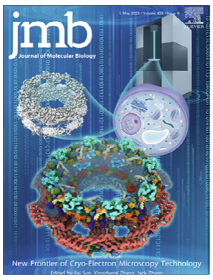
Review Article



# Computational Methods Toward Unbiased Pattern Mining and Structure Determination in Cryo-Electron Tomography Data

Hannah Hyun-Sook Kim [1], Mostofa Rafid Uddin [2], Min Xu [2⇑] and Yi-Wei Chang [1⇑]

1 - Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA
2 - Computational Biology Department, School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, USA

Correspondence to Min Xu and Yi-Wei Chang: mxu1@cs.cmu.edu (M. Xu), ywc@pennmedicine.upenn.edu (Y.-W. Chang)@hannahinthelab (H.H.-S. Kim), @duran_rafid (M.R. Uddin), @xumin100 (M. Xu), @yiweichang (Y.-W. Chang)
https://doi.org/10.1016/j.jmb.2023.168068
Edited by Xingzheng Zhang

## Abstract

Cryo-electron tomography can uniquely probe the native cellular environment for macromolecular structures. Tomograms feature complex data with densities of diverse, densely crowded macromolecular complexes, low signal-to-noise, and artifacts such as the missing wedge effect. Post-processing of this data generally involves isolating regions or particles of interest from tomograms, organizing them into related groups, and rendering final structures through subtomogram averaging. Template-matching and reference-based structure determination are popular analysis methods but are vulnerable to biases and can often require significant user input. Most importantly, these approaches cannot identify novel complexes that reside within the imaged cellular environment. To reliably extract and resolve structures of interest, efficient and unbiased approaches are therefore of great value. This review highlights notable computational software and discusses how they contribute to making automated structural pattern discovery a possibility. Perspectives emphasizing the importance of features for user-friendliness and accessibility are also presented.



## Introduction

Cryo-electron tomography (cryo-ET) enables in situ visualization of subcellular structures in their native states inside a cell, which is not possible by any other imaging technology.[1] X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and single particle cryo-electron microscopy (cryo-EM) can resolve atomic details of macromolecular structures but are in vitro methods that lack cell context. Conversely, fluorescence imaging methods can image an entire cell and the organelles within but cannot resolve macromolecular structures. Cryo-ET uniquely bridges these imaging methods by providing molecular resolution information of structures in a cellular context.[2] The absence of disruptive sample preparation, such as chemical fixation,[3] further enables the observation of spatial- and temporal-regulated protein–protein interactions, permitting the study of a cell's native molecular sociology.[4] Thus, cryo-ET can reveal complex and novel macromolecular structures inside the cells which represent active function-conducting cellular machines. In preferred scenarios, the resolution can achieve the sub-nanometer range, thereby directly revealing structural mechanisms in situ.[5–7]

In cryo-ET, specimens are immobilized via cryo-fixation and subsequently imaged as the sample is progressively tilted about an axis perpendicular to

the electron beam. This generates a series of 2D transmission electron micrographs that can then be used to reconstruct a 3D image, or tomogram.[8] The tomograms contain readily visible organelles and macromolecular structures, represented as recurring density patterns that appear throughout the volume. By recognizing and then averaging multiple instances of a pattern, the underlying structure can be determined. A complete molecular structural landscape visualization of a cell can be pursued by determining multiple structures and localizing them throughout the tomogram (Figure 1(a)–(e)). Yet, this is challenging given that raw tomograms contain large, complex data that is further complicated by artifacts such as the missing wedge effect and low signal-to-noise ratio (SNR).[9] Manual analysis of such data is not only cumbersome but at risk of bias. Hence, automated methods to extract and analyze unique cellular structures captured by cryo-ET are of great importance.

Tomogram processing suites such as IMOD,[10–12] EMAN2,[13–16] PyTom,[17–20] Protomo,[21–25] and others provide comprehensive tools for general tomogram analysis (see Table 1). In addition, the development of new methods that target specific tasks to advance cryo-ET analysis is also being pursued from many facets (see Table 2). Below we follow the general sequence for a structural determination that follows tomographic reconstruction, discussing the advantages, shortcomings, and accessibility of currently available methods to contextualize the current state of quantitative cryo-ET analysis. We primarily focus on approaches that reduce the manual intervention required by researchers and highlight algorithmic developments that we envision will advance the field towards mining structures de novo from cryo-ET tomograms.

## Segmentation

Segmentation is frequently the first step of examining the tomogram. Voxels that correspond to biological features of interest are annotated in tomogram for 3D illustration. Segmentations be performed manually in software such as where a user traces the region of interest slice, but this strategy is tedious and some- inconsistent.[61] Low SNR, low contrast, mis- wedge effects, and crowded densities exacerbate the difficulty of this task.[26,44,62] Develop- strategies for fast, reproducible, and high-quality segmentations is crucial for reliable data annotation and structural analyses.

Different methods have been developed to reduce the level of human intervention required for segmentation. Early approaches largely focused on boundary detection,[63–65] watershed algorithms,[66] and comparisons of regions to predefined models of structures.[62,67] Amira[26] is a popular graphics software by Thermo Fisher Scientific that implements watershed thresholding algorithms to semi-automatically segment biological structures such as membranes, nuclei, and filaments. Methods that unite watershed and boundary approaches have also been investigated.[68] Each approach has its own set of weaknesses, causing its performance to be highly dependent on the dataset. Moreover, they require significant input from the user for reasonable segmentations.

SuRVoS[44] is a more recent tool that employs machine learning on manually segmented slices to predict regions on subsequent slices or automatically segment boundaries by analyzing local density changes within the tomogram, effectively reducing the amount of annotation required. Deep learning methods have also been developed toward fully automated segmentation. These are capable of robustly segmenting multiple subcellular structures at once. For instance, the EMAN2 software suite[14–15] includes a deep learning-based automated segmentation method that requires only 10 manually annotated images to locate similar particles throughout the tomogram. This vastly reduces the need for manual annotation for feature extraction (Figure 1(f)–(i)). The software can also utilize a neural network trained on one tomogram to segment other tomograms, provided they have a similar imaging quality and voxel size. This semi-automated segmentation approach has supported the investigations of Trypanosoma brucei as well as thylakoid membranes.[69–70] In addition to annotation, segmentation is commonly used to facilitate other structural pattern-mining tasks. Many software introduced in later sections of the review can also perform automated segmentation to facilitate particle picking by restricting the search space (e.g., using membrane segmentations to locate membrane-bound complexes) and informing on particle orientation.[2,36,61]

## Particle picking

Particle picking can be considered a variation of segmentation. This process picks subvolumes, the also known as subtomograms, of equal size in the can tomogram and each contains a structure of IMOD,[10] interest (the "particle"). The subtomograms can slice-by- then be computationally cropped out and the collec-times tion of particles is overlaid to generate a subtomo-ing gram average with enhanced contrast and resolution of the structure of interest. While soft-ing ware packages often allow users to manually select positions to generate subtomograms,[29,72] this can be time-consuming, subjective, and incomplete.[73] Template matching and template-free approaches for particle picking have been developed to facilitate this task.

### Template matching

Templates are used to find corresponding density matches in the tomogram. They are generally

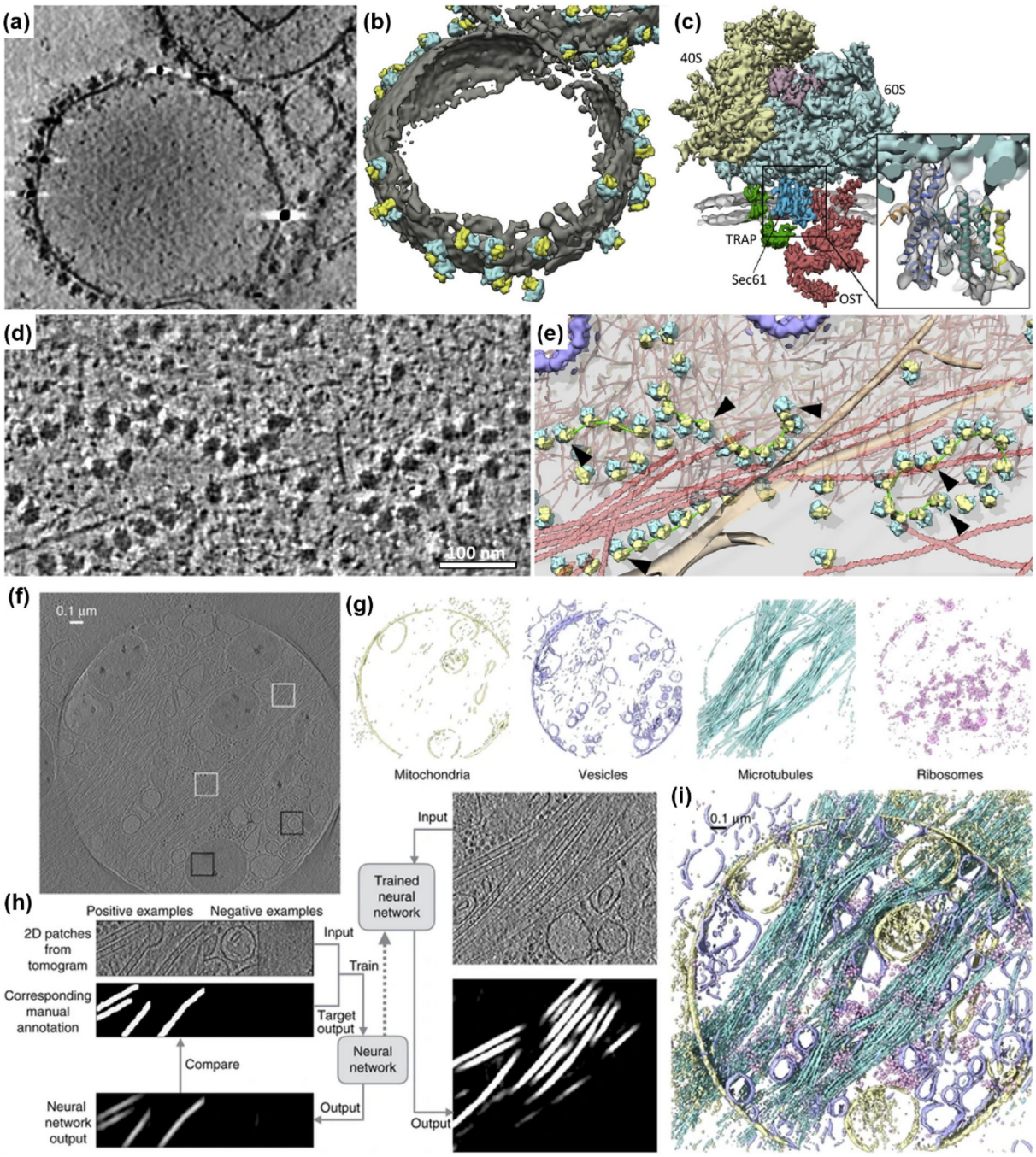



Figure 1. Exploring the diverse, crowded environment within cellular tomograms. A tomogram slice (a) and corresponding isosurface rendering (b) of an endoplasmic reticulum (ER)-derived vesicle decorated in ribosome-transolocon complexes. (c) The cryo-EM structure of the translocon complex bound to the translocon associated protein complex (TRAP) and oligosaccharyl-transferase (OST), determined from subtomogram averaging of the complex on vesicles. A tomogram slice (d) and the isosurface rendering of the determined complexes (e) in intact cells demonstrates the increased complexity of studying structures in the native cellular context and demonstrates the generation of a limited visual atlas.[78] A tomogram slice (f) of a PC12 cell and (g) the training and annotation of regions of interest in the EMAN2 analysis software. (h) The automated segmentation of four independently trained neural networks for each labeled structure of interest. (i) The final merged annotation of the four features identified by EMAN2 segmentation neural network.[14]

macromolecular structures resolved with X-ray crystallography, NMR spectroscopy, or single-particle cryo-electron microscopy (cryo-EM). These reference structures must be properly low-pass-filtered and adjusted to reflect the tomogram's resolution and distortions from the

