PAPER

Poisson–Gaussian Noise Removal for Low-Dose CT Images by Integrating Noisy Image Patch and Impulse Response of Low-Pass Filter in CNN

May Thet Tun, Yosuke Sugiura and Tetsuya Shimamura

Graduate School of Science and Engineering, Saitama University, 255 Shimo-Okubo, Sakura-ku, Saitama 338-8570, Japan
E-mail: may.t.t.218@ms.saitama-u.ac.jp, ysugiura@mail.saitama-u.ac.jp, shima@mail.saitama-u.ac.jp

Abstract　　In this paper, we propose the incorporation of noisy image patches and the impulse response of a low-pass filter (LPF) in a convolutional neural network (CNN) to denoise Poisson–Gaussian noise in low-dose computed tomography (LDCT) images. The approach is referred to as fast and flexible denoising CNN (FFDNet)-impulse response (FFDNet-IR) in this paper. The power spectrum sparsity LPF (SLPF) allows low-frequency components to pass through while suppressing higher frequency components by the sparsity approach of the power spectrum, and it is employed to determine the impulse response of LPF. Three well-known types of LPF, namely, Direct LPF, Gaussian LPF, and Butterworth LPF, are also considered to obtain the impulse response of LPF. In the FFDNet-IR, both the noisy image patches and the IR of the LPF are sequentially inputted into the FFDNet to eliminate the Poisson–Gaussian noise. This approach enhances the denoising performance in LDCT images compared with the conventional FFDNet in the evaluation metrics of the peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and feature similarity (FSIM). Moreover, the FFDNet-IR trained with the Poisson–Gaussian noise model demonstrates the generalization ability and effectively eliminates only Poisson or Gaussian noise. The experiments indicate that the FFDNet-IR more effectively suppresses the noise artifacts and preserves image details compared with the baseline FFDNet, as well as traditional methods such as block-matching and 3D filtering (BM3D) and nonlocal mean (NLM) for LDCT image denoising.

Keywords: image denoising, Poisson–Gaussian noise removal, convolutional neural network, FFDNet, FFDNet-IR, impulse response of low-pass filter, power spectrum sparsity

1. Introduction

Efficient image denoising methods play a pivotal role in enhancing the quality of images in diverse applications such as medical imaging, low-light conditions, and satellite imaging. Nowadays, the potential cancer risk associated with radiation exposure in computed tomography (CT) imaging has become a significant issue in medical diagnosis. Radiologic technologists reduce the number of projection views or lower the amount of dose to mitigate the risks in CT imaging. Nevertheless, employing the low-dose technique results in the occurrence of artifacts and degrades the quality of the resulting images. The greater the emphasis on minimizing radiation doses for patient safety, the higher the levels of noise and the lower the

quality of images. Thus, the accuracy of medical images is constrained by limitations in sharpness and quality: this highlights the pivotal significance of noise reduction. As prominently observed in low-dose computed tomography (LDCT) scans, there is a need to establish a balance between the image quality and the degree of radiation exposure.

When tackling issues associated with denoising of LDCT images, researchers have categorized their methodologies into three main approaches: (i) sinogram filtering, (ii) iterative reconstruction, and (iii) postprocessing methods [1]. The sinogram domain filtering methods, such as bilateral filters, adaptive filters, and penalized likelihood methods, effectively reduce the unwanted noise of sinogram data. However, these approaches face the challenge of oversmoothed edges owing to filtering. Over the

past few decades, iterative reconstruction algorithms such as total variation (TV), nonlocal means (NLM), and low-rank matrix decomposition have been utilized in LDCT denoising by integrating prior knowledge into the objective function. Subsequently, the NLM algorithm conducted by Chen et al. [2] focuses on the similarity between the central pixel and its neighborhood pixels to improve the quality of thoracic LDCT images. However, the significant drawback of this approach is the long computational time due to the iterative process involving both projection and back projection. Postprocessing techniques such as block-matching 3D (BM3D), the most popular postprocessing algorithm, have shown significant improvements in eliminating additive white Gaussian noise (AWGN) from CT images [3]. The k-mean singular value decomposition (K-SVD) method is based on iteratively updating a dictionary and finding sparse representation, thereby being effective in reducing AWGN and artifacts in abdominal tumor CT images [4].

Even though LDCT image denoising has been widely explored in numerous studies over the years, the diverse types of noise remain a challenge. Reduced radiation doses in CT scanning decrease the number of detected photons, consequently causing Poisson noise. In low-dose imaging scenarios, the statistical basis of photon detection commonly results in Poisson noise, while electronic and environmental factors during the image acquisition lead to Gaussian noise. The combination of these two types of noise, known as Poisson–Gaussian noise, presents a complex noise and is frequently encountered in astronomy, remote sensing, biology, and medical images including X-rays, CT, and positron emission tomography (PET) scans [5]. Furthermore, it affects the complexity of the denoising procedures and requires sophisticated approaches. Thus, Bahador et al. [5] proposed a time-space fractional differential equation to remove mixed Poisson–Gaussian noise from images. At present, traditional methods face a problem in precisely eliminating a variety of noise. The methods mentioned earlier are effective only for specific types of noise and yield poor results if the prior knowledge of noise type is not accurate.

Recently, researchers have explored deep learning techniques to remove artifacts in mixed noise in LDCT because deep learning techniques can learn complex noise patterns. However, they struggle for generalization ability in the case of mismatched distribution of noise. For example, a denoising model trained on mixed Poisson and Gaussian noise may not eliminate single Poisson noise. This point is a crucial consideration, not only in deep learning methods but also in traditional methods.

Generative adversarial networks (GANs) have emerged as a revolutionary technology in diverse applications; however, it has been known that residual learning in a convolutional neural network (CNN) exhibits a simpler representation architecture and superior adaptability to image-denoising tasks compared with GANs [1]. By incorporating residual learning and batch normalization (BN), denoising CNN (DnCNN) demonstrates higher performance when applied to denoise mixed Poisson–Gaussian noise in the Fluorescence Microscopy Denoising dataset [6]. Limited research has been conducted on the removal of Poisson–Gaussian noise by deep learning. The majority of existing studies focus on the removal of mixed impulse and Gaussian noise [7][8]. However, medical images are infrequently affected by impulse noise, which is corrupted by sudden spikes in intensity.

Zhang et al. [9] proposed a dual transformer and CNN-based strategy for LDCT image denoising and showed superior results to state-of-the-art methods, although the emphasis in [9] is only on Gaussian noise. The practical problem of the method in [9] is the use of the transformer module, where the number of parameters is much larger than that of CNN. This can increase the demand for computing resources and time in the training process. Then, the more complex the architecture grows, the larger the dataset required for effective training. Unfortunately, acquiring medical image datasets is limited by privacy constraints.

