DeepTracer-Denoising: Deep Learning for 3D Electron Density Map Denoising

Haowen Guan
Center of Data Science
New York University
New York City, US
haowen@nyu.edu

Dong Si*
Division of Computing and Software Systems
University of Washington Bothell
Bothell, US
dongsi@uw.edu

***Abstract*—Cryo-electron microscopy (Cryo-EM) is widely used in molecular structure determination and drug discovery. Experimental cryo-EM images suffer from the noises introduced by electron beam dose and sample preparation. Although many approaches have been proposed to improve the signal-to-noise ratio (SNR) for cryo-EM image denoising, the noises are still presented after 3D reconstruction and can obstruct the analysis and visualization of the 3D density map. Here we present DeepTracer-Denoising, a method for 3D electron density map denoising. We employ a 3D Neural Network to learn the pattern of noises and the biological structure from density maps. Our method is designed to work on medium to high-resolution maps ranging from 2.5 Å to 10.0Å. It is configurated with two modes to tackle both background noise and structural noise in a 3D density map. Our method can correctly identify 97.70% background noise while preserving 96.46% density of the native structure. For the maps that contain structural noise, DeepTracer-Denoising achieves an overall accuracy of 98.95%.**

Keywords—Cryo-EM, 3D electron density map, Deep learning, 3D CNNs, Denoising

I. INTRODUCTION

Cryo-electron microscopy (Cryo-EM) is an awarded technology of the 2017 Nobel Prize in Chemistry. Since the first three-dimensional (3D) image of protein was generated using electron microscopy in 1990, cryo-EM has been noticed for its potential for resolving biomolecules 3D structures. Recently, With the rapid advancement in electron detector and image processing, cryo-EM is now capable of resolving high-resolution 3D structures. Notably, in the past decade, the number of 3D biological structures resolved by cryo-EM has increased exponentially [1]-[2], and it has risen to one of the leading techniques for structural biology [3]-[5].

Cryo-EM 3D structure, also known as cryo-EM density map, is 3D reconstructed from 2D cryo-EM images taken from microscopy [6]. Raw cryo-EM images are suffered from the low signal-to-noise ratio (SNR) problem [7]. To reduce the noise, there have been many approaches proposed to enhance the SNR of 2D cryo-EM images in the past decade. The early conventional methods like BM3D [8], band-pass filter [9], and Wiener filter [10] approach the denoising problem through mathematical image restoration algorithms, which can improve the contrast and decrease the noise level. Recently, due to rapid growth in computation power and available data, learning-based algorithms, particularly deep convolutional neural networks (CNNs), have demonstrated a revolutionary performance in image processing and denoising [11]. Cryo-CARE [12] and Topaz-Denoise [13], two CNN-based 2D cryo-EM image denoising approach, was proposed by adopting the Noise2Noise [14] framework - a general deep CNNs framework for image restoration. CDAE [15] is a cascade of denoising autoencoder for single particle cryo-EM images. "Robust denoising for cryo-EM" [16] is a newly proposed model which combines autoencoder with robust generative adversarial networks (GANs) and achieves an efficient and robust cryo-EM denoising performance. Recently, a cryo-EM image denoising framework, called Noise-Transfer2Clean [17], was proposed; it used a GAN to synthesize noise with a CNN as denoiser and achieved a state-of-the-art result.

With extensive research from the past, the image restoration denoising algorithms for cryo-EM images have reached a very high level, and it greatly helps the cryo-EM particle picking and 3D reconstruction steps to produce the final 3D electron density map. However, denoising through improving SNR has a shortcoming, that it is unable to "completely" eliminate noises. Improving SNR can shrink noise signals to a low level while amplifying the true structure signals; however, it cannot modify the noise signals directly to 0. For this reason, we can still observe noises when viewing the cryo-EM density map with a low contour level after the 3D reconstruction (Fig. 1). According to previous research [18], the noise in cryo-EM comes from three aspects: (1) "structural noise" sourced from ice matrix on top of the molecule surface and a thin layer of the superimposed carbon film. Furthermore, the portion of molecule structure that is not reproducible due to conformational variations is also counted as "structural noise." (2) "shot noise" sourced from the quantum nature of the electron radiation. (3) "digital noise" sourced from the photograph recording and subsequent digitization. In the cryo-EM density map, all these noises are carried over from cryo-EM images and remain in existence. Shot noise and digital noise do not have a specific shape, and they distribute randomly like dust in the density map. We will summarize them as background noise in this paper. Structural noise, however, has a specific shape, and the density signal is often stronger than background noise, as in Fig. 1. For the cryo-EM density map, the noise voxel signals are troublesome and redundant. Not only do they shade the actual structure from observation, but they can also interfere with the prediction accuracy of automatic cryo-EM protein structure modeling algorithms such as DeepTracer [19] and other state-of-the-art structure analysis algorithms.

To eliminate noise in the cryo-EM density map, we build a convolutional neural network (CNN) denoising model to proceed with noise classification on the 3D electron density map. We designed a procedure for generating a 3D cryo-EM training dataset and, in addition, proposed a metric to measure the similarity between the cryo-EM density map and the corresponding solve biomolecule structure for dataset filtering. Our model's performance for both background noise

* Correspondence to Dong Si at dongsi@uw.edu.

