

Bioimage informatics

Noise-Transfer2Clean: denoising cryo-EM images based on noise modeling and transfer

Hongjia Li ^{1,2,†}, Hui Zhang^{2,3,†}, Xiaohua Wan¹, Zhidong Yang^{1,2}, Chengmin Li³, Jintao Li¹, Renmin Han^{4,*}, Ping Zhu^{2,3,*} and Fa Zhang^{1,*}

¹High Performance Computer Research Center, Institute of Computing Technology Chinese Academy of Sciences, Beijing 100190, China, ²University of Chinese Academy of Sciences, Beijing 100049, China, ³National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China and ⁴Research Center for Mathematics and Interdisciplinary Sciences, Shandong University, Qingdao 266237, China

*To whom correspondence should be addressed.

[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.

Associate Editor: Jinbo Xu

Received on October 29, 2021; revised on December 31, 2021; editorial decision on January 20, 2022

Abstract

Motivation: Cryo-electron microscopy (cryo-EM) is a widely used technology for ultrastructure determination, which constructs the 3D structures of protein and macromolecular complex from a set of 2D micrographs. However, limited by the electron beam dose, the micrographs in cryo-EM generally suffer from the extremely low signal-to-noise ratio (SNR), which hampers the efficiency and effectiveness of downstream analysis. Especially, the noise in cryo-EM is not simple additive or multiplicative noise whose statistical characteristics are quite different from the ones in natural image, extremely shackling the performance of conventional denoising methods.

Results: Here, we introduce the Noise-Transfer2Clean (NT2C), a denoising deep neural network (DNN) for cryo-EM to enhance image contrast and restore specimen signal, whose main idea is to improve the denoising performance by correctly learning the noise distribution of cryo-EM images and transferring the statistical nature of noise into the denoiser. Especially, to cope with the complex noise model in cryo-EM, we design a contrast-guided noise and signal re-weighted algorithm to achieve clean-noisy data synthesis and data augmentation, making our method authentically achieve signal restoration based on noise's true properties. Our work verifies the feasibility of denoising based on mining the complex cryo-EM noise patterns directly from the noise patches. Comprehensive experimental results on simulated datasets and real datasets show that NT2C achieved a notable improvement in image denoising, especially in background noise removal, compared with the commonly used methods. Moreover, a case study on the real dataset demonstrates that NT2C can greatly alleviate the obstacles caused by the SNR to particle picking and simplify the identifying of particles.

Availability and implementation: The code is available at <https://github.com/Lihongjia-ict/NoiseTransfer2Clean/>.

Contact: hanrenmin@gmail.com or zhup@ibp.ac.cn or zhangfa@ict.ac.cn

Supplementary information: [Supplementary data](#) are available at *Bioinformatics* online.

1 Introduction

Cryo-electron microscopy (cryo-EM) is a widely used technology that resolves high-resolution three-dimensional (3D) structures of protein and macromolecular complexes from a series of two-dimensional (2D) micrographs (Bai *et al.*, 2015). However, the signal-to-noise ratio (SNR) of raw cryo-EM images is estimated to be only as high as 0.01–0.1 (Bendory *et al.*, 2020), among the lowest in any imaging field, which extremely decreases the accuracy and efficiency in downstream analysis of cryo-EM images and reduces the

confidence of structures determination. Therefore, an image restoration operation is usually necessary before particle picking, structure segmentation and other cryo-EM data analysis processes to attain high-resolution cryo-EM 3D reconstructions.

A variety of conventional methods have been developed to improve the contrast and decrease the noise level in cryo-EM micrographs, such as BM3D (Dabov *et al.*, 2007), band-pass filter (Penczek, 2010) and Wiener filter (Sindelar and Grigorieff, 2011). For an image restoration algorithm, additional image prior knowledge will be introduced to repair the missing and degenerated

information, in which human knowledge concluded from the natural images is usually used. However, the noise model in the cryo-EM micrograph is usually unknown and varies in different data collection configurations. Therefore, the pre-defined image priors used in these conventional methods cannot correctly fit the noise model in cryo-EM, leading to a limited performance when the conventional methods are applied to cryo-EM data.

Recently, learning-based denoising methods have shown their advantages. [Mao et al. \(2016\)](#) proposed an encoding-decoding framework with symmetric convolutional-deconvolutional layers for image restoration. [Ledig et al. \(2017\)](#) presented a generative adversarial network (GAN) ([Goodfellow et al., 2020](#)) for image super-resolution (SR) which recovers photo-realistic textures from heavily downsampled images. However, most of these learning methods require a clean-noisy paired dataset for training, therefore, cannot be applied to cryo-EM, where ground truth is unavailable. To overcome this barrier, several methods learned from paired noisy images or single noisy images are proposed ([Krull et al., 2019](#)). [Lehtinen et al. \(2018\)](#) presented a general machine learning (ML) framework, called Noise2Noise (N2N), for learning denoising models from paired noisy images. [Chen et al. \(2018\)](#) proposed a GAN-CNN-based framework, GCBD, for learning denoising models from single noisy images where GAN is utilized to build paired training datasets and then convolutional neural network (CNN) is employed for denoising. Specifically, [Bepler et al. \(2020\)](#) proposed a denoiser called Topaz-Denoise for cryo-EM and cryo-electron tomography (cryo-ET), based on an N2N architecture and trained by thousands of cryo-EM micrographs. However, the N2N framework is based on a hypothesis that the noise is zero mean and independent and identically distributed. Therefore, theoretically, the noise removed by TopazDenoise is mainly digitization noise which is caused by the detector in cryo-EM.

