Self-Supervised Denoising with Noise Propagation Model: Improving Material Decomposition in Photon-Counting CT

Qianyu Wu, Xu Ji, Xiaoxue Lei, Xiaopeng Yu, Mengqing Su, Wenhui Qin, Yikun Zhang, Wenying Wang, Yanyan Liu, Guotao Quan, et al.

To cite this version:

Qianyu Wu, Xu Ji, Xiaoxue Lei, Xiaopeng Yu, Mengqing Su, et al.. Self-Supervised Denoising with Noise Propagation Model: Improving Material Decomposition in Photon-Counting CT. IEEE Transactions on Biomedical Engineering, 2025, pp.1-12. ⟨10.1109/TBME.2025.3620135⟩. ⟨hal-05471534⟩

HAL Id: hal-05471534

https://univ-rennes.hal.science/hal-05471534v1

Submitted on 13 Feb 2026

Self-Supervised Denoising with Noise Propagation Model: Improving Material Decomposition in Photon-Counting CT

Qianyu Wu, Xu Ji, Xiaoxue Lei, Xiaopeng Yu, Mengqing Su, Wenhui Qin, Yikun Zhang, Wenying Wang, Yanyan Liu, Guotao Quan, Gouenou Coatrieux, *Senior Member, IEEE*, Jean-Louis Coatrieux, *Life Fellow, IEEE*, Xiaochun Lai, Yang Chen, *Senior Member, IEEE*

***Abstract*—The inherent spectral properties of photon-counting computed tomography (PCCT) allow detailed material identification through decomposition techniques, but these methods often amplify image noise and artifacts. Current denoising approaches mainly focus on improving already degraded images, ignoring the fundamental noise caused by random variations in photon detection. To tackle these issues, we combine a physics-based noise analysis with deep learning to control noise during the material decomposition process. Our work has three key parts: (1) A noise analysis model that explains how random photon-count variations in the detector affect the noise levels in different materials after decomposition. This model connects the Poisson-distributed detector noise to material-specific noise patterns. (2) A self-supervised training method that combines the noise model with neural networks using probability-based optimization, allowing the system to learn from limited training data without needing high-quality data. (3) A flexible image improvement system that adapts to different body structures and noise conditions, ensuring reliable results across various scanning scenarios. Tests using real patient scan data show our method better preserves material accuracy and produces cleaner virtual monochromatic images compared to traditional approaches. Importantly, our solution works effectively with small training datasets and can be practically used in hospital settings without slowing down workflows. This research bridges the gap between theoretical noise analysis and clinical medical imaging needs, offering a balanced approach to improving PCCT technology.**

***Index Terms*—Photon-counting CT, computed tomography imaging, material decomposition, image denoising.**

I. INTRODUCTION

Photon Counting Computed Tomography (PCCT) is an emerging spectral CT technology, offering the potential for ultra-low dose and ultra-high resolution imaging [1]–[4]. Photon counting detectors (PCDs) in PCCT can directly measure x-ray photon energy, enabling precise differentiation of materials. This capability is of crucial significance in clinical diagnostics, particularly in applications such as kidney stone characterization [5] and iodine quantification [6].

One approach for extracting material property information involves decomposing the acquired projections or reconstructed image into a set of coefficients of basis materials, called material decomposition (MD) [7], [8]. Depending on the target of decomposition, this technique is generally categorized into two types: projection domain decomposition [9], [10] and image domain decomposition [11], [12]. Projection domain decomposition employs a nonlinear model to transform the measured projection values into line integrals representing the basis materials (such as soft tissue and bone). Then, the decomposed CT images are reconstructed using traditional CT reconstruction algorithms. Image domain decomposition entails performing a linear matrix inversion on conventional CT images reconstructed from measured counts of two or more energy levels. However, both methodologies tend to result in significant noise amplification [13], which has emerged as a pervasive challenge associated with MD. Noise issue is even worse, as event receive is only one over tenth of a conventional detector pixel, because of fin pitch design of PCD.