The LDCT image denoising method proposed in [1], known as transfer learning residual CNN (TLR-CNN), has a two-step training process. In the first step, TLR-CNN is pretrained with natural images, followed by fine tuning with CT images. The limitation of TLR-CNN is that it has a complex architecture and requires a large number of parameters, leading to a longer training time. Furthermore, the fine tuning of TLR-CNN requires the careful adjustment of hyperparameters, which is crucial to avoid overfitting.

The wavelet domain residual network (WavResNet) [1] provides clear low-dose X-ray CT reconstruction results, preserving image textures in the image reconstruction process. Nevertheless, WavResNet employs a high-dimensional embedding approach to both contourlet transform and residual learning. The incorporation enhances the reconstruction ability but increases the model's complexity and computational time. In [1], the execution time of TLR-CNN was compared with those of BM3D, K-SVD, FFDNet [10], WavResNet, and TLR-CNN. The results show that 1.49, 2.96, 0.09, 17.66, and 0.1 s are required for BM3D, K-SVD, FFDNet, WavResNet, and TLR-CNN, respectively. Among these approaches, WavResNet has the longest execution time, whereas FFDNet has the shortest execution time. Consequently, the authors in [1] recommend TLR-CNN and FFDNet as effective

methods for LDCT image denoising.

Motivated by Ref. [1], we consider the residual neural-network-based FFDNet, the state-of-the-art AWGN image denoiser, as a conventional method for Poisson–Gaussian noise removal in LDCT images. The joint training of the impulse response of the LPF and the noisy image patch into the FFDNet can eliminate the AWGN noise in natural images, leading to superior image denoising outcomes compared with the conventional FFDNet [11] [12]. In this paper, this impulse response embedded approach is denoted as FFDNet-IR, and the performance of FFDNet-IR for removing mixed Poisson–Gaussian noise in LDCT images is investigated in a comparative fashion with the conventional FFDNet.

Deep learning methods for LDCT image denoising such as those in [1] and [9] require paired training data of both full-dose CT (FDCT) and LDCT images. However, obtaining the paired training data may be difficult in medical imaging applications. According to the constraints of paired data in medical images, in this paper, we adopt unpaired learning without the need to align FDCT and LDCT images. The experiments in FDCT image denoising are also conducted. Moreover, a mixed Poisson–Gaussian noise model is also applied to images degraded by only Gaussian or Poisson noise.

The rest of the paper is structured as follows. Sect. 2 presents an overview of the Poisson–Gaussian noise model and describes the architecture and learning of FFDNet-IR. In Sect. 3, we elaborate on experimental findings and visual representations of CT image denoising. Conclusions are drawn in Sect. 4.

2. FFDNet-IR

2.1 Poisson–Gaussian noise model

Gaussian noise is widely employed in image-denoising applications, including natural image and medical image denoising, in which thermal electronic noise occurs in high- and medium-light imaging. However, in low-light scenarios, the Gaussian model fails to accurately capture shot noise. In low-light applications such as spectral imaging, medical imaging, fluorescence confocal microscope imaging, and photon counting imaging, images are corrupted by Poisson noise, which is also called shot or scatter noise. Its variance is equal to the signal intensity and hence is also known as signal-dependent noise. The simultaneous occurrence of Poisson noise and Gaussian noise is a result of the inherent randomness of photons and the fluctuation of electrons. The presence of these two different types of noise effect in medical images creates a challenging problem for mixed noise denoising. Figure 1 shows various types of noise
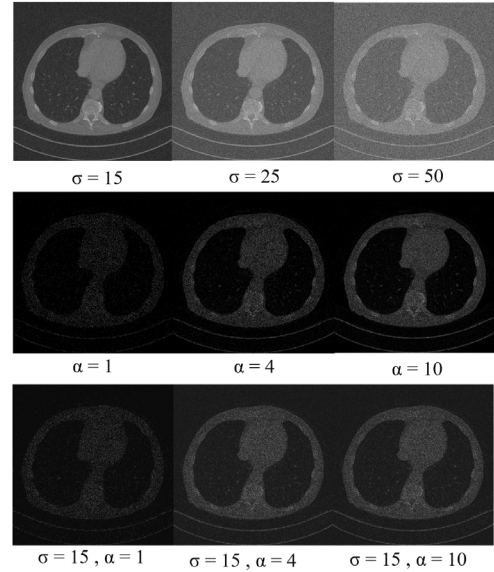

Fig. 1 Some examples of noise added to an LDCT image, with σ for Gaussian noise and α for Poisson noise

and its variances added to an LDCT image. The investigation of Poisson–Gaussian noise through deep learning has been limited in comparison with Gaussian noise. However, the Gaussian noise model encounters difficulties in effectively handling non-Gaussian models. The mixed Poisson–Gaussian process exhibits independence, with its variance resulting from combining two distinct noise variances encompassing signal-dependent Poisson noise and signal-independent Gaussian noise. In the context of the paper, mixed noise is defined as a combination of Poisson–Gaussian noise with varying noise levels. For the Poisson–Gaussian noise model, the i-th noise pixel value η_i can be mathematically represented as

$$\eta_i = \alpha P_i + \xi_i, i = 1, .., Q \tag{1}$$

where $P_i \sim Poisson(x_i)$ represents the Poisson noise corrupted to the i-th original (clean) image pixel value x_i, and α is a scaling factor that depends on the sensor intensity. $\xi_i \sim N(0, \sigma^2)$ is AWGN with zero mean, and σ is a standard deviation of random Gaussian noise [13]. Q corresponds to the number of pixels. The Poisson–Gaussian noise model describes the real source noise in medical images and consists of a diverse noise variance of two associated parameters (α, σ).

2.2 Network architecture and learning

In the preprocessing of FFDNet-IR and FFDNet, input images are downsampled into 70×70 patches in size in accordance with the formula $W \times H \times C$ into $W/2 \times H/2 \times 4C$. The width and height of the images
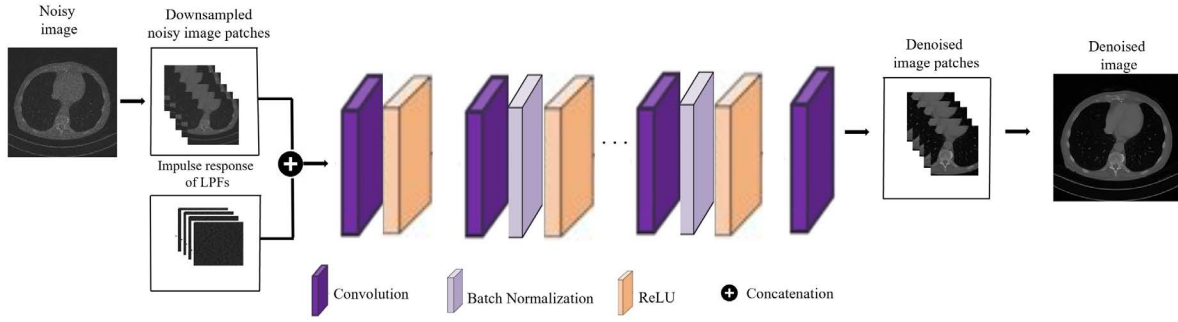
Fig. 2 FFDNet-IR architecture

are represented by W and H, respectively, while C specifies the number of channels (1 for grayscale and 3 for color). Then, image patches are randomly augmented through counterclockwise rotations of 90, 180, and 270 degrees and horizontal and vertical flips up and down. The network incorporates residual learning and BN to enhance denoising performance and accelerate the learning process. Additionally, rectified linear units (ReLUs) are employed for faster training. BN is integrated as a middle layer in the convolution (Conv), and ReLU facilitates a higher learning rate through BN. The Conv filter is implemented with the size of 3×3 and 64 feature maps.