Table 1 Cryo-ET image software for pattern-mining tasks.

| Software tool | GUI | Segmentation | Template matching | Template-free | Subtomogram averaging | Classification |
|---|---|---|---|---|---|---|
| Amira + EMPackage[26] | Yes | Yes | Yes (for fiber tracing) | No | No | No |
| Bsoft (w/ Jsubtomo)[27,28] | No | Yes | Yes | No | Yes | No |
| Dynamo[29–33] | Yes | No | Manual | No | Yes | Yes |
| EMAN2[13–16] | Yes | Yes | Yes | Yes | Yes | Yes |
| emClarity[34,35] | No | No | Yes | No | Yes | No |
| IMOD (w/ PEET)[10–12] | Yes | Yes | Manual | No | Yes | No |
| Protomo[21–25] | No | No | Yes | No | Yes | Yes |
| PySeg[36–38] | No | Yes | No | Yes | Yes | Yes |
| PyTom[17–20] | Partial | No | Yes | No | Yes | Yes |
| RELION[39–43] | Yes | No | No | No | Yes | Yes |
| SuRVoS[44] | Yes | Yes | No | No | No | No |
| Tomominer, TomoMinerCloud[45] | No | No | Yes | Yes | Yes | Yes |

cryo-ET imaging process.[74–75] Templates may also be generated de novo from subtomogram averages if densities are highly abundant and regular.[76] All possible orientations of the templates are scanned through the tomogram to calculate correlation maps. Statistical analyses, such as peak extraction, are applied to the map to discern macromolecule positions throughout the tomogram. The reference structure can even be mapped back to these positions to create a visual atlas of the macromolecule populating the tomogram[18] (Figure 2(a)).

Many software packages offer template matching, including Dynamo,[29–33] emClarity,[34–35] PyTom,[17–20] EMAN2,[13–16] and TomoMiner[45] (see Table 1). Template matching has enabled the determination of highly resolved structures of large macromolecular complexes, such as ribosomes, the 26S proteosome, and viral Gag assembles and is widely used to automatically identify particles of interest from cryo-ET tomograms.[75,77–79] Aside from discrete particles, it can be used in concert with tracing algorithms to segment filamentous structures, such as those present in the cytoskeleton.[80–82]

Despite its popularity, traditional template matching suffers from several drawbacks. It is low throughput and can be extremely time-consuming.[52] Performance depends on both the image quality of the tomogram and the reference structure. Molecular crowding in the tomogram is another major concern.[83] Dense packing of particles in tomograms may increase the presence of off-target densities, making it more challenging to locate densities of interest. While a tight mask enables algorithms to ignore extraneous densities and focus on the most relevant regions of the candidate object when calculating template matching scores, this solution can only be used in limited situations where molecules are rigid and well-resolved.[74] False positives are another issue, resulting in references being matched to not only the true location of structures but also regions of noise.[84–85]

The most inherent disadvantage of template matching is template bias, where detectable structures are limited to those with an existing template obtained from other structural biology methods.[2,61,86] Likewise, template-matching algorithms may suffer sample-specific biases that can occur if templates do not reflect unexpected binding partners or conformations. On the other hand, averaging manually selected subtomograms potentially introduces user-dependent bias. Consequently, template matching cannot be utilized effectively in studies involving flexible macromolecules or complex interactions, and alternative methods must be considered.

## Template-free particle picking

To forgo the need for preexisting structural references to recover cellular structures de novo for pattern recognition, template-free approaches are of growing interest for cryo-ET data analysis. To achieve the nonspecific extraction of subtomogram volumes, some approaches have turned to deep learning. Convolutional neural networks (CNNs) are deep learning architectures that are powerful tools especially adept at image processing, which is useful given that subtomograms are essentially 3D images representing particles.[87] Neural networks such as DeepPicker[73] and SPHIRE-crYOLO[88] exist for automated particle picking in single-particle cryo-EM micrographs. Performing the same task in cellular tomograms is a greater challenge because the native cellular environment is more crowded and contains a more diverse set of structures compared to cryo-EM micrographs.[15]

Neural networks for feature extraction have also been successfully implemented for cryo-ET. As previously noted, the EMAN2 package[14–15] implements deep learning for automated segmentation, but can also perform automatic particle picking. Although this approach can identify particles from more than just a single class, the computation time

Table 2 Recent notable technical developments for specific cryo-ET pattern-mining tasks.

| Software tool | Task | Notes |
|---|---|---|
| SuRVoS[44] | Segmentation | Allows users to vary the degree of manual segmentation required. |
| DSM-Net[46] | SegmentationClassification (supervised)<br>Structure determination | Multi-task deep learning model to train three networks simultaneously on (1) semantic segmentation, (2) classification, and (3) coarse structure determination. Outperformed single-task models and was able to identify structures absent in training data. |
| VP-Detector[47] | Segmentation<br>Particle pickingClassification (supervised) | Uses a CNN to take user-annotated data and automatically segment, localize, and classify particles of interest. |
| DeepFinder[48] | Particle picking | Efficient multiclass particle picking of structures with various sizes and shapes. |
| 3D-UCaps[49] | Particle picking | Efficient multiclass particle picking was given limited training data. |
| PySeg[36] | Particle pickingClassification (unsupervised) | Particle picking of small membrane-bound complexes. |
| DSRF3D-v2, RB3D, CB3D[50] | Classification (supervised) | Classifies complexes spanning a range of molecular weights. |
| DISCA[51] | Classification (unsupervised) | High-throughput, fully unsupervised approach to cluster particles using deep learning. |
| MPP[52 7] | Classification (unsupervised)<br>Structure determination | Iterative alignment to sort a heterogenous set of particles into structure patterns, capable of identifying patterns de novo. |
| Jim-Net[53] | Classification (unsupervised)<br>Structure determination | Uses pair-matching alignment and hierarchical clustering to achieve comparable accuracy with supervised method to determine structures de novo. |
| RELION-4.0 Bayesian single-particle tomography structure determination[43] | Structure determination | Structure determination via optimization of a regularized likelihood function. In contrast to previous RELION renditions, features a new weighting system and refinement methods. |
| FAML[54] | Structure determination | Combines features of FRM and maximum likelihood structure determination methods to achieve higher robustness to noise and artifacts compared to FRM methods and requires fewer subtomograms. |
| HEMNMA-3D[55] | Structure determination | Creates a conformational space of a complex that maps how a structure may transition between different configurations in situ. |
| -[56] | Data augmentation | High density packing by modeling structures with multiple spheres. |
| -[57] | Data augmentation | Computationally efficient packing by modeling structures as single spheres with gradient descent algorithm. |
| CryoETGAN[58] | Data augmentation | Deep learning model designed to train on tomograms to generate imitations that can be used to increased training data without the need to simulate artificial tomograms, which may not fully reflect experimental data. |
| 3D-ADA[59] | Domain adaptation | Utilizes deep learning to map particles from images with different parameters to an intermediate latent space. |
| Cryo-Shift[60] | Domain adaptation | Transforms training data to improve the generalization ability of classifiers to predict on different tomography datasets. |

scales linearly for each additional class. Because particles of different structures could be used to inform the recognition of each other, it is possible to develop a more efficient system to leverage this information. DeepFinder[48] is a CNN that can simultaneously segment multiple classes of varying shapes and sizes (Figure 2(b)–(h)). While the EMAN2 classifier has a shorter training time to identify a single class, DeepFinder outperforms EMAN2 in multi-class automatic particle picking. Moreover, it reduces the need for extensive postprocessing for reliable results. Improving on these attempts, 3D-UCaps[49] is a deep learning framework that can also perform binary and multi-class particle identification. Replacing the data-hungry CNN architecture with the data-efficient CapsNet, this method demonstrates comparable picking performance even in situations with limited training data (defined as less than 10,000 annotated particles).

