Multimodal registration of FISH and nanoSIMS images using convolutional neural networks

Xiaojia He[1,2,*], Suchendra M. Bhandarkar[3] and Christof Meile[1]

[1] Department of Marine Sciences, University of Georgia, Athens, GA 30602, USA

[2] Chemical Insights Research Institute, UL Research Institutes, Marietta, GA 30067, USA

[3] School of Computing, University of Georgia, Athens, GA 30602, USA

Abstract

Nanoscale secondary ion mass spectrometry (nanoSIMS) and fluorescence *in situ* hybridization (FISH) microscopy provide high-resolution, multimodal image representations of cell identity and cell activity, respectively, for studies of targeted microbial communities in microbiological research. Despite its importance to microbiologists, multimodal registration of FISH and nanoSIMS images is challenging given the morphological distortion and background noise in both image modalities. In this paper we propose a scheme for multimodal registration of FISH and nanoSIMS images that employs convolutional neural networks (CNNs) for multiscale feature extraction, shape context for computation of the minimum transformation cost feature matching and the thin-plate spline (TPS) model for the registration of the two image modalities. Registration accuracy is quantitatively assessed against manually registered images, at both the pixel and structural levels, using standard metrics. Experimental results show that among the six CNN models that were tested, *ResNet18* outperforms *VGG16, VGG19, GoogLeNet, ShuffleNet* and *ResNet101* based on most evaluation metrics. This study demonstrates the utility of CNNs in the registration of multimodal images with significant background noise and morphology distortion. We also show that the shape of microbial aggregates, preserved by binarization, to be a robust feature for registering multimodal microbiology-related images. The proposed multimodal image registration scheme can serve as a powerful tool in microbiological research.

1. Introduction

Nanoscale secondary ion mass spectrometry (nanoSIMS) is a powerful tool to quantify elemental distribution at nanometer-scale resolution [1]. Combining nanoSIMS imaging with fluorescence in situ hybridization (FISH) microscopy allows one to study microbial activity and correlate it with the identity of cells [2]. However, the nanoSIMS and FISH images display unequal magnification and distortion. Several image registration algorithms exploit geometrical information to align the input images [3]. Notably, feature-based registration methods rely on point- or shape-based correspondences between two images where the features, such as corners or contours of structures, are either derived automatically from the underlying image or from markers with known positions. Once the corresponding points are selected, their locations in the two images are used to reconstruct a spatial transformation [4, 5]. In contrast, in intensity-based methods, only pixel intensity values, instead of specific features, are considered to determine the spatial transformation.

Deep learning has been increasingly recognized as a powerful toolbox for multimodal image registration, especially in medical imaging [6, 7] and remote sensing [8, 9]. The convolutional neural network (CNN) is a widely used deep neural network (DNN) architecture comprising of convolutional layers, max-pooling layers and a softmax layer, in addition to problem-specific layers. CNNs have been used extensively for feature extraction in image classification [10, 11], image segmentation [12, 13] and image registration [14, 15], and several variants of the CNN architecture have been proposed for multimodal image registration [6, 16, 17], and have been shown to be successful in solving biomedical image registration problems [18-21].

In this paper, we present an automated scheme to register FISH and nanoSIMS images using multiple CNN models. Although images of neither microorganisms nor microbial aggregates are in the ImageNet database, deep CNN architectures that are pre-trained on ImageNet have been shown to be very effective at general image feature extraction. The convolutional feature map is extracted at multiple image resolutions and used for feature point selection. The shape context descriptor is used to identify matched features and the thin-plate spline (TPS) model is employed to register the FISH and nanoSIMS images by computing a transformation matrix [22]. The results obtained using the different CNNs, feature matching approaches and transformation computation and registration methods are compared and discussed. To the best of our knowledge, this is the first documented application of deep CNN models to extract features from multimodal microbial images and subsequently register them.

2. Materials and Methods

The FISH and nanoSIMS images were acquired using the protocol proposed by McGlynn et al. [23] and a detailed description of sample collection and preparation, measurement methodology and data analysis is given in [23]. In brief, anaerobic methane-oxidizing consortia were obtained from ocean sediment samples collected at Hydrate Ridge North (station HR-7) during the AT 18-10 Hydrate Ridge August/September 2011 expedition. Push core sediment samples were processed on ship and kept under an N_2 atmosphere at $4°C$. Slurry incubations were carried out with anoxic filtered seawater at elevated pressure. FISH and nanoSIMS images were then collected and manually aligned using the Matlab program Look@nanoSIMS as described in [24]. These manually aligned images were used as ground truth in this study.

