Automated Particle Picking in Cryo-Electron Micrographs using Deep Regression

Nguyen P. Nguyen  
Dept. of Electrical Engineering and Computer Science  
University of Missouri  
Columbia, MO, USA  
npntz3@mail.missouri.edu

Ilker Ersoy  
MU Informatics Institute  
University of Missouri  
Columbia, MO, USA  
ersoyi@health.missouri.edu

Tommi White  
Dept. of Biochemistry and Electron Microscopy Core  
University of Missouri  
Columbia, MO, USA  
whiteto@missouri.edu

Filiz Bunyak*  
Dept. of Electrical Engineering and Computer Science  
University of Missouri  
Columbia, MO, USA  
bunyak@missouri.edu

Abstract—Selection of good particles in cryo-electron micrographs is an important step in the reconstruction of high resolution 3D structures. In this study, we constructed a deep learning-based method to automatically detect particle centers from micrographs. This is a challenging task because of the low signal-to-noise ratio of cryo-EM micrographs, and the size, shape, and grayscale-level differences of particles. We proposed a Fully Convolutional Regression Network (FCRN) that maps the particle image to a continuous distance map that acts like a probability density function of particle centers. This approach is simple, but very effective in recognizing different grayscale patterns corresponding to 2D views of 3D particles. Our experimental results on dataset β-galactosidase (EMPIAR-10017) [1] showed that FCRN outperformed Faster-RCNN, Apple picker, and RELION’s particle picker. Compared to the ground truth of this dataset, FCRN achieved better picking performance, and 3D structure of those picked particles also had higher resolution.

Index Terms—convolutional neural networks, regression, deep learning, cryo electron microscopy, image segmentation, particle picking

I. INTRODUCTION

Although protein structure determination via single particle analysis using electron microscopy is widely used, it is still a challenging technique because the resulting micrograph is noisy, with low contrast intensity due to the radiation limitations in the applied electron beam exposure, particle images have a noisy background and low contrast intensity [2]. Thus, the proteins of interest, "particles", especially certain orientations of the particles, are sometimes challenging to identify in these 2D cryoEM micrographs. Several software solutions have been developed to address this problem. These solutions require a large number of particles to estimate accurately the relative angular orientations of these protein particles in 3D, which is used to create 3D reconstructions of averaged protein particle structures. One solution, RELION [1] uses the Bayesian approach and expectation maximization algorithm to solve for protein structures. RELION’s workflow allows users to identify and select particles in 2D cryoEM micrographs manually ("picking"); then aligns, classifies and averages those particles to make 2D class averages. Suitable 2D class averages are manually selected as templates, which are then used to perform automated particle selection ("autopicking") by assessing particle’s correlation to the template. EMAN2 [3], [4] is another popular solution to create 3D reconstructions from cryoEM micrographs. EMAN offers users three methods to pick particles from manual samples: local search, region search, and a convoulutional neural network for pattern recognition based on THEANO library [5]. One recent software called cryoSPARC [6] has a very promising approach, searching for a globally optimal model from the regularized likelihood function instead of using an expectation maximization algorithm.

Although these software solutions have different methods for the particle picking process, the common point is that a ground truth (a set of good particles selected by experts from the noisy cryo-EM micrographs) needs to exist. This manual picking task is time consuming, and often an expert is not available. After obtaining good 2D class average templates, these software packages rely on two main methods to auto-select particles: template matching by cross-correlation and pattern recognition by a simple deep neural network. Template matching is sensitive to noise, and it may cause strong bias [1], whereas neural networks are less sensitive and usually perform better. Unfortunately, this simple deep neural network is not robust enough to handle all the complicated noise patterns nor the illumination/transmission variations in cryoEM micrographs. Also, necessary particle orientations cannot be usually obtained with the necessary stringent constraints (such as high picking threshold) due to the low contrast nature of cryoEM data. If we implement low picking thresholds, it is likely that these programs pick not only particles, but also noise particles known as false positives [7]. There are trade-offs to have picking accuracy or false positive avoidance.