978-1-6654-6819-0/22/$31.00 ©2022 IEEE

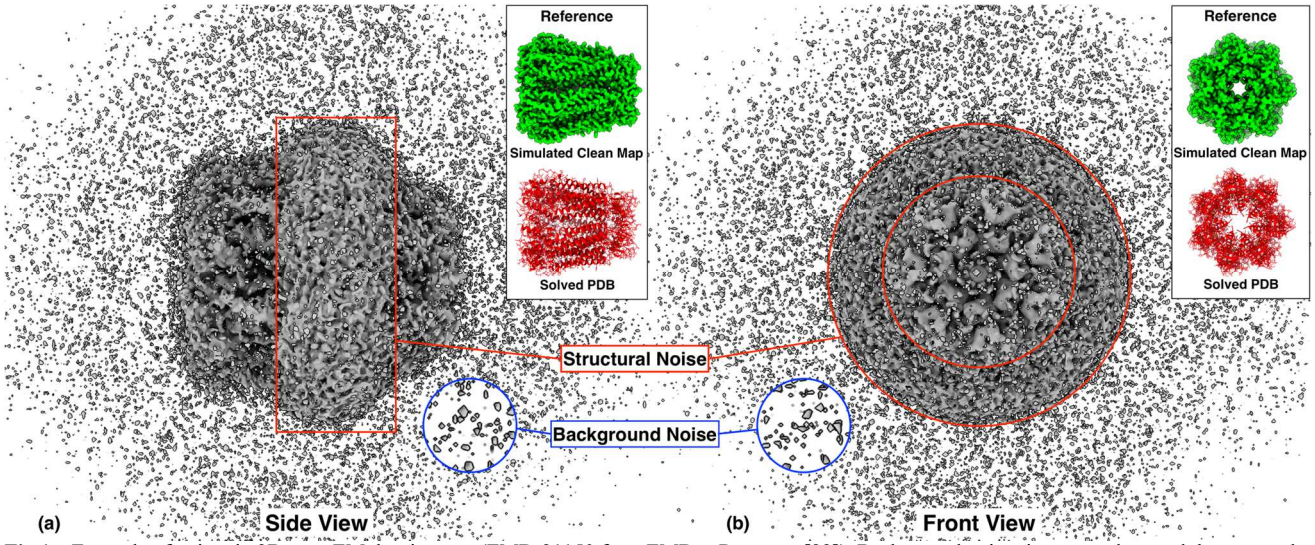

Fig.1. Example of noises in 3D cryo-EM density map (EMD-21150 from EMDataResource [22]). Background noises is scattered around the map and a ring shape structural noise is encloseing the center protein structure. Top right is the visualization of solve protein structure and a simulated map generated from the PDB served as a reference of the clean 3D electron density map. (a) Side view. (b) Front view.

and structural noise is evaluated on a test set of 304 cryo-EM density maps. Lastly, we compared our approach with ChimeraX's [20] hide-dust tool. To summarize, our contributions are as follows:

- We proposed a CNN-based voxel classification approach for 3D electron density map denoising, which can eliminate noise signals completely.
- We present the first deep learning denoising approach for 3D electron density maps, and it can denoise both structural and background noise efficiently and accurately.

II. Materials and Method

A. Method Overview

Classification of voxel's type is the key to our denoising approach. To take advantage of CNN's image recognition ability and learn the pattern of biological structure, we adopt U-Net [21], which is a fully convolutional network architecture designed for biomedical image recognition. The original U-Net was proposed as a 2D CNN architecture and suited for various kinds of 2D biological images taken by microscopies. However, it does not fit cryo-EM structure. Cryo-EM structure is in the format of 3D electron density map, which is 3D reconstructed from 2D micrographs. It carries all of the information in 2D images and in additionally records the spatial conformation of the biological, which could aid the identification of biological structure. To take advantage of this, we modify the original U-Net as a 3D U-Net to work on electron density map data [19]. Since electron density maps have various image sizes, to standardize the input of the model, we cut the map into multiple unitary sizes of (64,64,64) as input. Also, different from the original U-Net, we modify and add padding for each convolution layer to maintain the dimension. The output of the neural network has a shape of (64,64,64,2), which is the SoftMax prediction result for every voxel, indicating the binary classification probabilities to be either noise or structure. Finally, using the classification result as a mask, we filter out the noise voxels and get the denoised 3D cryo-EM structure. The prediction pipeline is shown in Fig. 2.

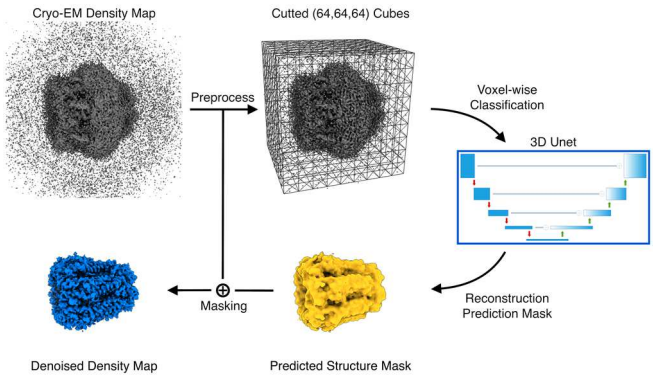

Fig. 2. DeepTracer-Denoising prediction pipeline

All of the electron density maps are preprocessed before training and prediction. We reset all of the voxels below 0 to 0 and normalized the range of voxel values to 0 ~ 1. Furthermore, all density maps are resampled on a grid voxel size of (1Å, 1Å, 1Å) to maintain the unitary.

B. Training Dataset Generation

To train a voxel-wised classification model, our dataset is built with data pairs of X (input) and Y (target), where X is the original cryo-EM structure (3D electron density map), and Y is the corresponding labeled ground truth map. Both X and Y data maps should have the exact dimensions, inferring a one-to-one mapping of each voxel's class (structure or noise). 3D electron density maps are collected from the EMDataResource website [22] alone with their solved biological structure (PDB file). Labeled ground truth maps are generated based on biological structure. This labeling procedure is the most important while also the most challenging step amount this denoising approach since the neural network's classification performance is highly correlated with the labeling style, yet the style can be arbitrary

as there is no such "correct label"[1]. In the following sections, we explained the steps to generate labeled ground truth map based on simulated electron density map.

1) Simulated Eelectron Density Map

Simulated electron density map is created by representing each atom in the biological structure as a 3D gaussian distribution centered on its location to simulate its electron density cloud. It is an ideal representation of proteins under the perfect cryo-electron microscopy without the interference of noises, which means that any nonzero reading in the simulated density map indicates the imaging of biological structure. By utilizing the simulated map, we could label all the nonzero voxels as 1 (indicating structure) and the rest of the area as 0 (indicating empty space or noise). As mentioned above, we know there is a boundary issue with the electron density map (footer. 1). Since our denoising goal is to mimic human perception and eliminate noises as much as possible while maintaining most of the structure information, we decided to use a 3σ threshold for the simulated map, which is guaranteed to cover 99.73% of total imaging of biological structure [23] (see discussion).

