing broader contextual information to enhance the accuracy of the deformation field. The Context Fusion Module is a simple U-net network and ultimately outputs the refined deformation field $\Phi_{m,f,t}$.

Loss Function. According to the optimization objective defined in Eq.(1), our loss function is composed of two primary components: similarity loss and regularization loss. The similarity loss is formulated to compute the weighted similarity across k adjacent slices. The regularization loss focuses on the deformation fields Φ_t obtained at each iteration, ensuring that the transformations are smooth and physically plausible. The final loss function is expressed as:

$$L_{total} = \sum_{i=-k, i\neq0}^{k} w_i ||\hat{I}_m - I_{m+i}||_2^2 + \lambda \sum_{t=1}^{T} ||\nabla \Phi_{m,f,t}||_2^2, \tag{11}$$

where k represents the number of neighboring slices and w_i represents the weight between $\hat{I}_m$ and I_{m+i},which is inversely proportional to the distance between adjacent slices. T is the total number of iterations.

Compared to the traditional similarity loss (Liu et al. 2023), our loss function L_s not only considers the similarity with the reference image I_{m-1} but also takes into account the weighted similarity between multiple adjacent slices, which ensures the axial structural continuity.

Experiments

Experimental Settings

We compared our proposed method with four widely used 3D registration methods, including a commonly applied traditional method, TrakEM2 (Cardona et al. 2012), and three recent deep learning methods, EMReg (Liu et al. 2023), EFSR (Xin et al. 2023), and SEMLeSS (Popovych et al. 2024). We evaluated the performance of 3D elastic registration from different perspectives by statistically analyzing the NCC (GT), SSIM (GT), and MI (GT) between the registered image stacks of each method and the ground truth.

Datasets. We evaluate our method on three datasets from the OpenOrganelle platform (Xu et al. 2021). These datasets include electron microscope images of various tissues such as mouse kidney[1], mouse liver[2], and mouse skin[3]. We cropped these images into slices of size 1184×1184 and added different degrees of elastic deformation to simulate nonlinear distortion. The lengths of the three datasets are 1000, 557, and 1230, respectively, which helps us assess performance in terms of error accumulation. These datasets provide us with diverse and representative biological samples, which are useful for comprehensively evaluating the applicability and robustness of our method across different biological tissues. More details about the data usage are provided in the appendix.

Implementation. Our method is implemented using PyTorch (Paszke et al. 2019), and all experiments were conducted on an NVIDIA A100 GPU. During training, we used the Adam optimizer with a learning rate of 5e-3, a random seed of 42. The weight of the regularization loss λ was set to 1.8. The network architecture included three cascading steps by default. The feature pyramid had three levels, corresponding to 1/4, 1/8, and 1/16 of the original size. In the loss function, we chose $k = 1$. In the Gaussian filtering strategy, the kernel radius was fixed at $r = 1$, σ was set to 3.0, and the maximum registration distance was $L = 1$.

[1]https://openorganelle.janelia.org/datasets/jrc_mus-kidney
[2]https://openorganelle.janelia.org/datasets/jrc_mus-liver
[3]https://openorganelle.janelia.org/datasets/jrc_mus-skin-1