The first layer, denoted as L_1, is constructed by Conv+ReLU, expressed as

$$L_1(y) = ReLU(Conv(y, W, b)) \qquad (2)$$

where $Conv(y, W, b)$ represents the Conv function applied to the input y with weights W and biases b, followed by the ReLU activation function. The subsequent layers are constructed as Conv+ BN+ReLU, formulated as

$$L_j(y) = ReLU(BN(\sum_j W_j * L_{j-1}(y) + b_j)) \qquad (3)$$

where $L_{j-1}(y)$ denotes the convolved output to the input y from the previous j-th layer (j is designated as 15 for grayscale CT images). W_j and b_j denote the weights and biases of the j-th layer, respectively, and $*$ represents the elementwise multiplication operation. After Conv extracts image features, the subsequent BN and ReLU functions enhance the learning process. Lastly, the final layer is constructed by the Conv function.

The mean square error loss (MSE) for the Poisson–Gaussian noise reduction

$$\text{MSE} = \frac{1}{2N} \sum_{k=1}^{N} \parallel \zeta(y_k, M_k; v_{(\alpha,\sigma)}) - x_k \parallel^2 \qquad (4)$$

is minimized, where N represents the number of CT image patches, ζ denotes the residual mapping
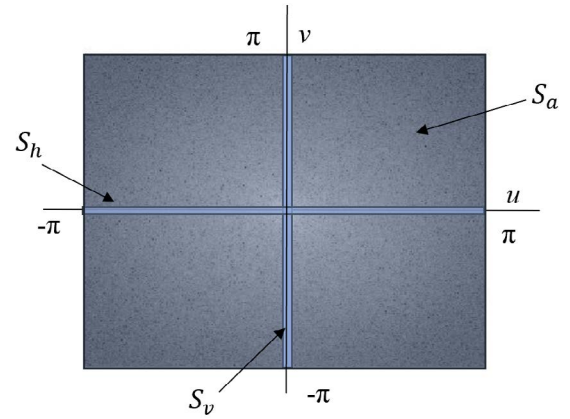

Fig. 3 Power spectrum sparsity of the image patch

function connecting noisy and clean image patches of y_k and x_k, and $\{(y_k, x_k)\}_{k=1}^{N}$ is a pair of degraded noisy and clean image patches. M_k represents the noise-level map, and $v_{(\alpha,\sigma)}$ is the variance of Poisson–Gaussian noise. FFDNet-IR constructs y_k by combining degraded noisy CT image patches with the impulse response of LPFs. Hence, there is a difference in the number of input features between FFDNet and FFDNet-IR. FFDNet incorporates five features, including the noise-level map M and downsampled subimages, whereas FFDNet-IR utilizes ten features. However, the output features of FFDNet and FFDNet-IR are identical in four features because both of them minimize the MSE loss between clean and denoised image patches.

FFDNet-IR is motivated as follows. The image downsampling method, commonly used in FFDNet and other denoisers for computational efficiency, leads to aliasing, which distorts image frequencies. Then, the utilization of the MSE loss or L2 loss function and placing a priority on learning the low-frequency components in FFDNet cause oversmoothed images. To tackle these problems, the impulse response of LPF as an extra image patch is incorporated in

FFDNet-IR. The total system of FFDNet-IR is shown in Fig. 2.

FFDNet-IR considers the four different impulse responses of LPFs because each has unique advantages depending on the characteristics of the image data. By considering the various impulse responses of LPFs, we can determine the most suitable types for specific situations.

The four different impulse responses for LPFs are constructed in three steps.

(1) Determine the size of each LPF as the identical size of the image patch.

(2) Establish the frequency response of each LPF.

For the power spectrum sparsity-based LPF (SLPF), the power spectrum sparsity value is calculated as

$$\rho = \frac{S_a}{S_h + S_v} \qquad (5)$$

where S_a denotes the sum of all power spectrum values, and S_h and S_v denote the totals of the horizontal and vertical power spectrum values, respectively, as shown in Fig. 3. The ρ in Eq. (5) specifies the concentration of frequency components in the image, and a lower value of ρ indicates a sparser power spectrum [14]. Then, the sparsity threshold φ is determined by

$$\varphi = \tau \rho \qquad (6)$$

where τ is a scaling parameter, which is adjusted with the ρ value. Then, the frequency response of SLPF, $H_s(u, v)$, is determined as

$$H_s(u, v) = \begin{cases} 1 & : \ S(u, v) \leq \varphi \\ 0 & : \ otherwise \end{cases} \qquad (7)$$

where $S(u, v)$ is the power spectrum of the image patch. A detailed explanation of SLPF is available in [11] and [12], where FFDNet-IR is developed as a denoising method for Gaussian noise on the images.

For the direct LPF (DLPF), the frequency response is given as

$$H_d(u, v) = \begin{cases} 1, & D(u, v) \leqslant \omega_c \\ 0, & otherwise \end{cases} \qquad (8)$$

where $D(u, v)$ represents the ideal LPF frequency response of the image and ω_c corresponds to the cut-off angular frequency.

For the Gaussian LPF (GLPF), it is

$$H_g(u, v) = e^{-D^2(u,v)/2\omega_c^2} \qquad (9)$$

For the Butterworth LPF (BLPF),

$$H_b(u, v) = \frac{1}{1 + [D(u, v)/\omega_c]^{2n}} \qquad (10)$$

where n (the order of the filter) is assigned the value of 1, resulting in the highest peak signal-to-noise ratio (PSNR) result. The cut-off angular frequency is set to $\frac{\pi}{2}$ for DLPF, GLPF, and BLPF.

(3) The impulse response of each LPF is obtained by the inverse fast Fourier transform ($IFFT$) of each frequency response.

FFDNet-IR integrates the impulse response of power spectrum sparsity-based LPF and noisy image patches into FFDNet, which is denoted as SLPF+FFDNet in this paper. When the impulse response of DLPF is utilized instead of that of SLPF in SLPF+FFDNet, it is denoted as DLPF+FFDNet. In the same way, when the impulse response of GLPF and that of BLPF are utilized, they are denoted as GLPF+FFDNet and BLPF+FFDNet, respectively.