In this article, we propose a novel denoising framework, the Noise-Transfer2Clean (NT2C), to restore the specimen signal and enhance the image contrast by learning the unknown noise distribution directly from the pure noise patches in the cryo-EM image. First, a coarse CNN denoiser is trained to enhance the contrast of a cryo-EM image, to distinguish the background and specimen signal. Then, the pure noise patches are extracted from the micrographs and fed into a GAN to estimate and simulate the noise distribution. Finally, a fine denoising network is able to be trained by the clean-noisy pairs simulated from the accurately estimated noise distribution in GAN. By completely mining the noise patterns from pure noise patches, our strategy is able to further decomplex the specimen signal from the noisy background. Especially, to cope with the complex noise model in cryo-EM, we design a contrast-guided noise and signal re-weighted algorithm to achieve clean-noisy data synthesis and data augmentation, making our method authentically achieve signal restoration based on the noise’s true properties. We have tested and compared our denoising model with several commonly used cryo-EM denoising algorithms on both synthetic and real datasets. The experiment results show that NT2C performs particularly excellent in background noise removal and the denoising performance of NT2C can be comparable to current state-of-the-art methods. A case study on particle picking further proves that our denoising method greatly simplifies the difficulty of particle identifying in micrographs with extremely low SNR.

2 Materials and methods

2.1 NT2C protocol

2.1.1 Overview of the procedure

The key idea of NT2C is to learn the noise patterns in cryo-EM images over pure noise patches and transfer the statistical nature of noise into the denoiser. As shown in [Figure 1](#), NT2C contains three modules:

- Noise extraction.** The noise extraction module takes raw micrographs as input, and output the pure noise patches of the background (see [Section 2.2.2](#)). Due to the extremely low SNR in

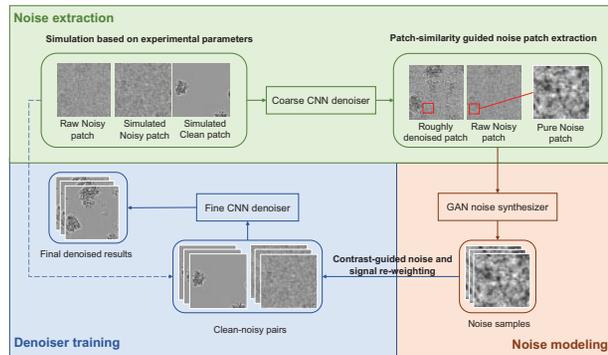


Fig. 1. The overall protocol of NT2C method for cryo-EM image denoising. There are three main modules in NT2C: (a) noise extraction, (b) noise modeling and (c) denoiser training

- Noise extraction.** Distinguishing the background from the particles is a hard task in raw noisy micrographs. Here, a coarse denoiser (see [Section 2.1.2](#)) is trained to roughly enhance the image quality and aid the extraction of noise patches, based on the simulated datasets with the same experimental parameters as the real-world dataset (see [Section 2.2.1](#)).
- Noise modeling.** The statistical properties of noise in cryo-EM micrograph change with different configurations during data collection. It is critical to correctly understand the statistical nature of the noise for image denoising. Here, a GAN noise synthesizer (see [Section 2.1.2](#)) is trained to learn the statistical properties of noise, with pure noise patches as input and simulated noise patches as output.
- Denoiser training.** The noise synthesizer poses the possibility of clean-noisy pair generation for cryo-EM, which is critical in denoiser training. However, as described in [Supplementary Section S1](#) (see [Supplemental Materials](#)), the noise pattern in cryo-EM is quite complex. Here, we design a contrast-guided noise and signal re-weighted algorithm (see [Section 2.2.3](#)) to transfer the non-additive noise to a clean image, to achieve clean-noisy pair synthesis and data augmentation. Based on the abundant synthesized clean-noisy pairs, a fine denoiser (see [Section 2.1.2](#)) is able to be trained to precisely restore specimen signal from the high-level noise.

2.1.2 Main components

(1) CNN denoiser

The CNN denoiser in NT2C is based on a U-net architecture ([Ronneberger et al., 2015](#)), which contains five max-pooling down-sampling blocks and five nearest-neighbor up-sampling blocks, with skip connections between down- and up-sampling blocks at each spatial resolution (shown in [Fig. 2](#)). Given the set of clean-noisy pairs $\{y$ and $x \sim \text{Noise}(y)\}$, a denoising function f with parameter θ can be learned. The loss function for our task is

$$\operatorname{argmin}_{\theta} E_{x \sim X} [\|f_{\theta}(x) - y\|_p] \quad (1)$$

where $p = 2$ is used in NT2C to find f with mean-seeking behavior.

The CNN denoiser has been called twice in NT2C’s procedure: (i) as a coarse denoiser trained with the simulated clean-noisy pairs (see [Section 2.2.1](#)) to roughly enhance the contrast of cryo-EM images, for the ease of noise patch extraction; (ii) as a fine denoiser trained with the clean-noisy pairs produced by the noise and signal re-weighted algorithm (see [GAN noise synthesizer](#) and [Section 2.2.3](#)) to capture the nature of noise statistical properties and restore the specimen signal in cryo-EM micrographs. Here, the model parameters of the coarse denoiser could be transferred to the fine denoiser to avoid retrain from scratch.

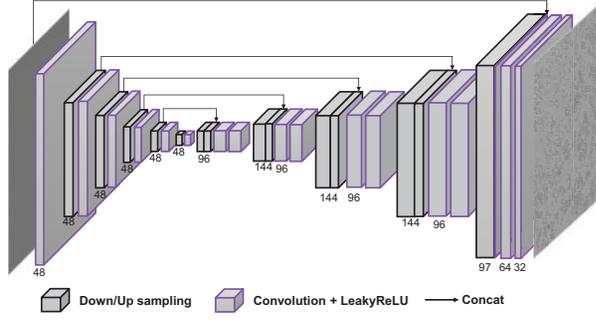


Fig. 2. The network architecture of the CNN denoiser. The U-net model consists of five convolutional and down-sampling blocks followed by five convolutional and up-sampling blocks. Skip connections link each down-sampling block to the mirrored up-sampling block