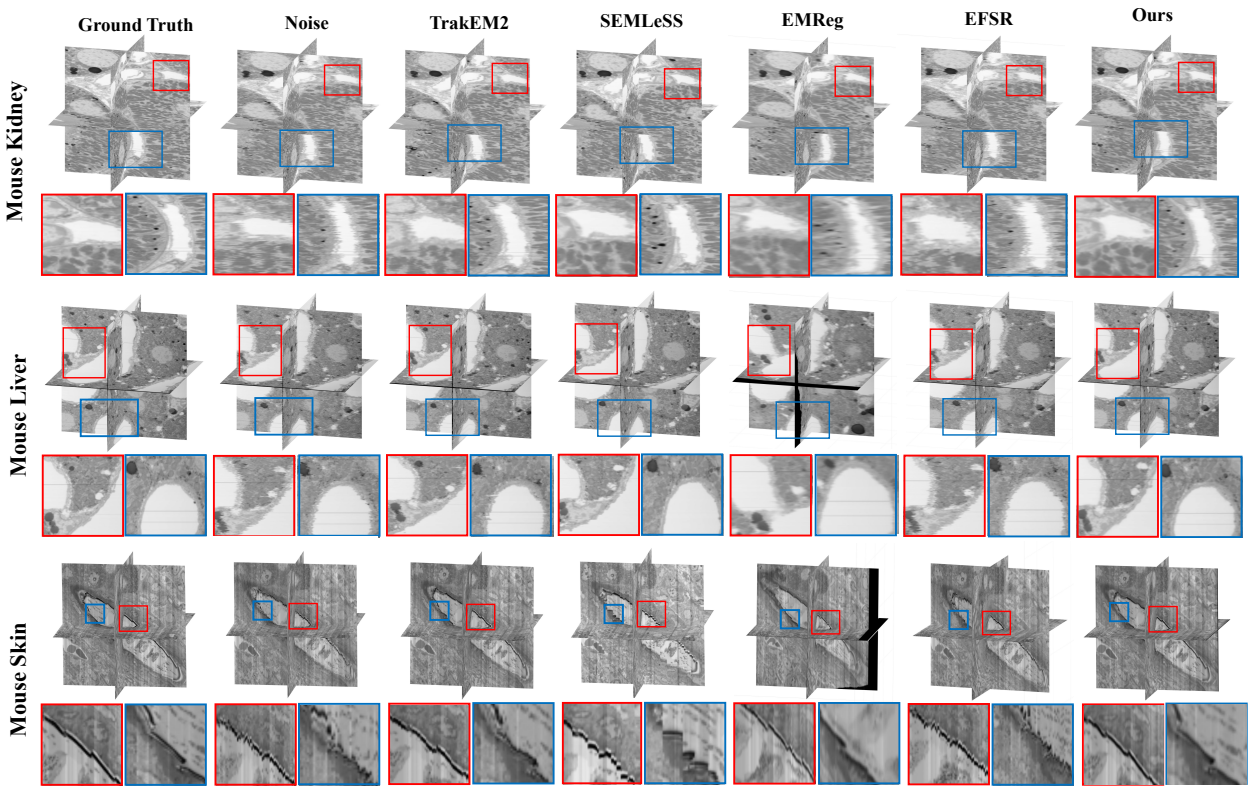

Figure 5: 3D visualization of registration results for various methods on the Mouse Kidney, Mouse Liver, and Mouse Skin datasets.

Method	Time cost (s)
TrakEM2	3.52
EFSR	3.38
EMReg	0.25
SEMLeSS	29.04
Ours	0.72

Table 2: Average time cost (s) by each method for registering a slice

Experimental Results

Results for synthetic data. We evaluated the registration results of various methods on mouse kidney, mouse liver, and mouse skin datasets using NCC (GT), SSIM (GT), and MI (GT), which assess the registration quality from different perspectives. As shown in the quantitative results in Table 1, our proposed method consistently demonstrates superior performance across all metrics on all datasets. This is attributed to our novel perspective on addressing the 3D elastic registration problem in the frequency domain. We extract and integrate displacement field information between adjacent slices and use Gaussian filters to decouple nonlinear distortion and natural deformation.

Figure 5 shows the 3D visualization results of various methods on three datasets. It can be seen that our method achieves better and smoother registration results, which are closest to the ground truth. This is especially evident in the mouse skin dataset, which involves cellular structures with significant inter-slice gaps. In this case, our method demonstrates smooth and continuous registration results even for structures with very long spans, highlighting its advantage in long-sequence registration. Additionally, we measured the average time required by ours and the comparison methods to register a single slice. As shown in Table 2, our method can complete the registration of a single slice in less than one second, demonstrating its potential for handling large-scale datasets. Additional experimental results are provided in the appendix.