3. Experimental Results

3.1 Data and experimental settings

In this section, we describe the experiment configuration employed for LDCT images. The first dataset selected for this study is from the Mayo Clinic as part of the 2016 NIH-AAPM Mayo Clinic Low Dose CT Grand Challenge, accessible at https://www.aapm.org/GrandChallenge/LowDoseCT in the form of DICOM images. In these LDCT images, specifically one-quarter of the full-dose images are utilized, and there is no connection between full-dose and low-dose images. For the experimental phase, a total of 916 LDCT images are utilized to serve as training images. The patient case L506, including 211 images, is designated for testing purposes. The image size is commonly 512×512.

Although LDCT is widely utilized in medical diagnostics, FDCT is still essential for specific diagnostic requirements, particularly in complex cases. FDCT ensures the highest diagnostic accuracy by producing high-quality images. Hence, we conducted the removal of image noise for FDCT using data sourced from the Cancer Image Archive (TCIA), which is available at https://wiki.nci.nih.gov/ in the form of DICOM images. TCIA is a service repository of various medical images available for researchers to download. Patient images within the TCIA repository are commonly associated with prevalent diseases, such as cancer images, and image morphology types of MRI, CT, and PETs are for research purposes. The specific medical lung CT image dataset extracted from the TCIA database is QIN_LUNG_CT. For the image denoising models, 800 FDCT images are utilized, with 600 images randomly chosen as the training set and the remaining 200 designated for testing. In this case, the image size is 256×256.

After the preprocessing step, 78,776 image patches are generated from the Mayo Clinic datasets and 51,600 patches from the QIN_LUNG_CT datasets. In the process of training for Poisson–Gaussian mixed

noise, the random noise levels within the ranges of $\alpha\epsilon[0, 20]$ and $\sigma\epsilon[0, 75]$ are added to generate distorted images. The minibatch learning rate was set to 128, the epoch was 50, and the Adam optimizer was employed at a learning rate of 10^{-4}. The experiments were conducted on one GPU Intel (R) Core i9 processor, Nvidia Quadro P2200, and 32 GB of CPU memory. The experiments were executed using Pytorch running on Windows 10. For a fair comparison, we employed identical parameter configurations for both FFDNet-IR and the conventional FFDNet.

3.2 Experimental evaluation

Experimental results were captured through visual representations and quantitative evaluations, including the PSNR, structural similarity (SSIM), and feature similarity (FSIM) [15]. To assess the effectiveness of eliminating Poisson–Gaussian noise, CT images were assessed by setting the noise pair distributions of σ values to 15, 25, and 50 and α values to 1, 4, and 10.

Tables 1–3 display the comparative PSNR, SSIM, and FSIM results between FFDNet-IR and the baseline FFDNet on the Mayo Clinic image dataset. Tables 4–6 show those on the QIN_LUNG_CT image dataset. From the results in Tables 1–6, we see that FFDNet-IR outperforms the baseline FFDNet in terms of PSNR, SSIM, and FSIM scores. The superior performance of FFDNet-IR over the baseline is due to the learning process of CNN incorporated by the additional valuable information with the impulse response of LPF.

In Table 1, Poisson–Gaussian noise is artificially added to LDCT images in the Mayo Clinic dataset. It can be seen that DLPF+FFDNet and SLPF+FFDNet achieve superior denoising performance compared with the others because LDCT images suffer from loss of low-frequency information owing to the reduced radiation dose. The impulse response in DLPF+FFDNet enables the preservation of essential low-frequency components, which helps to provide a clearer and more accurate representation of the structure details than other impulse response types. The power spectrum sparsity approach based on SLPF+FFDNet reconstructs the image using the dominant image components. Therefore, SLPF+FFDNet concentrates the image frequencies into a few dominant components within LDCT images. This approach aims to protect the low-frequency image components and suppress the high-frequency components, effectively eliminating noise in LDCT images. Consequently, SLPF+FFDNet results in high-frequency noise reduction and preserves crucial structural information for the precise diagnosis of noise in LDCT images.

Table 1 Average quantitative results (PSNR/SSIM/FSIM) considering Poisson–Gaussian denoising for the Mayo Clinic dataset [Bold: Best, Underline: Second Best]

σ	α	Method	PSNR	SSIM	FSIM
15	1	FFDNet	25.40	0.673	0.806
		SLPF+FFDNet	25.48	<u>0.701</u>	0.813
		DLPF+FFDNet	**25.91**	**0.711**	<u>0.814</u>
		GLPF+FFDNet	<u>25.66</u>	0.687	0.813
		BLPF+FFDNet	25.59	0.696	**0.815**
	4	FFDNet	<u>29.60</u>	0.817	<u>0.880</u>
		SLPF+FFDNet	**29.66**	<u>0.831</u>	**0.884**
		DLPF+FFDNet	29.55	**0.834**	0.878
		GLPF+FFDNet	29.21	0.800	0.863
		BLPF+FFDNet	29.34	0.809	0.866
	10	FFDNet	30.63	0.835	**0.911**
		SLPF+FFDNet	<u>31.07</u>	<u>0.863</u>	<u>0.909</u>
		DLPF+FFDNet	**31.14**	**0.876**	0.906
		GLPF+FFDNet	30.84	0.850	0.894
		BLPF+FFDNet	30.70	0.844	0.891
25	1	FFDNet	25.25	0.659	0.801
		SLPF+FFDNet	25.41	0.689	0.810
		DLPF+FFDNet	**25.78**	<u>0.698</u>	0.810
		GLPF+FFDNet	<u>25.57</u>	0.687	<u>0.811</u>
		BLPF+FFDNet	25.50	**0.699**	**0.812**
	4	FFDNet	29.27	0.780	0.873
		SLPF+FFDNet	**29.46**	<u>0.817</u>	**0.880**
		DLPF+FFDNet	<u>29.34</u>	**0.818**	<u>0.874</u>
		GLPF+FFDNet	29.01	0.786	0.858
		BLPF+FFDNet	29.12	0.791	0.862
	10	FFDNet	30.07	0.797	<u>0.903</u>
		SLPF+FFDNet	<u>30.60</u>	<u>0.844</u>	**0.904**
		DLPF+FFDNet	**30.88**	**0.860**	<u>0.903</u>
		GLPF+FFDNet	30.46	0.825	0.887
		BLPF+FFDNet	30.30	0.817	0.883
50	1	FFDNet	24.48	0.591	0.776
		SLPF+FFDNet	24.91	0.636	0.789
		DLPF+FFDNet	**25.17**	0.636	0.790
		GLPF+FFDNet	<u>25.14</u>	**0.663**	<u>0.797</u>
		BLPF+FFDNet	24.93	<u>0.661</u>	**0.798**
	4	FFDNet	28.06	0.700	0.839
		SLPF+FFDNet	**28.51**	**0.754**	**0.856**
		DLPF+FFDNet	<u>28.29</u>	<u>0.747</u>	<u>0.848</u>
		GLPF+FFDNet	28.14	0.730	0.836
		BLPF+FFDNet	28.27	0.734	0.841
	10	FFDNet	28.64	0.730	0.871
		SLPF+FFDNet	28.76	<u>0.741</u>	**0.875**
		DLPF+FFDNet	**29.54**	**0.785**	<u>0.874</u>
		GLPF+FFDNet	<u>28.96</u>	0.741	0.851
		BLPF+FFDNet	28.82	0.731	0.846