Over the past few decades, several conventional algorithms have been proposed to address this issue. Dobbins et al. [14] introduced a noise clipping algorithm designed to constrain the upper limit of pixel intensity in high-energy images,

This work was supported by the National Key Research and Development Program of China under Grant 2022YFC2401600. (Corresponding author: Yang Chen and Xiaochun Lai)

Yang Chen, Xu Ji and Yikun Zhang are with the Jiangsu Provincial Joint International Research Laboratory of Medical Information Processing, Southeast University, Nanjing China, and also with the Laboratory of Image Science and Technology, School of Computer Science and Engineering, Southeast University Nanjing 210096, China and also with the Key Laboratory of New Generation Artificial Intelligence Technology and Its Interdisciplinary Applications (Southeast University), Ministry of Education, Nanjing 210096, China. (e-mail: chenyang.list@seu.edu.cn, xuji@seu.edu.cn, 101013635@seu.edu.cn)

Qianyu Wu and Mengqing Su are with the Laboratory of Image Science and Technology, School of Computer Science and Engineering, Southeast University, Nanjing 210096, China. (e-mail: wqy021434@163.com, 230249715@seu.edu.cn)

Xiaochun Lai is with the Photonic Sensing and Imaging Lab, School of Biomedical Engineering, ShanghaiTech University, Shanghai 201210, China, and also with the Shanghai Clinical Research and Trial Center, Shanghai, China. (e-mail: laixch@shanghaitech.edu.cn)

Xiaoxue Lei, Xiaopeng Yu and Wenhui Qin are with the Photonic Sensing and Imaging Lab, School of Biomedical Engineering, ShanghaiTech University, Shanghai 201210, China. Qianyu Wu is also a visiting student in this lab. (email: leixx2022@shanghaitech.edu.cn, yuxp2022@shanghaitech.edu.cn, qinwh@shanghaitech.edu.cn)

Wenying Wang, Yanyan Liu and Guotao Quan are with the CT RPA Department, United Imaging Healthcare Co., Ltd., Shanghai 201807, China. (e-mail: wenying.wang@united-imaging.com, yanyan.liu@united-imaging.com, guotao.quan@united-imaging.com)

Gouenou Coatrieux is with the IMT Atlantique, Inserm, LaTIM UMR1101,Brest 29000, France. (e-mail: gouenou.coatrieux@imt-atlantique.fr)

Jean-Louis Coatrieux is with the Laboratoire Traitement du Signal et de l'Image, Universit de Rennes 1, F-35000 Rennes, France, with the Centre de Recherche en Information Biomedicale Sino-francais, 35042 Rennes, France, and also with the National Institute for Health and Medical Research, F-35000 Rennes, France. (e-mail: jean-louis.coatrieux@univ-rennes1.fr)

leveraging the characteristics of higher contrast observed in low-energy images. Leng et al. [15] proposed an energy-domain noise reduction algorithm that utilizes information redundancy to maximize iodide contrast-to-noise ratio (CNR) in adult abdominal CT. Warp et al. [16] developed a low-energy image weighted smoothing kernel aimed at preserving edge details within high-energy images. Advanced algorithms exploit the negative statistical correlation between basis images to mitigate noise effectively [12], [17]. Furthermore, iterative methods also contribute to noise reduction by incorporating various regularization terms. For instance, Niu et al. [18] proposed an iterative image domain MD method that capitalizes on statistical correlations among basis images for enhanced noise reduction. Xue et al. [19] formulated a statistical image-domain decomposition approach featuring a negative logarithm-likelihood term alongside an edge-preserving regularization technique. However, these methods heavily rely on manually designed regularization terms that require specific fine-tuning processes; thus, their robustness for practical applications requires further enhancement.

With the advancement of deep learning (DL) technology [20]–[22], an increasing body of researches have incorporated deep neural networks into the MD process to alleviate noise and artifacts. Supervised DL-based reconstruction demonstrates exceptional performance in terms of noise reduction when a considerable amount of high-quality labels are available for training [23]–[28]. Hu et al. [29] leveraged the non-local similarity inherent in spectral CT images to cluster similar patches into a fourth-order tensor group. Zhang et al. proposed a multi-level network architecture capable of restoring poly-energetic sinograms and estimating monoenergetic sinograms. However, due to the challenges associated with acquiring sufficient high-quality training data from clinical commercial CT scanners (this is even worse for PCCT since it is still in the early stage of clinical studies), recent studies have made strenuous efforts to reconstruct high-quality images without any reference. For instance, Niu et al. [30] achieved denoising effects comparable to those obtained through supervised learning by utilizing paired noisy-noisy datasets. Approaches such as Noise2Void [31] and Noise2Self [32] employed blind-spot learning by omitting pixels at the center of each receptive field and substituting them with neighboring pixels. Nevertheless, these approaches do not fully explore the noise characteristics associated with the decomposition process. They ignored the additional noise resulting from the the quantum fluctuation of the received photons, which consequently limits their applicability in clinical PCCT imaging tasks.

