doi:10.1017/S1431927621000908
Microsc. Microanal. 27 (Suppl 1), 2021

Automatic analysis of cryo-electron tomography using computer vision and machine learning

Min Xu
Carnegie Mellon University, Pittsburgh, United States

Cryo-electron tomography (cryo-ET) is an emerging technology for the 3D visualization of structural organizations and interactions of subcellular components at near-native state and sub-molecular resolution. Tomograms captured by cryo-ET contain heterogeneous structures representing the complex and dynamic subcellular environment. Since the structures are not purified or fluorescently labeled, the spatial organization and interaction between both the known and unknown structures can be studied in their native environment. The rapid advances of cryo-ET have generated abundant 3D cellular imaging data. However, the systematic localization, identification, segmentation, and structural recovery of the subcellular components have been very difficult due to the structural complexity and imaging limits. For example, the native macromolecular structures are very diverse. They can be inside a very crowded molecular environment. The tomograms are often of low signal-to-noise ratio and have missing values. Therefore, advanced efficient and accurate large-scale image analysis methods are needed for analyzing such tomogram data.

We developed and adopted a suite of computer vision and machine learning methods for such analysis. In particular, we have developed a set of methods for the *de novo* structural pattern mining among cryo-ET data for the discovery of macromolecular structures without using any external structural template. These methods can be divided into following categories: subtomogram classification, subtomogram alignment and averaging, subtomogram and tomogram segmentation. Subtomogram classification aims at efficiently separating millions of subtomograms into structurally homogeneous subsets. Subtomogram alignment finds optimal rotation and translation so that two subtomograms can achieve maximal overlap. Subtomogram averaging overlay the aligned subtomograms containing identical structures in order to obtain improved resolution representation of the structure. Tomogram segmentation identifies the sub-regions of the tomogram that occupied by different types of subcellular components.

Our efforts on the development of such structural pattern mining methods provide is an important step towards visual proteomics analysis of single cells using cryo-ET.

References

1. Du X, Wang H, Zhu Z, Zeng X, Chang Y, Zhang J, Xing E, Xu M. Active learning to classify macromolecular structures in situ for less supervision in cryo-electron tomography. Bioinformatics. doi:10.1093/bioinformatics/btab123
2. Li R, Yu L, Zhou B, Zeng X, Wang Z, Yang X, Zhang J, Gao X, Jang R, Xu M. Few-shot learning for classification of novel macromolecular structures in cryo-electron tomograms. PLOS Computational Biology. doi:10.1371/journal.pcbi.1008227
3. Zeng X, Xu M. Gum-Net: Unsupervised geometric matching for fast and accurate 3D subtomogram image alignment and averaging. CVPR 2020.
4. Du X, Zeng X, Zhou B, Singh A, Xu M. Open-set Recognition of Unseen Macromolecules in Cellular Electron Cryo-Tomograms by Soft Large Margin Centralized Cosine Loss. BMVC 2019.
5. Zhao G, Zhou B, Wang K, Jiang R, Xu M. Respond-CAM: Analyzing Deep Models for 3D Imaging Data by Visualizations. Medical Image Computing & Computer Assisted Intervention (MICCAI) 2018. arXiv:1806.00102
6. Zhao Y, Zeng X, Guo Q, Xu M. An integration of fast alignment and maximum-likelihood methods for electron subtomogram averaging and classification. ISMB 2018. Bioinformatics. 2018 Jul 1; 34(13): i227–i236. doi:10.1093/bioinformatics/bty267.