In Tables 2 and 3, the Poisson–Gaussian noise model is employed for the elimination of either Poisson or Gaussian noise individually. For the exclusive removal of Poisson noise, as seen in Table 2, FFDNet-IR exhibits superior performance compared with the baseline FFDNet when α is set to 1, the highest noisy peak value. However, for α values of

Table 2 Average quantitative results
(PSNR/SSIM/FSIM) considering Poisson
denoising for the Mayo Clinic dataset [Bold:
Best, Underline: Second Best]

α	Method	PSNR	SSIM	FSIM
1	FFDNet	25.43	0.678	0.808
	SLPF+FFDNet	25.49	0.708	0.815
	DLPF+FFDNet	**25.92**	**0.718**	**0.816**
	GLPF+FFDNet	25.68	0.686	0.815
	BLPF+FFDNet	25.60	0.689	**0.816**
4	FFDNet	**29.83**	0.837	0.884
	SLPF+FFDNet	29.78	0.840	**0.886**
	DLPF+FFDNet	29.65	**0.844**	0.879
	GLPF+FFDNet	29.33	0.809	0.865
	BLPF+FFDNet	29.46	0.823	0.868
10	FFDNet	31.11	0.880	**0.915**
	SLPF+FFDNet	**31.35**	0.872	0.912
	DLPF+FFDNet	31.29	**0.886**	0.908
	GLPF+FFDNet	31.10	0.869	0.898
	BLPF+FFDNet	30.96	0.864	0.895

Table 3 Average quantitative results
(PSNR/SSIM/FSIM) considering Gaussian
denoising for the Mayo Clinic dataset [Bold:
Best, Underline: Second Best]

σ	Method	PSNR	SSIM	FSIM
15	FFDNet	26.26	0.788	0.839
	SLPF+FFDNet	26.06	0.803	0.841
	DLPF+FFDNet	27.16	0.787	0.834
	GLPF+FFDNet	27.50	0.803	**0.859**
	BLPF+FFDNet	**27.51**	**0.817**	0.847
25	FFDNet	26.00	0.752	0.835
	SLPF+FFDNet	25.96	0.777	0.839
	DLPF+FFDNet	26.93	0.759	0.833
	GLPF+FFDNet	**27.44**	**0.811**	**0.858**
	BLPF+FFDNet	27.36	0.807	0.847
50	FFDNet	25.03	0.635	0.805
	SLPF+FFDNet	25.31	0.673	0.818
	DLPF+FFDNet	26.04	0.670	0.821
	GLPF+FFDNet	**26.91**	**0.760**	**0.848**
	BLPF+FFDNet	26.29	0.695	0.837

Table 4 Average quantitative results
(PSNR/SSIM/FSIM) considering
Poisson–Gaussian denoising for the
QIN_LUNG_CT dataset [Bold: Best,
Underline: Second Best]

σ	α	Method	PSNR	SSIM	FSIM
15	1	FFDNet	23.72	0.490	0.708
		SLPF+FFDNet	24.06	0.489	0.741
		DLPF+FFDNet	24.41	**0.545**	0.739
		GLPF+FFDNet	24.08	0.514	**0.754**
		BLPF+FFDNet	**24.57**	0.527	0.742
	4	FFDNet	27.40	0.671	0.819
		SLPF+FFDNet	27.60	0.686	0.837
		DLPF+FFDNet	27.71	0.720	0.827
		GLPF+FFDNet	27.97	0.730	0.847
		BLPF+FFDNet	**28.27**	**0.753**	**0.848**
	10	FFDNet	28.98	0.757	0.850
		SLPF+FFDNet	28.91	0.755	0.857
		DLPF+FFDNet	29.13	0.777	0.854
		GLPF+FFDNet	29.28	0.766	**0.870**
		BLPF+FFDNet	**29.63**	**0.799**	**0.870**
25	1	FFDNet	23.63	0.478	0.703
		SLPF+FFDNet	23.97	0.475	0.737
		DLPF+FFDNet	24.32	**0.535**	0.735
		GLPF+FFDNet	23.98	0.494	**0.749**
		BLPF+FFDNet	**24.49**	0.512	0.738
	4	FFDNet	27.22	0.664	0.812
		SLPF+FFDNet	27.42	0.672	0.832
		DLPF+FFDNet	27.49	0.707	0.822
		GLPF+FFDNet	27.74	0.711	**0.841**
		BLPF+FFDNet	**28.01**	**0.731**	**0.841**
	10	FFDNet	28.62	0.734	0.838
		SLPF+FFDNet	28.63	0.736	0.852
		DLPF+FFDNet	28.76	0.751	0.845
		GLPF+FFDNet	28.92	0.742	**0.863**
		BLPF+FFDNet	**29.30**	**0.777**	**0.863**
50	1	FFDNet	23.17	0.410	0.676
		SLPF+FFDNet	23.44	0.409	0.706
		DLPF+FFDNet	23.90	**0.476**	0.713
		GLPF+FFDNet	23.56	0.422	**0.725**
		BLPF+FFDNet	**24.04**	0.447	0.713
	4	FFDNet	26.43	0.617	0.788
		SLPF+FFDNet	26.55	0.613	0.806
		DLPF+FFDNet	26.61	**0.655**	0.797
		GLPF+FFDNet	26.73	0.631	**0.814**
		BLPF+FFDNet	**26.92**	0.650	0.807
	10	FFDNet	27.18	0.640	0.794
		SLPF+FFDNet	27.42	0.658	0.818
		DLPF+FFDNet	27.41	0.669	0.812
		GLPF+FFDNet	27.52	0.655	**0.828**
		BLPF+FFDNet	**27.90**	**0.691**	0.827

4 and 10, the baseline FFDNet has higher denoising capabilities than FFDNet-IR. Table 3 focuses solely on Gaussian noise removal in the Mayo Clinic dataset. GLPF+FFDNet outperforms others because it attenuates the high frequency while preserving the low-frequency components related to image details, resulting in higher Gaussian noise reduction.