2) Standardizing Voxel Size

The data of electron density map is essentially a 3D array; it is difficult to identify the actual size of the biological structure without knowing the voxel size of the map (the physical size of each data grid). Furthermore, density maps have various different resolutions due to equipment limitations. It is almost impossible to resolve the molecule structure size with varied electron density resolution and unknown voxel size. This challenges the neural network; for example, an alpha helix in a low-resolution map can be too blurred to distinguish between being a straight backbone chain in a high-resolution map. To tackle this problem, we standardize the voxel size of all input density maps by resampling them on the grid size of 1.0 Å. With the fixed voxel size, the primary and secondary structures can be distinguished by counting the number of voxels it occupies since their physical size is different.

3) Step by Step Procedure

The detailed steps to generate a labeled map are as follows: (1) Create a blank voxel grid with 1.0 Å voxel size and the same spatial dimension as the input 3D electron density map. (2) Simulating each atom's electron density imaging by 3D gaussian distribution according to the simulating resolution and number of electrons in the atom. Then, embedding atoms into the blank voxel grid using their coordinate information. (Note that we used a 3-sigma threshold for the 3D gaussian distribution; values beyond 3-sigma are set to 0). (3) Label all the nonzero voxels in the density map as 1 (indicating structure) and the remaining as 0 (indicating empty space or noise).

C. Filtering Dataset

The denoising model is only successful if it is trained with correct data. However, we have seen many cases in that data pairs are not complete/matched. For example, the solved protein structure explains only a part of the biological structure in the experimental 3D electron density map, as EMD-9339 shown in figure 3-I. The reason for this is that the complete cryo-EM structure is a repeating version of a certain structure. Researchers will tend to skip matching the repeating pattern because it is redundant work. Even though these data pairs are not technically "wrong", we are unable to use them to generate the correct labeled map since we need to simulate electron density from PDB files. As in figure 3-a.(ii), we can see that the simulated map (red) only matches the middle portion of the electron density map (gray), and the left and right portions are left empty. To filter out such data pairs, we proposed a metric based on correlation coefficient to measure the similarity between the experimental 3D electron density map and the PDB file.

1) Correlation-Based Similarity Metric

Correlation coefficient is a statistical measure of the strength of the relationship between the relative movements of two variables, which range from -1 to 1. In other words, when two variables increase and decrease together and show a similar signal pattern, the correlated coefficient between them will be prominent and approach the maximum value of 1. This property is very useful since it can measure the similarity of two datasets. Precisely, we could calculate the correlation coefficient between voxels in a 3D electron density map and the simulated map generated from the corresponding PDB file. If the solve biological structure covers most part of the cryo-EM density map and resolves correctly, the simulated map created from it would appear to have a similar pattern as the cryo-EM density map and result in a high correlation. Figure 3-II. is an example of a complete/matched data pair.

2) Definition of the Similarity Metric

In this section, we will define the portion of data to calculate the correlation coefficient. This may seem confusing since one could simply use all of the data (the whole experimental density map) to do the calculation, but it is not ideal. According to the stats of our dataset, 97.6% of data points in experimental density maps belong to empty space. If we calculate the correlation using such data, every map will end up with a coefficient close to 0 since the empty space data points are constantly 0, which indicates no correlation, and they dominate the experimental density map.

To find the precise and informative portion of data for the calculation of similarity, the recommended contour level is used to identify the core part of the map. Contour level is a viewing option and indicates the surface level of the density map for visualization; it could be compared to the contour line in a 2D image. When researchers deposit a new experimental density map into the database, they are required to give a recommended contour level based on their perception to guide other people viewing the structure. Using this property, we can classify the core part of a map as any voxel greater than the contour level. Since the recommended contour level is widely accessible and it mimics general human perceptions, it serves as an excellent resource for identifying the core part of maps automatically.

Sometimes, the recommended contour level is not ideal as it is a human subjective value. It may be biased and only focus on a tiny high-density area. An alternative approach,

[1] Under the Cryo-EM microscope, the imaging of each atom is shown as an electron density cloud (a 3D gaussian distribution in theory). Gaussian distribution does not have a boundary since it extends to infinity from the center. To define a boundary for our purpose, we will pick a sigma value as a boundary threshold. For example, 2σ boundary correspond to 95.45% of total value and 3σ boundary correspond to 99.73% of total value [23].