In our workflow, depicted in Figure 1, 41 raw RGB images and their binarized versions were used as input. In brief, the input images were preprocessed to remove background noise and then fed to the chosen CNN models with pretrained weights. Features were then extracted at desired predetermined layer depths (scales) using the CNN architectures *ShuffleNet* [25], *GoogLeNet* [26], *ResNet-18* and *ResNet-101* [27], *VGG16* and *VGG19* [28], with pretrained weights derived from the several million training images in the ImageNet database (http://www.image-net.org). A subset of the extracted features was selected and further constrained to generate a 2-D array of matched feature points using shape context and bipartite graph matching algorithms [22]. Finally, the matched feature points were used for image transformation computation and image registration using the thin-plate spline (TPS) model. Quantitative registration accuracy metrics such as the *root mean squared error* (RMSE), *structural similarity index* (SSIM), and *average absolute intensity difference* (AAID) were computed at both the pixel and structural levels. Additional details on the above- mentioned of methods are available at https://doi.org/10.6084/m9.figshare.26321587.v3.
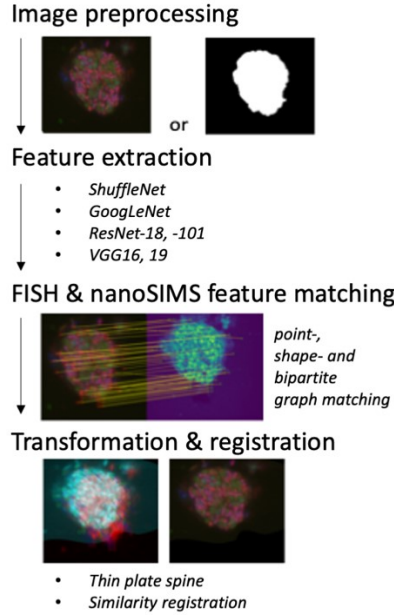
Figure. 1: Workflow for multimodal registration of FISH and nanoSIMS images.

2.1. Image preprocessing

FISH images are intensity measurements represented in their respective coordinate systems in the individual RGB channels, whereas nanoSIMS images represent ion counts at each pixel location. A global threshold was first generated using Otsu's method [29] to minimize the intra-class variance (i.e., weighted sum of variances of black and white pixels in a binary image) and was modified manually based on trial and error to preserve aggregate morphology. Aggregate(s) from the FISH image were then chosen and cropped to best match the nanoSIMS image. The resulting input images to the CNN were either raw RGB or preprocessed binary FISH and nanoSIMS images. All input images were rescaled to a size of 224×224 pixels and fed through the convolutional layers in the CNN.

2.2. Feature extraction and matching

For the FISH and nanoSIMS images, features were extracted from the final layer of each individual module in the CNN architecture starting with a layer size of 28×28 and proceeding to layer sizes of 14×14 and 7×7. The selection of convolutional layers was heuristic and aimed to include both high- and low-level features. The feature maps obtained from each layer was normalized by applying the transformation $z = (x\text{-}\mu)/\sigma$, where the feature x in each feature map is assumed to be normally distributed with mean μ and standard deviation σ. Next, we generated the feature distance map by computing the symmetric matrix of pairwise feature distance values. We concatenated the feature distance maps from each layer to yield a single feature distance map for each FISH and nanoSIMS image pair and processed the concatenated feature distance map by selecting the smallest value from each row and using the match threshold to select the top 20% matched features.

2.3. Shape context descriptor

After selecting the preliminary matching features, we used the shape context descriptor to determine the feature correspondence that minimizes a transformation cost function. The transformation cost function quantifies the shape similarity based on the neighborhood structure of a feature point on a shape contour. The shape context descriptor at feature point p_i is defined as a histogram h_i of the relative coordinates q of the remaining n-1 feature points [22]:

$$h_i(k) = \#\{q \neq p_i : |(q - p_i)| \in bin(k)\} \tag{1}$$

where the bins are designed to uniformly partition the log-polar ($\log\rho$, θ) space (ρ is the radial distance and θ is the polar angle). To generate a shape context descriptor, we first computed the Euclidian distance values between points in the matched feature map and normalized them by the mean. Next, we computed the shape context descriptor by directly counting the points within each radial and angular region (bin) as described above.

2.4. Bipartite graph matching

We consider minimizing the total cost of matching given by
$$H(\pi) = \sum_i C\left(p_i, q_{\pi(i)}\right) \tag{2}$$
where π denotes a permutation, and C is the cost function defined as $C_{i,j} = \frac{1}{2}\sum_{k=1}^{K}\frac{(h_i(k)-h_j(k))^2}{h_i(k)+h_j(k)}$, where h_i and h_j are the obtained shape context descriptors (normalized k-bin histograms) for the matched feature points p_i and q_j on the FISH and nanoSIMS images, respectively. The resulting weighted bipartite graph matching problem based on $H(\pi)$ was solved using the efficient Jonker-Volgenant algorithm [30]. Finally, we computed the Euclidian distance between each matched feature pair and only retained the matches that fall between the 25% and 75% quantiles as inliers. The values of the matching threshold were chosen based on trial and error.

2.5. Transformation and Registration

Given a finite set of point correspondences between two shapes, the image transformation and registration function $T: \mathbb{R}^2 \to \mathbb{R}^2$ can be realized using the TPS model [31] which performs non-rigid registration or alignment of deformed images. The underlying transformation was modeled as a radial-basis function where the foreground pixels of the moving image deform under the influence of the control points p_i, $i = 1, \ldots n$.