In the QIN_LUNG_CT image dataset, where Poisson–Gaussian noise is added to the FDCT images, the use of BLPF+FFDNet stands out as an effective method not only in Poisson–Gaussian denoising but also in the singular task of Poisson denoising, as shown in Tables 4 and 5. By modifying the cut-off angular frequency in BLPF+FFDNet, the noise characteristics are finely tuned, leading to the preservation of

the structural details in medical images during the denoising process. For Gaussian denoising in FDCT images, DLPF+FFDNet and GLPF+FFDNet stand out as more effective methods than others, as shown in Table 6.

Tables 1–6 provide comprehensive results of FFDNet-IR and FFDNet, corresponding to each type

Table 5 Average quantitative results (PSNR/SSIM/FSIM) considering Poisson denoising for the QIN_LUNG_CT dataset [Bold: Best, Underline: Second Best]

α	Method	PSNR	SSIM	FSIM
1	FFDNet	23.73	0.495	0.710
	SLPF+FFDNet	24.06	0.494	0.741
	DLPF+FFDNet	<u>24.43</u>	**0.543**	0.739
	GLPF+FFDNet	24.11	0.523	**0.755**
	BLPF+FFDNet	**24.59**	<u>0.534</u>	<u>0.744</u>
4	FFDNet	27.51	0.672	0.823
	SLPF+FFDNet	27.71	0.692	0.840
	DLPF+FFDNet	27.85	0.727	0.830
	GLPF+FFDNet	<u>28.09</u>	<u>0.733</u>	<u>0.851</u>
	BLPF+FFDNet	**28.43**	**0.766**	**0.852**
10	FFDNet	29.23	0.772	0.857
	SLPF+FFDNet	29.08	0.767	<u>0.861</u>
	DLPF+FFDNet	29.38	<u>0.796</u>	0.860
	GLPF+FFDNet	<u>29.49</u>	0.781	**0.874**
	BLPF+FFDNet	**29.80**	**0.813**	**0.874**

Table 6 Average quantitative results (PSNR/SSIM/FSIM) considering Gaussian denoising for the QIN_LUNG_CT dataset [Bold: Best, Underline: Second Best]

σ	Method	PSNR	SSIM	FSIM
15	FFDNet	24.96	0.532	0.783
	SLPF+FFDNet	25.24	0.519	0.799
	DLPF+FFDNet	**25.96**	**0.678**	<u>0.812</u>
	GLPF+FFDNet	<u>25.66</u>	0.562	**0.819**
	BLPF+FFDNet	25.29	<u>0.572</u>	0.802
25	FFDNet	24.83	0.494	0.759
	SLPF+FFDNet	25.01	0.474	0.780
	DLPF+FFDNet	**25.92**	**0.636**	<u>0.798</u>
	GLPF+FFDNet	<u>25.51</u>	0.519	**0.807**
	BLPF+FFDNet	25.23	<u>0.534</u>	0.789
50	FFDNet	24.18	0.378	0.701
	SLPF+FFDNet	24.06	0.364	0.722
	DLPF+FFDNet	**25.29**	**0.492**	0.746
	GLPF+FFDNet	24.66	0.403	**0.759**
	BLPF+FFDNet	<u>24.75</u>	<u>0.448</u>	<u>0.747</u>

Table 7 Average quantitative results (PSNR/SSIM/FSIM) for the different types of noise in CT image denoising [Bold: Best, Underline: Second Best]

Noise	Method	PSNR	SSIM	FSIM
Gaussian	FFDNet	25.21	0.597	0.787
	SLPF+FFDNet	25.23	0.602	0.800
	DLPF+FFDNet	<u>26.22</u>	**0.670**	0.807
	GLPF+FFDNet	**26.28**	0.643	**0.825**
	BLPF+FFDNet	26.01	<u>0.646</u>	<u>0.812</u>
Poisson	FFDNet	27.81	0.722	0.833
	SLPF+FFDNet	27.91	0.623	**0.843**
	DLPF+FFDNet	<u>28.09</u>	**0.753**	0.839
	GLPF+FFDNet	27.97	0.734	**0.843**
	BLPF+FFDNet	**28.14**	<u>0.748</u>	<u>0.842</u>
Poisson–Gaussian	FFDNet	27.10	0.634	0.814
	SLPF+FFDNet	27.33	0.687	<u>0.827</u>
	DLPF+FFDNet	<u>27.52</u>	**0.711**	0.825
	GLPF+FFDNet	27.38	0.690	**0.828**
	BLPF+FFDNet	**27.54**	<u>0.704</u>	0.826

Table 8 Execution time in seconds

Method	NLM	BM3D	FFDNet	FFDNet-IR
256×256	8.8	**0.6**	**0.6**	**0.6**
512×512	35.7	2.6	**0.7**	**0.7**

of noise in the Mayo-Clinic and QIN_LUNG_CT image datasets. Using these results, we conduct a qualitative comparison of different types of noise. The average scores of both datasets are shown in Table 7. Specifically, GLPF+ FFDNet can remove Gaussian noise, whereas BLPF+FFDNet and DLPF+FFDNet are effective specifically for Poisson and Poisson–Gaussian denoising. SLPF+FFDNet achieves superior FSIM results for Poisson and Poisson–Gaussian denoising. Overall, FFDNet-IR can not only eliminate noise but also preserve the image texture details better than the conventional FFDNet in the three types of noise removal, as shown in Table 7.

In addition to the high quality of the image,

another important aspect of medical image denoising is execution time. Table 8 presents the execution times of various methods for images with sizes of 256×256 and 512 × 512 with Poisson–Gaussian noise (σ=15 and α=10) on CPU. It can be seen that BM3D performs competitively with FFDNet-IR and FFDNet for the 256×256 image size. However, for images with a size of 512 × 512, FFDNet and FFDNet-IR, which include SLPF+FFDNet, DLPF+FFDNet, GLPF+FFDNet, and BLPF+FFDNet, have superior computational efficiency compared with BM3D and NLM.

To illustrate the effectiveness of denoising results, the Poisson–Gaussian noise removal results and magnified views of selected regions from the LDCT images from the Mayo Clinic dataset and the FDCT images from the QIN_LUNG_CT dataset are shown in Figs. 4 and 5. It should be noted that FFDNet and FFDNet-IR have higher visual performance for Poisson–Gaussian denoising than BM3D and NLM. Although BM3D and NLM effectively eliminate Gaussian noise, as seen in Figs. 4 and 5, they struggle with the Poisson noise of LDCT from the Mayo Clinic dataset and that of the FDCT image from QIN_LUNG_CT dataset.