Advances in machine learning, specifically deep learning, hold much promise to improve and automate biomedical image analysis [8]–[10]. In microscopic image analysis, [11] described many machine learning approaches from support vector machines (SVM) to convolutional neural network (CNN) for segmentation.

A recent study of Shin et al. [12] has employed deep learning models to learn semantics in MRI scans, and extract

*Corresponding Author
features to detect different organs. Xie [13] proposed a novel deep neural network for robust nucleus localization, where instead of using a pixel-wise classifier or a regressor, they combined CNN with a nonlinear voting transformation. Inspired by the success of these examples, we sought to apply these deep learning methods to cryo-EM image analysis, specifically particle picking, which is the most tedious step in most cryoEM image analysis workflows.

In this paper, we introduce a fully convolutional regression network (FCRN) to predict the probability map of particle centers and then classify those candidates to pick particles. The paper is organized as follows. In Section 2, we review recent advanced methods to detect particles in microscopy images. In Section 3, we explain all the methods we used in this study and how FCRN works. In the final section, we present our experiment results and discuss their related issues.

II. RELATED WORK

DeepPicker [14] is a fully automated particle picking program based on deep neural networks. Before DeepPicker, other packages such as RELION and EMAN2 used an algorithm to collect particles with semi-automated workflows. DeepPicker slides a window per micrograph with the same dimensions as a specified particle (box) with a default step size of four pixels to collect candidate patches. All image patches are normalized and fed into a convolutional neural network to determine whether they belong to qualified particles or not. In the fully automated mode, DeepPicker has a pretrained network to pick particles for the first time as positive training samples. The negative samples are then used to train a new convolutional network to pick the final particles.

However, in our experiment with β-galactosidase dataset (EMPIAR-10017), the pretrained mode did not obtain satisfactory results. Next, we worked with DeepPickers training mode. In training mode, DeepPicker can adopt particles from RELION (using a manually picked .star file) to train the DeepPicker network. This second approach provided suitable results with higher particle picking accuracy and better capture of different particle orientations. In summary, DeepPicker worked quite well in the semi-automated mode. In fully automated mode, because of the simple network configuration with four layers, DeepPicker was not powerful enough to acquire appropriate results with low contrast particles in our hands (EMPIAR-10061).

Zhu et-al. [7] introduced a convolutional neural network (CNN) composed of eight layers called DeepEM to pick particles. DeepEM works with special training and a prediction procedure including data augmentation. The experiment was performed on a keyhole limpet haemocyanin (KLH) dataset and showed improved detection performance avoiding unwanted particles and non-particles. This framework, however, was only tested on one dataset (KLH collected with a Tecnai F30 Twin on a Gatan Ultrascan 4000 CCD camera). If one required a fully automated detection framework, this algorithms complicated training process needs to be more versatile to adapt to a variety of differently shaped particles to create a suitable pretrained network.

Xiao and Yang [15] employed Fast-RCNN to select particles from micrographs. Instead of using a default region proposal function for Fast-RCNN, they employed a sliding window to provide candidate regions for the detection CNN. This approach is equivalent to a normal CNN used in DeepPicker. The strength of this method is that the network was trained with three classes to discriminate false positive samples.

Heimowitz et al. [16] introduced a new picker called APPLE based on cross-correlation. It automatically selects reference windows to compare, and eliminate the need to prepare manual templates. Compared to RELIONs manual template generation, APPLE picker worked faster, accurately, and automatically. However, APPLE picker did not work well with high concentration, hollow cylindrical particles (KLH, GroEL).

crYOLO [17] utilizes the first generation YOLO network [18] to detect particles. YOLO is based on a fixed-grid regression method. Compared to other region-based methods, this fixed-grid helps YOLO to work faster in real-time [19]. However, it has a difficulty to deal with small objects or objects with unusual aspect ratios [19]. Also, YOLO has high localization error because of its regression and classification approach [19]. crYOLO performed detection very fast, up to six micrographs per second with image dimensions of 1024x1024 pixels [17]. crYOLO attained higher recall, precision and AUC (Area Under Curve) compared to the original YOLO model when working with small objects as particles in those test datasets. The experiments do not mention how this architecture performs on particles of other sizes and aspect ratios.