2.6. Similarity registration

Similarity registration was used as a comparison to our proposed non-rigid, TPS-based registration scheme. Using the features extracted from CNN models, similarity registration allows for alignment of images via a combination of globally applied rigid-body translation, rotation, and scaling operations [32, 33].

2.7. Comparison to a state-of-the-art registration method and non-CNN feature extraction-based registration methods

The Contrastive Multimodal Image Representations (*CoMIR*) scheme [34; https://github.com/MIDA-group/CoMIR], which is shown to outperform several other state-of-the-art image registration methods in biomedical and remote sensing applications, was used in this study for the purpose of comparison. To further evaluate the performance of our CNN-based feature extraction schemes, we also implemented and evaluated a variety of traditional non-CNN-based feature extraction methods such as the similarity-invariant, fast and robust algorithm for local feature extraction SURF [35], scale- and rotation-invariant, fast multiscale feature detection and description approach for nonlinear scale spaces KAZE [36], scale- and rotation-invariant, fast feature point extraction algorithm BRISK [37], and Harris corner feature detector [38] and features from accelerated segment test (FAST) [39].

2.8. Quantitative image registration assessment

The results of automated registration were compared to manually registered images (i.e., the ground truth). Three different error metrics were employed to assess registration accuracy at

the pixel and structural levels: root mean squared error (RMSE), structural similarity index (SSIM), and average absolute intensity difference (AAID). RMSE quantifies the difference between registered images $(\hat{y}, y)$ by computing the square root of the mean square error of pixel values over the RGB channels between the two images (40):

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N}} \tag{3}$$

The SSIM metric measures the perceived similarity in structural information between two images and entails computing a weighted combination of *the luminance index l*, the *contrast index c* and the *structural index s* (41):

$$SSIM = [l(\hat{y}, y)]^\alpha [c(\hat{y}, y)]^\beta [s(\hat{y}, y)]^\gamma \tag{4}$$

Here $l(\hat{y}, y) = \frac{2\mu_{\hat{y}}\mu_y + c_1}{\mu_{\hat{y}}^2 + \mu_y^2 + c_1}$, $c(\hat{y}, y) = \frac{2\sigma_{\hat{y}}\sigma_y + c_2}{\sigma_{\hat{y}}^2 + \sigma_y^2 + c_2}$, and $s(\hat{y}, y) = \frac{\sigma_{\hat{y}y} + c_3}{\sigma_{\hat{y}}\sigma_y + c_3}$ where $\mu_{\hat{y}}$, μ_y are the local means; $\sigma_{\hat{y}}$, σ_y the standard deviations; and $\sigma_{\hat{y}y}$ the cross-covariance for images $\hat{y}$ and y, respectively. The weights α, β and γ were set to 1.

The AAID metric is based on the absolute intensity difference between the two images $(\hat{y}, y)$ [42]:

$$AAID = \frac{1}{MNQ}\sum_{i=1}^{M}\sum_{j=1}^{N}\sum_{k=1}^{Q}|\hat{y}_{i,j,k} - y_{i,j,k}| \tag{5}$$

where *M*, *N*, and *Q* represent the dimension of images. Smaller RMSE and AAID values represent a better registration result, whereas the SSIM value is larger for better aligned images.

3. Results

Visual (Figures 2 and 3) and quantitative (Table 1) comparison of our results with manually registered images shows good agreement, signifying the advantages of the automated registration. Image preprocessing with binary thresholding significantly improved the accuracy of image registration compared to the raw input RGB images (the left panels in Figures 2 and 3). It also yielded substantially better quantitative results compared to than when analyzing the raw RGB images, as reflected in the smaller the small pixel differences in RMSE and AAID values (TPS registration, Table 1), and the larger SSIM indices, which exceeded 0.8 for a significantly deformed FISH image and 0.78 for a deformed FISH image with multiple connected components, respectively (TPS registration). Additional details on the aforementioned results are available at https://doi.org/10.6084/m9.figshare.26321587.v3.

The additional intra-aggregate features during RGB image registration, which may differ between the FISH and nanoSIMS images, results in deterioration of the registration results (Figures 2 and 3). It is also noted that residual component(s) in the FISH and nanoSIMS images outside the region of interest (ROI) did not match well even after several exhaustive trial and error iterations (see binary images in Figures 2 and 3). However, mismatches between the small connected components due to the binarization preprocessing did not impact the registration of our microbial aggregate images and hence there is no need to first remove the small connected components in the two images before proceeding to align them.

Our results show that TPS-based registration outperforms registration based on similarity metrics (Table 1). With radial basis functions, TPS-based registration is capable of locally transforming and warping the target FISH image onto the nanoSIMS image. In contrast, similarity-based registration involving only global linear rigid-body transformations, i.e., rotation, scaling, and translation [43], leads to significant disparity in registration results between TPS-based and similarity-based registration (Figure 2).