On the other hand, FFDNet and FFDNet-IR effectively eliminate both Poisson noise and Gaussian noise while preserving the image's structural details compared with BM3D and NLM. However, FFDNet introduces blurred edges in smooth regions, resulting
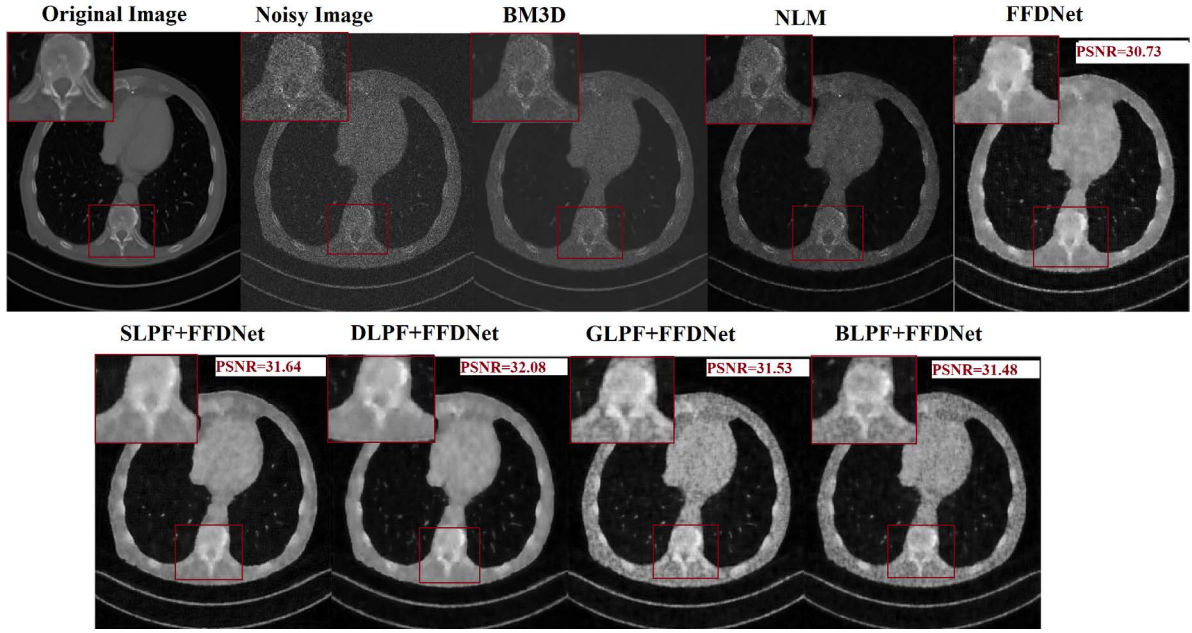
Fig. 4 Noise reduction for LDCT images affected by Poisson–Gaussian noise from the Mayo Clinic dataset (Poisson–Gaussian noise with $\sigma=15$ and $\alpha=10$)
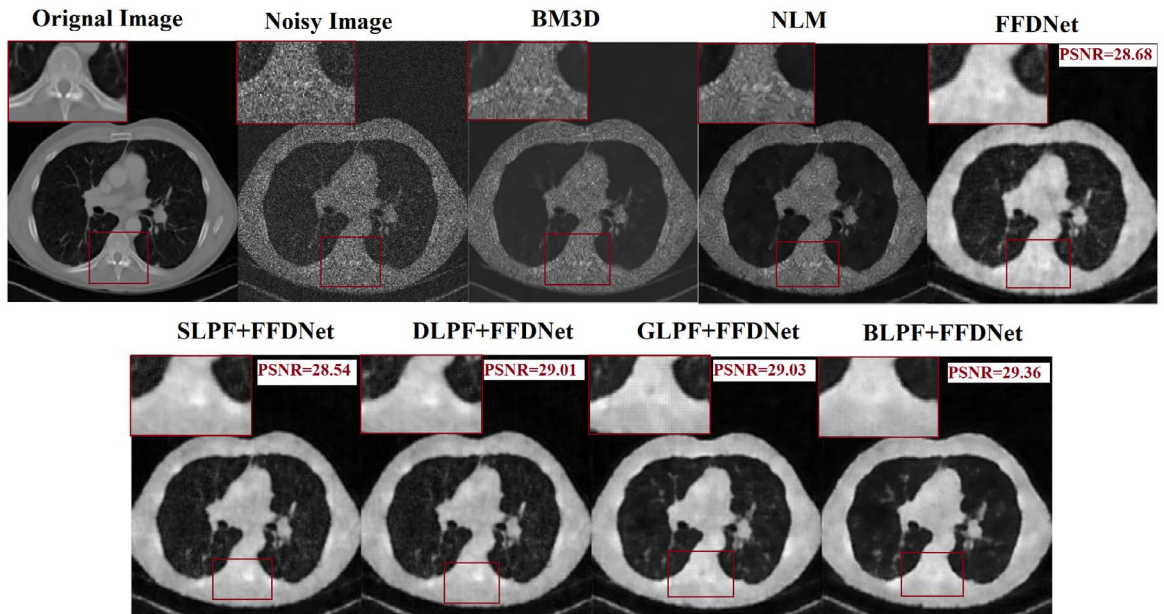

Fig. 5 Noise reduction for FDCT images affected by Poisson–Gaussian noise from the QIN_LUNG_CT dataset (Poisson–Gaussian noise with $\sigma=15$ and $\alpha=10$)

in a loss of detailed information, as shown in Fig. 5. Subsequently, FFDNet-IR shows more structure and image details, particularly in the white regions of LDCT images, and successfully eliminates Poisson–Gaussian noise, as shown in Figs. 4 and 5.

Figure 6 shows the improvement of FFDNet-IR specifically for Poisson denoising with α set to 10. BM3D and NLM exhibit the lowest performance in removing Poisson noise. In contrast, both FFDNet-IR and FFDNet demonstrate the ability to eliminate Poisson noise while preserving certain parts of the structural information. Visual evaluation performed by experts before clinical application is a future research topic.
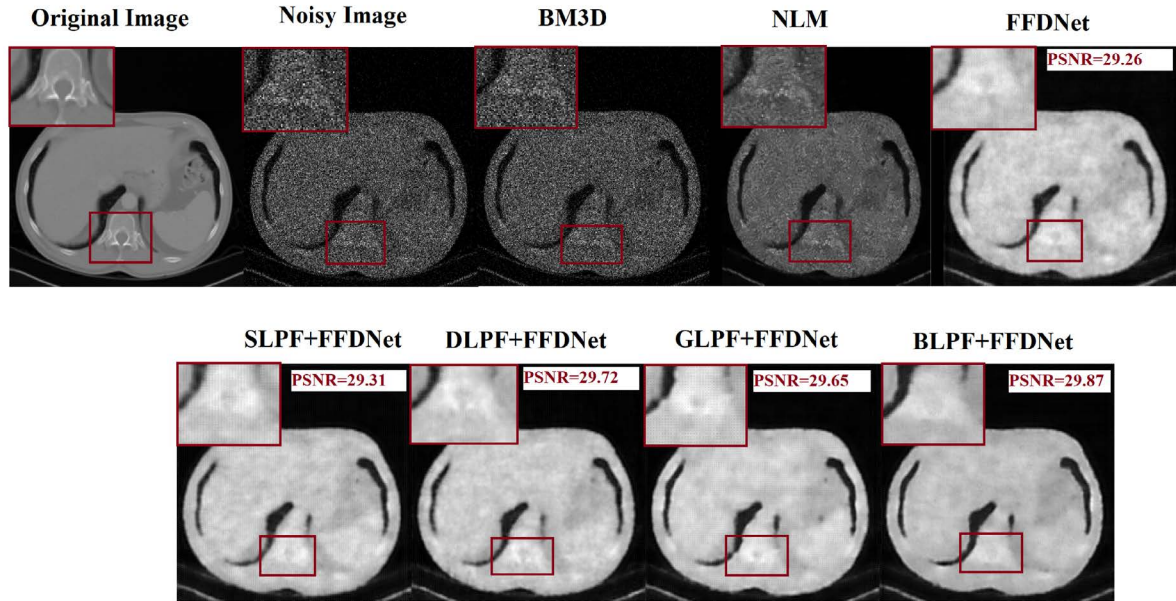
Fig. 6 Noise reduction for FDCT images affected by Poisson noise (α=10) from the QIN_LUNG_CT dataset