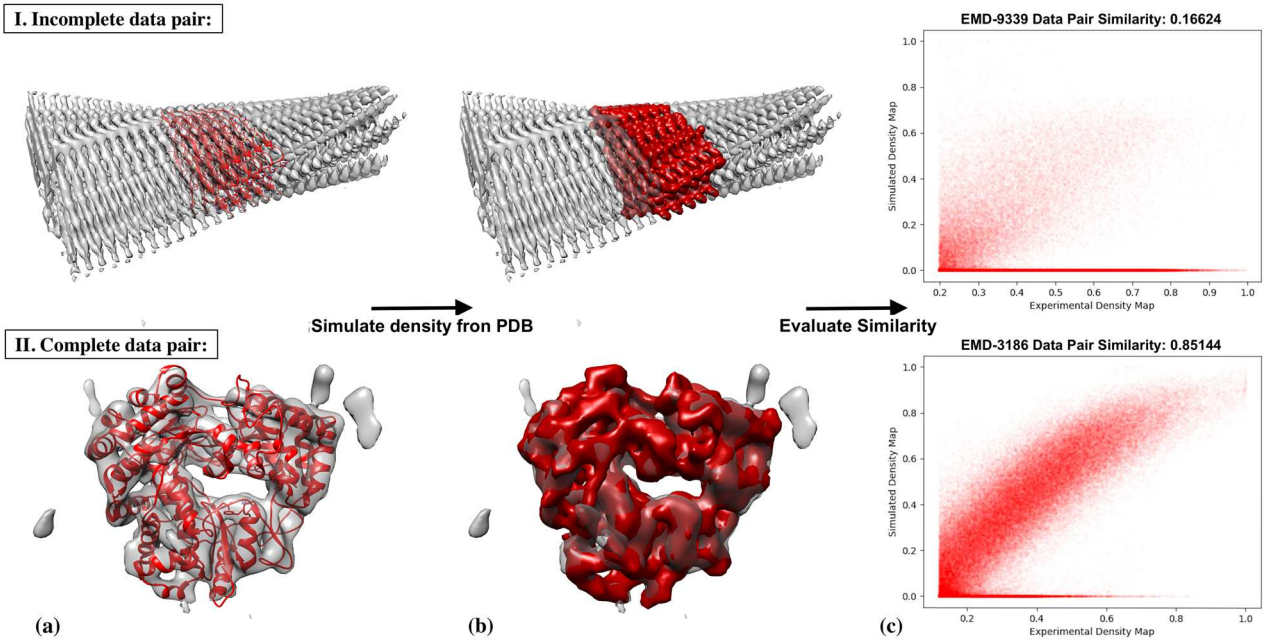

Fig.3. Incomplete and complete data pair example and similarity evaluation illistration. Imcomplete example is EMD-9339 with PDB ID code 6N37 from EMDataResource[22]; Complete example is EMD-3186 with PDB ID code 5FJ6. (a) The transparent gray shape is experimental 3D electron density map. Red is the corresponding solved structure. (b) The red figure superimposed on experimental map is the PDB simulated map. (iii) The scatter plot of simulated map and experimental map. The complete example's similarity 0.85144 is significantly higher than incomplete example's similarity 0.16624.

which uses the simulated density map, can sometimes give a better result. As we know, the simulated map is an ideal representation of proteins; any nonzero values in the simulated map infers the imaging of biological structure. We can use the nonzero values in the simulated map to represent the core part of the structure, which can serve as a percentage lower bound of our approach to prevent bias. Specifically, when the percentage of experimental density map core voxels is less than the percentage of simulated map core voxels, we will use the percentage of the simulated map as the core data percentile. The equation to calculate the core part percentile is,

$$P_{core} = MAX\left(1 - F_{Exp}(C), 1 - F_{Sim}(0)\right) \quad (1)$$

Where P_{core} is the core data percentile, Exp is the experimental density map, Sim is the simulated map, C is the recommended contour level, and $F_X(x)$ is the cumulative distribution function equivalent to $P(X \leq x)$. Combining the percentile found in equation (1), the equation to find the best portion of data is,

$$Exp \times Sim \in \{Exp > [1 - 2 \times P_{core}]^{th}\ of\ Exp\} \quad (2)$$

Where the $Exp \times Sim$ indicates experimental and simulated maps voxel-wise data pairs. In short, we first find the maximum percentage of experimental map core data and simulated map. Then multiplying it by two [2], it is the percentage of the top values in the experimental map we will use to calculate the correlation coefficient. Finally, the correlation coefficient value indicates the similarity of data pairs, which range from 0 to 1 (We consider negative correlation as no correlation).

$$Simularity = corr(Exp \times Sim) \quad (3)$$

In figure 3. (c), we present the scatter plots for correlation inspection. The first example (EMD-9339) is an incomplete data pair; Most of the structure voxels in the electron density map are paired with empty voxels in the simulated map. For this reason, we can see a wide and dense bar lying under the figure (not correlated data points) and give an overall correlation of 0.16624. The second example (EMD-3186) is a complete data pair; it shows a strong correlation and gives a value of 0.85144.

3) Training Dataset Analysis

We collected 1889 experimental Cryo-EM density maps from the EMDataResource website [22]. The resolution of maps is ranged from 2.5 ~ 10 Å, and the distribution of the resolutions is intentionally left uncontrolled to accommodate the preference in the cryo-EM community. We applied the correlation-based similarity metric defined above to evaluate the completeness of data pairs of the dataset. The similarity distribution of the dataset is shown in fig. 4.

From the figure, we can observe a gaussian distribution on the right side of the histogram; these data pairs are inferred as the complete group. The data pairs scattered on the left of the red threshold fall far outside of the distribution; thus, they are identified as flawed or incomplete data pairs. After filtering out unwanted data, our final dataset consists of 1739 experimental density maps, and it is split randomly into training and validation sets with an 80:20 ratio.

[2] Timing the percentage by 2, we intentionally add an extra amount of noise voxels into the correlation coefficient calculation. According to the nature of correlation coefficient, two uncorrelated cluster of points can yield a high correlate if they are spatially separated. Therefore, adding an extra amount of noise voxel can boost the correlation of complete data pairs while merely affects the incomplete data pairs, since structure data pairs are spatially separated from noise data pairs.

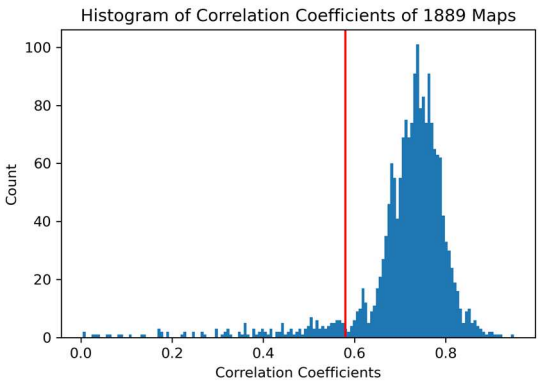

Fig. 4. Correlation Coefficients (similarity) of data pairs in training dataset. Low similaity (incomplete) data pairs on the right of red threshold is filtered out from training dataset.

D. Training Procedure Hyperparameters

We used Python TensorFlow as the deep learning framework. The loss was calculated based on categorical cross-entropy, and the model was trained with an Adam optimizer with a learning rate of 0.0002. Since the number of structure data points was way less than the number of noise data points in the training dataset, we applied a weight ratio of 40: 1 to the cross-entropy loss function to balance the loss between the two classes. Our model was trained with mini-batch gradient descent. Each mini-batch consisted of 256 sliced data cubes of size (64,64,64). There were and in total 2405 mini-batches in each epoch and iterated for ten epochs to arrive at the final model. We used TensorFlow distributed strategy to do multi-GPUs training across $8 \times RTX\ 8000\ GPUs$. The training procedure took about 24 hours.