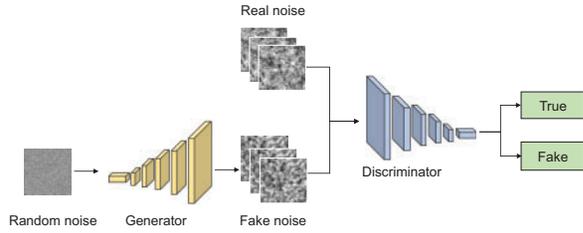


Fig. 3. The network architecture of the GAN noise synthesizer. The generative network is trained to generate noise samples, while the discriminative network is trained to determine whether a noise sample is from real data or the generative network

(2) GAN noise synthesizer

The patterns and quantity of noise blocks extracted from raw micrographs are limited, especially when input noisy images is not enough or the biological samples are densely distributed in micrographs. Moreover, the noise model in cryo-EM is too complex to be explicitly described by an analytical expression. Therefore, we build a GAN framework to implicitly learn the latent noise model in cryo-EM micrographs and generate more noise samples. In this framework, the convolutional layer, which can learn image features better than the multi-layer perceptron used in standard GAN, is adopted as the basic network structure of GAN. And to solve the unstable training problem of GAN, we use Wasserstein distance and gradient penalty to construct a cost function to achieve stable convergence of the network. Moreover, the Batch Normalization (Ioffe and Szegedy, 2015) is used to further secure the model stability and Leaky Rectified Linear Unit (LeakyReLU) (Maas et al., 2013) activation is used to ensure fast learning.

As shown in Figure 3, this framework contains two components, a generative network that consists of six transposed convolutional layers, and a discriminative network that consists of six convolutional layers. The generative network is trained to generate noise samples while the discriminative network is trained to determine whether a sample is from real data or the generative network. After the convergence of adversarial learning, the generative network will be able to produce noise patches hard to be distinguished from real noise patches. The loss function for our task is

$$\mathcal{L}_{GAN} = \mathbb{E}_{\tilde{x} \sim \mathbb{P}_{\tilde{x}}} [D(\tilde{x})] - \mathbb{E}_{x \sim \mathbb{P}_r} [D(x)] + \lambda \mathbb{E}_{\tilde{x} \sim \mathbb{P}_{\tilde{x}}} \left[\left(\|\nabla_{\tilde{x}} D(\tilde{x})\|_2 - 1 \right)^2 \right] \quad (2)$$

where \mathbb{P}_r is the distribution over noise patches, $\mathbb{P}_{\tilde{x}}$ is the generator distribution, $\mathbb{P}_{\tilde{x}}$ is defined as a distribution sampling uniformly along straight lines between pairs of points sampled from \mathbb{P}_r and $\mathbb{P}_{\tilde{x}}$.

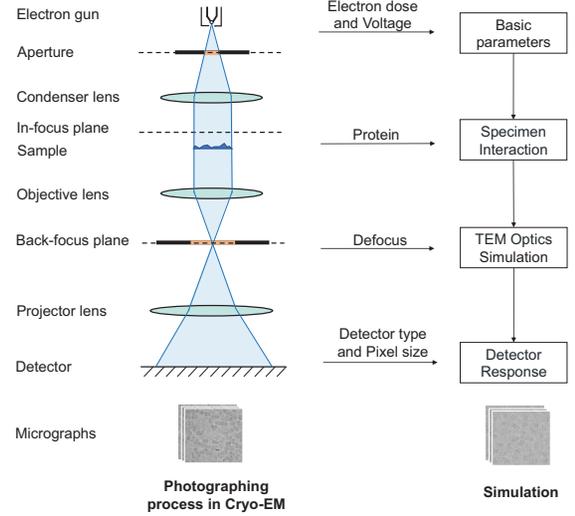


Fig. 4. Simulation based on experimental parameters. The experimental parameters in photographing process (left) are used for simulation (right), including electron dose, voltage, defocus, pixel size and detector type. The homologous protein with similar size is utilized for sample simulation

2.2 Detailed algorithm

2.2.1 Simulation based on experimental parameters

The lack of supervised training data hampers the application of learning-based denoising method in cryo-EM. Here, we adopt the simulation software InSilicoTEM (Vulović et al., 2013), which can simulate the photographing process in cryo-EM based on physical principles, to generate paired clean-noisy datasets.

We set the simulation according to the experimental parameters used in data collection (shown in Fig. 4), including the pixel size, defocus, voltage, electron dose and detector type, which make the limited resolution, contrast transfer function (Wade, 1992) and modulation in the simulation very close to the real-world data. The proteins downloaded from Protein Data Bank (PDB) (Burley et al., 2017) are used to produce clean ground truth. Such a simulation is possible to generate datasets with statistical properties of noise close to the ones in real cryo-EM micrographs. Consequently, the composed clean-noisy pairs could be fed into the coarse CNN denoiser to produce a model that achieves roughly denoising and contrast enhancement on the cryo-EM image, for the ease of noise patch extraction. Subsequently, the paired simulated data will be reweighted with the noise samples generated by GAN noise synthesizer to construct the final training set for the fine CNN denoiser.

2.2.2 Patch-similarity guided noise patch extraction

If we divide a cryo-EM micrograph into patches with suitable size, these patches could be classified into two categories: patch containing specimen signal or patch of background with pure noise (shown in Fig. 5). Naturally, the background patches are homogeneous to each other while the patches containing specimen signals have different patterns. Because the ice is almost transparent, the background patch represents pure noise.

Here, we proposed a patch-similarity guided algorithm to extract the noise patches in a cryo-EM micrograph:

1. Given a micrograph I , denoise I with the CNN coarse denoiser trained by the simulated data in Section 2.2.1 to get an enhanced image I' ;
2. Divide I' into a set of overlapping patches $\Theta = \{P_i\}$ ($d \times d$ pixel² per patch) with a step size of s ;
3. For each patch $P_i \in \Theta$, further divide P_i into N local patches $\{P_{i,k}\}$ and calculate the structural similarity (SSIM) between each P_{i,k_1} - P_{i,k_2} pairs ($k_1 \neq k_2$).