It is clear that a robust deep-learning tool is needed for cryoEM particle picking. We present an approach to develop a deep learning algorithm for particle picking that would work on multiple types of data (vitreous ice and negative stain), with various particle size and aspect ratios, and collected from different detectors (direct detection and CCD) with different defocus ranges. Below, we present our work on Deep Regression particle picking (fully convolutional regression network – FCRN). This method works on multiple types of data (tested on four protein datasets), from various detectors, with improved speeds (30s/micrograph) compared to manual template selection (30 min to select 1000 particles) required by RELION. The resulting outputs of Deep Regression were compared with the outputs of the RELION autopicking and produced final maps with similar resolution.

III. METHODS

A. Particle picking pipeline

Particle picking involves localization (detection) and classification processes. In classical computer vision, object detection and classification processes often rely on carefully hand-crafted image features and descriptors such as HOG (Hist-
When using different anchor sizes to generate Faster-RCNN model [25] to pick particles from electron (support-film edges and surface ice contamination). Scale appearance, low contrast and contain noise and artifacts objects in these datasets, electron micrographs have a gray-color, texture, and shape features. Unlike color images and (RGB, HSV, CYMK) images and larger objects with rich these models were originally designed to work with color performance in detection and classification tasks on image classes. Faster-RCNN and Mask-RCNN have shown great that classify each candidate as foreground versus background [26] first produce a sparse set of candidates (region proposals) stage detectors such as FasterRCNN [25] and Mask-RCNN can be coarsely grouped as single-stage and two-stage. Single-stage detectors such as Yolov3 [23] and SSD [24] rely on regular dense sampling of objects, scales, and aspect ratios, and perform detection and classification in one step. Two-stage detectors such as FasterRCNN [25] and Mask-RCNN [26] produce a sparse set of candidates (region proposals) that classify each candidate as foreground versus background classes. Faster-RCNN and Mask-RCNN have shown great performance in detection and classification tasks on image datasets such as Pascal VOC and MS COCO. However, these models were originally designed to work with color (RGB, HSV, CYMK) images and larger objects with rich color, texture, and shape features. Unlike color images and objects in these datasets, electron micrographs have a gray-scale appearance, low contrast and contain noise and artifacts (support-film edges and surface ice contamination).

Initially, we chose to experiment with the highly accurate Faster-RCNN model [25] to pick particles from electron micrographs. When using different anchor sizes to generate region proposals, the detector and classifier could be confused about target objects. Figure 1 shows that, indeed, Faster-RCNN could not detect particles correctly since region proposals had different box sizes, including boxes bigger than the particle size. We had to maintain an anchor scale to fit the proposed regions with particle size to remedy this issue. However, with fixed anchor scales, Faster-RCNN can only detect certain particles; we noted the detected particles usually had similar dimensions to particles used for training, and thus retraining Faster-RCNN model is necessary if the particle of interest has different dimensions than that of the training data. This approach is not efficient for an automatic particle picker.

A more versatile network than Faster-RCNN’s Region Proposal Network (RPN) [25] is needed to perform automated particle detection in cryoEM micrographs. In order to overcome the problem with RPN, we propose the Fully Convolutional Regression Network (FCRN), which is based on the blob detection concept. After the initial training with one dataset, FCRN can work with multiple particles having various sizes and shapes from different datasets without need for retraining.

Our particle picking pipeline (Figure 2) also includes a micrograph preprocessing step to enhance contrast and to correct transmission/illumination artifacts in order to ensure optimal performance by both networks. Picked particles are then fed to RELION 2.1 software for 3D reconstruction. Processing steps involved in the proposed automatic particle picking pipeline is described in the following subsections.

B. Fully Convolutional Regression Network (FCRN)

To address challenges of automated particle picking, we have treated the problem as a blob detection problem. Additionally, dimensions and shapes of particles are relatively similar. With a rough size estimate, the problem becomes the detection of particle centers. For the purposes of this study, segmentation of particle regions from image background is unnecessary. Our proposed model is the Fully Convolutional Regression Network (FCRN), allowing a prediction of the particle centers by producing a probability map of center locations.