Specialized segmentation methods exist for membrane-bound complexes. Pyto[89] is specifically designed to detect small, flexible complexes linking larger structures such as membranes, organelles, or filaments. Multiple connectivity-based segmentation at different thresholds is organized into a hierarchy to establish a final segmentation. This is in turn used to identify particles. Pyto has a high detection rate and offers statistical analyses to characterize the segmentations. Another software for particle picking is PySeg,[36] a software package that also focuses on the automatic localization of small, membrane-bound complexes. It incorporates

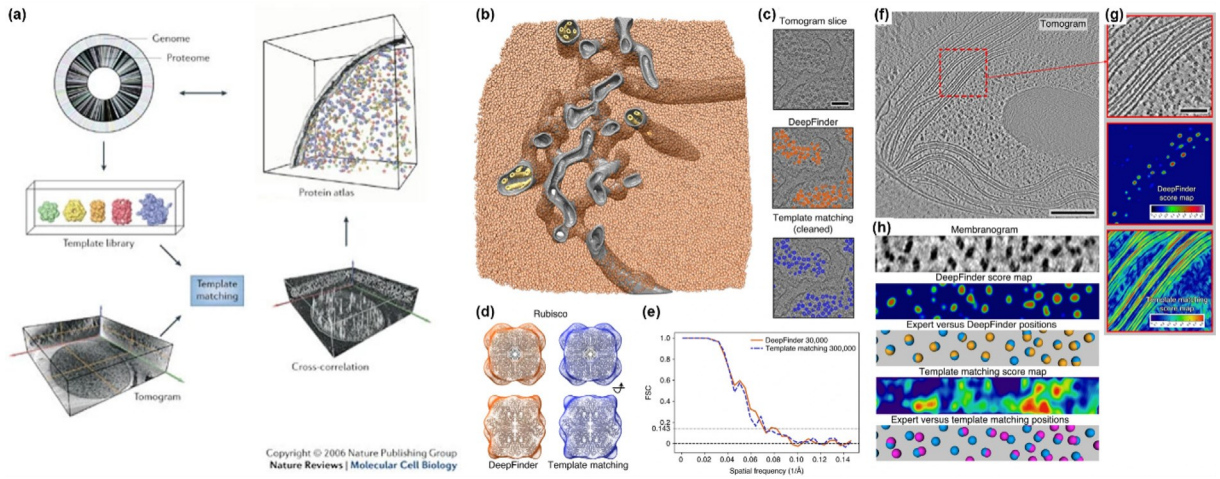



Figure 2. Particle picking via template matching or template-free approaches. (a) Template mapping to achieve a visual protein atlas. Template matching structures from a template library identifies the cross-correlation values of each reference against regions of a tomogram, ultimately identifying positions that most likely represent particles from the library. The rendered structure models at the calculated positions and orientations forms a protein atlas, providing information on the distribution of these structures throughout the cell and their relation to other structures.[91] (b) manually segmented membranes (gray and yellow) and particle picking positions Rubisco holoenzymes (orange) in native Chlamydomonas pyrenoid determined by DeepFinder.[48] (c) A tomogram slice (top) with the segmentations as performed by DeepFinder (middle; orange) and template matching (cleaned with CPCA classification). Scale bar, 50 nm. Subtomogram averages of holoenzymes generated from each particle picking method with Rubisco molecular structure (d) and FSC curves showing final resolution of 15 Å (FSC > 0.143; e). Slice of tomogram (f) and zoom-in comparing DeepFinder to template matching score map (g). (h) Membranogram (topological view of membrane surface by projecting tomogram densities onto segmented membrane surface) and corresponding comparisons between DeepFinder-, template matching-, and expert-derived particle picking positions.

TomoSegMemTV,[37] a package that utilizes tensor voting to achieve robust segmentation of membranes, as well as PyOrg,[38] which offers statistical analyses to characterize the segmentations.

## Structure determination

Once particles have been identified from a tomogram, the underlying structures must be recovered by further processes. The particles in subtomograms are typically positioned in various orientations but should contain common features with other particles of the same structural class. Thus, subtomograms are aligned and averaged to recover the high-resolution structure shared by the individual particles within each class. The better-resolved structures can then be re-embedded into the original tomogram space to obtain the visual landscape. This is similar to template matching but replaces the reference with the obtained structure. Several well-established analysis packages provide subtomogram averaging and manipulation protocols (see Table 1). Note that these strategies are designed to analyze a relatively homogenous particle set. The following section (2. Structure determination) discusses strategies to handle heterogeneous particles, which is especially relevant for de novo structure mining.

There are two primary approaches to subtomogram averaging. The first is alignment-based determination, where an angular search is performed to align the subtomogram to a reference structure. This is the subtomogram averaging approach utilized by software such as Dynamo.[29] The reference structure is first masked and rotated in real space. To account for the missing wedge, the reference is then masked in the Fourier space. The 3D cross correlation values are calculated between the reference and subtomogram. The algorithm identifies the transformations that produce highest cross correlation score. These transformations are applied to the subtomograms, a weighted average is performed to account for the previous masking to yield a final structure. However, the search risks converging to local minima, resulting in a suboptimal structure. Without a priori information on the search parameters, the algorithm may also be computationally intensive due to the Fourier transforms and rotations, which require the costly step of linear interpolation.[71] Fast rotational matching (FRM) is a form of alignment-based determination that forgo explicit rotation of volumes by decomposing rotational and translational search to achieve faster global alignment,[9,50] and has been implemented into tomogram analysis packages.[12,91] However, this comes at the cost of reaching suboptimal solutions

and is prone to be biased by image noise and the missing wedge effect.[54]

The second approach to subtomogram averaging, maximum likelihood-based determination, is more robust to the abovementioned effects, and which is used by software such as RELION.[43,92] Expectation-maximization algorithms[93] optimize a probability function of observing the experimental data given the initial reference structure.[71,92,94] The probability-weighted average of each subtomogram in each orientation becomes the new reference and the process is repeated. This is repeated in an iterative process as the reference and probability function are updated with a more refined sampling of possible orientations. The process converges into a well-refined high-resolution and identifiable structural map. This approach also overcomes the need to extensively tune parameters, as they are instead inferred during estimation.[40] Thus, maximum likelihood-based determination is more amenable for non-experts. However, the convergence is relatively slow and requires significant computational resources. For recovering structures from a very large data set, this may not be practical.[54,71]

Efforts have been made to improve subtomogram averaging methods. One technique that has been developed to combine these methods to overcome the weaknesses of FRM and maximum likelihood-based determination is the Fast Alignment Maximum Likelihood (FAML)[54] method. This uses fast alignment to sample transformations; the sampled transformations then approximate integrals to update the maximum likelihoods of subtomogram averages using the expectation–maximization algorithm. Compared to FRM alone, FAML exhibits higher robustness to noise and missing wedge effects, and it performs better with fewer subtomograms.

If the underlying structure of subtomograms is flexible, they will exhibit densities with a degree of variation. HEMNMA-3D[55] is an analysis method that maps the low-dimensional representation of heterogeneity within a set of subtomograms. This technique relies on an initial atomic or pseudo-atomic reference structure of the complex to perform elastic deformations and achieve the mapping. The resulting conformational space allows for grouping and averaging, identification of unanticipated conformations, and animations of the reference structure deforming along the densest regions of the space (i.e., conformations shared by the most subtomograms) or following alternative, rarer trajectories.

## Particle classification

As previously noted, if the total set of subtomograms contains particles with different structures, it must then be sorted into homogenous groups with similar features, known as structure patterns. Each structure pattern should correspond to a distinct macromolecule or alternate conformation of the same macromolecule. For de novo structure mining, the classification of particles is an inevitable hurdle to overcome. It can also serve to identify false positives from particle picking.

A common approach to classification is multireference alignment.[71,95] This slightly modifies the structure determination strategies discussed in the previous section (2. Structure determination). Rather than being compared to a single reference, the particle is aligned against a set of different references. The particle is ultimately assigned to the reference that yields the closest similarity. This roughly separates the initial particles into separate classes for each reference, which are then averaged to generate a new reference for subsequent iterations. This approach is implemented in popular software packages such as Dynamo and EMAN2.[29,72] These strategies require knowledge of particles a priori, in the form of initial references, so are not conducive to identifying novel structures that populate the tomogram. Classification approaches that forgo the use of references are therefore especially desirable for end-to-end pattern mining.

Reference-free classification can either be supervised or unsupervised. Supervised classification aims to group subtomograms into homogeneous structure patterns based on knowledge provided by the user. Instead of providing a reference structure, a user manually annotates a set of training subtomograms, assigning a label to the subtomogram based on the structure present. With the labeled training data, the classification algorithm aims to imitate the labeling scheme for new data. Deep learning models are powerful tools for this task, especially the CNN architecture, which as noted before, is especially adept at image processing. Many of these frameworks localize particles prior to classification, and so are also capable of automated particle picking.

To this end, Che et al. investigated three CNN architectures (DSRF3D-v2, RB3D, and CB3D).[50] Compared to previous, non-deep learning approaches, the CNNs demonstrate improved discrimination and scalability for structure determination. They found CB3D displays the best performance with a high level of accuracy, even with low SNR (0.01) and the missing wedge effect present. However, the classifier's performance was dependent on the size of the macromolecules, as larger complexes were associated with higher accuracies.

Additional supervised deep learning networks have been employed to locate and classify heterogeneous particles, many of which are introduced in the Shape Retrieval Challenge (SHREC) for the classification of cryo-ET data.[96–98] Such methods include 3D MS-D, 3D ResNet,[99–100]

DN3DUnet,[101–103] YOPO, and UMC[102,104–105] (introduced in SHREC 2020, see more details in[97], as well as URFinder,[99–100,106–107] U-CLSTM,[101–102,108–111] MC DS Net,[102,112] and CFN[99,113–115] (introduced in SHREC 2020, see more details in[98]. The performance of these techniques is evaluated on a fully annotated simulated tomogram dataset containing particles of various molecular weights (Figure 3(a)) and compared to the baseline of template matching methods performed in PyTom.[19] This assessment measured training time, prediction time, performance compared to network architecture size, and observed general trends for these approaches. The SHREC evaluations found that neural networks are more robust to noise than template matching while also significantly reducing computation time, depending on the architecture.[96–98] Of these methods, YOPO does not require semantic segmentation, and instead only relies on the coordinate and particle class for training making it more accessible for users compared to the other learning-based methods introduced in SHREC 2020.[97]

Similar to the case of CB3D, the performance of neural networks featured in SHREC tends to increase with size of the particle, a trend that is especially drastic for smaller particles[96–98] (Figure 3 (b) and (c)). VP-detector[47] is a deep learning framework designed to better predict on complexes across a wide range of molecular weights and features high accuracy, fewer training parameters, and good performance despite small training sets or class imbalances.

In contrast to fully or partially supervised classification, unsupervised approaches do not rely on any manual annotation inform on structure pattern and therefore recover novel or unlabeled structures de novo. k-means is a popular machine learning approach for clustering that is the basis for several approaches to unsupervised subtomogram classification. This strategy was often employed in early unsupervised classification strategies. For instance, Förster et al. developed the CPCA method[116] to perform binary classification while also reducing the impact of the missing wedge effect. A principal component analysis is performed on pair-wise constrained correlation values and used in conjunction with k-means clustering. This procedure discriminated between GroEL and GroEL/GroES complexes (Figure 3(d)–(h)) but was sensitive to SNR and the computational cost of calculating the correlation matrix increased quadratically with the number of subtomograms. AC3D[85] is an algorithm based upon a modified k-means clustering algorithm. This classification scheme is designed to focus on subtomogram regions with large structural discrepancies between class averages and successfully discriminates between structures with and without cofactors (Figure 3(i)). Both the CPCA and AC3D methods are implemented in PyTOM.[19,85]

One drawback of k-means is that the method requires the user to estimate the initial number of structural patterns. It also assumes the classes are evenly populated and well-distributed. A machine learning method capable of forming structure patterns without an initial cluster number is affinity propagation, which evaluates similarities between subtomograms by transferring information between data points. PySeg[36] implements this technique to achieve the automated detection and classification of heterogeneous membrane-bound complexes.