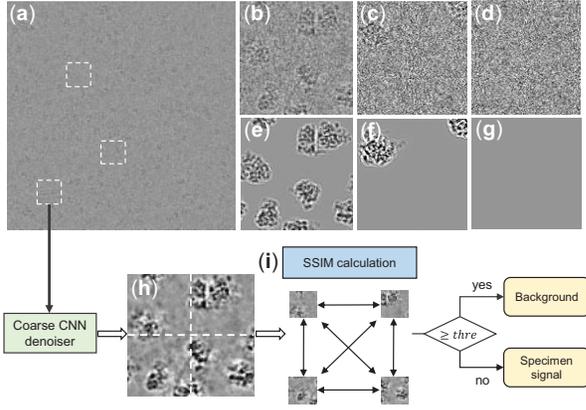


Fig. 5. The illustration of the noise extraction algorithm based on patch similarity. (a) An illustration of the raw micrograph. (b–d) Three representative patches selected from the raw micrograph. (e–g) The clean corresponding specimen signals of the selected patches. (h) A patch enhanced by the coarse CNN denoiser trained with simulator; the patch is divided into N local patches with $N = 4$ in here. (i) The calculation of structural similarity (SSIM) and the determination of pure noise patches

4. For the patch P_i , if $\forall k_1, k_2 (k_1 \neq k_2)$, $SSIM(P_{i,k_1}, P_{i,k_2})$ is large than a given threshold $thre$, determine P_i as a background patch;
5. Repeat 1–3 until all the background patches are identified, extract the exact patches in the original micrograph I .

Here, the commonly used value of patch size d is 320, the step size is set to half of the patch size, the value of N is set to 4, and the default value of $thre$ is set to 0.7. All the hyperparameters can be adjusted according to the properties of the datasets. It should be noted that the similarity determination is operated on micrograph I' but the noise patch is extracted from the original micrograph I . The similarity measurement $SSIM(P_{i,k_1}, P_{i,k_2})$ used in our algorithm is defined as

$$SSIM(P_{i,k_1}, P_{i,k_2}) = \frac{(2\mu_{P_{i,k_1}}\mu_{P_{i,k_2}} + c_1)(2\sigma_{P_{i,k_1}P_{i,k_2}} + c_2)}{(\mu_{P_{i,k_1}}^2 + \mu_{P_{i,k_2}}^2 + c_1)(\sigma_{P_{i,k_1}}^2 + \sigma_{P_{i,k_2}}^2 + c_2)} \quad (3)$$

where $\mu_{P_{i,k_1}}$ and $\sigma_{P_{i,k_1}}$ are the mean and standard deviation of P_{i,k_1} , $\mu_{P_{i,k_2}}$ and $\sigma_{P_{i,k_2}}$ are the mean and standard deviation of P_{i,k_2} , $\sigma_{P_{i,k_1}P_{i,k_2}}$ are the cross-covariance between patch P_{i,k_1} and P_{i,k_2} , c_1 and c_2 are the regularization constants with very small values to avoid the extreme small denominator. According to the study of Wang *et al.* (2004), $c_1 = (K_1L)^2$, $c_2 = (K_2L)^2$ and $L = 255$ for 8-bit images, with $K_1 = 0.01$ and $K_2 = 0.03$. In NT2C, the pixel values of an image are normalized to $[0, 1]$, and L is set to 1 to get the same result.

2.2.3 Contrast-guided noise and signal re-weighting

Though we have designed an experimental parameters-based simulation to produce simulated clean-noisy pairs (Section 2.2.1), these simulated data are still not good enough to present the noise pattern in real cryo-EM images, which is more complex than the principles in the simulation. On the contrary, the GAN noise synthesizer (Section 2.1.2) is able to produce noise patches with almost the same statistical properties as the real noise in the cryo-EM image. Here, we design a contrast-guided noise and signal re-weighted algorithm to transfer the noise pattern produced in the GAN noise synthesizer to the simulated clean data, to produce sophisticated clean-noisy pairs for fine CNN denoiser training.

Figure 6 shows the detailed process of the re-weighted algorithm. The algorithm accepts the pure noise patches $\{V_i\}$ generated from GAN and the clean signal patches $\{S_i\}$ generated from the simulation as input, utilizing the simulated patches $\{X_i\}$ corresponding to the exact signal patches from the simulation as a reference, and

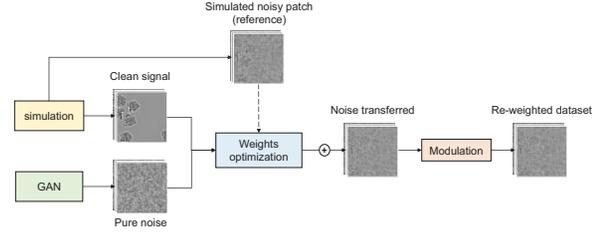


Fig. 6. Contrast-guided noise and signal re-weighting. The noise pattern produced in the GAN noise synthesizer and the simulated clean signal are re-weighted with optimal coefficients and then modulated to produce clean-noisy pairs

outputs a re-weighted dataset $\{Y_i\}$ with the GAN-synthesized noise transferred to the clean simulated signal. Here, we use the contrast of X_i as a baseline. Denote $Y_i = \mathcal{F}(\alpha * S_i + \beta * V_i + \gamma)$ ($\mathcal{F}(\cdot)$ is a modulation function), the noise transfer should comply with the following objective function:

$$\min_{\alpha, \beta, \gamma} \sum_{m=0}^{d-1} \sum_{n=0}^{d-1} \left\| \left(\alpha * S_i(m, n) + \beta * V_i(m, n) + \gamma \right) - X_i(m, n) \right\|_2^2 \quad (4)$$

where α , β and γ are scalar coefficients and d is the patch size. Such a minimizing problem can be easily solved by the least-square method. Then, with the solved coefficients, the signal will be re-weighted and modulated, to produce clean-noisy pairs. The constructed clean-noisy pairs are fed into a fine CNN denoiser to train the model that captures the true noise statistics and restores specimen signals from noise.