4. Conclusion

In this paper, we proposed FFDNet-IR as a technique of Poisson–Gaussian noise removal for LDCT and FDCT images. In FFDNet-IR, the noisy image patches and the impulse response of LPF are integrated into FFDNet. FFDNet-IR is effective not only for Poisson–Gaussian noise but also for a single track of either Gaussian or Poisson noise. FFDNet-IR provides higher performance than FFDNet and other state-of-the-art image denoisers on LDCT images from both quantitative and visual aspects. The experimental findings indicate that FFDNet-IR retains textural details, eliminates noise, and enhances LDCT and FDCT denoising results. In the future, the proposed idea will be extended to the edge detection of noisy images.

References

[1] A. Zhong, B. Li, N. Luo, Y. Xu, L. Zhou and X. Zhen: Image restoration for low-dose CT via transfer learning and residual network, IEEE Access, Vol. 8, pp. 112078-112091, 2020.

[2] Y. Chen, Z. Yang, Y. Hu, G. Yang, Y. Zhu, Y. Li, L. Luo, W. Chen and C. Toumoulin: Thoracic low-dose CT image processing using an artifact suppressed large-scale nonlocal means, Physics in Medicine and Biology, Vol. 57, No. 9, pp. 2667-2688, 2012.

[3] D. Kang, P. Slomka, R. Nakazato, J. Woo, D. S. Berman, C. Kuo and D. Dey: Image denoising of low-radiation dose coronary CT angiography by an adaptive block-matching 3D algorithm, Medical Imaging 2013: Image Processing, Vol. 8669, pp. 671-676, 2013.

[4] Y. Chen, X. Yin, L. Shi, H. Shu and C. Toumoulin: Improving abdomen tumor low-dose CT images using a fast dictionary learning based processing, Physics in Medicine and Biology, Vol. 58, No. 16, pp. 5803-5820, 2013.

[5] F. G. Bahador, P. Mokhtary and M. Lakestani: Mixed Poisson Gaussian noise reduction using time-space fractional differential equations, Information Sciences, Vol. 647, pp. 1-15, 2023.

[6] Y. Zhang, Y. Zhu, E. Nichols, Q. Wang, S. Zhang, C. Smith and S. Howard: A Poisson-Gaussian denoising dataset with real fluorescence microscopy images, Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition, pp. 11710-11718, 2019.

[7] M. T. Islam, S. M. Rahman and M. N. S. Swamy: Mixed Gaussian-impulse noise reduction from images using convolutional neural network, Signal Processing: Image Communication, Vol. 68, pp. 26-41, 2018.

[8] M. Mafi, W. Izquierdo, H. Martin, M. Cabrerizo and M. Adjouadi: Deep convolutional neural network for mixed random impulse and Gaussian noise reduction in digital images, IET Image Processing, Vol. 14, pp. 3791-3801, 2020.

[9] J. Zhang, Z. Shangguan, W. Gong and Y. Cheng: A novel denoising method for low-dose CT images based on the transformer and CNN, Computers in Biology and Medicine, Vol. 163, pp. 1-13, 2023.

[10] K. Zhang, W. Zuo and L. Zhang: FFDNet: Toward a fast and flexible solution for CNN-based image denoising, IEEE Trans. Image Processing, Vol. 27, pp. 4608-4622, 2018.

[11] M. T. Tun, Y. Sugiura and T. Shimamura: Image denoising by incorporating noisy image patch and impulse response of the low-pass filter in CNN learning, Proc. IEEE Global Conference on Consumer Electronics, pp. 721-722, 2023.

[12] M. T. Tun, Y. Sugiura and T. Shimamura: Joint training of noisy image patch and impulse response of the low-pass filter in CNN for image denoising, J. Signal Process., Vol.28, No. 1, pp. 1-17, 2024.

[13] J. Byun, S. Cha and T. Moon: Fbi-denoiser: Fast blind image denoiser for Poisson–Gaussian noise, Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition, pp. 5768-5777, 2021.

[14] N. J. Nyunt, Y. Sugiura and T. Shimamura: Parametric Wiener filter based on image power spectrum sparsity, J. Signal Process., Vol. 22, No. 6, pp. 287-297, 2018.

[15] L. Zhang, L. Zhang, X. Mou and D. Zhang: FSIM: A feature similarity index for image quality assessment, IEEE Trans. Image Processing, Vol. 20, No. 8, pp. 2378-2386, 2011.

May Thet Tun received her B.C.Sc. (Hons.) and M.C.Sc. degrees from the University of Computer Studies, Yangon, Myanmar, in 2015 and 2018, respectively. She is currently pursuing her Ph.D. degree in the field of digital image processing at the Graduate School of Science and Engineering, Saitama University, Saitama, Japan.

Yosuke Sugiura received his B.E., M.E., and Ph.D. degrees from Osaka University, Osaka, Japan in 2009, 2011 and 2013 respectively. In 2013, he joined Tokyo University of Science, Tokyo, Japan. In 2015, he joined Saitama University, Saitama, Japan, where he is currently an Assistant Professor. His research interests include digital signal processing, adaptive filter theory, and speech information processing.

Tetsuya Shimamura received his B.E., M.E., and Ph.D. degrees in electrical engineering from Keio University, Yokohama, Japan, in 1986, 1988, and 1991, respectively. In 1991, he joined Saitama University, Saitama, Japan, where he is currently a Professor. He was a visiting researcher at Loughborough University, U.K., in 1995 and at Queen's University of Belfast, U.K., in 1996, respectively. Professor Shimamura is an author and coauthor of six books. He serves as an editorial member of several international journals and is a member of the organizing and program committees of various international conferences. His research interests are in digital signal processing and its application to speech, image, and communication systems.