From the ground truth as provided by the chosen datasets in EMPIAR, we create binary masks for particle-containing electron micrographs. Distance transforms of these binary masks are the target ground truth for our FCRN model. As described above, distance transform represents the probability distribution function of the particle centers. Starting from the boundaries of the mask, the distance grows towards the particle center. Background pixels are set to zero. In other words, distance transform is a proxy for the probability of particles center represented by continuous values. Therefore, continuous regression values are the output of our FCRN model, instead of discrete class labels. Figure 3 shows a sample image region and the corresponding ground truth distance transform used for training. Using a continuous map also makes it possible to detect partial particles if they overlap. It can work with multiple sizes of input micrographs and particles without the
additional requirement to retrain the model, a feature missing in the existing algorithms reported to date.

The FCRN network has seven layers including input layer, five convolutional layers, and one max pooling layer. With a simple structure like this, our approach has smaller computational cost than a complex model such as Faster-RCNN, or a correlation method with a sliding window scanning over whole micrographs to compare with particle templates.

### TABLE I: FCRN Architecture

| Layer | Type       | Filter | Dimensions |
|-------|------------|--------|------------|
| 1     | Convolutional | 32     | 9x9x1      |
| 2     | Convolutional | 32     | 3x3x32     |
| 3     | Convolutional | 32     | 1x7x32     |
| 4     | Convolutional | 32     | 7x1x32     |
| 5     | Max pooling | 1      | 2x2        |
| 6     | Convolutional | 1      | 3x3x32     |

Our implementation of FCRN utilizes Matconvnet [27] deep learning framework. To train our network, we first cropped input images around particle regions and cropped their corresponding ground truth maps as patches having the same dimensions with FCRN input layer. The number of patches from each image is equal to the number of particles in those images. We then trained FCRN model to let convolutional filters learn weights for mapping input images to distance transform maps. Since our network does not have fully connected layers, it does not require fixed sizes of input images. In the testing stage, we applied this FCRN model directly to input images which can have different sizes to predict probability maps of particle centers. Table I shows the network architecture of the FCRN. Figure 4 shows samples of input images and their network outputs as probability maps of particle centers.

### C. Performance Metrics

We applied two types of metrics to measure our FCRN-based particle picking performance. If the test set has full ground truth, we compute recall, precision, and F-measure to see how accurately our model works. As a second form
of performance metric, for all of the datasets, we construct 3D structures and measure the resolution of those structures. Additionally, we analyzed angular distributions of those structures and the efficiency $E_{od}$ [28] to understand the effect of orientations from those particles picked by FCRN and RELION for the final resolution.

1) Recall, Precision, and F-measure: The overall detection performance of all particles are evaluated in three common metrics: recall, precision, and F-measure. We collected the detection results in confusion matrix which has the columns indicating the ground truth, and the rows indicating prediction results.

\[
Recall = \frac{TP}{TP + FN} \quad Precision = \frac{TP}{TP + FP} \\
F = \frac{2 \times (Recall \times Precision)}{Recall + Precision}
\]

2) Resolution: The second metric we apply to evaluate performance is the resolution of 3D reconstruction models built from the picked particles. However, the optical resolution is not applicable in this case. Instead of that, we rely on the concept of signal to noise ratio, and measure the internal consistency of each dataset containing picked particles. We take advantage of some characteristics of Fourier transform to apply to our data: correlation between to volumes of 3d models can be represented by a product in Fourier space, and then we split that product into shells by radial frequency to compute Fourier Shell Correlation (FSC) [29]–[31] as below.

\[
FSC(k, \Delta k) = \frac{\text{Re}\left\{ \sum_{(k, \Delta k)} F_1(K)F_2(K) \right\}}{\sqrt{\sum_{(k, \Delta k)} |F_1(K)|^2|F_2(K)|^2}}^{1/2}
\]

where $K$ is spatial frequency vector, $k = |K|$ is the size of spatial frequency, $\Delta K$ is the ring width or shell thickness, and $F_1(K), F_2(K)$ are Fourier transforms of two half set reconstruction. As a result, we can construct a plot of Fourier Shell Correlation vs. frequency. The relationship between FSC and spectral signal to noise ratio (SSNR) is explained by [32] as following:

\[
SSNR = \frac{FSC}{1 - FSC}, \quad FSC = \frac{SSNR}{SSNR + 1}
\]