Deep learning approaches are another alternative for unsupervised classification. PySeg,[36] previously noted to be capable of locating membrane-bound complexes in a template-free manner, is also able to perform unsupervised classification on the acquired subtomograms. This approach was validated in microsomal membranes and subsequently tested on tomograms of P19 embryonic carcinoma cells (Figure 4(a) and (b)). Of the determined structure patterns, many were consistent with known structures (Figure 4(c)), such as the ribosome-bound translocon, while other averages were not immediately recognizable (Figure 4(d)). However, the particles contained in these classes nonetheless appeared to be homogenous, suggesting that new structures may have been identified de novo.

DISCA[51] is a fully unsupervised method to identify structure patterns. It first uses template-free particle picking to collect a large set of heterogenous particles ($10^5$ subtomograms from 10 tomograms; Figure 4(e)–(g)). The approach makes use of the YOPO CNN,[97–98] which achieved the third highest accuracy of the tested frameworks and outperformed template matching for classification on the SHREC 2020 dataset. In DISCA, YOPO is used to extract second-order statistics on translation and rotation invariant features. These are then modeled as multivariate Gaussian distributions, effectively clustering the particles according to their underlying structure. DISCA is a time-efficient analysis, where recovering structures of multiple classes only takes a day or two of computation for what would take about a month using template matching.

Classification and structure determination can also be implemented concurrently, as subtomograms that fail to align well with each other can be indicative of particles from distinct classes. A template-free framework known as multi-pattern pursuit (MPP)[52] structure patterns from a large, highly heterogeneous collection of particles from a crowded environment (Figure 4(h)–(k)). This strategy uses a series of alignments based on constrained correlation and FRM to generate structures aligned to common frames. It discards redundant patterns until high-quality patterns remain. The results of MPP depend on the shape and size of the complex, image resolution, and levels of crowding. Although the final struc-

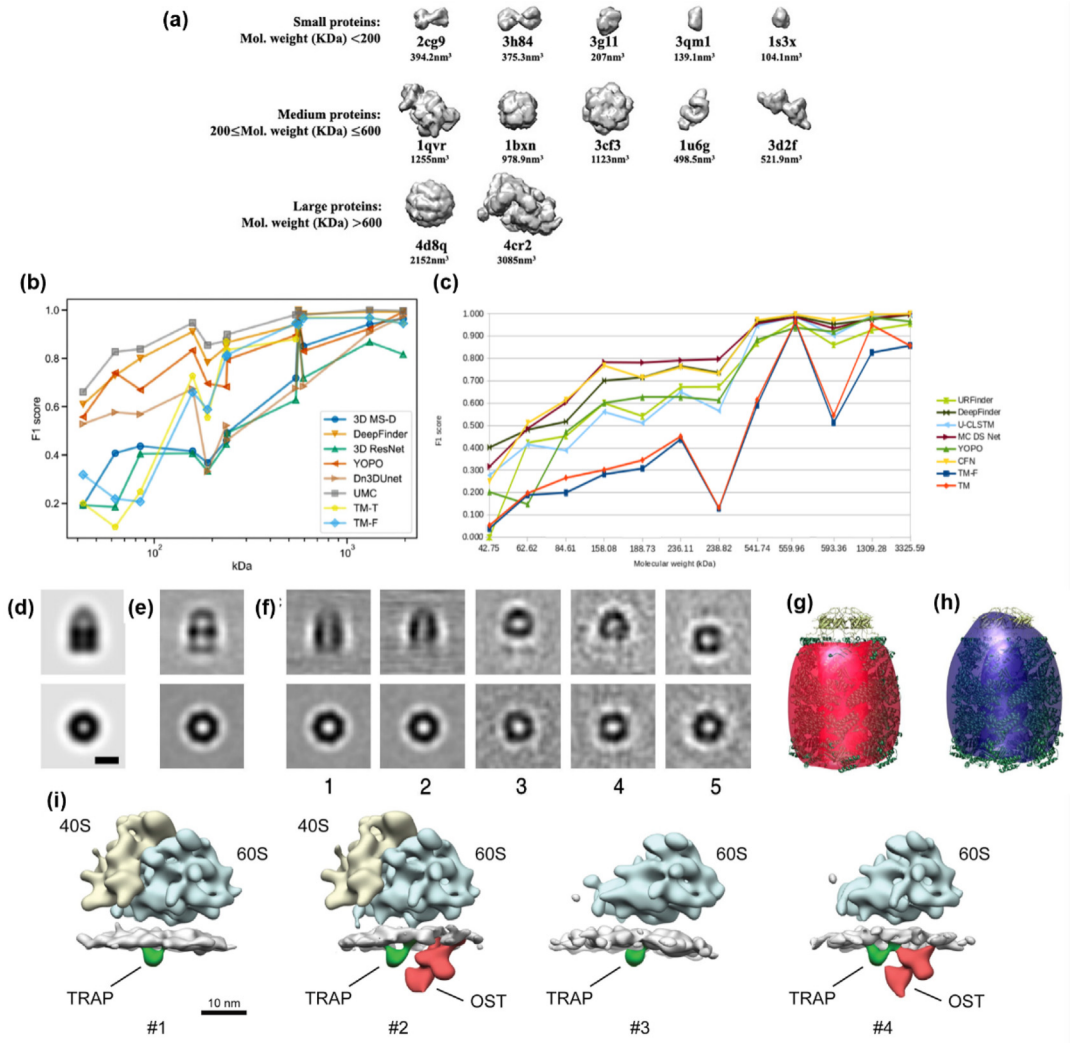



Figure 3. Recent supervised deep learning methods to locate and classify particles from tomograms and early work on unsupervised classification. (a) Isosurface rendering of protein structures across a range of molecular weights in the SHREC data.[47] Neural network performance was evaluated for each particle and PyTom template matching (TM/TM-F) acted as a baseline. These approaches were evaluated based on their F1 score (i.e., the harmonic average of precision, the proportion of true positives to all positives, and recall, the proportion of true positives to all possible relevant elements). (b,c) F1 score versus the molecular weight of the particle for deep learning localization and classification frameworks assessed in SHREC 2020[97] and SHREC 2021,[98] respectively. Classification by CPCA[116] using the reference structure shown in (d), generated from the GroEL14GroES7 complex on tomograms containing GroEL and GroEL/GroES. (e) The average of all subtomograms gathered. (f) Class averages produced from CPCA method. Isosurface rendering of subtomogram averages of two classes determined by CPCA corresponding to (g) GroEL and (h) GroEL/ES. (i) A schematic demonstrating the classification of mammalian ribosomes bound to native endoplasmic reticulum translocons by AC3D.[85] Four classes were identified: 80S ribosomes bound to translocons with TRAP only (class #1), 80S ribosomes bound to translocons with TRAP and OST (class #2), 60S ribosomes with TRAP only (class #3), and finally 60S ribosomes with both TRAP and OST.

tures are coarsely defined, they may serve as references for multireference alignment to yield higher resolution structures.

Simultaneous alignment and classification are also possible through multi-task learning, where a single deep learning model is trained to perform multiple tasks at once to optimize the entire pipeline as opposed to each step. DSM-Net[46] is a multi-task model for simultaneously performing semantic segmentation, classification, and coarse structure determination. Three networks each perform one of these tasks individually, and to achieve multi-task learning, the losses of these networks are combined into a global loss function through linear combination. Each task reinforces the others, and DSM-Net demonstrates improved feature

extraction, reaches greater accuracy, and outperforms single-task models. The model even identified new structures that were not present in the training data. Jim-Net[53] is an end-to-end model for simultaneous, unsupervised alignment and clustering. After hierarchical clustering of subtomograms, a pair-matching alignment algorithm couples and aligns image pairs from a given branch. This architecture performed better than when each task was completed alone, and its accuracy was on par with state-of-the-art supervised methods.

## Challenges & future directions

Although there have been significant advancements in cryo-ET data mining methods, there remain some outstanding challenges that need to be overcome to establish the approach's routine use.

### Limitation of annotated data

Models perform better when trained on large quantities of unbiased data, but obtaining tomograms is labor and time intensive, and more so if manual annotations are required. Although developments in cryo-ET technology expedite the imaging process,[78] issues remain in objectively and efficiently labeling tomograms. Data augmentation addresses the issue by simulating realistic data, while domain adaptation extrapolates information in previously labeled data sources to predict new tomograms.

Data augmentation via simulated tomograms and image synthesis. Via simulated tomograms, data augmentation generates substantial amounts of objectively labeled training data. In contrast, annotating real tomograms suffers from subjectivity, as users may locate and label particles differently. Because simulations offer a ground truth, they can benchmark the performance of data analysis methods. Simulations should closely resemble real tomograms, considering factors such as crowding conditions, image distortions, and noise.

To replicate the heterogeneous, crowded environment of tomograms, Liu et al.[56] simulated macromolecular complexes and ultra-structures as multiple spheres with fixed relative positions that could deform under a force field. This achieves a high density of packing with considerations of structural diversity and deformation but is time-consuming and unsuitable for large data generation. A more efficient approach[57] packs macromolecules, represented as a single sphere, in a noise-free density map that is then used to simulate a tomographic image.

Simulating realistic data must incorporate a variety of factors, and neural networks trained on such data tend to exhibit reduced performance once applied to real data. Rather than creating tomogram simulations from scratch, CryoETGAN[58] is a generative adversarial network designed to use real density maps of particles to synthesize imitations of subtomograms and efficiently infer undiscovered elements present in a tomogram of interest.

Domain adaptation to facilitate model performance between distinct data sets. Domain adaptation is another strategy to produce more training data by permitting a model to train on a set of annotated data and apply it to a target set. The primary challenge is overcoming the domain shift between the annotated data and targets – where the cross-data prediction is biased due to a difference in the imaging quality of the two image sets, or domains. 3D-ADA[59] is an adversarial domain adaptation framework that overcomes this by mapping subtomograms into a latent space existing between the domains. This setup improves cross-data source prediction and even the determination of novel structures absent from the training data. Cryo-Shift[60] is a framework to address domain shifts using multi-adversarial domain adaptation. Simulated tomograms are used as training data and are transformed by Cryo-Shift through domain randomization. This randomization improves the generalization ability of the classifier to better predict in the case of the experimental data. Despite being fully unsupervised and not training on any labeled experimental samples, Cryo-Shift still outperforms other cross-domain subtomogram classification approaches.

### Accessibility

For pattern mining methods to be used extensively, they must be convenient for a wide variety of users, including those without substantial computational backgrounds or resources.