III. Results

The performance of DeepTracer-denoising is promising. It is capable of denoise an experimental cryo-EM density map efficiently while maintaining high accuracy. The model comes with two modes for identifying background noise and structural noise. In the following section, we demonstrate the operation of DeepTracer-denoising, study our model statistically using a validation set and evaluate the accuracy of DeepTracer-denoising on a test set composed of 304 newly deposited Cryo-EM density maps (deposit after 2021) in the EMDataResource website, which is independent of the train set. Lastly, we compare our approach with the existing algorithm.

A. Example

Fig. 5 shows an example of a cryo-EM density map before and after the application of DeepTracer-Denoising. In this example, the background noises display as density values randomly distributed around the molecule, and we can also observe a ring shape structural noise surrounding the "waist" of the protein. In fig. 5 (ii), we demonstrate how DeepTracer-Denoising on classifies the type of voxels. Masking out the noises, the denoised protein structure is displayed in fig. 5 (iii). Compared with the solved structure, the denoised electron density map matches perfectly with no redundant density.

B. Evaluation of the Model

The evaluation metric of DeepTracer-Denoising is different from the mainstream metrics used in the past. Unlike those cryo-EM image (2D) denoising approaches that focus on the signal-to-noise ratio (SNR), we will be evaluating and studying our model based on confusion matrix and precision-recall curve since the core of our approach is a classification CNN model.

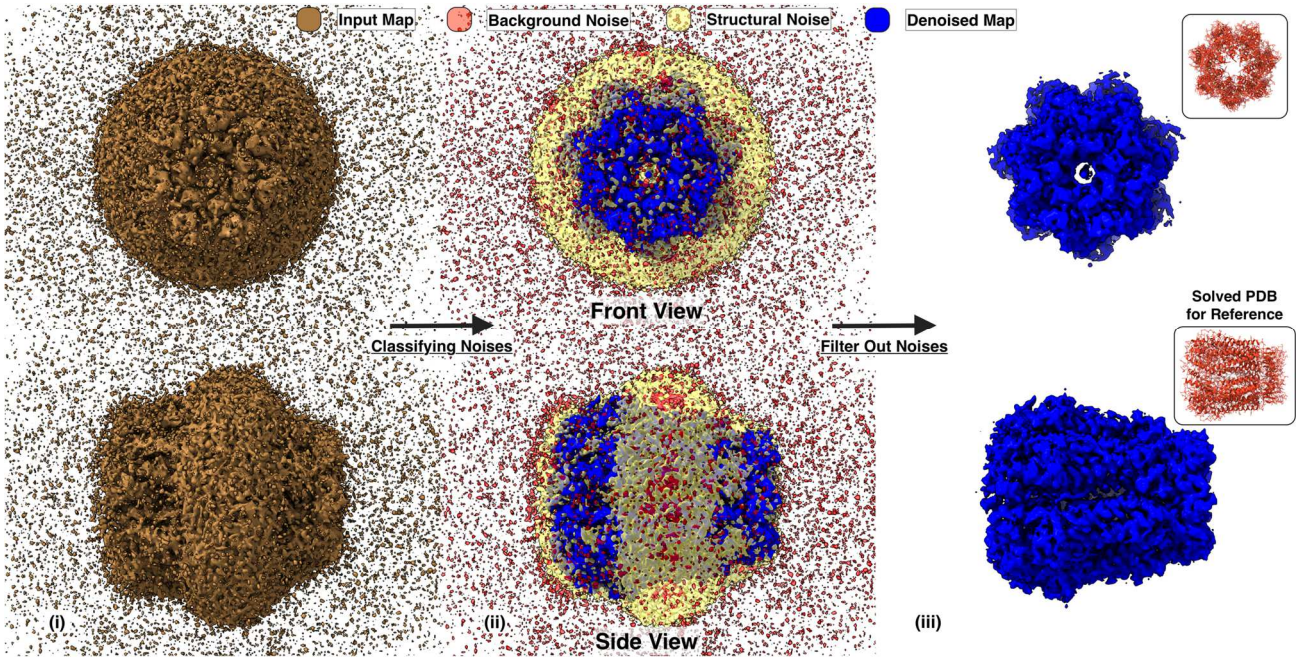

Fig. 5. DeepTracer-Denoising Example: EMD-21150 before and after denoising. (i) EMD-21150 density map displayed with a contour level of 1.6. (ii) The same map as input density map, while different region is labeled as different color based on its classification result. Background noise is labeled as tranparent red, ring shape structual noise is labeled as transparent yellow, and the structure is label as blue. (iii) Denoised electron density map (displayed at the contour level of 1.6) and the solve biological structure for comparison.

1) Definition of the Confusion Matrix

We define the type I error (False Positive) as classifying the structure voxel as noise since the primary goal of DeepTracer-Denoising is eliminating noise as much as possible while preserving the main structure untouched. The type II error (False Negative) is classifying noise voxel as structure. To be concise, we refer to the noise prediction as positive and structure prediction as negative from now on. In the following sections, all empty space voxels, 0 value voxel in input map, are (by default) excluded from the evaluation; their prediction result is trivial as the empty space voxel are assumed as noise. Only those which start with a nonzero value and predicted as noise will be counted toward the total true positive (TP) in the evaluation.

2) Optimal Threshold of Classification

For a normal classification model that is trained with a balanced dataset (or weight balanced), threshold is usually chosen at midpoints of the interval, for example, 0.5 in terms of binary classification. However, this threshold is not ideal for our purpose. We will find thresholds for different cases.