FSC is biased estimate of SSNR. For large number of images, variance(SSNR) is equal to variance(FSC), and the bias is negligible. When FSC is calculated for a data set split into halves:

\[
SSNR = \frac{2FSC}{1 - FSC}
\]

We use FSC to measure the internal consistency of our data. For a specific level of FSC, we can derive an equivalent frequency. That frequency has a unit of 1/A which is the inverse of the resolution we need to measure. In order to divide our data into two volumes to calculate FSC, we can select odd particles for the first volume, and the even particles for the second volume after we construct the 3D model. This is a simple way of selection, but the two halves of our data are not independent. Another way is to split particles from the beginning and build two 3D models. This method is often referred as Gold standard Fourier Shell Correlation. RELION uses this standard to compute resolution. After computing FSC, we also need a cut-off threshold to identify the resolution for our model. A reasonable criterion is to include only Fourier information that is above the noise level. Typical cut-off level for Gold standard split is 0.143 [31], [33], and for odd-even split, it is 0.5 [31], [34].

IV. EXPERIMENTAL RESULTS

Taking cryoEM datasets available in EMPIAR or locally generated negative stain micrographs, we applied both FCRN and Faster-RCNN to generate datasets with roughly similar numbers of particles. Both particles of FCRN and Faster-RCNN were then processed through RELIONs 2D classification to identify good 2D class averages [35]. After that,
FCRN and RELION-autopicked 2D class average templates were further classified in 3D (5 possible classes) using the corresponding deposited, low-resolution (60Å) filtered 3D reconstruction as the initial model. All 3D classes were refined to generate the final 3D reconstruction, and the gold standard resolution FSC curves are computed.

We selected 10 micrographs from the EMPIAR-10017 dataset (β-galactosidase) to train our FCRN model and Faster-RCNN. Scheres [1] collected this dataset using FEI Falcon II camera and a pixel size of 1.77Å. EMPIAR-10017 contains 84 micrographs and 40,863 manually selected particles. The micrographs were used to test and compare precision, recall, and F-measure performance between FCRN and Faster-RCNN.

With a provided ground truth (particles picked by an expert), we computed detection accuracy for this dataset. Table II shows that our recall has a success rate on particles of 78.6%, a precision rate of 70.8%, and F-measure of 74.5%. Faster-RCNN achieved a recall of 77.6%, precision of 70.0%, and F-measure of 73.6%. Figure 5 shows sample detection results from FCRN (left) and Faster-RCNN (right). Faster-RCNN picked more low-quality particles near the boundaries and missed overlapped particles.

In order to compare the reconstruction resolution performance of our FCRN, we used the whole dataset instead of testing set only. With the total number of picked particles around 44,000 particles for FCRN and Faster-RCNN, 2D classification was performed to select 2D class averages. According to visual inspection, we kept good templates after 2D and 3D classification to build the 3D structure. The percentages of particles kept after 2D and 3D classification correspond to their good 2D class average, given in Table III. FCRN obtained 85% after 2D classification. The final number of particles kept from FCRN is 37,566. Faster-RCNN also obtained 86% after 2D classification. The final number of particles kept from Faster-RCNN is 38,148. Number of particles and the percentage of kept particles as reported by APPLE picker [16] and RELION [1] are also given in Table III. This experiment showed that the 3D structure obtained from FCRN particles had a resolution of 4.1Å, while the structure of Faster-RCNN particles had a resolution of 4.4Å. Apple picker [16] and RELION [1] had resolutions of 4.5Å and 4.2Å, respectively. Figure 6 shows the reconstruction results for our FCRN (left) and Faster-RCNN (right). First row shows the 2D average templates, second row shows the reconstruction results.