User convenience. While publicly releasing software code is now routine for reproducibility and to foster research developments, there are additional considerations that can increase their ease of use for researchers. For instance, including a graphic user interface attracts a broader audience to utilize the protocols, as in the case of IMOD and Dynamo.[10,29] Consolidating scripts into a single suite is convenient for a user to select analyses of interest. Similarly, ScipionTomo[117] is a framework that enables a user to select protocols from multiple suites into a single analysis pipeline. Computational speed-ups such as GPU acceleration and parallel processing should also be considered, especially for machine learning processes, which typically require GPUs to achieve a reasonable timeframe.[61] Finally, cloud support or web-based portals enable research labs without in-

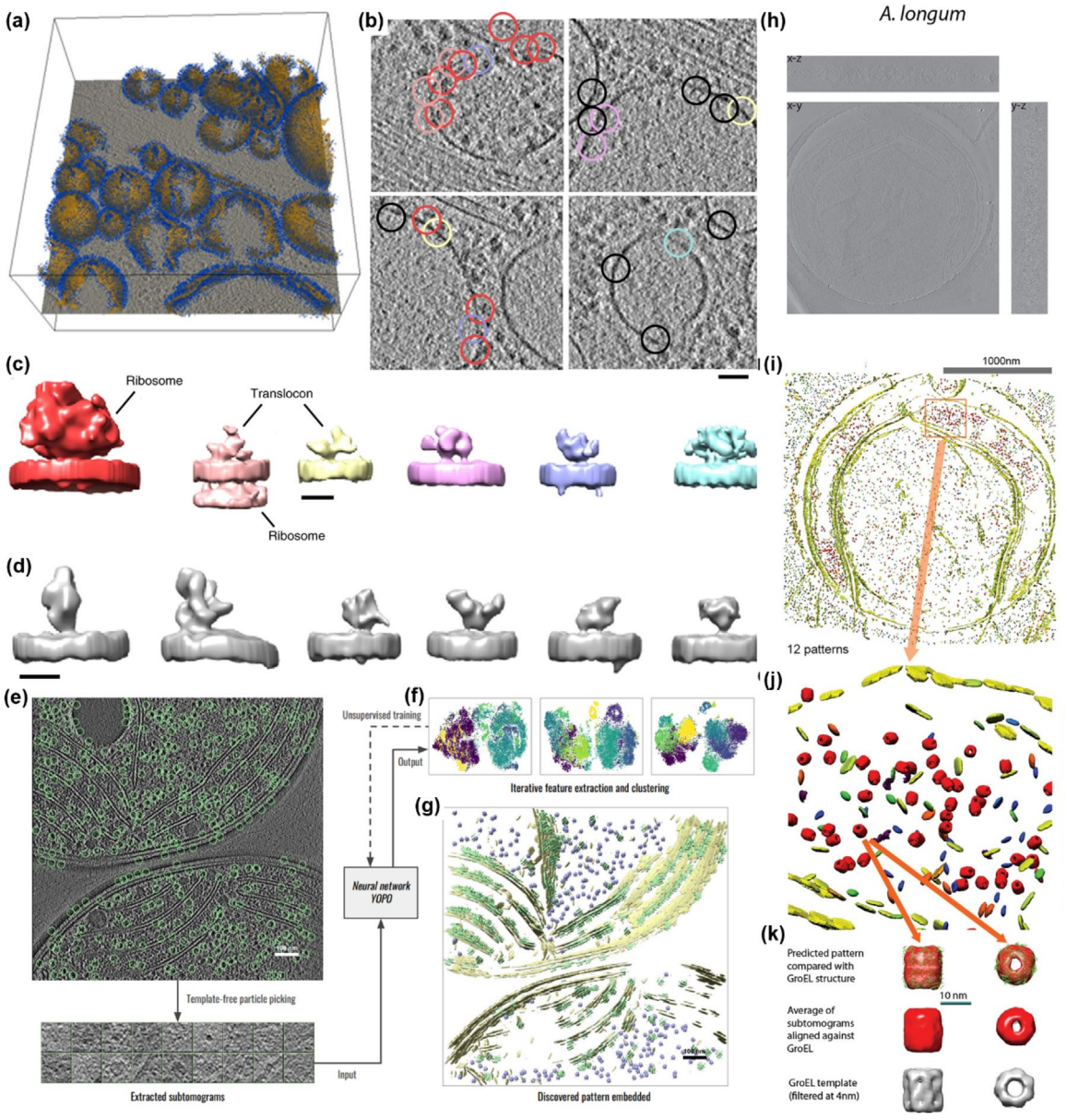



Figure 4. Template-free pattern mining approach examples. (a) A P19 cell tomogram slice superimposed with membranes (yellow) and particles (blue) identified by PySeg.[36] Tomogram slices (b) where circles encompass the x-y position of particles identified for classes shown in (c). Scale bar, 50 nm. (c) Average densities of identified particle classes. From left to right: membrane-associated ribosome (red), ribosome-associated translocon (light red), ribosome-free translocon (yellow), two putative PLC complexes (magenta and purple), putative IP3 receptor (blue). Scale bar, 10 nm. (d) Examples of additional class averages identified by PySeg. (e) A tomogram slice of intact Synechocystis cell with circles depicting particle picking results. (f) Unsupervised, iterative clustering of extracted features, plotted in 2D, performed by YOPO neural network.[51] (g) Isosurface rendering of discovered patterns. A tomogram slice (h) of intact Acetonema longum and (i) rendered structure models of patterns identified via multi-pattern pursuit (MPP). (j) Magnified region from (i) showing a close-up of the embedded structures. (k) A comparison of an acquired GroEL-like pattern compared to the known atomic model. The average of all subtomograms of this structure pattern aligned to the known structure of GroEL; template is shown.[52]

house hardware resources to utilize analysis software.[29,34,45]

Model size and efficiency considerations for tool design. Deep learning models such as CNNs are advantageous for many cryo-ET pattern mining tasks but also can require considerable amounts of state-of-the-art hardware to train and store, diminishing their accessibility.[118] Thus, it is ideal to reduce the computational load of software when

possible. To this end, Matuszewski et al.[118] aimed to reduce the size of CNN models, specifically the popular U-Net architecture[102] (utilized by the UMC, U-CLSTM, and MC DS Net supervised classifiers.[97–98] The study characterized the impact of a deep learning model's size on its accuracy. The resulting lighter U-Net featured a parameterized architecture with fewer trainable weights and achieved comparable accuracy in a particle recognition task with four times fewer weights but at the cost of more sensitive hyperparameters. Deeper architectures tended to feature better performance at the cost of longer training, more required training data, and susceptibility to overfitting. Thus, researchers must make practical design considerations to avoid making overly complex systems while also retaining some level of robustness.

### Performance benchmarks

Evaluating proposed frameworks is commonly achieved by comparing their performance against current standard approaches. However, researchers may apply different evaluation methods or standards, and the performance of methods against each other may be ambiguous. Thus, implementing benchmark studies that examine several models at once using the same data set and assessment serves to properly determine the performance of cutting-edge methods.

The previously noted SHREC track[96–98] is an annual study to evaluate learning-based localization and classification methods developed by various groups using the same set of validation data. Studies such as these not only introduce new quantitative cryo-ET analysis strategies, but also clarify their strengths, weaknesses, and promising directions compared to current state-of-the-art technologies. Benchmark studies should aim to assess new methods with ground-truth simulated data that closely reflects experimental data, well-validated experimental data, or both. Furthermore, factors such as runtime and user accessibility should also be important considerations in evaluating the strengths and weaknesses of each approach.

## Perspectives

In terms of studying protein structure and interactions, cryo-ET offers a large-scale perspective, presenting a unique opportunity to investigate the molecular sociology within a cell. To this end, computational pattern mining methods can achieve visual proteomics, mapping positions of complexes in tomograms to elucidate the spatial and population distributions of identified macromolecules.[78] Given that a tomogram is both information dense and costly to image and annotate, effective pattern mining is critical for high-throughput analysis. Ideal methods should perform segmentation, classification, and structure determination with great speed, spatial resolution, and structure identification accuracy.

Template-free, unsupervised approaches are especially promising, as they overcome the substantial annotation requirements that render supervised methods laborious and introduce a pathway towards de novo structure discovery as they forego a priori knowledge for complex identification. Aside from generating structure patterns, unsupervised classification can facilitate supervised methods. Given sets of patterns, a user only needs to label groups of subtomograms rather than each subtomogram individually. This could be exploited to generate training data for supervised classification with much less effort than fully manual approaches. The advancement of these methods will address a major bottleneck in the cryo-ET pipeline. Moreover, it will open new opportunities to revisit existing cryo-ET data and search for previously overlooked biomolecular interactions. Such a development would enable exciting discoveries of protein–protein interactions and structural conformations at a rapid rate, significantly accelerating research in numerous biological contexts.

However, to achieve this goal, it is not sufficient to simply develop highly effective algorithms. To ensure that these developments are widely disseminated, software packages should encourage implementation of cutting-edge cryo-ET analysis techniques in the broader research community. Cross-comparative studies of the latest methods should regularly assess the state of the field. As methods become more advanced, integrative modeling becomes essential, as pattern mining results should be validated or informed with other techniques such as quantitative or cross-linking mass spectrometry and computational structural predictions.[2,78,119–122] Advancements in tomogram processing for pattern mining will translate into a deeper and more complete understanding of countless cellular processes.

### CRediT authorship contribution statement

Hannah Hyun-Sook Kim: Conceptualization, Writing – original draft. Mostofa Rafid Uddin: Conceptualization, Writing – original draft, Writing – review & editing. Min Xu: Conceptualization, Writing – review & editing. Yi-Wei Chang: Conceptualization, Writing – review & editing, Supervision.