2.3 Datasets

Four simulated datasets and three real-world datasets are used to evaluate the performance of NT2C. We generated four simulated datasets by 2wrj.pdb, 1kd1.pdb, 5lzf.pdb and 1gr5.pdb to demonstrate the performance of NT2C on simulated datasets and denoted them as SIM1, SIM2, SIM3, SIM4. The real-world datasets are collected from public repositories. We downloaded the EMPIAR-10025 (abbr. EM25) (Campbell *et al.*, 2015), EMPIAR-10028 (abbr. EM28) (Wong *et al.*, 2014) and EMPIAR-10077 (abbr. EM77) (Fischer *et al.*, 2016) from EMPIAR and generated three corresponding simulated datasets by 1iru.pdb, 1i96.pdb and 6w6p.pdb, marked as SIM-EM25, SIM-EM28 and SIM-EM77. These three simulated datasets are used for the training of coarse CNN denoiser and real-world datasets are used to demonstrate the performance of NT2C on real-world noise. For each dataset, we randomly extracted 10% of images as validation images, 10% of images as test dataset and the remaining as training dataset. The detailed information of these datasets is summarized in Supplementary Table S1.

3 Experiments and results

We designed a series of experiments to demonstrate the effectiveness of NT2C on the simulated datasets (see Supplementary Materials) and real-world datasets (see Section 3.1), and compared it with several commonly used cryo-EM denoising methods. We further explored the effect of denoising on particle picking tasks with a real-world dataset (see Section 3.2). In addition, we provide an experiment to verify the robustness of simulation-based noise pattern discovery, which is a critical part of NT2C (see Supplementary Materials). Finally, an ablation study is provided to illustrate the important contributions from individual components of NT2C (see Supplementary Materials).

3.1 Evaluation with real noise

To demonstrate NT2C's ability to deal with complex latent noise in real cryo-EM images, we evaluated the performance of NT2C on two real-world datasets (EM25 and EM28). We compared the performance of NT2C with four mainstream cryo-EM denoising

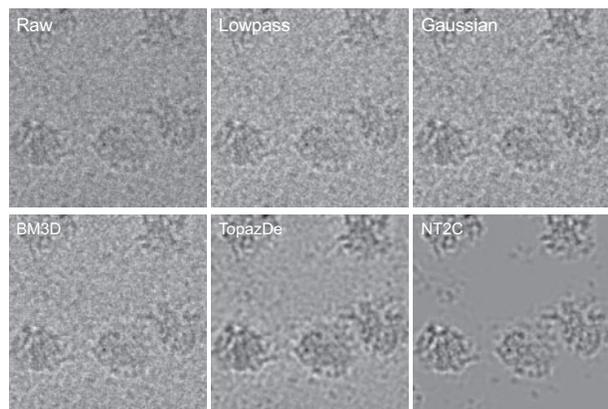


Fig. 7. Denoising with NT2C improves micrograph interpretability in real-world datasets. Comparison among different denoising methods is carried out on real-world datasets EM28. A small region is selected to illustrate that NT2C performs better in noise smoothing and signal enhancing than both conventional (Lowpass filter, Gaussian filter, BM3D) and learning-based methods (TopazDe)

methods, including three conventional methods, i.e. Low-pass filter, Gaussian filter (Haddad and Akansu, 1991), BM3D and learning-based methods, i.e. Topaz-Denoise (The pre-trained model provided by Topaz-Denoise is denoted as 'TopazDe-gen' and the Topaz-Denoise model retrained with a specific dataset is denoted as 'TopazDe'. Since EM25 and EM28 are included in the training dataset for Topaz-Denoise general model, the TopazDe is adopted for EM25 and EM28. While EM77 is not included and the raw movies are unavailable, the TopazDe-gen is adopted for EM77.).

Denoising with NT2C improves micrograph interpretability and SNR. Figure 7 shows a representative region selected from EM28 and denoised results of Low-pass filter (2x binning), Gaussian filter, BM3D, TopazDe and NT2C. It can be seen that two learning-based methods, TopazDe and NT2C are better than conventional methods in noise smoothing and specimen signal enhancing. Moreover, NT2C achieves a stronger noise removal performance and reserved clearer specimen signals than N2N, therefore providing better interpretability. The whole-image results are shown in Supplementary Figure S1.

Furthermore, we quantitatively assessed denoising performance by measuring the SNR of raw micrographs and micrographs denoised with different methods. Due to the non-existent ground truth, the SNR is estimated in a similar way to Bepler et al. (2020). First, we select 10 paired signal and background regions across 10 micrographs where the background regions are as close as possible to the corresponding signal region. Given N signal and background pairs, x_s^i, x_b^i , the mean and variance of each background region is marked as μ_b^i, v_b^i . We define the signal for each region as $s^i = x_s^i - \mu_b^i$ and calculate the mean and variance of signal region, μ_s^i, v_s^i . The average SNR in dB for the regions is defined as:

$$\text{SNR} = \frac{10}{N} \sum_{i=1}^N \log_{10}(v_s^i) - \log_{10}(v_b^i) \quad (5)$$

This SNR has no physical meaning, just criteria for comparison. As shown in Table 1, the conventional methods only improve roughly 0.1 dB over raw micrographs. NT2C method improves SNR by 8 dB over raw micrographs and roughly 6 dB over TopazDe methods. It should be noted that since the SNR metrics and the signal and noise regions selected by our method and TopazDenoise for SNR calculations are different, the SNR values of the same dataset reported in our work may be different from that reported in Bepler et al. (2020). However, this does not affect the evaluation of denoising performance and the comparison among the denoising algorithms.