Robustness to error accumulation. Minor deviations can inevitably accumulate into significant errors during the registration of long-sequence image stacks, ultimately leading to short structure truncations and sequence drift (Xin et al. 2023). Figure 6 illustrates the average registration accuracy across different slice intervals as the number of slices increase. The results demonstrate that our method consistently achieves higher accuracy compared to other methods across all intervals, highlighting its robustness against error accumulation. Notably, even in datasets with a thousand slices, our method does not exhibit significant accuracy decline, demonstrating its robustness to error accumulation and its

Figure 6: Scatter plot showing the registration accuracy across different slice intervals.

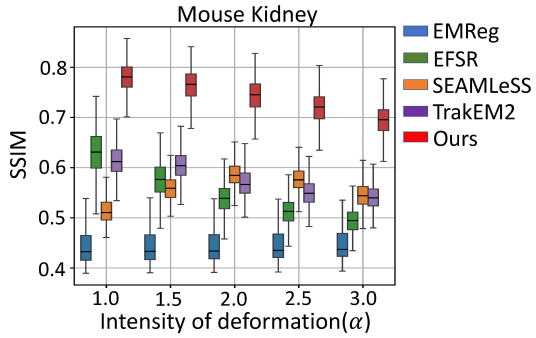

Figure 7: Box plot of the registration accuracy for various methods under different deformation levels.

potential to scale to large datasets.

Robustness to complex deformations. To assess the robustness of our method under different levels of deformation, we applied varying degrees of deformation to mouse kidney dataset by fixing $\sigma = 0.08$ and varying α from 1.0 to 3.0. For each α, we registered the datasets and computed the average registration accuracy. Figure 7 presents box plots of the average registration accuracy for various methods across datasets with different deformation levels. Our method achieves the highest accuracy across all levels of elastic deformation, consistently maintaining stable performance without noticeable drops. This demonstrates its strong robustness against complex nonlinear distortions, ensuring reliable and accurate results even under challenging conditions.

Ablation Studies

The effectiveness of Gaussian filter strategy. To verify the effectiveness of our core algorithm: the Gaussian filter strategy, we show the registration accuracy with and without the Gaussian filter strategy in Table 3. Meanwhile, in order to demonstrate the universal adaptability of our algorithm, we replaced the optical flow estimation network with other models. We chose the LKUnet (Jia et al. 2022) as the registration module in the Gaussian filter strategy and trained it according to the training strategy of our network. As shown in Table 3, when the Gaussian filter strategy was not applied,

	Mouse Kidney	Mouse Skin	Mouse Liver
w/o Gauss filter	0.63	0.67	0.66
w/ Gauss filter	0.78	0.82	0.80
LKUnet	0.43	0.47	0.41
LKUnet+Gauss filter	0.75	0.80	0.77

Table 3: Ablation study on the Gaussian filter strategy.

$r \setminus T$	1	2	3
1	0.83	0.82	0.83
2	0.72	0.65	0.92
3	0.58	0.55	0.58

Table 4: Ablation study on the radius of receptive field (r) and the number of iterations (T)

the network merely served as a simple registration network, and the registration accuracy decreased significantly. However, after introducing the Gaussian filter strategy, the registration accuracy was greatly improved. This fully demonstrates the effectiveness of the core algorithm as well as its universality in being compatible with advanced networks.

Hyperparameters in Gaussian filter. In order to select the optimal hyperparameters for the Gaussian filter strategy, we conducted further ablation experiments on the receptive field radius r and the number of iterations T. As can be seen from Table 4, changes in the number of iterations T did not cause significant changes in accuracy, but as the receptive field radius r increased, the registration accuracy dropped sharply. This may be because a larger receptive field increases the complexity of integrating the deformation field, leading to the inability of the Gaussian kernel parameters to cope. It is worth noting that in order to handle large-scale datasets, we avoid complex parameter designs and instead adopt a simple Gaussian filtering model. Therefore, future work may explore learnable adaptive filtering methods to integrate a wider range of slice information more efficiently. Further analysis is provided in the appendix.