Table 1. Image registration accuracy for a significantly deformed FISH image and a deformed FISH image with multiple connected components. RMSE and AAID values are smaller when the registration result is better, whereas the SSIM value is larger for a better aligned image. The best performance measure compared across CNNs is highlighted in gray shade. The best performance measure for a registration scheme and CNN architecture for a given image type is indicated in bold font. The best performance measure for each aggregate is indicated with an asterisk (*).

	Significantly deformed										
	Binary						RGB				
	Similarity			TPS			Similarity			TPS		
	RMSE	AAID	SSIM	RMSE	AAID	SSIM	RMSE	AAID	SSIM	RMSE	AAID	SSIM
GoogLeNet	19.26	6.41	0.83	19.48	2.68	0.82	75.05	60.91	0.2	61.03	11.4	0.48
ResNet101	19.66	4.93	0.814	13.97	**1.66***	0.88	75.07	63.37	0.21	49.53	19.77	0.53
ResNet18	25.01	8.66	0.75	**12.44***	1.67	**0.89***	54.80	32.06	0.51	**32.45**	**6.65**	**0.7**
ShuffleNet	15.52	4.68	0.85	13.23	1.83	0.88	54.80	33.93	0.48	44.7	9.36	0.62
VGG16	39.04	17.75	0.67	14.38	2.25	0.87	66.49	51.63	0.24	45.74	10.59	0.6
VGG19	15.68	4.4	0.86	19.84	2.1	0.81	67.67	51.45	0.23	54.67	23.39	0.53

|	Deformed with multiple components										
	Binary						RGB				
	Similarity			TPS			Similarity			TPS		
	RMSE	AAID	SSIM	RMSE	AAID	SSIM	RMSE	AAID	SSIM	RMSE	AAID	SSIM
GoogLeNet	38.69	**2.5**	0.75	27.18	4.02	0.8	39.96	3.06	0.74	52.73	7.53	0.62
ResNet101	38.78	2.56	0.74	31.15	4.56	0.78	38.81	4.28	0.74	45.07	6.57	0.71
ResNet18	43.34	3.09	0.71	29.92	3.56	0.79	40.98	4.8	0.719	52.69	7.51	0.63
ShuffleNet	40.97	3.2	0.72	29.50	5.43	0.78	39.12	**2.49***	0.75	34.48	2.8	**0.76**
VGG16	41.47	3.41	0.72	29.29	5.27	0.78	39	3.74	0.75	44.37	4.19	0.69
VGG19	41.08	3.17	0.73	**25.62***	5.22	**0.81***	39.18	3.1	0.75	55.21	6.07	0.63
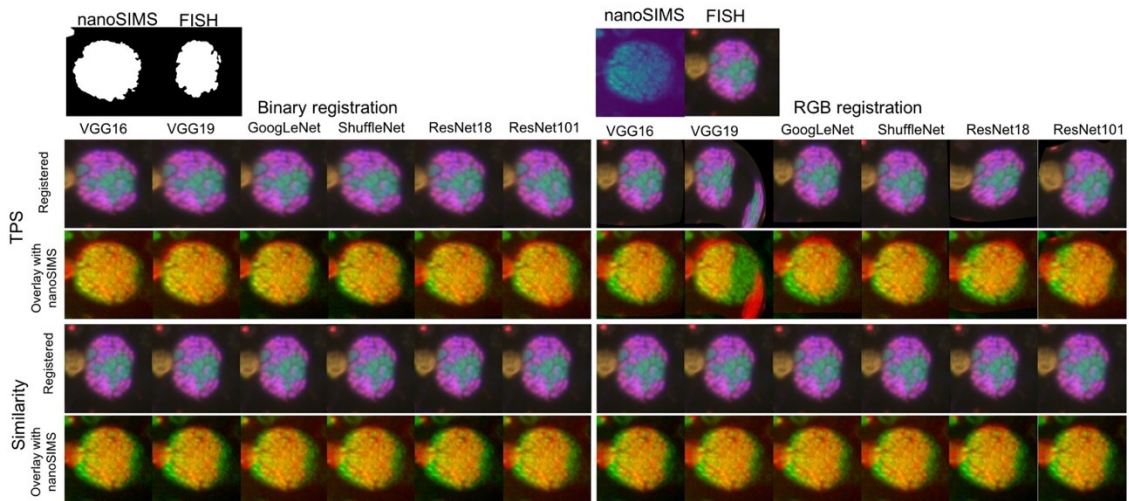
Figure 2: Registration of a significantly deformed FISH image and nanoSIMS image using binary (left) and RGB (right) images as input with TPS-based (upper two panels) and similarity-based (lower two panels) schemes.
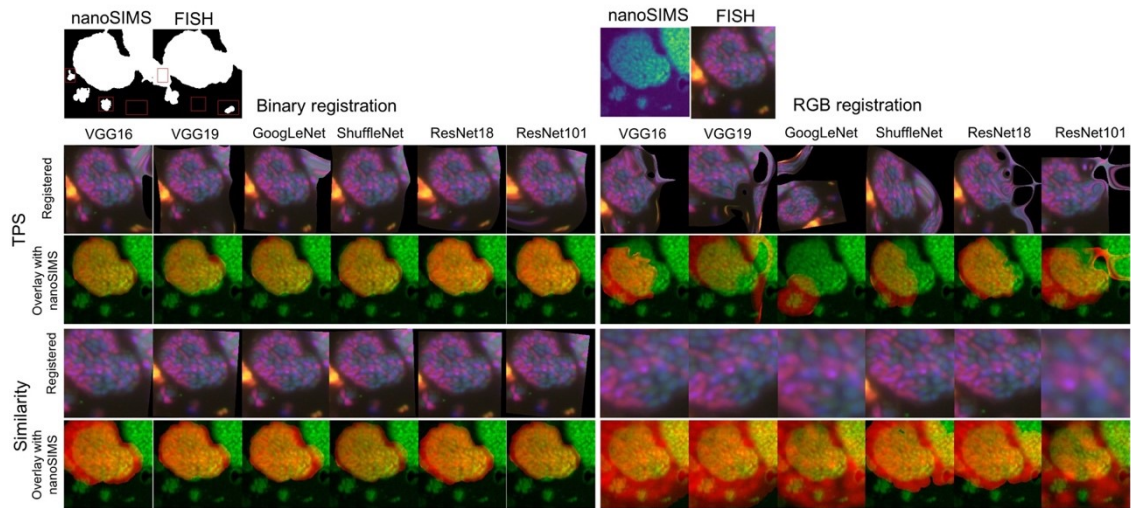
Figure 3: Same as Figure 2, but for a deformed FISH image with multiple components. Orange rectangles indicate missing components in either nanoSIMS or FISH images after binarization.