NT2C accurately resolve complex noise model and restore clear specimen signal. To further study the NT2C's performance on noise removal and specimen signal restoration, we selected two

Table 1. Comparison of denoising methods based on estimated SNR (in dB, larger is better)

Methods	Dataset		
	EM25	EM28	EM77
Raw	-0.16	-0.34	0.14
Lowpass	-0.04	-0.24	0.30
Gaussian	-0.14	-0.18	0.33
BM3D	0.07	-0.27	0.53
TopazDe-gen	—	—	1.24
TopazDe	0.21	1.04	—
NT2C	4.88	7.65	6.29

representative regions from dataset EM25, one containing specimen signals and the other containing pure noise. Figure 8 presents the raw micrograph and denoised results of Lowpass filter, Gaussian filter, BM3D, TopazDe and NT2C. It can be found that NT2C performs outstandingly in background noise removing, where the pure noise region denoised by NT2C is cleaner than all other methods. TopazDe removes most of the noise and the conventional methods achieve the poorest performance on noise smoothing. Moreover, as shown in the region containing specimen signal, while removing noise thoroughly, NT2C correctly decomplex structured features from complex noise. The signal restored by NT2C presents distinct structures and the particles with different projections are easily distinguished. Table 1 gives quantitative analysis on SNR, which further proves the notable performance achieved by NT2C. It improves SNR by 4.91 dB, 4.54 dB and >4.6 dB over raw micrograph, TopazDe and conventional methods.

3.2 Denoising with NT2C makes the task of particle picking much easier

To free researchers from laborious particle picking work, a number of fully automatic and semi-automatic particle picking approaches have been proposed (Bepler et al., 2019; Zhang et al., 2017). However, the performance of particle picking is limited by many factors, among which the extremely low SNR of the micrographs is an important one, especially for the fully automatic or conventional particle picking algorithms. It can be seen from the above experiments that one prominent advantage of the NT2C denoising algorithm is that it can remove background noise thoroughly. Therefore, the particles and the background can be easily distinguished. This provides a possibility that if combined with NT2C denoising, the particle picking may become a very simple task.

Here, we designed a set of experiments to combine the denoising algorithm with a fully automatic particle picking algorithm, PIXER (Zhang et al., 2019) and choose TopazDe as a comparison. In PIXER, the author designed a segmentation network to convert the noisy micrographs to probability density maps and the probability indicates the likelihood of one pixel belonging to a particle. Then the preliminary particle coordinates can be generated from probability density maps. The real data EM77, where the SNR is extremely low and the particles are difficult to identify, is adopted here. Figure 9 shows the raw and denoised micrographs (the first column), the heat map (probability density maps) generated by PIXER (the second column), the clusters calculated by meanshift algorithm (Comaniciu and Meer, 2002) based on the segmented heat map (the third column), and the particle picking results (the last column).

It can be found that the particles in the raw micrograph can hardly be recognized from noisy backgrounds. TopazDe-gen greatly enhances the image contrast, while the contour of the particles is not clear enough and the background is not clean. However, the NT2C significantly increases protein density confidence and almost completely removes the background noise, making the particles easy to identify. It can be seen from the heatmap and clustered results, the particle and the background cannot be accurately segmented, making the particles in the raw micrograph difficult to be identified.

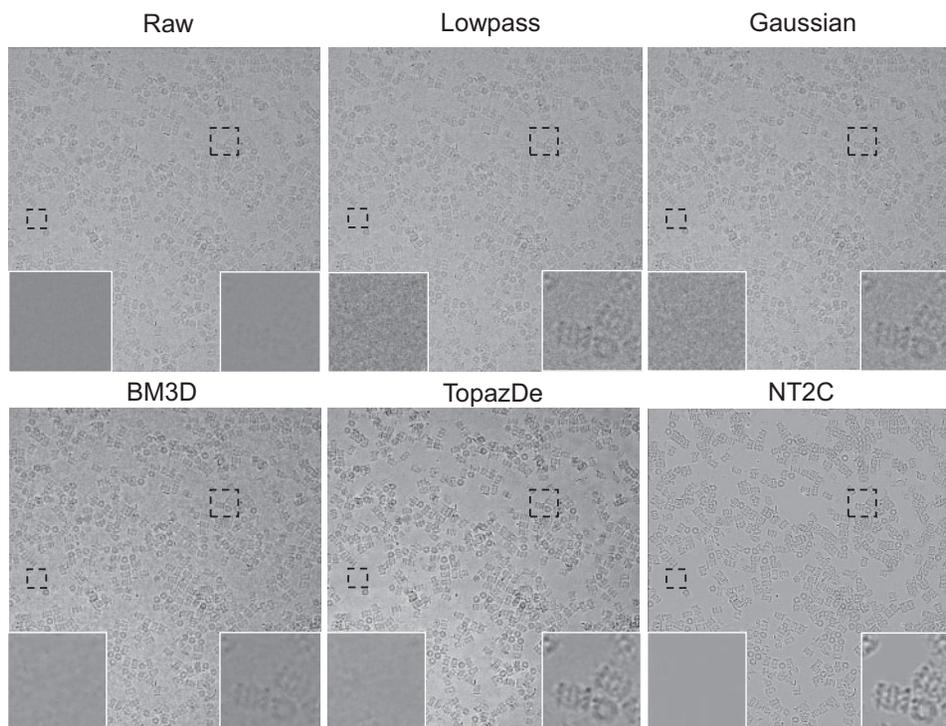


Fig. 8. NT2C accurately resolve complex noise model and restore clear specimen signal. Two regions, one containing pure noise and the other containing specimen signal, are selected from one micrograph of EM25 to illustrate the performance of NT2C. Comparing with Lowpass filter, Gaussian filter, BM3D and TopazDe, NT2C can thoroughly remove noise and decomplex structured features from noise

Although the particles denoised by TopazDe can be segmented out by PIXER, affected by residual background noise, there are many false-positive particles picked, as the red cross labeled ones shown in Figure 9. On the contrary, the particles in the micrograph denoised by NT2C are easily and accurately identified.

Inspired by the excellent denoising performance of NT2C, we tested the effect of directly combining the denoising algorithm with the meanshift clustering algorithm to locate particles (see Supplementary Materials). As shown in Supplementary Figure S3, it is difficult to obtain particles directly from the raw micrograph and the micrograph denoised by TopazDe with the meanshift clustering algorithm. On the contrary, most of the particles in micrograph denoised by NT2C can be accurately identified. Although there are false-positive ones, the accuracy of particle locating is even better than the result of the combination of TopazDe and PIXER (Fig. 9H). It can be proved that NT2C provides great possibilities for simplifying the task of particle picking.