## Acknowledgements

YWC is supported by the David and Lucile Packard Fellowship for Science and Engineering

(2019-69645), Mark Foundation for Cancer Research ASPIRE Award (19-044-ASP), Pennsylvania Department of Health FY19 Health Research Formula Fund, NIH R01GM134020, RF1AG065341, R01AI127798 and RM1GM126511. HK is supported by the Biomedical Graduate Studies at University of Pennsylvania Perelman School of Medicine. MX and MRU are supported by NIH R01GM134020 and P41GM103712, NSF DBI-1949629, IIS-2007595, IIS-2211597 and MCB2205148, Mark Foundation for Cancer Research 19-044-ASP, and AMD COVID-19 HPC Fund.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Received 19 October 2022;
Accepted 26 March 2023;
Available online 31 March 2023

Keywords:
Cryo-ET;
image pattern recognition;
data mining;
machine learning;
cellular structural biology

## References

1. Oikonomou, C.M., Jensen, G.J., (2017). Cellular electron cryotomography: Toward structural biology in situ. Annu. Rev. Biochem. 86 https://doi.org/10.1146/annurev-biochem-061516-044741.
2. Turk, M., Baumeister, W., (2020). The promise and the challenges of cryo-electron tomography. FEBS Lett. 594, 3243–3261. https://doi.org/10.1002/1873-3468.13948.
3. Adrian, M., Dubochet, J., Lepault, J., McDowall, A.W., (1984). Cryo-electron microscopy of viruses. Nature 308 https://doi.org/10.1038/308032a0.
4. Robinson, C.V., Sali, A., Baumeister, W., (2007). The molecular sociology of the cell. Nature 450 https://doi.org/10.1038/nature06523.
5. Schur, F.K., (2019). Toward high-resolution in situ structural biology with cryo-electron tomography and subtomogram averaging. Curr. Opin. Struct. Biol. 58, 1–9. https://doi.org/10.1016/j.sbi.2019.03.018.
6. Koning, R.I., Koster, A.J., Sharp, T.H., (2018). Advances in cryo-electron tomography for biology and medicine. Ann. Anat. - Anat. Anz. 217, 82–96. https://doi.org/10.1016/j.aanat.2018.02.004.
7. Frangakis, A.S., (2021). It's noisy out there! A review of denoising techniques in cryo-electron tomography. J. Struct. Biol. 213, https://doi.org/10.1016/j.jsb.2021.107804 107804.
8. Lučić, V., Rigort, A., Baumeister, W., (2013). Cryo-electron tomography: The challenge of doing structural biology in situ. J. Cell Biol. 202 https://doi.org/10.1083/jcb.201304193.
9. Bartesaghi, A., Sprechmann, P., Liu, J., Randall, G., Sapiro, G., Subramaniam, S., (2008). Classification and 3D averaging with missing wedge correction in biological electron tomography. J. Struct. Biol. 162, 436–450. https://doi.org/10.1016/J.JSB.2008.02.008.
10. Kremer, J.R., Mastronarde, D.N., McIntosh, J.R., (1996). Computer visualization of three-dimensional Image data using IMOD. J. Struct. Biol. 116, 71–76. https://doi.org/10.1006/JSBI.1996.0013.
11. Nicastro, D., Schwartz, C., Pierson, J., Gaudette, R., Porter, M.E., McIntosh, J.R., (2006). The molecular architecture of axonemes revealed by cryoelectron tomography. Science (1979) 313 https://doi.org/10.1126/science.1128618.
12. Heumann, J.M., Hoenger, A., Mastronarde, D.N., (2011). Clustering and variance maps for cryo-electron tomography using wedge-masked differences. J. Struct. Biol. 175 https://doi.org/10.1016/j.jsb.2011.05.011.
13. Tang, G., Peng, L., Baldwin, P.R., Mann, D.S., Jiang, W., Rees, I., Ludtke, S.J., (2007). EMAN2: An extensible image processing suite for electron microscopy. J. Struct. Biol. 157, 38–46. https://doi.org/10.1016/J.JSB.2006.05.009.
14. Chen, M., Dai, W., Sun, S.Y., Jonasch, D., He, C.Y., Schmid, M.F., Chiu, W., Ludtke, S.J., (2017). Convolutional neural networks for automated annotation of cellular cryo-electron tomograms. Nat. Methods 14, 983–985. https://doi.org/10.1038/nmeth.4405.
15. Bell, J.M., Chen, M., Durmaz, T., Fluty, A.C., Ludtke, S.J., (2018). New software tools in EMAN2 inspired by EMDatabank map challenge. J. Struct. Biol. 204 https://doi.org/10.1016/j.jsb.2018.09.002.
16. Chen, M., Bell, J.M., Shi, X., Sun, S.Y., Wang, Z., Ludtke, S.J., (2019). A complete data processing workflow for cryo-ET and subtomogram averaging. Nat. Methods 16 https://doi.org/10.1038/s41592-019-0591-8.
17. Nickell, S., Förster, F., Linaroudis, A., del Net, W., Beck, F., Hegerl, R., Baumeister, W., Plitzko, J.M., (2005). TOM software toolbox: acquisition and analysis for electron tomography. J. Struct. Biol. 149, 227–234. https://doi.org/10.1016/J.JSB.2004.10.006.
18. Förster, F., Han, B.G., Beck, M., (2010). Visual proteomics. Methods Enzymol. 483, 215–243. https://doi.org/10.1016/S0076-6879(10)83011-3.
19. Hrabe, T., Chen, Y., Pfeffer, S., Kuhn Cuellar, L., Mangold, A.-V., Förster, F., (2012). PyTom: A python-based toolbox for localization of macromolecules in cryo-electron tomograms and subtomogram analysis. J. Struct. Biol. 178, 177–188. https://doi.org/10.1016/j.jsb.2011.12.003.
20. Chen, Y., Pfeffer, S., Hrabe, T., Schuller, J.M., Förster, F., (2013). Fast and accurate reference-free alignment of subtomograms. J. Struct. Biol. 182, 235–245. https://doi.org/10.1016/J.JSB.2013.03.002.
21. Taylor, K.A., Liu, J., Winkler, H., (2006). Localization and classification of repetitive structures in electron tomograms of paracrystalline assemblies. in: Electron Tomography: Methods for Three-Dimensional Visualization of Structures in the Cell. https://doi.org/10.1007/978-0-387-69008-7_16.
22. Winkler, H., (2007). 3D reconstruction and processing of volumetric data in cryo-electron tomography. J. Struct. Biol. 157 https://doi.org/10.1016/j.jsb.2006.07.014.

23. Winkler, H., Taylor, K.A., (1999). Multivariate statistical analysis of three-dimensional cross-bridge motifs in insect flight muscle. Ultramicroscopy 77 https://doi.org/10.1016/S0304-3991(99)00035-2.
24. Winkler, H., Zhu, P., Liu, J., Ye, F., Roux, K.H., Taylor, K. A., (2009). Tomographic subvolume alignment and subvolume classification applied to myosin V and SIV envelope spikes. J. Struct. Biol. 165 https://doi.org/10.1016/j.jsb.2008.10.004.
25. Wu, S., Liu, J., Reedy, M.C., Winkler, H., Reedy, M.K., Taylor, K.A., (2009). Methods for identifying and averaging variable molecular conformations in tomograms of actively contracting insect flight muscle. J. Struct. Biol. 168 https://doi.org/10.1016/j.jsb.2009.08.007.
26. Pruggnaller, S., Mayr, M., Frangakis, A.S., (2008). A visualization and segmentation toolbox for electron microscopy. J. Struct. Biol. 164, 161–165. https://doi.org/10.1016/j.jsb.2008.05.003.
27. Heymann, J.B., Belnap, D.M., (2007). Bsoft: Image processing and molecular modeling for electron microscopy. J. Struct. Biol. 157 https://doi.org/10.1016/j.jsb.2006.06.006.
28. Heymann, J.B., (2021). High resolution electron tomography and segmentation-by-modeling interpretation in Bsoft. Protein Sci. 30 https://doi.org/10.1002/pro.3938.
29. Castaño-Díez, D., Kudryashev, M., Arheit, M., Stahlberg, H., (2012). Dynamo: A flexible, user-friendly development tool for subtomogram averaging of cryo-EM data in high-performance computing environments. J. Struct. Biol. 178, 139–151. https://doi.org/10.1016/J.JSB.2011.12.017.
30. Castaño-Díez, D., Kudryashev, M., Stahlberg, H., (2017). Dynamo Catalogue: Geometrical tools and data management for particle picking in subtomogram averaging of cryo-electron tomograms. J. Struct. Biol. 197, 135–144. https://doi.org/10.1016/J.JSB.2016.06.005.
31. Castaño-Díez, D., (2017). The Dynamo package for tomography and subtomogram averaging: Components for MATLAB, GPU computing and EC2 Amazon Web Services. Acta Crystallogr. D Struct. Biol.. https://doi.org/10.1107/S2059798317003369.
32. Navarro, P.P., Stahlberg, H., Castaño-Díez, D., (2018). Protocols for subtomogram averaging of membrane proteins in the Dynamo software package. Front. Mol. Biosci. 5 https://doi.org/10.3389/fmolb.2018.00082.
33. Scaramuzza, S., Castaño-Díez, D., (2021). Step-by-step guide to efficient subtomogram averaging of virus-like particles with Dynamo. PLoS Biol. 19 https://doi.org/10.1371/journal.pbio.3001318.
34. Himes, B.A., Zhang, P., (2018). emClarity: software for high-resolution cryo-electron tomography and subtomogram averaging. Nat. Methods 15 https://doi.org/10.1038/s41592-018-0167-z.
35. Ni, T., Frosio, T., Mendonça, L., Sheng, Y., Clare, D., Himes, B.A., Zhang, P., (2022). High-resolution in situ structure determination by cryo-electron tomography and subtomogram averaging using emClarity. Nat. Protoc. 17 https://doi.org/10.1038/s41596-021-00648-5.
36. Martinez-Sanchez, A., Kochovski, Z., Laugks, U., Meyer zum Alten Borgloh, J., Chakraborty, S., Pfeffer, S., Baumeister, W., Lučić, V., (2020). Template-free detection and classification of membrane-bound complexes in cryo-electron tomograms. Nat. Methods 17, 209–216. https://doi.org/10.1038/s41592-019-0675-5.
37. Martinez-Sanchez, A., Garcia, I., Asano, S., Lucic, V., Fernandez, J.J., (2014). Robust membrane detection based on tensor voting for electron tomography. J. Struct. Biol. 186, 49–61. https://doi.org/10.1016/J.JSB.2014.02.015.
38. Martinez-Sanchez, A., Baumeister, W., Lučić, V., (2022). Statistical spatial analysis for cryo-electron tomography. Comput. Methods Programs Biomed. 218, https://doi.org/10.1016/J.CMPB.2022.106693 106693.
39. Scheres, S.H.W., (2012). RELION: Implementation of a bayesian approach to cryo-EM structure determination. J. Struct. Biol. 180 https://doi.org/10.1016/j.jsb.2012.09.006.
40. Scheres, S.H.W., (2012). A bayesian view on cryo-EM structure determination. J. Mol. Biol. 415 (2), 406–418. https://doi.org/10.1016/j.jmb.2011.11.010.
41. Bharat, T.A.M., Russo, C.J., Löwe, J., Passmore, L.A., Scheres, S.H.W., (2015). Advances in single-particle electron cryomicroscopy structure determination applied to sub-tomogram averaging. Structure 23 https://doi.org/10.1016/j.str.2015.06.026.
42. Bharat, T.A.M., Scheres, S.H.W., (2016). Resolving macromolecular structures from electron cryo-tomography data using subtomogram averaging in RELION. Nat. Protoc. 11, 2054–2065. https://doi.org/10.1038/nprot.2016.124.
43. Zivanov, J., Otón, J., Ke, Z., von Kügelgen, A., Pyle, E., Qu, K., Morado, D., Castaño-Díez, D., et al., (2022). A bayesian approach to single-particle electron cryo-tomography in RELION-4.0. Elife 11 https://doi.org/10.7554/elife.83724.
44. Luengo, I., Darrow, M.C., Spink, M.C., Sun, Y., Dai, W., He, C.Y., Chiu, W., Pridmore, T., et al., (2017). SuRVoS: Super-Region volume segmentation workbench. J. Struct. Biol. 198, 43–53. https://doi.org/10.1016/J.JSB.2017.02.007.
45. Frazier, Z., Xu, M., Alber, F., (2017). TomoMiner and TomoMinerCloud: A software platform for large-scale subtomogram structural analysis. Structure 25 https://doi.org/10.1016/j.str.2017.04.016.
46. Liu, C., Zeng, X., Wang, K., Guo, Q., Xu, M., (2018). Multi-task learning for macromolecule classification, segmentation and coarse structural recovery in cryo-tomography. http://arxiv.org/abs/1805.06332 (accessed May 30, 2022).
47. Hao, Y., Wan, X., Yan, R., Liu, Z., Li, J., Zhang, S., Cui, X., Zhang, F., (2022). VP-Detector: A 3D multi-scale dense convolutional neural network for macromolecule localization and classification in cryo-electron tomograms. Comput. Methods Programs Biomed. 221, https://doi.org/10.1016/J.CMPB.2022.106871 106871.
48. Moebel, E., Martinez-Sanchez, A., Lamm, L., Righetto, R. D., Wietrzynski, W., Albert, S., Lariviamp, D., Fourmentin, E., et al., (2021). Deep learning improves macromolecule identification in 3D cellular cryo-electron tomograms. Nat. Methods 18, 1386–1394. https://doi.org/10.1038/s41592-021-01275-4.
49. Hajarolasvadi, N., Sunkara, V., Khavnekar, S., Beck, F., Brandt, R., Baum, D., (2022). Volumetric macromolecule identification in cryo-electron tomograms using capsule networks. BMC Bioinf. 23, 360. https://doi.org/10.1186/s12859-022-04901-w.