The CNN models also performed well with deformed FISH images containing multiple connected components. The analysis of a significantly deformed image (Figure 4), and a deformed image with multiple connected components (Figure 5) revealed that *ResNet* and *ShuffleNet* often outperformed *VGG* and *GoogLeNet* implementations. Registration using a fine-tuned CNN in which the weights of a pre-trained CNN are refined by training with new data [44] produced almost the same registration results using binary images as input (not shown). Using raw (RGB) images as input improved the registration slightly, but the differences in registration performance were minimal.
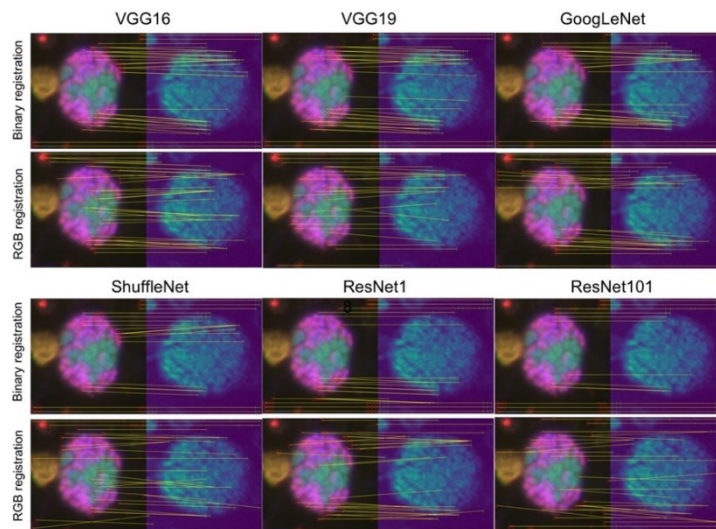

Figure 4: Feature correspondences after final thresholding during registration of a significantly deformed FISH image and nanoSIMS image using binary (upper row for each method) and RGB inputs (lower row for each method).

For validation purposes, we first compared our CNN-based methods to other well-recognized traditional feature extraction-based methods that employ SURF, KAZE, BRISK, Harris corner detector and FAST features (see Figures S15-S16 in the supplemental materials available at https://doi.org/10.6084/m9.figshare.26321587.v3). None of the aforementioned traditional feature extraction-based methods produced satisfactory registration results in our tests for a modestly deformed, a significantly deformed, and a multiple-component deformed FISH image

with a nanoSIMS image. With RGB images as input, all the aforementioned traditional feature extraction-based methods failed completely to register the FISH images with the nanoSIMS image due to the inherent shortcomings of the extracted and matched features.
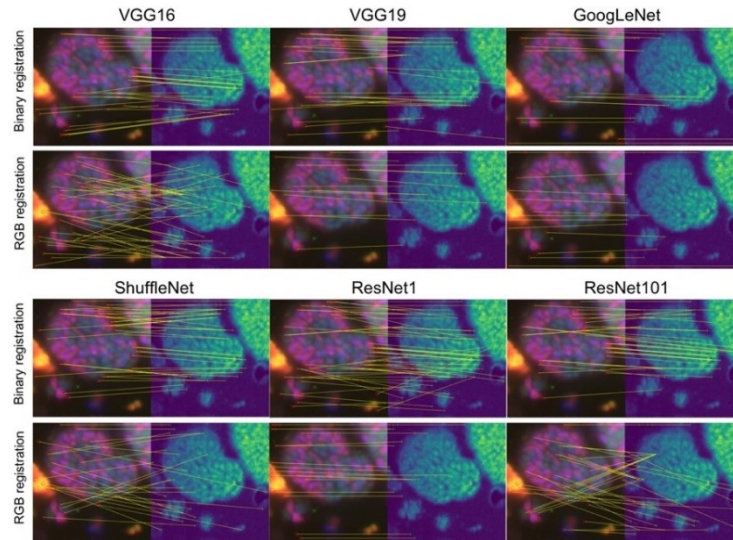

Figure 5: Same as Figure 4, but for a deformed FISH image with multiple components.