4 Discussion

In this work, we proposed a new idea for cryo-EM image denoising which learns the unknown noise distribution directly from the pure noise patches in cryo-EM and transfer the statistical nature of noise into the denoiser. It has achieved impressive denoising performance and performs especially outstanding in noising removal.

As described in Baxter et al. (2009), the noise in cryo-EM comes from three aspects: (i) structural noise which is caused by the ice matrix around molecules and the superimposed carbon film. Conceptually, any part of the molecule structure that is not reproducible due to conformational variations is also counted as structural noise. (ii) Shot noise caused by the quantum nature of the electron radiation. (iii) Digitization noise caused by the photographic recording and subsequent digitization. However, the currently most commonly used denoising frameworks for cryo-EM, such as Noise2Noise which was adopted by TopazDenoise or Noise2Void, are based on a hypothesis: the noise is zero mean and independent and identically distributed. Therefore, theoretically, this kind of

algorithm will regard structural noise and shot noise as signals and mainly remove digitization noise which is caused by the detector. Conversely, our algorithm directly learns the noise distribution from the noise patches in cryo-EM images, therefore, it can mine noise patterns more completely and implement modeling for all types of noise, achieving more thorough removal of noise. This is the reason our algorithm achieved outstanding performance on noise removal and signal recovery.

We believe that our method is inspirational and has great potential, however, the performance of NT2C is closely related to the difference between simulated data and real data. Therefore, we are working hard to find optimal simulation software. At present, the simulation software adopted in this work makes our algorithm practical, however, it brings two limitations. First, since our algorithm requires the generation of simulated clean-noisy pairs, it brings a problem that our algorithm takes more time than other denoising algorithms. Second, the real noise model in cryo-EM images is rather complicated and affected by various factors, such as uneven ice thickness, uneven distribution of samples and different under-focus values. However, the existing simulation software generally generates micrographs based on the physical model under ideal conditions which limit our method to mainly perform excellently on samples with relatively uniform backgrounds. Therefore, our following work includes further exploring the complex relationship between image formation theory adopted in simulation and the realistic imaging process of cryo-EM to develop a more professional simulation software. We are also attempting to adopt deep learning techniques to solve the problem of simulated data generation, thus improving NT2C's ability on the cryo-EM data collected in various complex situations.

It should be noted that the framework is flexible, therefore the components, such as GAN noise synthesizers and coarse CNN denoiser, can be replaced by better models or methods in the future. Our study is designed to focus specifically on denoising ability this time and we will continuously optimize the algorithm in the future to provide a more general model for cryo-EM denoising.

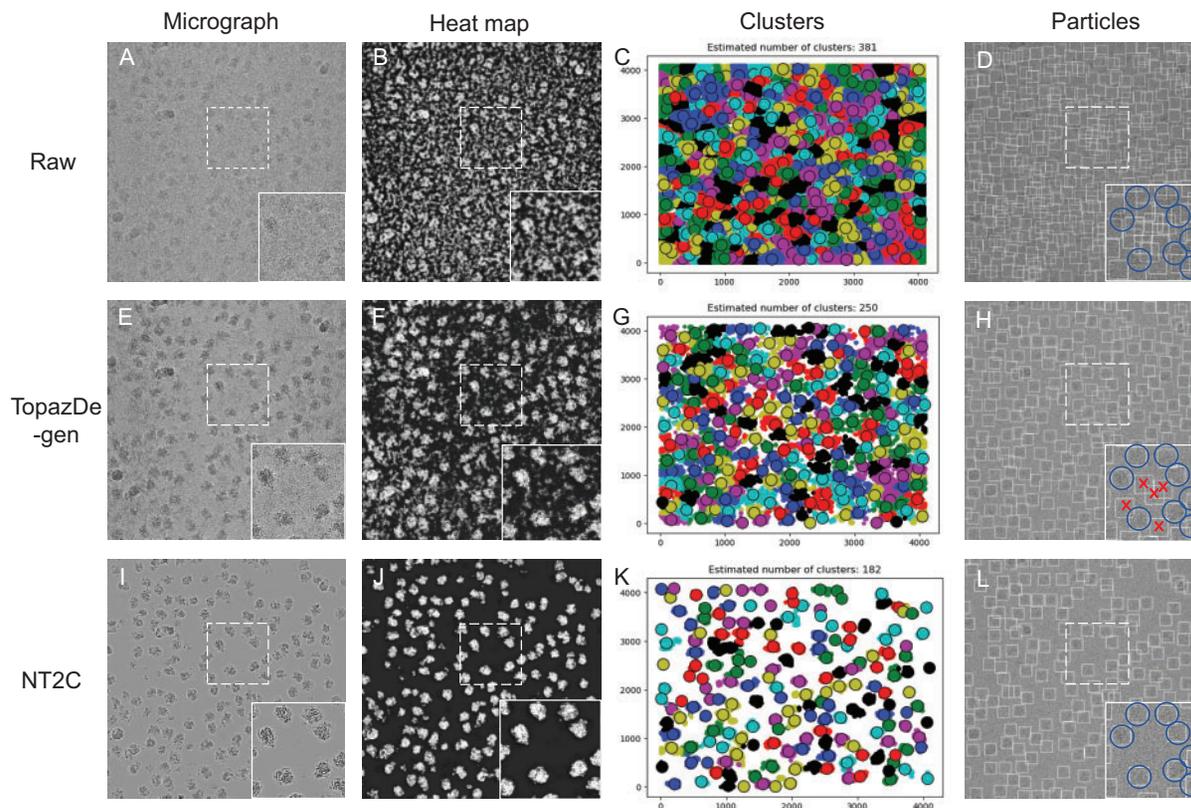


Fig. 9. The particle picking results on EM77 when combine denoising algorithm (TopazDe and NT2C) with particle picking algorithm (PIXER). The micrograph, heat map generated by PIXER, the clusters calculated by meanshift algorithm and the picked particles are shown in columns one to four. The one to three rows correspond to raw micrograph, micrograph denoised by TopazDe-gen and NT2C, respectively. The region boxed by the white solid line is an enlargement of the region boxed by white dotted line. The particles boxed by blue circle are true particles, and the particles labeled by red cross are false-positive ones