**TABLE II: Recall, Precision and F-measure of EMPIAR-10017 (β-galactosidase) dataset.**

| Method             | # Particles | Recall | Precision | F-measure |
|--------------------|-------------|--------|-----------|-----------|
| FCRN               | 44,136      | 78.6%  | 70.8%     | 74.5%     |
| Faster-RCNN        | 44,368      | 77.8%  | 70.0%     | 73.6%     |

V. CONCLUSION

We proposed an automated particle picking system based on Fully Convolutional Regression Network (FCRN). We have successfully demonstrated its performance on EMPIAR-10017 dataset (β-galactosidase), and compared it to Faster-RCNN, APPLE picker, and RELION’s autopick. FCRN could work better with different shapes and orientations of particles in micrographs. As a result, FCRN attained higher particle detection recall and precision, and higher 3D structure resolution. For the future work, we will develop an unsupervised method to cluster good and bad particles after detection to further improve reconstruction resolution. Quality of the 3D reconstruction relies on various factors such as orientation distribution of the detected particles. We will incorporate these factors in analyzing particles picked by FCRN to improve resolution and compare our results to RELION and other software.

**REFERENCES**

[1] S. H. Scheres, “Semi-automated selection of cryo-em particles in relion-1.3 (dataset empiar 10017),” *J Struct Biol*, vol. 189, no. 2, pp. 114–22, 2015. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/25486611

[2] R. Norousi, S. Wickles, C. Leidig, T. Becker, V. J. Schmid, R. Beckmann, and A. Tresch, “Automatic post-picking using mappers improves particle image detection from cryo-em micrographs,” *J Struct Biol*, vol. 182, no. 2, pp. 59–66, 2013. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/23454482

[3] J. M. Bell, M. Chen, P. R. Baldwin, and S. J. Ludkte, “High resolution single particle refinement in eman2.1,” *Methods*, vol. 100, pp. 25–34, 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/26931650

[4] S. J. Ludkte, “Single-particle refinement and variability analysis in eman2.1,” Methods Enzymol., vol. 579, pp. 159–89, 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/27577277

[5] Theano Development Team, “Theano: A Python framework for fast computation of mathematical expressions,” *arXiv e-prints*, vol. abs/1605.02688, May 2016. [Online]. Available: http://arxiv.org/abs/1605.02688

[6] A. Punjani, J. L. Rubinstein, D. J. Fleet, and M. A. Brubaker, “cryosparc: algorithms for rapid unsupervised cryo-em structure determination,” *Nat Methods*, vol. 14, no. 3, pp. 290–296, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/28165473

[7] Y. Zhu, Q. Ouyang, and Y. Mao, “A deep convolutional neural network approach to single-particle recognition in cryo-em microscopy,” *BMC Bioinformatics*, vol. 18, no. 1, p. 348, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/28732461

[8] S. K. Zhou, *Medical Image Recognition, Segmentation and Parsing*. Academic Press, 2016.

[9] S. K. Zhou, H. Greenspan, and D. Shen, *Deep Learning for Medical Image Analysis*. Academic Press, 2017.

[10] G. Wu, D. Shen, and M. R. Sabuncu, *Machine Learning and Medical Imaging*. Academic Press, 2016.

[11] L. F.Xing, chapter 4 - Machine learning and its application in microscopic image analysis. *Academic Press*, 2016, p. 31.

[12] H.-C. Shin, M. Orton, D. J. Collins, and M. Leach, *ORGAN DETECTION USING DEEP LEARNING*. Academic Press, 2016, book section 7, pp. 123–153.

[13] Y. Xie, F. Xing, and L. Yang, *Deep Voting and Structured Regression for Microscopy Image Analysis*. Academic Press, 2017, book section 7.
Fig. 5: Sample detection results. Left column (a,c) is our FCRN’s detections, right column (b,d) is Faster-RCNN’s detections.

[14] F. Wang, H. Gong, G. Liu, M. Li, C. Yan, T. Xia, X. Li, and J. Zeng, “Deeppicker: A deep learning approach for fully automated particle picking in cryo-em,” J Struct Biol, vol. 195, no. 3, pp. 325–36, 2016. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/27424268

[15] Y. Xiao and G. Yang, “A fast method for particle picking in cryo-electron micrographs based on fast R-CNN,” in American Institute of Physics Conference Series, ser. American Institute of Physics Conference Series, vol. 1836, Jun. 2017, Conference Proceedings, p. 020080.