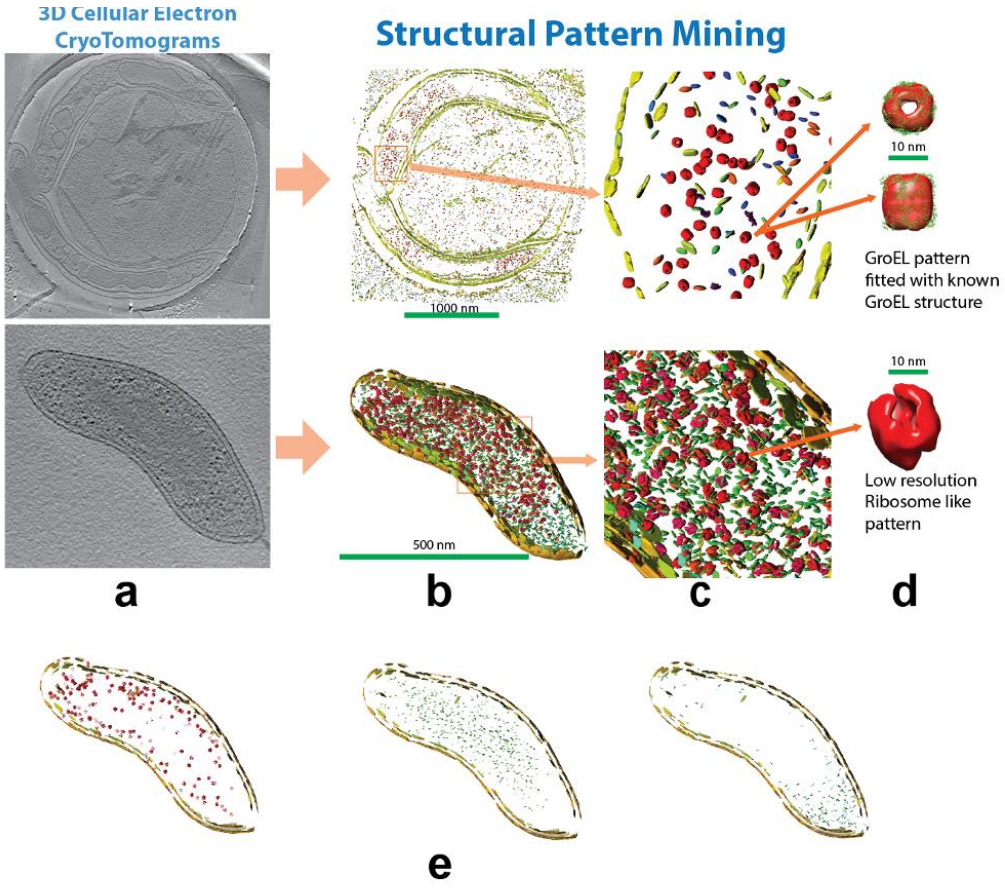

Figure 1. The figure above shows extracted structural patterns in cellular tomograms: (a) Slices of 3D tomogram images of two bacterial cells. Image data are from the Jensen Lab at Caltech. (b) Isosurfaces of instances of extracted structural patterns embedded into the original images. (c) Embedded instances, zooming in on a particular region. (d) Isosurfaces of one example structural pattern from each experiment. (e) Spatial distributions of instances of different structural patterns: left: the Ribosome like patterns distributed outside the nucleoid region; middle: patterns distributed on the nucleoid region; right: patterns distributed at the tip of the cell. For details, see our paper on Structure.

7. Zeng X, Leung M, Zeev-Ben-Mordehai T, Xu M. A convolutional autoencoder approach for mining features in cellular electron cryo-tomograms and weakly supervised coarse segmentation . Journal of Structural Biology. 2018 May;202(2):150-160. doi:10.1016/j.jsb.2017.12.015
8. Xu M, Singla J, Tocheva E, Chang Y, Stevens R, Jensen G, Alber F. De novo structural pattern mining in cellular electron cryo-tomograms. Structure. 2019 Apr 2;27(4):679-691.e14. doi:10.1016/j.str.2019.01.005. (Appeared on Structure volume cover and highlighted in Nature Methods 16, page 285 (2019), doi:10.1038/s41592-019-0382-2)
9. Xu M, Beck M, Alber F. High-throughput subtomogram alignment and classification by Fourier space constrained fast volumetric matching. Journal of Structural Biology. 2012 May;178(2):152-64.

https://doi.org/10.1017/S1431927621000908 Published online by Cambridge University Press