Denoising Natural and Cryo-EM Images without Clean Data

Zhenyu Tan
Life Sciences Institute
University of Michigan
Ann Arbor, MI 48105
zhenytan@umich.edu

Liangqing Cui
Department of Electrical Engineering
University of Michigan
Ann Arbor, MI 48105
licu@umich.edu

Xin Luo
Department of Electrical Engineering
University of Michigan
Ann Arbor, MI 48105
xinluo@umich.edu

***Abstract*— We consider the problems of image denoising using Convolutional Neural Network(CNN). We apply basic machine leaning statistical reasoning to mapping corrupted data to clean version with a assumption that: clean images can be retrieved by only using corrupted data. We validate the potential of the noise2noise method through two different denoising cases: the natrual images and cryo-EM images. We can achieve the good denoising performance by only using the noisy data.**

I. INTRODUCTION

Images captured by daily photography or scientific experiment are corrupted by additive noise. Reducing the noise level in image is critical for the interpretation of the data. Specifically, we are focusing on the denoising the cryo-EM images. Cryo-electron microscopy (cryo-EM) is a powerful structural biology technique capable of determining atomic-resolution structures of biological macromolecules. Despite this ability, the low signal-to-noise ratio of cryo-EM image data continues to remain a hurdle for assessing raw cryo-EM micrographs and subsequent image analysis. Recently advance in deep learning, which train a regression model with large numbers of pairs (corrupted images, clean targets) has shown the ability to recover the information with a performance which far exceeds simple deconvolution. However, it is impossible to obtain the clean data with long exposure in cryo-EM due to the radiation damage to the sample from the electron beam. To address this longstanding problem, we are working to finding a feasible method for denoising images by only use the noisy images. Inspired by the paper: 'Noise2Noise: Learning Image Restoration without Clean Data', we are going to adapt the method to denoise the cryo-EM image data.

II. IMAGE DENOISING WITHOUT CLEAN DATA

A. Noise2Noise Theoretical Background

The classical way of training a neural network is by using the input and the desired output pair in the training process. However, clean images (noisy-free) are difficult to obtain, like in cryo-electron microscopy. Lack of clean data was expected to greatly hinder the performance of the image denoising neural network, since no clean reference is set as target for neural network output. This problem was solved by a recent research[1], which has demonstrated the feasibility of denoising images using only noisy data. They has mathematically proved the feasibility of this method, and such model is called noise2noise. To help understanding the idea behind noise2noise, one of examples is estimating room temperature y from given multiple measurements $(y_1, y_2, y_3, ...)$.Since estimate prediction is a random variable, the most common strategy is to find a number that has smallest average deviation from given measurement according to some loss function L:

$$\underset{z}{\operatorname{argmin}} \mathbb{E}_y[L(z, y)] \tag{1}$$

L2 loss:$L(z, y) = (z-y)^2$ is the most common used loss and the corresponding optimal solution has form of $z^* = \mathbb{E}_y[y]$. Now suppose we added some noise N to the measurement$y$, then the measurement becomes $(\hat{y_1}, \hat{y_2}, \hat{y_3}, ...)$, where $\hat{y} = y + n$ is a new random variable. Note that if the noise has zero value expectation and is independent of data y, then the optimal solution for noise added measurement become $z^*_{new} = \mathbb{E}_{\hat{y}}[\hat{y}] = \mathbb{E}_y[y]$, which is exactly the same solution as the previous one in no noise added training problem! This amazing result inspire us a way of estimating ground truth without using clean data.

In image denoising task, the general idea is that by inspecting noisy image as $\hat{x} = y + n$, we could train a neural network capable of retrieve clean data y from given data pairs$(\hat{x_i}, y_i)$. Generally, the training process can be mathematically represented as:

$$\underset{\theta}{\operatorname{argmin}} \mathbb{E}_{(\hat{x},y)}[L(f_\theta(\hat{x}, y))] \tag{2}$$

By applying law of iterated expectation, the above problem can be decomposed into point estimation problem like:

$$\underset{\theta}{\operatorname{argmin}} \mathbb{E}_{\hat{x}}[\mathbb{E}_{y|\hat{x}}[L(f_\theta(\hat{x}), y)]] \tag{3}$$

Since for each training data $\hat{x_i}$, the corresponding optimum has form of $f^*_\theta(\hat{x}) = \mathbb{E}_{y|\hat{x}}[y]$ for the same reason discussed above. If we replace target image y with corrupted version $\hat{y}$: $\hat{y} = y + n$, as long as the expectation of the noise is zero, the new optimal solution has form of $f^*_\theta(\hat{x}) = \mathbb{E}_{y|\hat{x}}[y + n] = \mathbb{E}_{y|\hat{x}}[y]$. The solution is not changed, even if the noisy data is used as the target for training.