5 Conclusion

In this article, we presented a denoising framework for image contrast enhancement and specimen signal restoration in cryo-EM. The key idea of NT2C is discovering the noise model of cryo-EM images over pure noise patches and transferring the statistical nature of noise into the denoiser, making the denoising based on noise's true properties. To cope with the complex noise model in cryo-EM, we further design a contrast-guided noise and signal re-weighted algorithm to achieve clean-noisy data synthesis and data augmentation for denoiser. Our work has verified the feasibility of denoising based on capturing the nature of noise statistical properties directly from noise patches in cryo-EM images. Comprehensive experiments on both simulated and real-world datasets demonstrate that NT2C is able to deal with high-level complex noise in cryo-EM images. A case study further demonstrates that NT2C can significantly reduce the difficulty of particle picking and even make it possible to use conventional image processing algorithms to achieve particle identifying.

Funding

This research was supported by the National Key Research and Development Program of China [2017YFA0504700, 2020YFA0712401 and 2021YFF0704300], National Natural Science Foundation of China projects Grant [61932018, 62072280, 62072441, 31730023 and 31521002], the Chinese Academy of Sciences (CAS) [XDB37010100] and the National Laboratory of Biomacromolecules of China [2019KF07].

Data availability statement

All experimental data underlying this article are freely available in EMP IAR at <https://www.ebi.ac.uk/empiar/>.

Conflict of Interest: none declared.

References

- Bai,X. et al. (2015) How cryo-EM is revolutionizing structural biology. *Trends Biochem. Sci.*, **40**, 49–57.
- Baxter,W.T. et al. (2009) Determination of signal-to-noise ratios and spectral SNRs in cryo-EM low-dose imaging of molecules. *J. Struct. Biol.*, **166**, 126–132.
- Bendory,T. et al. (2020) Single-particle cryo-electron microscopy: mathematical theory, computational challenges, and opportunities. *IEEE Signal Process. Mag.*, **37**, 58–76.
- Bepler,T. et al. (2019) Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat. Methods*, **16**, 1153–1160.
- Bepler,T. et al. (2020) Topaz-Denoise: general deep denoising models for cryo-EM and cryoET. *Nat. Commun.*, **11**, 1–12.
- Burley,S.K. et al. (2017) Protein Data Bank (PDB): the single global macromolecular structure archive. *Protein Crystallogr.*, 627–641.
- Campbell,M.G. et al. (2015) 2.8 Å resolution reconstruction of the *Thermoplasma acidophilum* 20S proteasome using cryo-electron microscopy. *Elife*, **4**, e06380.
- Chen,J. et al. (2018) Image blind denoising with generative adversarial network based noise modeling. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Salt Lake City, Utah, pp. 3155–3164.
- Comaniciu,D. and Meer,P. (2002) Mean shift: a robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, **24**, 603–619.
- Dabov,K. et al. (2007) Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.*, **16**, 2080–2095.
- Fischer,N. et al. (2016) The pathway to GTPase activation of elongation factor SelB on the ribosome. *Nature*, **540**, 80–85.

- Goodfellow, I. et al. (2020) Generative adversarial networks. *Commun. ACM*, **63**, 139–144.
- Haddad, R.A. and Akansu, A.N. (1991) A class of fast Gaussian binomial filters for speech and image processing. *IEEE Trans. Signal Process.*, **39**, 723–727.
- Ioffe, S. and Szegedy, C. (2015) Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. In: International Conference on Machine Learning. PMLR, Lille, France, pp. 448–456.
- Krull, A. et al. (2019) Noise2void-learning denoising from single noisy images. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Long Beach, CA, pp. 2129–2137.
- Ledig, C. et al. (2017) Photo-realistic single image super-resolution using a generative adversarial network. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, Hawaii, pp. 4681–4690.
- Lehtinen, J. et al. (2018) Noise2noise: Learning image restoration without clean data. In: International Conference on Machine Learning. PMLR, Stockholm, Sweden, pp. 2965–2974.
- Maas, A.L. et al. (2013) Rectifier nonlinearities improve neural network acoustic models. *Proc. ICML*, **30**, 3.
- Mao, X.J. et al. (2016) Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In: Advances in neural information processing systems. Barcelona, Spain, pp. 2802–281.
- Penczek, P.A. (2010) Image restoration in cryo-electron microscopy. *Methods Enzymol.*, **482**, 35–72.
- Ronneberger, O. et al. (2015) U-net: convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany, pp. 234–241.
- Sindelar, C.V. and Grigorieff, N. (2011) An adaptation of the Wiener filter suitable for analyzing images of isolated single particles. *J. Struct. Biol.*, **176**, 60–74.
- Vulović, M. et al. (2013) Image formation modeling in cryo-electron microscopy. *J. Struct. Biol.*, **183**, 19–32.
- Wade, R.H. (1992) A brief look at imaging and contrast transfer. *Ultramicroscopy*, **46**, 145–156.
- Wang, Z. et al. (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.*, **13**, 600–612.
- Wong, W. et al. (2014) Cryo-EM structure of the *Plasmodium falciparum* 80S ribosome bound to the anti-protozoan drug emetine. *Elife*, **3**, e03080.
- Zhang, F. et al. (2017) A two-phase improved correlation method for automatic particle selection in Cryo-EM. *IEEE/ACM Trans. Comput. Biol. Bioinf.*, **14**, 316–325.
- Zhang, J. et al. (2019) PIXER: an automated particle-selection method based on segmentation using a deep neural network. *BMC Bioinformatics*, **20**, 1–14.