

Real-Time Image Registration via A Deep Learning Approach for Correlative X-ray and Electron Microscopy

Yanqi Luo¹, Nestor Zaluzec², Mathew Cherukara¹, Xiaolan Wu³ and Si Chen⁴

¹Argonne National Laboratory, United States, ²Argonne National Laboratory / Photon Science Directorate, Bolingbrook, Illinois, United States, ³Faculty of Materials and Manufacturing, Key Laboratory of Advanced Functional Materials, Education Ministry of China, Beijing University of Technology, Beijing, China, United States, ⁴Argonne National Laboratory, Lemont, Illinois, United States

Image registration is often a required step for correlative analysis, which involves aligning two or more mono- or multi- modal image data to relate functional properties and other material characteristics. In this study, a deep learning approach is explored to enable real-time automatic image registration without human intervention. We train a deep convoluted neural network (CNN) to extract important features from X-ray and electron beam (e-beam) microscopic images, and predict the affine transformation parameters such as rotation, scaling and shifts on x- and y- axes between a pair of fixed and moving images.

The goal of image registration is either finding a coordinate transformation matrix H that aligns a pairs of images globally,[1] or predicting a non-rigid deformation field ϕ for aligning features with changes in geometry locally.[2] Coordinate transformation matrix H is often needed to align data with different field of views (FOVs), while deformable registration assumes FOVs are the same but needs local adjustments. Due to resolution limits and specimen preparation difference between X-ray and e-beam measurements, aligning images from the two often involves finding the scaling, translation, and rotation factors, which are the parameters in transformation matrix H and the main focus of this study. Traditionally, intensity- and feature- based approaches have been applied to finding this coordinate transform matrix . Intensity-based approach is straightforward, in which the algorithm stops updating H when minimum pixel intensity difference is reached,[3] while feature-based approach generally has better performance in handling images with acceptable difference in contrast, signal to noise ratio (S/R) and field of view difference, because it solves H directly using coordinates of similar features identified by feature detectors such as SIFT (scale invariant feature transform) and ORB (Oriented FAST and Rotated BRIEF).[4]

Convoluted neural networks (CNNs) have excelled at various computer vision tasks including feature classification, object detection and segmentation.[5] In this work, we utilize two-headed CNN architecture, that takes as input two images and predicts the parameters of the affine transformation. Thus, solely using the two images as input, the CNN can learn how to solve the registration problem. Image augmentation is also utilized to obtain a trained CNN model that is more tolerant of noise or blurring effect from both mono- or multi- modal measurement.[6]

Figure 1 illustrates the initial performance of the trained model on the test dataset. These testing images, also known as “Target” images, are generated by applying the known transformation on the image data (shown in Figure 2). These known transformations are also normalized and standardized using their average and standard deviation value for CNN training purposes. Figure 1 suggests a larger difficulty in training the rotation and scaling parameter because larger prediction variation is observed among all the parameters. This is also reflected on transformed transmission electron microscopic (TEM) image as shown in Figure 2, where the transformed image (Data-tform) zooms in a bit more than the target image. This result demonstrates the potential of using deep learning to predict affine transformation, and the likelihood of estimating real-time alignment once the CNN is trained, which allows multimodal alignment during data collection and systematic correlative microscopic analysis for large dataset.

Acknowledgements. This research used resources of the Advanced Photon Source, and the Center for Nanoscale Materials both U.S. Department of Energy (DOE) Office of Science User Facilities operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. The work is also funded by Laboratory Directed Research and Development 2020-0038.