One conclusion drawn from this analysis is that minimizing L2 loss is equal to retrieving expectation of target, and we can simply use the noisy data as the optimization target in our

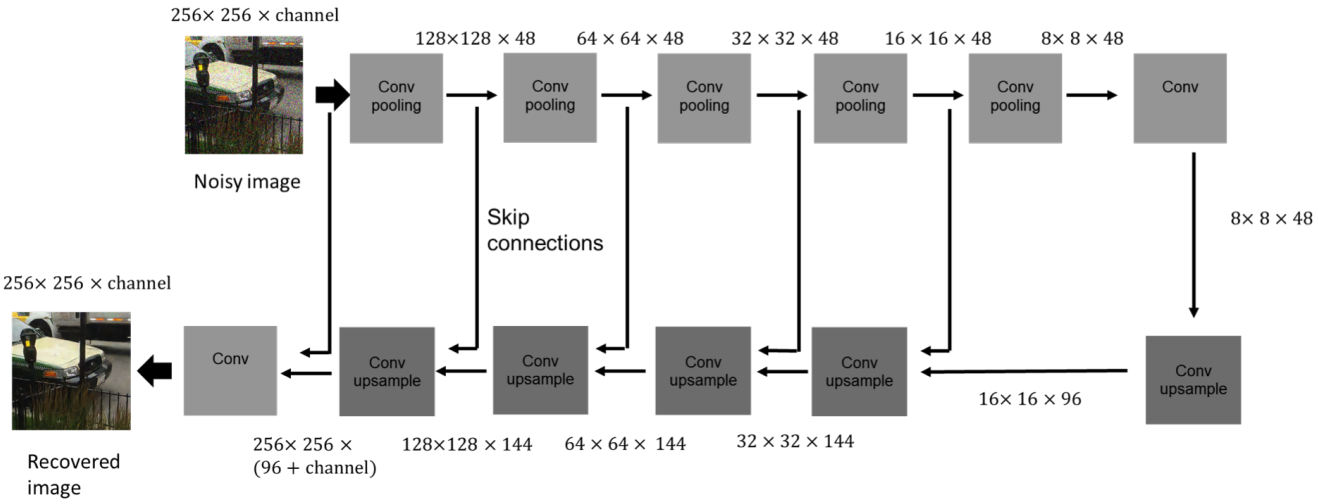

Fig. 1: The U-Net architecture used in our methods

convolutional neural network without worrying about wrong prediction.

B. Neural Network Architecture and Training Details

We implement noise2noise method using Pytorch and applied it to two different data set. One is the normal photography images. We intentionally synthesize different additive gaussian white noise on it and test the denoising performance of the noise2noise neural network. After that we apply the same idea to experiment-obtained cryo-EM images, which ground truth is not readily available. As for architecture, the neural network we are using is a generic U-Net (Fig.1). For testing on the natural RGB images, the input and output channels were both 3, while on cryo-EM images denoising task, we use 1 for input and output channel. Input images were normalized to $[0, 1]$. No batch normalization, dropout or other regularization techniques were used. In training procedure, ADAM was used with parameter values $\beta_1 = 0.9$,$\beta_2 = 0.99$,$\epsilon = 10^{-8}$. We set learning rate as 0.001. Minibatch size of 4 was used for all experiments.

C. Performance Evaluation

For the first case(natural image denoising), we evaluate our neural network performance by inspecting peak signal to noise ratio. Peak signal to noise ratio(PSNR) represent a measure of the peak error and is defined as $PSNR = 10log_{10}(\frac{R^2}{MSE})$ where R is the maximum pixel value and MSE is the mean square error between input and output. For RGB images, its MSE is the MSE sum of three channels divided by three. As for cryo-EM images, since there is no ground truth available, PSNR evaluation is not applicable in this case.

III. EXPERIMENTs

A. Natrual Images Denoising

To validate our method, we first test it on the natural images. We train the network using 256x256 pixel crops drawn from the 5k images in the COCO 2017 validation set for 120 epochs. The 256x256 patches was added with different gaussian noise as input and target. We furthermore randomize the noise standard deviation $\sigma = [0, 50]$ separately for each training example. So, the neural network has to tune the parameter and estimate noise magnitude while removing additive noise. We can see from the results that the denoising performance is pretty good (Fig.2).

These results demonstrate that the clean targets are unnecessary in this application. We also noticed that the recovered images become blurry when the noise increase. We further feed an example with noise standard deviation $\sigma = 70$, which the network has not seen. Surprisingly, the network still can output reasonable results, although the details are somewhat lost. We can conclude that for the zero-mean noise, using noisy images as targets can obtain the same performance as using the clean images as targets.