Firstly, we analyzed the general performance of DeepTracer-Denoising on the validation set, which was split from the train set and contained 347 experimental maps. Using the validation set, we gathered all negative class voxel's predictions into a group and all positive class voxel's predictions into another group and plotted the histograms in Fig. 7.I.(a). The histogram is a way to inspect the scale of the confusion matrix. The positive class (red) will be true positive (TP) if on the left of the threshold and false negative (FN) if on the right; the Negative class (blue) will be false positive (FP) if on the left and true negative (TN) if on the right. Furthermore, we generate the precision-recall curve of the validation set (Fig. 6.I.(b)) with G-mean to find the optimal threshold, where the G-mean is an unbiased evaluation metric for imbalanced classification. The formula is,

$$GMean = \sqrt{\frac{TP}{TP+FN} * \frac{TN}{FP+TN}} \tag{4}$$

The general performance of this model is significant as shown in fig. 6.I.(b). The G-mean metric gives an optimal threshold of 0.035 for classification, which coincide with two classes' relative histogram intersection (as in fig. 6.I.(a)). Based on our experiment, we conclude that **0.035 is the optimal threshold for identifying background noise**.

We also studied our model's structural noise classification accuracy. Since background noise voxels dominate the positive class, to reveal the performance of structural noise denoising, we exclude all the data points classified as background noise – prediction values smaller than 0.035 - from the positive class. After that, we used the remaining data to generate the precision-recall curve (Fig. 6.

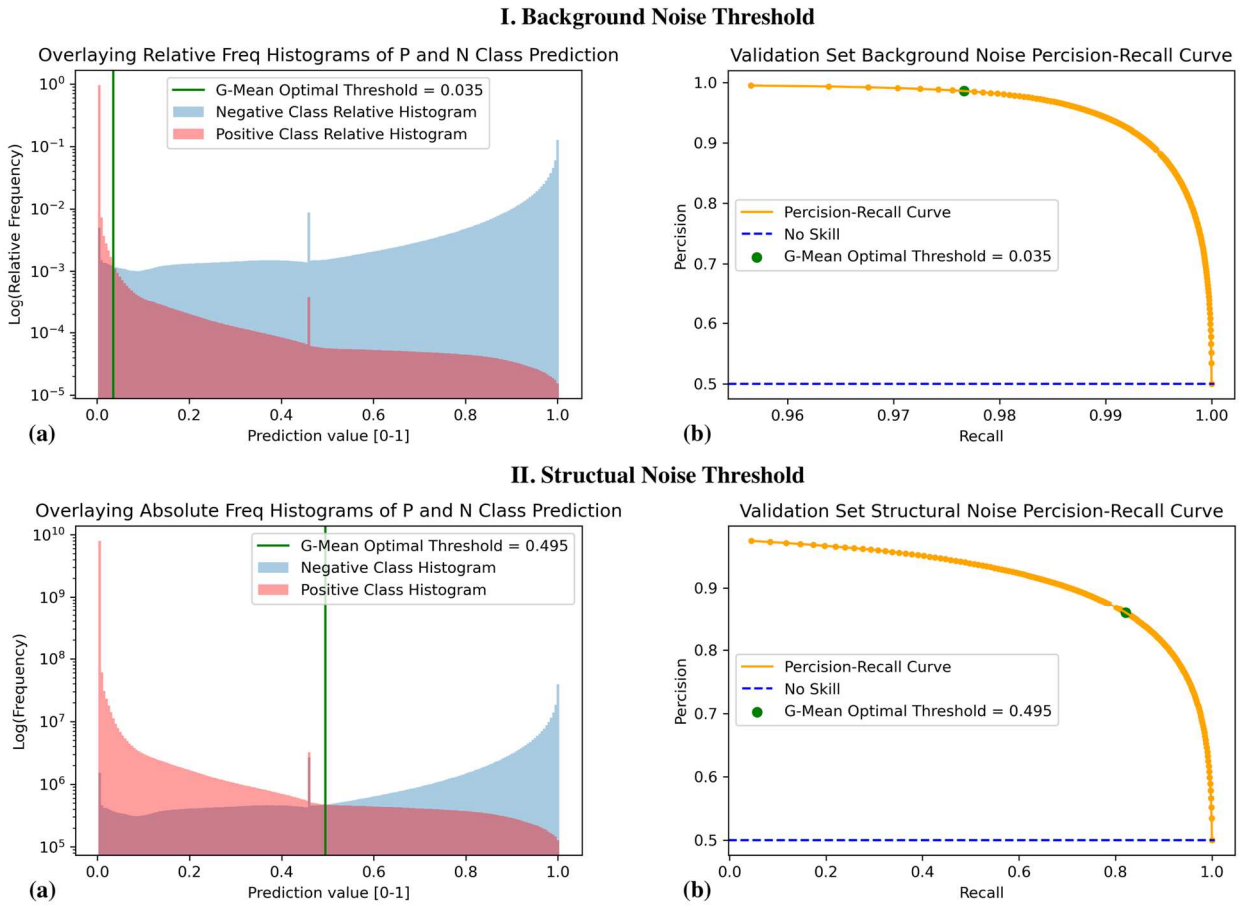

Fig. 6. Threshold analysis for background noise and strctural noise. Optimal threshold for background noise and strctural noise are 0.035 and 0.495 respectively. I.(a) The relative frequency histogram of positive and negative classes' probability predictions. They are overlayed for comparison. I.(b) The percision-recall curve representing the performance of background noise classification. II.(a) The absolute frequency histogram of positive and negative classes' probability predictions. II.(b) The percision-recall curve of the validation set excluding the background noise, representing the structural noise classification performance.

TABLE I. ACC, FP, FN FOR DIFFERENT RESOLUTION AND DENOISING MODE ON VAL & TEST SET

Resolution (Å)		[2.5 – 3)	[3 – 4)	[4 – 5)	[5 - 6)	[6 - 7)	[7 - 8)	[8 - 9)	[9 – 10]
Background Noise Mode	**ACC %**	94.5	93.8	89.2	88.1	87.9	87.7	92.3	89.6
	FPR %	2.81	2.67	0.93	0.71	1.31	0.93	2.33	4.49
	FNR %	6.31	7.19	21.15	24.55	25.20	18.66	11.34	27.46
Structural Noise Mode	**ACC %**	97.7	96.8	88.7	87.8	86.9	87.7	91.4	86.8
	FPR %	13.35	13.31	3.72	3.11	7.45	0.93	7.06	5.24
	FNR %	1.56	1.75	17.64	20.85	20.5	18.66	10.44	22.22
Dataset Fraction %		13.8%	53.9%	17.8%	4.3%	4.5%	2.3%	1.4%	2.0%

II. (b)) for structural noise. The **optimal threshold** given by the G-mean metric for **classifying structural noise is 0.495**. This value coincides with the intersection of the two classes' absolute frequency histogram (Fig. 6. II. (a)).