To assess the quality of our CNN-based implementations, we compared them with the results of the state-of-the-art, pretrained *CoMIR* method [34] that is based on contrastive learning. Here we considered three distinct types of deformed FISH images: a moderately deformed FISH image (Figure 6A), significantly deformed FISH image (Figure 6B), and multiple-component deformed FISH image (Figure 6C). *CoMIR* registered these FISH images with the corresponding nanoSIMS images with high accuracy. Our proposed CNN-based methods performed comparably to the state-of-the-art *CoMIR* method, while significantly outperforming the rigid-body registration methods.
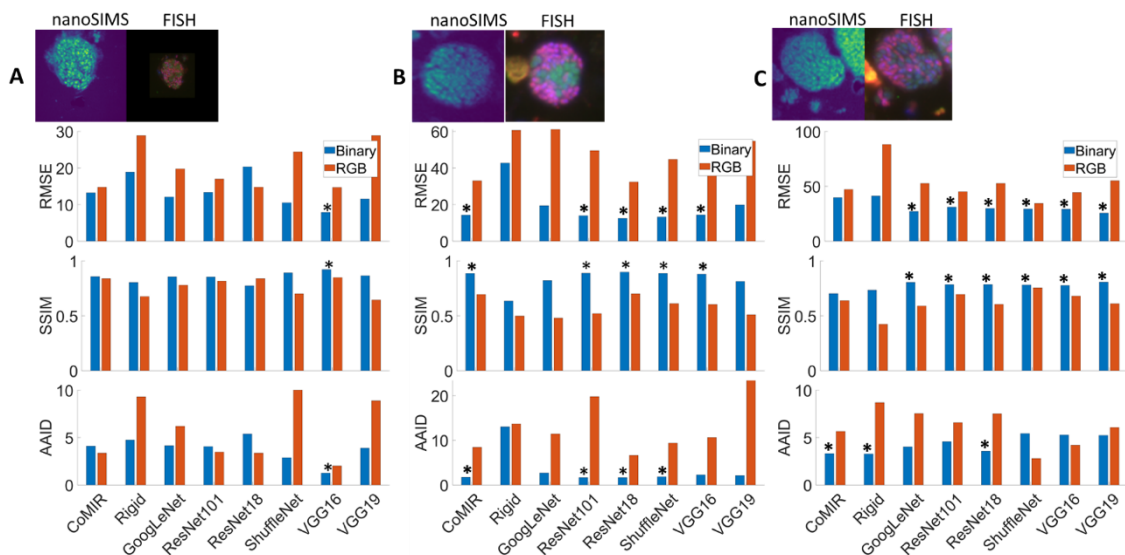

Figure 6: Comparison of state-of-art method to our proposed CNN models. Image registration accuracy for a moderately deformed FISH image (A), a significantly deformed FISH image (B), and a deformed FISH image with multiple components (C), quantified by the RMSE (upper panel), AAID (middle panel), and SSIM (lower panel). RMSE and AAID values are smaller when the registration result is better, whereas higher SSIM values denote a better aligned image. The asterisks indicate better performance with either smaller RMSE and AAID values, or higher SSIM values.

4. Discussion

The integration of multiple multimodal data streams is critical to gain new insights into the functioning of microbial communities. Here we present the results of a processing pipeline that merges spatially explicit data sets on the identity and activity of microorganisms in the form of images. The alignment of such multimodal images to resolving individual cells can be challenging due to image distortion.

We successfully used CNNs that are pretrained on the ImageNet database to replace tedious manual alignment and image registration. Our results indicate that all six CNN models yield high registration accuracy at both, the pixel and structural levels (Figs. 2 and 3, Table 1), even though the ImageNet database does not contain microbial imagery. Nevertheless, our pipeline produces results that compare favorably with manually registered images. This good agreement illustrates that automated registration is a valuable tool for microbial image analysis.

The finding that binary thresholding significantly improved image registration shows that aggregate shape is a useful characteristic or feature to employ and that alignment of (deformed) aggregate contours that are consistent between image modalities yields robust results (Figures 2 and 3; Table 1). Our analysis also shows that the selection of regions of interest is not extremely critical and that the results are not sensitive to small mismatches. This is largely due to the observation that features extracted using the CNNs were mostly found to be from the dominant object in the image (Figure 5). This facilitates the alignment of FISH and nanoSIMS images, with the former covering larger areas compared to the more detailed, high-resolution nanoSIMS observations. However, the registration performance is negatively impacted when objects in the image are fragmented resulting in the absence of a dominant object.