Conclusion

We provide a new perspective on the 3D registration problem in series section electron microscopy and propose a Gaussian filtering-based 3D registration method. We examine the differences between natural deformation and nonlinear distortion and have successfully decoupled these two in the frequency domain. Our method integrates deformation field information within the receptive field by combining Gaussian filtering and optical flow networks, achieving state-of-the-art performance. We conducted experiments on simulated and real datasets of three cell types, validating our method's effectiveness in reducing error accumulation and managing complex elastic deformations, highlighting its potential for large-scale applications.

Acknowledgments

This research was supported by the National Key Research and Development Program of China (2021YFF0704300), the Natural Science Foundation of Ningxia Provinve (2024AAC03247), and the Shanghai Municipal Science and Technology Major Project (No.2018SHZDZX01), Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (LCNBI) and ZJLab.

References

Arganda-Carreras, I.; Sorzano, C. O.; Marabini, R.; Carazo, J. M.; Ortiz-de Solorzano, C.; and Kybic, J. 2006. Consistent and elastic registration of histological sections using vector-spline regularization. In *Computer Vision Approaches to Medical Image Analysis: Second International ECCV Workshop, CVAMIA 2006 Graz, Austria, May 12, 2006 Revised Papers 2*, 85–95. Springer.

Baena, V.; and Terasaki, M. 2019. Three-dimensional organization of transzonal projections and other cytoplasmic extensions in the mouse ovarian follicle. *Scientific reports*, 9(1): 1262.

Balakrishnan, G.; Zhao, A.; Sabuncu, M. R.; Guttag, J.; and Dalca, A. V. 2018. An unsupervised learning model for deformable medical image registration. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 9252–9260.

Cardona, A.; Saalfeld, S.; Schindelin, J.; Arganda-Carreras, I.; Preibisch, S.; Longair, M.; Tomancak, P.; Hartenstein, V.; and Douglas, R. J. 2012. TrakEM2 software for neural circuit reconstruction. *PloS one*, 7(6): e38011.

Chen, J.; Wang, L.; Li, X.; and Fang, Y. 2019. Arbicon-net: Arbitrary continuous geometric transformation networks for image registration. *Advances in neural information processing systems*, 32.

Dalca, A.; Rakic, M.; Guttag, J.; and Sabuncu, M. 2019. Learning conditional deformable templates with convolutional networks. *Advances in neural information processing systems*, 32.

Dalca, A. V.; Balakrishnan, G.; Guttag, J.; and Sabuncu, M. R. 2018. Unsupervised learning for fast probabilistic diffeomorphic registration. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I*, 729–738. Springer.

de Senneville, B. D.; Khoubai, F. Z.; Bevilacqua, M.; Labedade, A.; Flosseau, K.; Chardot, C.; Branchereau, S.; Ripoche, J.; Cairo, S.; Gontier, E.; et al. 2021. Deciphering tumour tissue organization by 3D electron microscopy and machine learning. *Communications biology*, 4(1): 1390.

Gaffling, S.; Daum, V.; Steidl, S.; Maier, A.; Köstler, H.; and Hornegger, J. 2014. A Gauss-Seidel iteration scheme for reference-free 3-D histological image reconstruction. *IEEE transactions on medical imaging*, 34(2): 514–530.

Haskins, G.; Kruger, U.; and Yan, P. 2020. Deep learning in medical image registration: a survey. *Machine Vision and Applications*, 31(1): 8.

Hering, A.; Hansen, L.; Mok, T. C.; Chung, A. C.; Siebert, H.; Häger, S.; Lange, A.; Kuckertz, S.; Heldmann, S.; Shao, W.; et al. 2022. 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.

Jia, X.; Bartlett, J.; Zhang, T.; Lu, W.; Qiu, Z.; and Duan, J. 2022. U-net vs transformer: Is u-net outdated in medical image registration? In *International Workshop on Machine Learning in Medical Imaging*, 151–160. Springer.

Johnson, B. E.; Creason, A. L.; Stommel, J. M.; Keck, J. M.; Parmar, S.; Betts, C. B.; Blucher, A.; Boniface, C.; Bucher, E.; Burlingame, E.; et al. 2022. An omic and multidimensional spatial atlas from serial biopsies of an evolving metastatic breast cancer. *Cell Reports Medicine*, 3(2).