[16] A. Heimowitz, J. Anden, and A. Singer, “Apple picker: Automatic particle picking, a low-effort cryo-em framework,” J Struct Biol, vol. 204, no. 2, pp. 215–227, 2018. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/30134153

[17] T. Wagner, F. Merino, M. Stabrin, T. Moriya, C. Gatsogiannis, and S. Raunser, “Sphire-cryolo: A fast and well-centering automated particle picker for cryo-em,” biorxiv, 2018.

[18] J. Redmon and A. Farhadi, “Yolo9000: Better, faster, stronger,” arXiv, 2016.

[19] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, “Object detection with deep learning: A review,” arXiv:1807.05511v1, 2017.

[20] P. Hough, “A method and means for recognizing complex patterns,” 1962.

[21] D. G. Lowe, “Distinctive image features from scale-invariant keypoints (sift),” International Journal of Computer Vision, vol. 60, no. 2, p. 28, 2004.

[22] H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, “Speeded-up robust features (surf),” Computer Vision and Image Understanding, vol. 110, no. 3, p. 14, 2008.

[23] J. Redmon and A. Farhadi, “Yolov3: An incremental improvement,” arXiv, 2018.

[24] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, “Ssd: Single shot multibox detector,” arXiv, 2016.

[25] S. Ren, K. He, R. Girshick, and J. Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” in Advances in Neural Information Processing Systems 28, C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, Eds. Curran Associates, Inc., 2015, pp. 91–99.

[26] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask r-cnn,” arXiv, 2017.

[27] A. Vedaldi and K. Lenc, “Matconvnet – convolutional neural networks for matlab,” in Proceeding of the ACM Int. Conf. on Multimedia, 2015.
Fig. 6: Reconstruction on EMPIAR-10017 (β-galactosidase) dataset. top left: 2D average templates by FCRN. top right: 2D average templates by Faster-RCNN. bottom left: 3D structure by FCRN. bottom right: 3D structure by Faster-RCNN. Red scale bar corresponds to 10 Å.

[28] K. Naydenova and C. J. Russo, “Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy,” *Nat Commun*, vol. 8, no. 1, p. 629, 2017. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/28931821

[29] J. Frank, *Three-Dimensional Electron Microscopy of Macromolecular Assemblies: Visualization of Biological Molecules in Their Native State*. OXFORD UNIVERSITY PRESS, 2006.

[30] P. A. Penczek, “Chapter three - resolution measures in molecular electron microscopy,” in *Cryo-EM, Part B: 3-D Reconstruction*, ser. Methods in Enzymology, G. J. Jensen, Ed. Academic Press, 2010, vol. 482, pp. 73–100.

[31] S. H. Scheres and S. Chen, “Prevention of overfitting in cryo-em structure determination,” *Nat Methods*, vol. 9, no. 9, pp. 853–4, 2012.

[32] M. UNSER, B. L. TRUS, and A. C. STEVEN, “A new resolution criterion based on spectral signal-to-noise ratios,” *Ultramicroscopy*, vol. 23, p. 4, 1987.

[33] R. Henderson, S. Chen, J. Z. Chen, N. Grigorieff, L. A. Passmore, L. Ciccarelli, J. L. Rubinstein, R. A. Crowther, P. L. Stewart, and P. B. Rosenthal, “Tilt-pair analysis of images from a range of different specimens in single-particle electron cryomicroscopy,” *J Mol Biol*, vol. 413, no. 5, pp. 1028–46, 2011. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/21939668

[34] B. Bottcher, S. Wynne, and R. Crowther, “Determination of the fold of the core protein of hepatitis b virus by electron cryomicroscopy,” *Nature*, vol. 386, p. 4, 1997.

[35] S. H. Scheres, *Chapter Six - Processing of structurally heterogeneous cryo-EM data in RELION*, ser. Methods in Enzymology. Academic Press, 2016, vol. 579, pp. 125 – 157.