TABLE I: PSNR Comparison between noisy input and denoised result

Noise level	Noisy input	Prediction
$\sigma = 10$	25.58dB	28.14dB
$\sigma = 20$	23.94dB	27.20dB
$\sigma = 40$	23.38dB	26.08dB
$\sigma = 70$	23.26dB	24.52dB

B. Cryo-EM Images Denoising

1) Current bottleneck in the field: Cryo-electron microscopy (cryo-EM) is an ensemble of techniques allowing the observation of biological specimens in their native environment at cryogenic temperatures in EM. Despite the power of cryo-EM to determine new and challenging macromolecular structures, there remain a number of significant bottlenecks that continue to slow the adoption of Cryo-EM and the throughput of structural data analysis. One such barrier is the inherent low signal-to-noise ratio (SNR) of individual micrographs due to the sensitivity of biological specimens to an incident electron beam. In single particle image processing a large number of particle images are collected during one

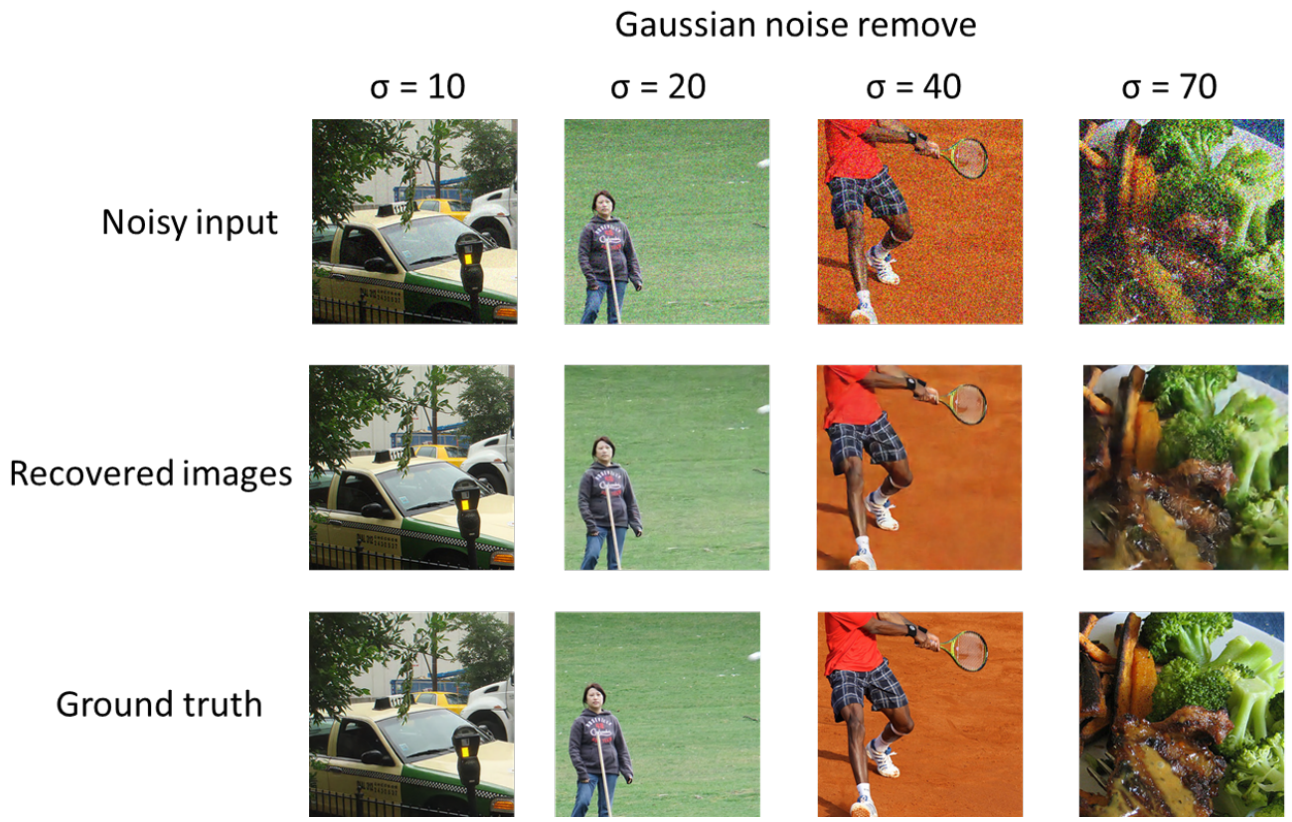

Fig. 2: Denoised result for different gaussian noise

or several TEM sessions, corresponding to different angle views of the protein molecule. An iterative mathematical 3D reconstruction schemes can then be applied to obtain a 3D model from these data and refine it to higher and higher resolution. However, the low SNR prevent the assessment of the raw micrograph, and inspection of the protein complex integrity. One can only judge the data quality when they went through the whole data processing pipeline and got the 3d reconstruction. Once the data is bad, they need to collect more data and repeat the step until obtaining the high-resolution structure. This iterative process is very time consuming. So, we need to way to denoising the raw micrograph, letting us evaluate the data quality without the needed for going through the whole pipeline, which will save a huge amount of time.