We further demonstrate that the use of more involved registration methods can improve the results substantially. While computationally more intensive, TPS-based registration introduces smooth, elastic deformations, producing a reasonably well-aligned image even for a significantly deformed FISH image (Figure 2). This finding is consistent with the reported high accuracy and robustness of TPS in data interpolation and image registration [45].

While all CNNs performed well – better than several standard extraction methods, and comparable to *CoMIR*, *there* are some differences between them. Notably, features extracted by *ResNet* and *ShuffleNet* were generally more complex than those extracted by their *VGG* and *GoogLeNet* counterparts; thus, potentially contributing to slightly better registration results for a significantly deformed FISH image (Figure 2 and Table 1), or a deformed FISH image with multiple components (Figure 3). Moreover, we found that fine-tuning did not improve registration significantly. As it consumes significantly more computing power and takes substantially longer to finish, we deemed that fine-tuning is not necessary for this type of image registration task. Lastly, graph theory-based [46] and phase-based [47] image registration techniques have also demonstrated promising registration accuracy for multimodal images. Future avenues of work will include the incorporation of these techniques. The code for our processing pipeline is publicly available on the Bitbucket repository at https://bitbucket.org//MeileLab/he_imageregistration/src/master.

5. Conclusions

Our workflow employed advanced CNN models to successfully extract shared feature points in FISH and nanoSIMS images for multimodal image registration. CNN-derived, feature-based non-rigid TPS registration methods significantly outperformed conventional similarity-based rigid-body registration methods and produced registration results that were very comparable to those of the state-of-art method *CoMIR* method that is based on contrastive learning. We tested six CNN models using TPS-based non-rigid registration for different FISH and nanoSIMS images. The differences between the registration results obtained from the different CNN models considered

in this study were minor. We demonstrated that image preprocessing with binarization is critical for final image registration and aggregate shape is a robust feature for microbiology-derived images such as FISH and nanoSIMS images. This may be largely owing to the significant differences in intra-aggregate patterns between the FISH and nanoSIMS images, leading to poor registration performance when using raw RGB images as input. It is also noted that images with significant background noise (non-ROI components) that cannot be easily removed via simple thresholding and binarization still pose a significant challenge. This highlights the importance of aggregate morphology and reducing background noise in images with multiple aggregates.

Acknowledgements

References

[1] S.G. Boxer, M.L. Kraft, and P.K. Weber, Advances in imaging secondary ion mass spectrometry for biological samples. Annual Review of Biophysics, 2009. 38: p. 53-74.

[2] A.E. Dekas, et al., Activity and interactions of methane seep microorganisms assessed by parallel transcription and FISH-NanoSIMS analyses. The ISME Journal, 2015. 10: p. 678–92.

[3] L.G. Brown, A survey of image registration techniques. ACM Computing Surveys, 1992. 24(4): p. 325-76.

[4] P.S. Heckbert, Fundamentals of texture mapping and image warping. MS Thesis in Electrical Engineering and Computer Science. 1989, University of California, Berkeley: Berkeley, CA. p. 88.

[5] N. Arad and D. Reisfeld, Image warping using few anchor points and radial functions, In the Proceedings of Computer Graphics Forum. 1995, Blackwell Science Ltd. p. 35-46.

[6] H. Hermessi, O. Mourali, and E. Zagrouba, Convolutional neural network-based multimodal image fusion via similarity learning in the shearlet domain. Neural Computing and Applications, 2018. 30(7): p. 2029-2045.

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

[8] A. Zampieri, et al., Multimodal image alignment through a multiscale chain of neural networks with application to remote sensing. In the Proceedings of the European Conference on Computer Vision (ECCV). 2018.

[9] H. Zhang, et al., Registration of multimodal remote sensing image based on deep fully convolutional neural network. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2019. 12(8): p. 3028-3042.

[10] M. Zhang, W. Li, and Q. Du, Diverse region-based CNN for hyperspectral image classification. IEEE Transactions on Image Processing, 2018. 27(6): p. 2623-2634.

[11] D. Han, Q. Liu, and W. Fan, A new image classification method using CNN transfer learning and web data augmentation. Expert Systems with Applications, 2018. 95: p. 43-56.

[12] B. Kayalibay, G. Jensen, and P. van der Smagt, CNN-based segmentation of medical imaging data. arXiv preprint arXiv:1701.03056, 2017.

[13] S. Bao and A.C. Chung, Multi-scale structured CNN with label consistency for brain MR image segmentation. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, 2018. 6(1): p. 113-117.

[14] H. Sokooti, et al., Nonrigid image registration using multi-scale 3D convolutional neural networks. In the Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2017. Springer.

[15] P. Jiang and J.A. Shackleford. CNN driven sparse multi-level b-spline image registration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2018.