Jud, C.; Möri, N.; Bitterli, B.; and Cattin, P. C. 2016. Bilateral regularization in reproducing kernel hilbert spaces for discontinuity preserving image registration. In *Machine Learning in Medical Imaging: 7th International Workshop, MLMI 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 17, 2016, Proceedings 7*, 10–17. Springer.

Kievits, A. J.; Lane, R.; Carroll, E. C.; and Hoogenboom, J. P. 2022. How innovations in methodology offer new prospects for volume electron microscopy. *Journal of Microscopy*, 287(3): 114–137.

Krebs, J.; Delingette, H.; Mailhé, B.; Ayache, N.; and Mansi, T. 2019. Learning a probabilistic model for diffeomorphic registration. *IEEE transactions on medical imaging*, 38(9): 2165–2176.

Kremer, J. R.; Mastronarde, D. N.; and McIntosh, J. R. 1996. Computer visualization of three-dimensional image data using IMOD. *Journal of structural biology*, 116(1): 71–76.

Lee, W.-C. A.; Bonin, V.; Reed, M.; Graham, B. J.; Hood, G.; Glattfelder, K.; and Reid, R. C. 2016. Anatomy and function of an excitatory network in the visual cortex. *Nature*, 532(7599): 370–374.

Liu, X.; Zhang, Y.; Zhou, S.; Xiong, Z.; and Sun, X. 2023. Electron Microscopy Image Registration Using Correlation Volume. In *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, 1–5. IEEE.

Morgan, J. L.; and Lichtman, J. W. 2020. An individual interneuron participates in many kinds of inhibition and innervates much of the mouse visual thalamus. *Neuron*, 106(3): 468–481.

Müller, A.; Schmidt, D.; Xu, C. S.; Pang, S.; D'Costa, J. V.; Kretschmar, S.; Münster, C.; Kurth, T.; Jug, F.; Weigert, M.; et al. 2021. 3D FIB-SEM reconstruction of microtubule–organelle interaction in whole primary mouse β cells. *Journal of Cell Biology*, 220(2).

Paszke, A.; Gross, S.; Massa, F.; Lerer, A.; Bradbury, J.; Chanan, G.; Killeen, T.; Lin, Z.; Gimelshein, N.; Antiga, L.; et al. 2019. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32.

Pielawski, N.; Wetzer, E.; Öfverstedt, J.; Lu, J.; Wählby, C.; Lindblad, J.; and Sladoje, N. 2020. CoMIR: Contrastive

multimodal image representation for registration. *Advances in neural information processing systems*, 33: 18433–18444.

Popovych, S.; Macrina, T.; Kemnitz, N.; Castro, M.; Nehoran, B.; Jia, Z.; Bae, J. A.; Mitchell, E.; Mu, S.; Trautman, E. T.; et al. 2024. Petascale pipeline for precise alignment of images from serial section electron microscopy. *Nature Communications*, 15(1): 289.

Roberts, R. A.; and Mullis, C. T. 1987. *Digital signal processing*. Addison-Wesley Longman Publishing Co., Inc.

Ryan, K.; Lu, Z.; and Meinertzhagen, I. A. 2016. The CNS connectome of a tadpole larva of Ciona intestinalis (L.) highlights sidedness in the brain of a chordate sibling. *Elife*, 5: e16962.

Saalfeld, S.; Fetter, R.; Cardona, A.; and Tomancak, P. 2012. Elastic volume reconstruction from series of ultra-thin microscopy sections. *Nature methods*, 9(7): 717–720.

Sandkühler, R.; Andermatt, S.; Bauman, G.; Nyilas, S.; Jud, C.; and Cattin, P. C. 2019. Recurrent registration neural networks for deformable image registration. *Advances in Neural Information Processing Systems*, 32.

Shen, Z.; Vialard, F.-X.; and Niethammer, M. 2019. Region-specific diffeomorphic metric mapping. *Advances in Neural Information Processing Systems*, 32.