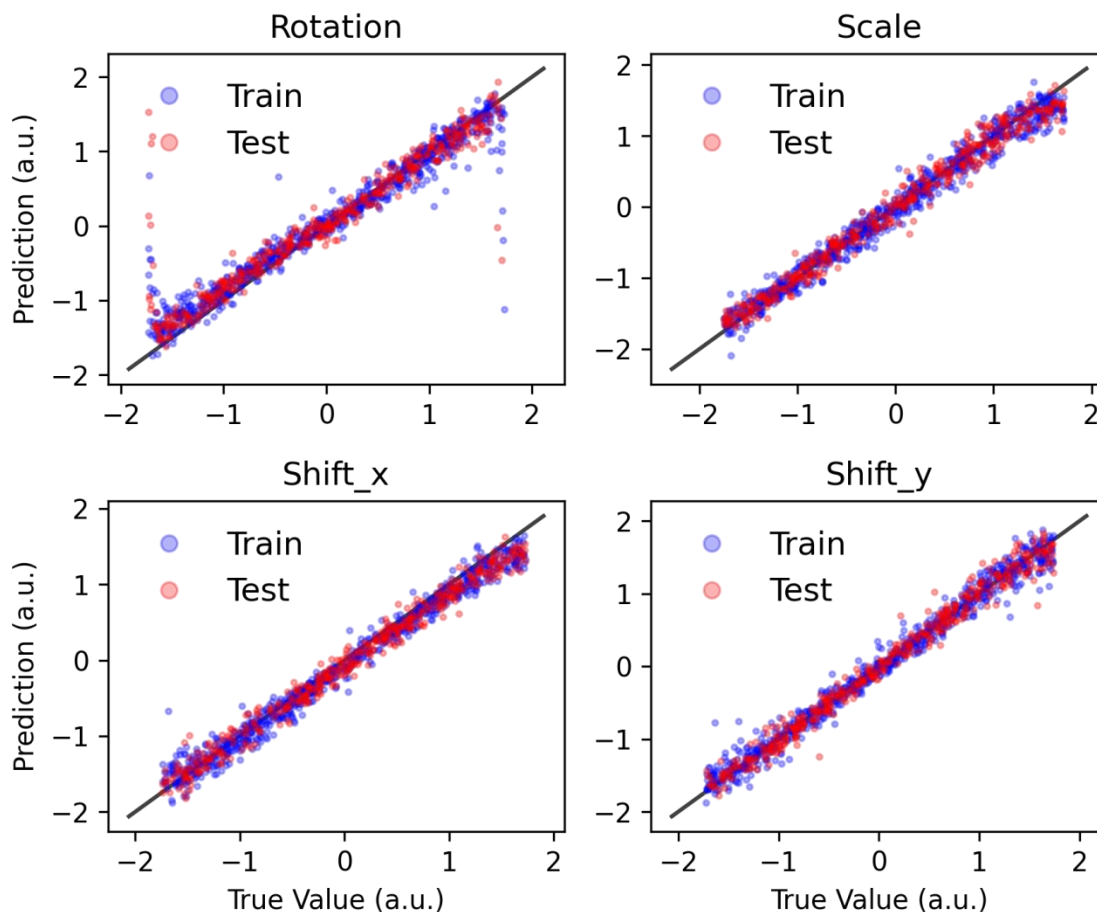


Figure 1. Figure 1. Parameter predictions of the trained CNN. Given a pair of test images, CNN predicts four parameters: angle of rotation, scaling fraction, shifts on both x- and y- axes. Correlation plots for training and testing dataset are shown where the ground truth and predicted values are plotted on x- and y- axis, respectively. A guided black line with $y = x$ is also plotted. For each parameter, both target and predicted values are standardized before CNN training. It helps the network to consider the impact from each parameter evenly.

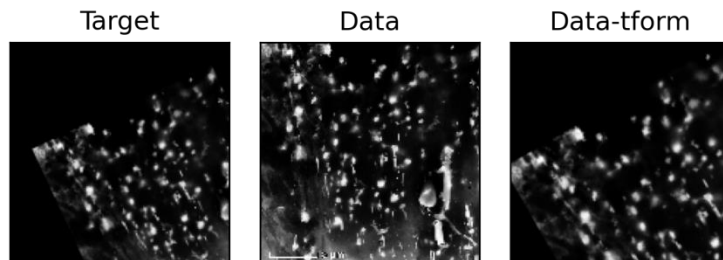


Figure 2. Figure 2. Image transformation using CNN predicted value. Target (left), Data (middle), and Data-tform (right) is fixed, moving, and registered image, respectively. Scale bar in data image is 5 μm .

References

- [1] T. Nguyen, S. W. Chen, S. S. Shivakumar, C. J. Taylor, and V. Kumar, “Unsupervised Deep Homography: A Fast and Robust Homography Estimation Model,” *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 2346–2353, 2018, doi: 10.1109/LRA.2018.2809549.
- [2] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, “VoxelMorph: A learning framework for deformable medical image registration,” *IEEE Trans. Med. Imaging*, vol. 38, no. 8, pp. 1788–1800, 2018.
- [3] P. Thévenaz, U. E. Ruttimann, and M. Unser, “A pyramid approach to subpixel registration based on intensity,” *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 27–41, 1998, doi: 10.1109/83.650848.
- [4] D. G. Low, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, pp. 91–110, 2004, [Online]. Available: <https://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf>.
- [5] A. Zhao, G. Balakrishnan, F. Durand, J. V. Guttag, and A. V. Dalca, “Data augmentation using learned transformations for one-shot medical image segmentation,” *arXiv*, 2019.
- [6] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, and A. A. Kalinin, “Albumentations: Fast and Flexible Image Augmentations,” *Information*, vol. 11, no. 2, pp. 1–20, 2020.
- [1] P. Thévenaz, U. E. Ruttimann, and M. Unser, “A pyramid approach to subpixel registration based on intensity,” *IEEE Trans. Image Process.*, vol. 7, no. 1, pp. 27–41, 1998, doi: 10.1109/83.650848.
- [2] D. G. Low, “Distinctive image features from scale-invariant keypoints,” *Int. J. Comput. Vis.*, pp. 91–110, 2004, [Online]. Available: <https://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf>.