[16] H. Uzunova, et al., Training CNNs for image registration from few samples with model-based data augmentation. In the Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2017. Springer.

[17] E. Ferrante, et al., On the adaptability of unsupervised CNN-based deformable image registration to unseen image domains. In the Proceedings of the International Workshop on Machine Learning in Medical Imaging. 2018. Springer.

[18] J.A. Lee, et al., A deep step pattern representation for multimodal retinal image registration. In the Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019: p. 5077-5086.

[19] J. Hu, et al., End-to-end multimodal image registration via reinforcement learning. Medical Image Analysis, 2021(68): p. 101878.

[20] A. Hering, et al., CNN-based Lung CT Registration with Multiple Anatomical Constraints. Medical Image Analysis, 2021: p. 102139.

[21] H.R. Boveiri, et al., Medical image registration using deep neural networks: A comprehensive review. Computers & Electrical Engineering, 2020. 87: p. 106767.

[22] S. Belongie and J. Malik. Matching with Shape Contexts. In the 2000 Proceedings Workshop on Content-based Access of Image and Video Libraries. 2000. Head Island, SC, USA.

[23] S.E. McGlynn, et al., Single cell activity reveals direct electron transfer in methanotrophic consortia. Nature, 2015. 526(7574): p. 531-535.

[24] L. Polerecky, et al., Look@NanoSIMS-a tool for the analysis of nanoSIMS data in environmental microbiology. Environmental Microbiology, 2012. 14(4): p. 1009-1023.

[25] X. Zhang, et al., ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

[26] C. Szegedy, et al., Going deeper with convolutions. In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.

[27] K. He, et al., Deep residual learning for image recognition. In the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

[28] K. Simonyan and A. Zisserman, Very deep convolutional networks for large-scale image recognition. In the Proceedings of the 3rd International Conference on Learning Representations. 2015: San Diego, CA, USA.

[29] N. Otsu, A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics. 1979, 9(1), 62–66.

[30] R. Jonker and A. Volgenant, A shortest augmenting path algorithm for dense and sparse linear assignment problems. Computing, 1987. 38(4): p. 325-340.

[31] M.J.D. Powell, A thin plate spline method for mapping curves into curves in two dimensions. Computational Techniques and Applications (CTAC '95), 1995.

[32] A. Goshtasby, Image registration by local approximation methods. Image and Vision Computing, 1988. 6: p. 255-261.

[33] A. Goshtasby, Piecewise linear mapping functions for image registration. Pattern Recognition, 1986. 19: p. 459-466.

[34] N. Pielawski, et al., CoMIR: Contrastive multimodal image representation for registration. In the Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS 2020). 2021: Vancouver, Canada.

[35] H. Bay, et al., SURF:Speeded Up Robust Features. Computer Vision and Image Understanding (CVIU), 2008. 110(3): p. 346–359.

[36] P.F. Alcantarilla, A. Bartoli, and A.J. Davison, KAZE Features. In the Proceedings of the Computer Vision – ECCV 2012. Lecture Notes in Computer Science, A. Fitzgibbon, et al., Editors. 2012, Springer: Berlin, Heidelberg. p. 7577.

[37] S. Leutenegger, M. Chli, and R. Siegwart, BRISK: Binary Robust Invariant Scalable Keypoints. In the Proceedings of the 2011 International Conference on Computer Vision. 2011. Barcelona, Spain.

[38] C. Harris and M. Stephens, A Combined Corner and Edge Detector. In the Proceedings of the 4th Alvey Vision Conference, 1988: p. 147-151.

[39] E. Rosten and T. Drummond, Fusing Points and Lines for High Performance Tracking. In the Proceedings of the IEEE International Conference on Computer Vision, 2005. 2: p. 1508–1511.

[40] Y. Bentoutou, et al., An automatic image registration for applications in remote sensing. IEEE Transactions on Geoscience and Remote Sensing, 2005. 43(9), pp.2127-2137.

[41] Z. Wang, et al., Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 2004. 13(4), pp.600-612.

[42] Z. Zhang, et al., A new image registration algorithm based on evidential reasoning. Sensors, 2019. 19(5): p. 1091.

[43] A. Goshtasby, 2-D and 3-D Image Registration for Medical, Remote Sensing, and Industrial Applications. 2005: John Wiley & Sons.

[44] N. Tajbakhsh, et al., Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging, 2016. 35(5): p. 1299-1312.

[45] R. Sprengel, et al., Thin-plate spline approximation for image registration. In the Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 1996. 3: p. 1190-1191.

[46] B.W. Papież, et al., Non-local graph-based regularization for deformable image registration. In the Proceedings of the Medical Computer Vision and Bayesian and Graphical Models for Biomedical Imaging: MICCAI 2016 International Workshops, MCV and BAMBI, Athens, Greece, October 21, 2016, pp. 199-207. Springer International Publishing.

[47] L. Tautz, et al., 2010, April. Phase-based non-rigid registration of myocardial perfusion MRI image sequences. In the Proceedings of the 2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 516-519. IEEE.