Shoemark, A.; Boon, M.; Brochhausen, C.; Bukowy-Bieryllo, Z.; De Santi, M. M.; Goggin, P.; Griffin, P.; Hegele, R. G.; Hirst, R. A.; Leigh, M. W.; et al. 2020. International consensus guideline for reporting transmission electron microscopy results in the diagnosis of primary ciliary dyskinesia (BEAT PCD TEM Criteria). *European Respiratory Journal*, 55(4).

Terasaki, M.; Shemesh, T.; Kasthuri, N.; Klemm, R. W.; Schalek, R.; Hayworth, K. J.; Hand, A. R.; Yankova, M.; Huber, G.; Lichtman, J. W.; et al. 2013. Stacked endoplasmic reticulum sheets are connected by helicoidal membrane motifs. *Cell*, 154(2): 285–296.

Thevenaz, P.; Ruttimann, U. E.; and Unser, M. 1998. A pyramid approach to subpixel registration based on intensity. *IEEE transactions on image processing*, 7(1): 27–41.

Vishnevskiy, V.; Gass, T.; Szekely, G.; Tanner, C.; and Goksel, O. 2016. Isotropic total variation regularization of displacements in parametric image registration. *IEEE transactions on medical imaging*, 36(2): 385–395.

Weidner, N.; Cote, R. J.; Suster, S.; and Weiss, L. M. 2009. *Modern surgical pathology e-book*. Elsevier Health Sciences.

Wilson, A.; and Babadi, M. 2023. SynapseCLR: Uncovering features of synapses in primary visual cortex through contrastive representation learning. *Patterns*, 4(4).

Wilson, C. E.; Lasher, R. S.; Yang, R.; Dzowo, Y.; Kinnamon, J. C.; and Finger, T. E. 2022. Taste bud connectome: implications for taste information processing. *Journal of Neuroscience*, 42(5): 804–816.

Xin, T.; Lv, Y.; Chen, H.; Li, L.; Shen, L.; Shan, G.; Chen, X.; and Han, H. 2023. A novel registration method for long-serial section images of EM with a serial split technique based on unsupervised optical flow network. *Bioinformatics*, 39(8): btad436.

Xu, C. S.; Pang, S.; Shtengel, G.; Müller, A.; Ritter, A. T.; Hoffman, H. K.; Takemura, S.-y.; Lu, Z.; Pasolli, H. A.; Iyer, N.; et al. 2021. An open-access volume electron microscopy atlas of whole cells and tissues. *Nature*, 599(7883): 147–151.

Yin, W.; Brittain, D.; Borseth, J.; Scott, M. E.; Williams, D.; Perkins, J.; Own, C. S.; Murfitt, M.; Torres, R. M.; Kapner, D.; et al. 2020. A petascale automated imaging pipeline for mapping neuronal circuits with high-throughput transmission electron microscopy. *Nature communications*, 11(1): 4949.

Yoo, I.; Hildebrand, D. G.; Tobin, W. F.; Lee, W.-C. A.; and Jeong, W.-K. 2017. ssEMnet: Serial-section electron microscopy image registration using a spatial transformer network with learned features. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings 3*, 249–257. Springer.

Zhao, S.; Dong, Y.; Chang, E. I.; Xu, Y.; et al. 2019. Recursive cascaded networks for unsupervised medical image registration. In *Proceedings of the IEEE/CVF international conference on computer vision*, 10600–10610.

Zheng, Z.; Lauritzen, J. S.; Perlman, E.; Robinson, C. G.; Nichols, M.; Milkie, D.; Torrens, O.; Price, J.; Fisher, C. B.; Sharifi, N.; et al. 2018. A complete electron microscopy volume of the brain of adult Drosophila melanogaster. *Cell*, 174(3): 730–743.

Zhou, S.; Hu, B.; Xiong, Z.; and Wu, F. 2023. Self-Distilled Hierarchical Network for Unsupervised Deformable Image Registration. *IEEE Transactions on Medical Imaging*.

Zhou, S.; Xiong, Z.; Chen, C.; Chen, X.; Liu, D.; Zhang, Y.; Zha, Z.-J.; and Wu, F. 2019. Fast and accurate electron microscopy image registration with 3D convolution. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22*, 478–486. Springer.