50. Che, C., Lin, R., Zeng, X., Elmaaroufi, K., Galeotti, J., Xu, M., (2018). Improved deep learning-based macromolecules structure classification from electron cryo-tomograms. Mach. Vis. Appl. 29, 1227–1236. https://doi.org/10.1007/s00138-018-0949-4.
51. Zeng, X., Kahng, A., Xue, L., Mahamid, J., Chang, Y.-W., Xu, M., (2023). High-throughput cryo-ET structural pattern mining by unsupervised deep iterative subtomogram clustering. Proceedings of the National Academy of Sciences 120, https://doi.org/10.1073/pnas.2213149120 e2213149120.
52. Xu, M., Singla, J., Tocheva, E.I., Chang, Y.W., Stevens, R.C., Jensen, G.J., Alber, F., (2019). De novo structural pattern mining in cellular electron cryotomograms. Structure 27, 679–691.e14. https://doi.org/10.1016/J.STR.2019.01.005.
53. Zeng, X., Howe, G., Xu, M., (2021). End-to-end robust joint unsupervised image alignment and clustering. In: 2021 IEEE/CVF International Conference on Computer Vision (ICCV). IEEE, pp. 3834–3846. https://doi.org/10.1109/ICCV48922.2021.00383.
54. Zhao, Y., Zeng, X., Guo, Q., Xu, M., (2018). An integration of fast alignment and maximum-likelihood methods for electron subtomogram averaging and classification. Bioinformatics 34, i227–i236. https://doi.org/10.1093/bioinformatics/bty267.
55. Harastani, M., Eltsov, M., Leforestier, A., Jonic, S., (2021). HEMNMA-3D: Cryo electron tomography method based on normal mode analysis to study continuous conformational variability of macromolecular complexes. Front. Mol. Biosci. 8 https://doi.org/10.3389/fmolb.2021.663121.
56. Liu, S., Ban, X., Zeng, X., Zhao, F., Gao, Y., Wu, W., Zhang, H., Chen, F., et al., (2020). A unified framework for packing deformable and non-deformable subcellular structures in crowded cryo-electron tomogram simulation. BMC Bioinf. 21 https://doi.org/10.1186/s12859-020-03660-w.
57. Liu, S., Ma, Y., Ban, X., Zeng, X., Nallapareddy, V., Chaudhari, A., Xu, M., (2020). Efficient cryo-electron tomogram simulation of macromolecular crowding with application to SARS-CoV-2. In: Proceedings - 2020 IEEE International Conference on Bioinformatics and Biomedicine, BIBM 2020. Institute of Electrical and Electronics Engineers Inc., pp. 80–87. https://doi.org/10.1109/BIBM49941.2020.9313185.
58. Wu, X., Li, C., Zeng, X., Wei, H., Deng, H.-W., Zhang, J., Xu, M., (2022). CryoETGAN: Cryo-Electron tomography image synthesis via unpaired image translation. Front. Physiol. 1, 760404. https://doi.org/10.3389/fphys.2022.760404. Www.Frontiersin.Org.
59. Lin, R., Zeng, X., Kitani, K., Xu, M., (2019). Adversarial domain adaptation for cross data source macromolecule in situ structural classification in cellular electron cryo-tomograms. Bioinformatics 35, i260–i268. https://doi.org/10.1093/bioinformatics/btz364.
60. Bandyopadhyay, H., Deng, Z., Ding, L., Liu, S., Uddin, M. R., Zeng, X., Behpour, S., Xu, M., (2022). Cryo-shift: reducing domain shift in cryo-electron subtomograms with unsupervised domain adaptation and randomization. Bioinformatics 38, 977–984. https://doi.org/10.1093/bioinformatics/btab794.
61. Navarro, P.P., (2022). Quantitative cryo-electron tomography. Front. Mol. Biosci. 9 https://doi.org/10.3389/fmolb.2022.934465.
62. Martinez-Sanchez, A., Garcia, I., Fernandez, J.-J., (2011). A differential structure approach to membrane segmentation in electron tomography. J. Struct. Biol. 175, 372–383. https://doi.org/10.1016/j.jsb.2011.05.010.
63. Whitaker, R.T., Elangovan, V., (2002). A direct approach to estimating surfaces in tomographic data. Med. Image Anal. 6, 235–249. https://doi.org/10.1016/S1361-8415(02)00082-8.
64. Bartesaghi, A., Sapiro, G., Subramaniam, S., (2005). An energy-based three-dimensional segmentation approach for the quantitative interpretation of electron tomograms. IEEE Trans. Image Process. 14, 1314–1323. https://doi.org/10.1109/TIP.2005.852467.
65. Ress, D.B., Harlow, M.L., Marshall, R.M., McMahan, U.J., (2004). Methods for generating high-resolution structural models from electron microscope tomography data. Structure 12, 1763–1774. https://doi.org/10.1016/j.str.2004.07.022.
66. Volkmann, N., (2002). A novel three-dimensional variant of the watershed transform for segmentation of electron density maps. J. Struct. Biol. 138, 123–129. https://doi.org/10.1016/S1047-8477(02)00009-6.
67. Moussavi, F., Heitz, G., Amat, F., Comolli, L.R., Koller, D., Horowitz, M., (2010). 3D segmentation of cell boundaries from whole cell cryogenic electron tomography volumes. J. Struct. Biol. 170, 134–145. https://doi.org/10.1016/J.JSB.2009.12.015.
68. Nguyen, H., Ji, Q., (2008). Shape-Driven Three-Dimensional watersnake segmentation of biological membranes in electron tomography. IEEE Trans. Med. Imaging 27, 616–628. https://doi.org/10.1109/TMI.2007.912390.
69. Sun, S.Y., Kaelber, J.T., Chen, M., Dong, X., Nematbakhsh, Y., Shi, J., Dougherty, M., Lim, C.T., et al., (2018). Flagellum couples cell shape to motility in Trypanosoma brucei. PNAS 115 https://doi.org/10.1073/pnas.1722618115.
70. Huokko, T., Ni, T., Dykes, G.F., Simpson, D.M., Brownridge, P., Conradi, F.D., Beynon, R.J., Nixon, P.J., et al., (2021). Probing the biogenesis pathway and dynamics of thylakoid membranes. Nat. Commun. 12 https://doi.org/10.1038/s41467-021-23680-1.
71. Wan, W., Briggs, J.A.G., (2016). Cryo-Electron tomography and subtomogram averaging. In: Methods Enzymol. Academic Press, pp. 329–367. https://doi.org/10.1016/bs.mie.2016.04.014.
72. Galaz-Montoya, J.G., Flanagan, J., Schmid, M.F., Ludtke, S.J., (2015). Single particle tomography in EMAN2. J. Struct. Biol. 190, 279–290. https://doi.org/10.1016/J.JSB.2015.04.016.
73. Wang, F., Gong, H., Liu, G., Li, M., Yan, C., Xia, T., Li, X., Zeng, J., (2016). DeepPicker: A deep learning approach for fully automated particle picking in cryo-EM. J. Struct. Biol. 195, 325–336. https://doi.org/10.1016/J.JSB.2016.07.006.
74. Frangakis, A.S., Förster, F., (2004). Computational exploration of structural information from cryo-electron tomograms. Curr. Opin. Struct. Biol. 14, 325–331. https://doi.org/10.1016/j.sbi.2004.04.003.
75. Böhm, J., Frangakis, A.S., Hegerl, R., Nickell, S., Typke, D., Baumeister, W., (2000). Toward detecting and identifying macromolecules in a cellular context: Template matching applied to electron tomograms. PNAS 97 https://doi.org/10.1073/pnas.230282097.