2) Data prepossessing: The cryo-EM data we collected are high frame rate movies, which will be aligned and averaged to produce the final micrograph. We take the aligned sums of all odd and all even frames of a movie as the input and target. All acquisition-related noise will be independent between the two, while the underlying image content is almost identical.

3) Neural network training and results: We trained the network using 640x640 crops drawn from the 250 even and odd images for 500 epochs. We used two protein dataset we collected, one is aldolase, the other one is apoferritin. Aldolase is a $150KD$ homo tetrameric enzyme. Its smaller size makes harder for human visualization, thus represents a challenge case for micrograph assessment and picking individual particle images from the micrograph. The goal of the assessment is to gather all the micrographs that have no contamination and evenly distributed protein molecules. By using our method, we can clearly see the individual protein molecules in the images comparing to the raw data(Fig.3), making image assessment and particle picking much easier.

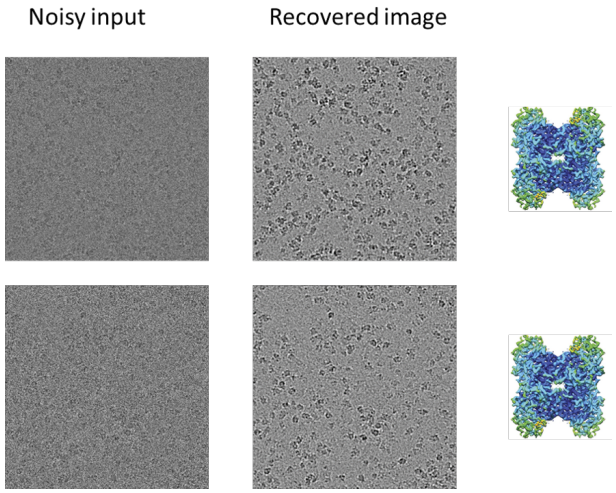

Fig. 3: Removing noise from the raw Cryo-EM images from the aldolase dataset

We then test on the apoferritin data set. Apoferritin is a ball shape protein with a molecular weight of $480KD$. Benefit from its size, we can see the individual particles in

the raw data. But some protein may be damaged or broken up, so we just want to pick the intact protein particle from the raw micrograph for the further processing. From the denoising results(Fig.4), we can clear distinguish the good protein particle from the broken one and only pick the intact protein particles.

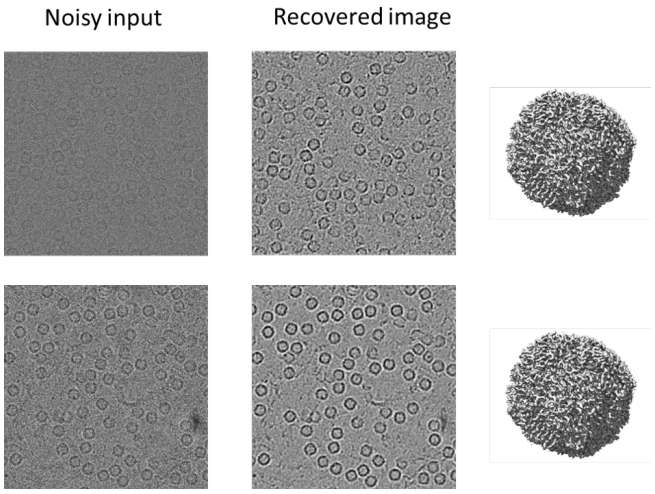

Fig. 4: Removing noise from the raw Cryo-EM images from the apoferritin dataset

IV. CONCLUSIONS

This work represents an important step proof-of-principle studies in the implementation of artificial intelligence-driven denoising of cryo-EM data. In all cases we tested, we observed our model is capable of recovering information from real cryo-EM data. The denoising method presented here can facilitate particle selection and raw particle interpretation in the raw micrograph. Moreover, it is able to return information to the user regarding the state of the protein complex. This suggests that our method can help to increase the throughput of cryo-EM if raw micrographs can be used for sample assessment. Given the denoising methods worked on 2D images, it is naturally to build a similar architecture for denoising 3D tomographic volume for the tomography cryo-EM data.

Code is available at https://github.com/ZhenyuTan/Noise2Noise-Cryo-EM-image-denoising

REFERENCES

[1] Lehtinen, Jaakko, et al. "Noise2Noise: Learning Image Restoration without Clean Data." arXiv preprint arXiv:1803.04189 (2018).

[2] Krull, Alexander, Tim-Oliver Buchholz, and Florian Jug. "Noise2Void-Learning Denoising from Single Noisy Images." arXiv preprint arXiv:1811.10980 (2018).