3) Test Set Evaluating

The test set is composed of 304 experimental cryo-EM density maps deposited in the database after we build our train set. It is filtered by our correlation-based metric with the same threshold, and it shares a similar resolution distribution as the train set, which ranges from 2.5Å to 10.0Å. When using **background noise mode**, among the 304 test maps, the average false positive rate (FPR) is 3.537%, the false negative rate (FNR) is 2.298%, and precision is 99.87%. When using **structural noise mode**, the average false positive rate is 19.788%, the false negative rate is 0.360 %, and precision is 99.27%. At first glance, this performance is a lot worse than the background noise above since the type I error FPR is much higher than the background noise. However, if we calculate the accuracy (ACC) for two modes, which the equation is $ACC = \frac{TP+TN}{TP+FP+TN+FN}$, we will find $ACC_{0.035} = 97.658\%$ and $ACC_{0.495} = 98.950\%$. A reason for FPR "conflicts" with ACC is the fact that negative and positive classes are imbalanced. In the test set, the number of negative class voxel is ~27 times of the positive class. Furthermore, in test set evaluation, we are applying the structural noise classification indiscriminately to all of the maps, while some maps do not have structural noise. In reality, users are able to inspect the map before the application, and they can choose the most suitable mode to do denoising. The detail statistics for different class is listed in TABLE I.

C. Comparing with Existing Algorithms

The DeepTracer-Denoising model is highly efficient. Our method took less than half mins to denoising one map with one RTX 8000. Unfortunally, there is no comparable algorithm in the past, so we only compared our method with ChimeraX's [20] hide-dust tool, which is a popular 3D electron density map application. In fig. 7, ChimeraX can denoise only part of the background noise while still suffers from the exploding noise when viewing at a low contour level. In contrast, DeepTracer-Denoising can view the structure down to level 0 without the interference of noises.

IV. DISCUSSION

In this work, we proposed a new approach for cryo-EM denoising through approaches the task as a classification problem. Unlike most of the mainstream methods in the denoising field, which focus on improving the signal-to-noise (SNR) of cryo-EM images, we target eliminating noise completely invisible. Our approach is a successful attempt for this purpose; it has achieved an impressive denoising performance on both background noise and structural noise.

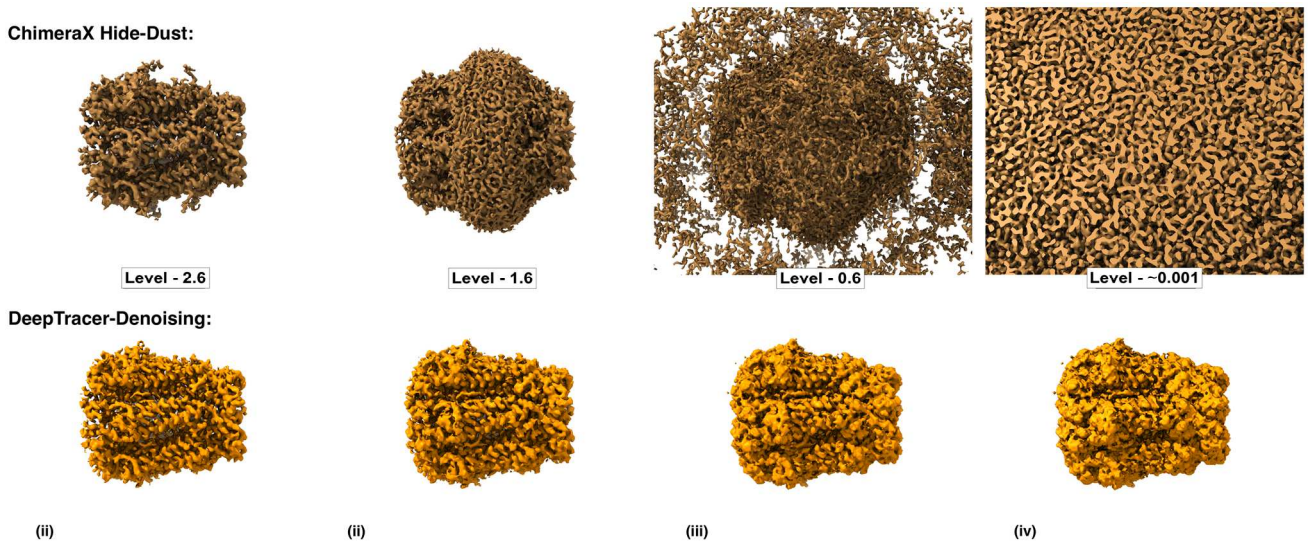

Fig .7. Comparison of ChimeraX hide-dust tool and DeepTracer-Denoising. ChimeraX unable to denoise structual noise, and the background noise still shows up when countour level goes down. DeepTracer-Denoising can visualize the structure at a countour level as low as 0 without interference.

Though the overall performance of DeepTracer-Denoising is great, there are a few limitations presented. First of all, the labels we generated in the training dataset are not truly correct, or to be precise, there are no such "correct" labels. The simulated density map is distributed as gaussian distribution, and there is no correct threshold to use to determine which portions of density should be classified as structure. We choose a 3-sigma threshold, which is equivalent to 99.73% of the theoretical portion of density. However, choosing a 2-sigma threshold, which is equivalent to 95.45% of the theoretical portion of density, is also a great option, and it is very like to give a better denoising performance. Secondly, different maps suit different denoising modes, and users sometimes have to inspect the denoising results to get the best one manually. In most cases, our model is guaranteed to find an ideal result from one of the two modes.