76. Asano, S., Engel, B.D., Baumeister, W., (2016). In situ cryo-electron tomography: A post-reductionist approach to structural biology. J. Mol. Biol. 428, 332–343. https://doi.org/10.1016/j.jmb.2015.09.030.
77. Frangakis, A.S., Böhm, J., Förster, F., Nickell, S., Nicastro, D., Typke, D., Hegerl, R., Baumeister, W., (2002). Identification of macromolecular complexes in cryoelectron tomograms of phantom cells. PNAS 99 https://doi.org/10.1073/pnas.172520299.
78. Beck, M., Baumeister, W., (2016). Cryo-Electron Tomography: Can it reveal the molecular sociology of cells in atomic detail? Trends Cell Biol. 26, 825–837. https://doi.org/10.1016/J.TCB.2016.08.006.
79. Lucas, B.A., Himes, B.A., Xue, L., Grant, T., Mahamid, J., Grigorieff, N., (2021). Locating macromolecular assemblies in cells by 2D template matching with cisTEM. Elife 10 https://doi.org/10.7554/eLife.68946.
80. Rigort, A., Günther, D., Hegerl, R., Baum, D., Weber, B., Prohaska, S., Medalia, O., Baumeister, W., et al., (2012). Automated segmentation of electron tomograms for a quantitative description of actin filament networks. J. Struct. Biol. 177, 135–144. https://doi.org/10.1016/J.JSB.2011.08.012.
81. Rusu, M., Starosolski, Z., Wahle, M., Rigort, A., Wriggers, W., (2012). Automated tracing of filaments in 3D electron tomography reconstructions using sculptor and situs. https://doi.org/10.1016/j.jsb.2012.03.001.
82. Weber, B., Greenan, G., Prohaska, S., Baum, D., Hege, H.C., Müller-Reichert, T., Hyman, A.A., Verbavatz, J.M., (2012). Automated tracing of microtubules in electron tomograms of plastic embedded samples of caenorhabditis elegans embryos. J. Struct. Biol. 178, 129–138. https://doi.org/10.1016/J.JSB.2011.12.004.
83. Grünewald, K., Medalia, O., Gross, A., Steven, A.C., Baumeister, W., (2002). Prospects of electron cryotomography to visualize macromolecular complexes inside cellular compartments: Implications of crowding. Biophys. Chem. 100 https://doi.org/10.1016/S0301-4622(02)00307-1.
84. Shatsky, M., Hall, R.J., Brenner, S.E., Glaeser, R.M., (2009). A method for the alignment of heterogeneous macromolecules from electron microscopy. J. Struct. Biol. 166, 67–78. https://doi.org/10.1016/j.jsb.2008.12.008.
85. Chen, Y., Pfeffer, S., Fernández, J.J., Sorzano, C.O.S., Förster, F., (2014). Autofocused 3D classification of cryoelectron subtomograms. Structure 22, 1528–1537. https://doi.org/10.1016/J.STR.2014.08.007.
86. Henderson, R., (2013). Avoiding the pitfalls of single particle cryo-electron microscopy: Einstein from noise. PNAS 110 https://doi.org/10.1073/pnas.1314449110.
87. Chen, Y., Hrabe, T., Pfeffer, S., Pauly, O., Mateus, D., Navab, N., Förster, F., (2012). Detection and identification of macromolecular complexes in cryo-electron tomograms using support vector machines. 10.1109/ISBI.2012.6235823.
88. Wagner, T., Merino, F., Stabrin, M., Moriya, T., Antoni, C., Apelbaum, A., Hagel, P., Sitsel, O., et al., (2019). SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. Commun Biol. 2 https://doi.org/10.1038/s42003-019-0437-z.
89. Lučić, V., Fernández-Busnadiego, R., Laugks, U., Baumeister, W., (2016). Hierarchical detection and analysis of macromolecular complexes in cryo-electron tomograms using Pyto software. J. Struct. Biol. 196, 503–514. https://doi.org/10.1016/J.JSB.2016.10.004.
90. Kovacs, J.A., Wriggers, W., (2002). Fast rotational matching. Acta Crystallogr. D Biol. Crystallogr. 58 https://doi.org/10.1107/S0907444902009794.
91. Nickell, S., Kofler, C., Leis, A.P., Baumeister, W., (2006). A visual approach to proteomics. Nat. Rev. Mol. Cell Biol. 7, 225–230. https://doi.org/10.1038/nrm1861.
92. Scheres, S.H.W., Melero, R., Valle, M., Carazo, J.M., (2009). Averaging of electron subtomograms and random conical tilt reconstructions through likelihood optimization. Structure 17, 1563–1572. https://doi.org/10.1016/J.STR.2009.10.009.
93. Dempster, A.P., Laird, N.M., Rubin, D.B., (1977). Maximum likelihood from Incomplete data Via the EM algorithm. J. Roy. Stat. Soc.: Ser. B (Methodol.) 39 https://doi.org/10.1111/j.2517-6161.1977.tb01600.x.
94. Scheres, S.H.W., Gao, H., Valle, M., Herman, G.T., Eggermont, P.P.B., Frank, J., Carazo, J.M., (2007). Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. Nat. Methods 4 https://doi.org/10.1038/nmeth992.
95. Castaño-Díez, D., Zanetti, G., (2019). In situ structure determination by subtomogram averaging. Curr. Opin. Struct. Biol. 58, 68–75. https://doi.org/10.1016/J.SBI.2019.05.011.
96. Gubins, I., van der Schot, G., Veltkamp, R.C., Förster, F., Du, X., Zeng, X., Zhu, Z., Chang, L., et al., (2019). SHREC'19 track: Classification in cryo-electron tomograms. Eurographics Workshop on 3D Object Retrieval. https://doi.org/10.2312/3dor.20191061.
97. Gubins, I., Chaillet, M.L., van der Schot, G., Veltkamp, R.C., Förster, F., Hao, Y., Wan, X., Cui, X., et al., (2020). Classification in cryo-electron tomograms. Comput. Graph. 91, 279–289. https://doi.org/10.1016/J.CAG.2020.07.010.
98. Gubins, I., Chaillet, M.L., van der Schot, G., Cristina Trueba, M., Veltkamp, R.C., Förster, F., Wang, X., Kihara, D., et al., (2021). SHREC 2021: Classification in cryo-electron tomograms, Eurographics Workshop on 3D Object Retrieval. 10.2312/3dor.20211307.
99. He, K., Zhang, X., Ren, S., Sun, J., (2016). Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778. https://doi.org/10.1109/CVPR.2016.90.
100. Hara, K., Kataoka, H., Satoh, Y., (2017). Learning Spatio-Temporal features with 3D residual networks for action recognition. In: 2017 IEEE International Conference on Computer Vision Workshops (ICCVW). IEEE, pp. 3154–3160. https://doi.org/10.1109/ICCVW.2017.373.
101. Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L., (2017). Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. IEEE Trans. Image Process. 26, 3142–3155. https://doi.org/10.1109/TIP.2017.2662206.
102. Ronneberger, O., Fischer, P., Brox, T., (2015). U-net: Convolutional networks for biomedical image segmentation. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). https://doi.org/10.1007/978-3-319-24574-4_28.
103. Sudre, C.H., Li, W., Vercauteren, T., Ourselin, S., Jorge Cardoso, M., (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced

segmentations. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 240–248. https://doi.org/10.1007/978-3-319-67558-9_28.
104. Dai, J., He, K., Sun, J., (2016). Instance-Aware semantic segmentation via multi-task network cascades. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, pp. 3150–3158. https://doi.org/10.1109/CVPR.2016.343.
105. Salehi, S.S.M., Erdogmus, D., Gholipour, A., (2017). Tversky loss function for image segmentation using 3D fully convolutional deep networks. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 379–387. https://doi.org/10.1007/978-3-319-67389-9_44.
106. Huang, H., Lin, L., Tong, R., Hu, H., Zhang, Q., Iwamoto, Y., Han, X., Chen, Y.-W., Wu, J., (2020). UNet 3+: A full-scale connected UNet for medical image segmentation. In: ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 1055–1059. https://doi.org/10.1109/ICASSP40776.2020.9053405.
107. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J., (2018). UNet++: A Nested U-Net Architecture for medical image segmentation. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 3–11. https://doi.org/10.1007/978-3-030-00889-5_1.
108. Milletari, F., Rieke, N., Baust, M., Esposito, M., Navab, N., (2018). CFCM: Segmentation via coarse to fine context memory. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 667–674. https://doi.org/10.1007/978-3-030-00937-3_76.
109. Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W., Woo, W., (2015). Convolutional LSTM Network: A machine learning approach for precipitation nowcasting. Adv. Neural. Inf. Process. Syst. 2015-January. https://doi.org/10.48550/arxiv.1506.04214.
110. Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., (2018). DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. 40, 834–848. https://doi.org/10.1109/TPAMI.2017.2699184.
111. Nguyen, N.P., Ersoy, I., Gotberg, J., Bunyak, F., White, T. A., (2021). DRPnet: automated particle picking in cryo-electron micrographs using deep regression. BMC Bioinf. 22, 55. https://doi.org/10.1186/s12859-020-03948-x.
112. Gubins, I., Veltkamp, R.C., (2020). Deeply cascaded U-Net for Multi-Task image processing. http://arxiv.org/abs/2005.00225 (accessed February 14, 2023).
113. He, K., Zhang, X., Ren, S., Sun, J., (2016). Identity mappings in deep residual networks. In: Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 630–645. https://doi.org/10.1007/978-3-319-46493-0_38.
114. Yu, F., Koltun, V., (2015). Multi-Scale context aggregation by dilated convolutions. In: 4th International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings. http://arxiv.org/abs/1511.07122.
115. Lin, T.-Y., Goyal, P., Girshick, R., He, K., Dollar, P., (2020). Focal loss for dense object detection. IEEE Trans. Pattern Anal. Mach. Intell. 42, 318–327. https://doi.org/10.1109/TPAMI.2018.2858826.
116. Förster, F., Pruggnaller, S., Seybert, A., Frangakis, A.S., (2008). Classification of cryo-electron sub-tomograms using constrained correlation. J. Struct. Biol. 161, 276–286. https://doi.org/10.1016/J.JSB.2007.07.006.
117. Jiménez de la Morena, J., Conesa, P., Fonseca, Y.C., de Isidro-Gómez, F.P., Herreros, D., Fernández-Giménez, E., Strelak, D., Moebel, E., et al., (2022). ScipionTomo: Towards cryo-electron tomography software integration, reproducibility, and validation. J. Struct. Biol. 214, https://doi.org/10.1016/J.JSB.2022.107872 107872.
118. Matuszewski, D.J., Sintorn, I.M., (2019). Reducing the U-Net size for practical scenarios: Virus recognition in electron microscopy images. Comput. Methods Programs Biomed. 178, 31–39. https://doi.org/10.1016/J.CMPB.2019.05.026.
119. Baek, M., DiMaio, F., Anishchenko, I., Dauparas, J., Ovchinnikov, S., Lee, G.R., Wang, J., Cong, Q., et al., (2021). Accurate prediction of protein structures and interactions using a three-track neural network. Science 373, 5–6. https://doi.org/10.1126/science.abj8754.
120. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., et al., (2021). Highly accurate protein structure prediction with AlphaFold. Nature 596, 583. https://doi.org/10.1038/s41586-021-03819-2.
121. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., Steinegger, M., (2022). ColabFold: making protein folding accessible to all. Nat. Methods 19 https://doi.org/10.1038/s41592-022-01488-1.
122. Rantos, V., Karius, K., Kosinski, J., (2022). Integrative structural modeling of macromolecular complexes using Assembline. Nat. Protoc. 17 https://doi.org/10.1038/s41596-021-00640-z.