We believe our method has great potential. It can be adopted to denoise other 3D biological images and solve problems other than denoising. One application is designing an automated bounding box finding algorithm for electron density maps, which allows users to cut density maps to smaller sizes and save storage space and computation power.

V. Conclusion

In this paper, we present a process to construct a Convolutional Neural Network (CNN) model for 3D electron density map denoising. The key idea of DeepTracer-Denoising is considering the denoising task as a classification problem, separating the noise and structure voxels into two classes, and masking out noise based on the classification result. In order to build the training dataset, we proposed a metric to measure the similarity between the 3D electron density map and the corresponding solved biomolecule structure for dataset filtering. Our model is designed to work on medium to high-resolution maps ranging from 2.5 Å to 10.0Å, and it is configurated with two modes to tackle both background noise and structural noise. When testing on 304 cryo-EM density maps, the background noise mode correctly identifies 97.70% of background noise (TPR) while preserving 96.46% density of the native structure (TNR). For maps that contain structural noise, DeepTracer-Denoising achieves an overall accuracy (ACC) of 98.95%.

Acknowledgment

This work was supported by NSF grant 2030381, the SRCP Seed Grant at the University of Washington Bothell, and the CoMotion Gap fund at the University of Washington (D.S.). Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases. We thank Michael Chavez, Andrew H. Nakamura, and Kiran Mitra for sharing the ideas and helping with the review and test.

References

[1] R. Fernandez-Leiro and S. H. Scheres, "Unravelling biological macromolecules with Cryo-electron microscopy," Nature, vol. 537, no. 7620, pp. 339–346, Sep. 2016.

[2] A. Patwardhan, "Trends in the electron microscopy data bank (emdb)," Acta Crystallographica Section D: Structural Biology, vol. 73, no. 6, pp. 503–508, 2017.

[3] E. Nogales, "The development of cryo-em into a mainstream structural biology technique," Nature methods, vol. 13, no. 1, pp. 24–27, 2016.

[4] J. Frank, "Advances in the field of single-particle cryo-electron microscopy over the last decade," Nature protocols, vol. 12, no. 2, pp. 209–212, 2017.

[5] Y. Cheng, "Single-particle cryo-em—how did it get here and where will it go," Science, vol. 361, no. 6405, pp. 876–880, 2018.

[6] X.-chen Bai, G. McMullan, and S. H. W. Scheres, "How cryo-EM is revolutionizing structural biology," Trends in Biochemical Sciences, vol. 40, no. 1, pp. 49–57, 2015.

[7] T. Bendory, A. Bartesaghi, and A. Singer, "Single-particle cryo-electron microscopy: Mathematical theory, computational challenges, and opportunities," IEEE Signal Processing Magazine, vol. 37, no. 2, pp. 58–76, Feb. 2020.

[8] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," in IEEE Transactions on Image Processing, vol. 16, no. 8, pp. 2080-2095, Aug. 2007.

[9] P. A. Penczek, "Image restoration in cryo-electron microscopy," Methods in Enzymology, pp. 35–72, 2010.

[10] C. V. Sindelar and N. Grigorieff, "An adaptation of the Wiener filter suitable for analyzing images of isolated single particles," Journal of Structural Biology, vol. 176, no. 1, pp. 60–74, 2011.

[11] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C.-W. Lin, "Deep learning on image denoising: An overview," Neural Networks, vol. 131, pp. 251–275, 2020.

[12] T.-O. Buchholz, M. Jordan, G. Pigino, and F. Jug, "Cryo-care: Content-aware image restoration for cryo-transmission electron microscopy data," 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pp. 502–506, Jul. 2019.

[13] T. Bepler, K. Kelley, A. J. Noble, and B. Berger, "Topaz-denoise: General deep denoising models for cryoEM and cryoET," Nature Communications, vol. 11, no. 1, Oct. 2020.

[14] J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2Noise: Learning Image Restoration without Clean Data," Computer Vision and Pattern Recognition, Mar. 2018.

[15] H. Lei and Y. Yang, "CDAE: A cascade of denoising autoencoders for noise reduction in the clustering of single-particle cryo-EM images," Frontiers in Genetics, vol. 11, Jan. 2021.

[16] P. Peng, K. Zhang, and H. Wei, "Robust epileptic seizure prediction with missing values using an improved denoising adversarial autoencoder," 2022 4th International Conference on Image, Video and Signal Processing, Aug. 2022.

[17] H. Li, H. Zhang, X. Wan, Z. Yang, C. Li, J. Li, R. Han, P. Zhu, and F. Zhang, "Noise-transfer2clean: Denoising cryo-EM images based on noise modeling and transfer," Bioinformatics, vol. 38, no. 7, pp. 2022–2029, Feb. 2022.

[18] W. T. Baxter, R. A. Grassucci, H. Gao, and J. Frank, "Determination of signal-to-noise ratios and spectral snrs in cryo-em low-dose imaging of molecules," Journal of Structural Biology, vol. 166, no. 2, pp. 126–132, 2009.

[19] J. Pfab, N. M. Phan, and D. Si, "Deeptracer for fast de novo cryo-EM protein structure modeling and special studies on cov-related complexes," Proceedings of the National Academy of Sciences, vol. 118, no. 2, Dec. 2020.

[20] E. F. Pettersen, T. D. Goddard, C. C. Huang, E. C. Meng, G. S. Couch, T. I. Croll, J. H. Morris, and T. E. Ferrin, "UCSF ChimeraX: Structure visualization for researchers, educators, and developers," Protein Science, vol. 30, no. 1, pp. 70–82, Oct. 2020.

[21] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Lecture Notes in Computer Science, pp. 234–241, Nov. 2015.

[22] EMDataResource. [Online]. Available: https://www.emdataresource.org/. [Accessed: 17-Aug-2022].

[23] Wikipedia contributors, "68–95–99.7 rule," Wikipedia, 08-Jul-2022. [Online]. Available: https://en.wikipedia.org/wiki/68%E2%80%9395%E2%80%9399.7_rule. [Accessed: 18-Aug-2022].