In this paper, we proposed a novel noise propagation model to transform the detector noise, which follows a Poisson distribution, into the noise of basis materials. Furthermore, we integrated this model with DL to effectively mitigate noise in MD processes. Specifically, we derived a formulation that decomposes the material covariance matrix, which can convert the fluctuations of photons measured by each detector pixel into the corresponding noise of basis materials. Based on this formulation, we incorporated the noise propagation model into a self-supervised learning framework to generate high-quality images of the basis materials. Finally, we utilized an all-in-one image restoration framework to improve the robustness of our method for different clinical settings.

The main contributions of our work can be summarized as follows: **(i)** A novel noise propagation model was proposed to derive the MD coefficient noise from the detector noise based on Poisson distribution, which is widely used in our field. **(ii)** A self-supervised noise reduction network was developed to incorporate the noise propagation model and effectively mitigate the basis material noise without requiring high-quality training labels. **(iii)** An all-in-one restoration framework was introduced to enhance the robustness of our method across different phantoms and noise levels. Consequently, there is no requirement for additional parameter settings during the inference phase. **(iv)** Extensive analyses conducted on real data have demonstrated that the proposed method significantly improved both the accuracy of MD in PCCT and the quality of virtual monochromatic imaging (VMI).

II. METHODOLOGY

The proposed method, illustrated in Fig. 1, consists of three main components. The first component is the proposed noise propagation model. PCCT counts transmitted photons separately, and raw measurements are assumed to follow a Poisson distribution. By employing the noise propagation model, the noise obtained on a single detector pixel can be converted to the noise corresponding to the MD coefficient. The second component is based on a self-supervised denoising network, which first utilizes traditional method [9] to decompose the original data into the projections of basis materials, and then uses the denoising network to generate a high-quality restored images. The third component is used for robustness training to improve the robustness of the method to different clinical settings. The following briefly introduces each component mentioned above.

A. Noise Propagation Model in Material Decomposition

The linear attenuation coefficient of matter for X-rays is mainly constituted by the Compton scattering and photoelectric effect terms:

$$\mu(E) = \alpha_1 \mu_p(E) + \alpha_2 \mu_c(E), \tag{1}$$

where $\mu\left(E\right)$ is the distribution of the linear attenuation coefficient in terms of energy E. α is the decomposition coefficient. μ_p and μ_c are the coefficients of the photoelectric effect and Compton scattering, respectively. A significant defect of the above decomposition model is its low accuracy in depicting actual substances. From the perspective of clinical application, decomposing the attenuation coefficient into images of different physical effects cannot directly provide anatomical structure and physiological information, which makes it difficult for radiologists to interpret the decomposed images. Another decomposition method uses the actual material attenuation coefficient instead of the photoelectric effect and Compton scattering:

$$\mu(E, \vec{r}) = a_1(\vec{r}) \mu_1(E) + a_2(\vec{r}) \mu_2(E), \tag{2}$$

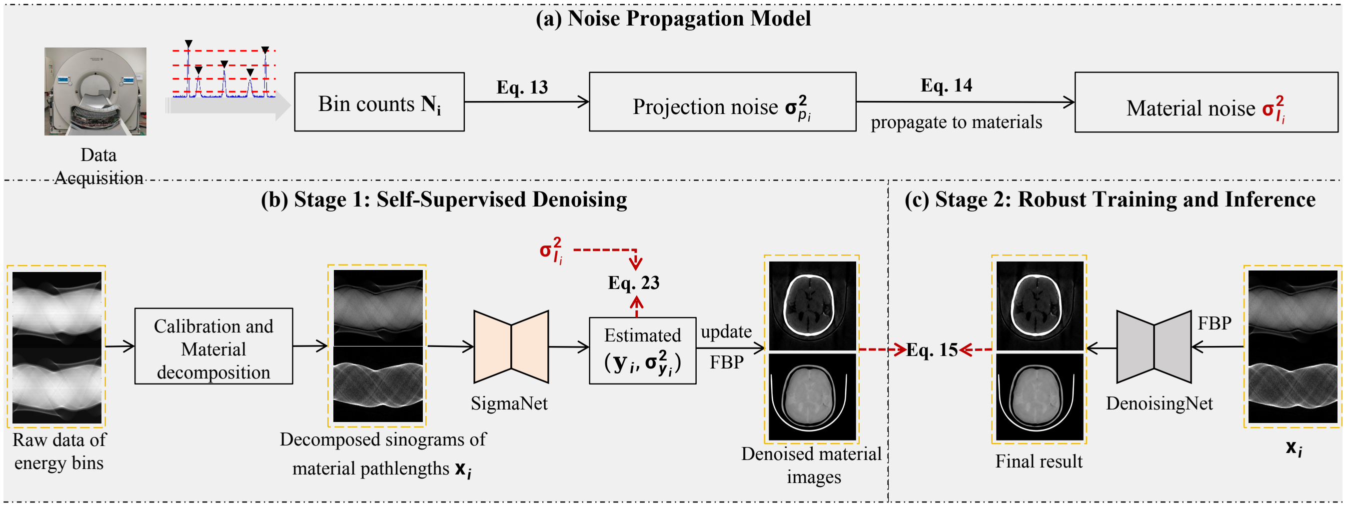


Fig. 1: Overall flowchart of the proposed method. (a) demonstrates the corresponding noise variances based on the counts of each detector pixel, and then converts them into the noise of the basis materials. (b) is the first stage that incorporates the noise model with self-supervised denoising. (c) is an all-in-one network to improve the robustness of the model to different phantoms and noise levels. This section is also used for the inference.

where $\mu(E, \vec{r})$ represents the distribution of the linear attenuation coefficient as a function of energy E and position vector $\vec{r}$. $\mu_1(E)$ and $\mu_2(E)$ are the specific attenuation of two materials under energy E (this paper mainly studies the dual-material denoising algorithm). $a_1(\vec{r})$ and $a_2(\vec{r})$ are the corresponding coefficients of the two material bases related to position $\vec{r}$.

In real scanning scenarios, the PCCT spectrum response model can be expressed in a MD form:

$$N_i = N_{0i} \int_{E_{\min}}^{E_{\max}} \Omega_i(E) \exp\left(-\mu_1(E) l_1) - \mu_2(E) l_2\right) dE, \tag{3}$$

where $[E_{\min}, E_{\max}]$ denotes the entire energy range. $l_1 = \int a_1(\vec{r}) \mathrm{d}\vec{r}$ and $l_2 = \int a_2(\vec{r}) \mathrm{d}\vec{r}$ are the effective lengths of corresponding materials along the integration path. $\Omega_i(E)$ represents the normalized X-ray spectrum distribution function of the i_{th} energy bin and can be expressed as follows:

$$\Omega_i(E) = \int_{E \in S_i} R(E'|E) \Omega_0(E) \, dE', \tag{4}$$

where $R(E'|E)$ is the response function of each detector, representing the probability that the input energy E is detected as E'. $\Omega_0(E)$ denotes the spectrum generated from the X-ray source. S_i denotes the energy range of the i_{th} energy bin (e.g., [30, 60) keV or [60, 120] keV in experimental setups). The integral along E denotes that within the i_{th} energy bin, the photon count N_i received by the detector is jointly determined by the unattenuated photon count N_{0i}, the spectral distribution $\Omega(E)$, and the material attenuation term. The corresponding post-logs projection value is

$$\begin{aligned} p_i &= -\ln \frac{N_i}{N_{0i}} \\ &= -\ln \int_{E_{\min}}^{E_{\max}} \Omega_i(E) \exp(-\mu_1(E) l_1 - \mu_2(E) l_2) \mathrm{d}E. \end{aligned} \tag{5}$$

Since p_i is continuously differentiable within its domain, the Taylor expansion of p_i at (k_1, k_2) is,

$$\begin{aligned} p_i(l_1, l_2) = p_i(k_1, k_2) + \left.\frac{\partial p_i(l_1, l_2)}{\partial l_1}\right|_{\substack{l_1 = k_1 \\ l_2 = k_2}} (l_1 - k_1) \\ + \left.\frac{\partial p_i(l_1, l_2)}{\partial l_2}\right|_{\substack{l_1 = k_1 \\ l_2 = k_2}} (l_2 - k_2) + R_n(l_1, l_2) \end{aligned} \tag{6}$$

where $R_n(\cdot)$ represents the remaining higher-order terms. Then, Eq. 6 is developed into:

$$\begin{aligned} & p_i(l_1, l_2) \\ &= \frac{\int_{E_{\min}}^{E_{\max}} \Omega_i(E) \mu_1(E) \exp(-\mu_1(E) k_1 - \mu_2(E) k_2) \mathrm{d}E}{\int_{E_{\min}}^{E_{\max}} \Omega_i(E) \exp(-\mu_1(E) k_1 - \mu_2(E) k_2) \mathrm{d}E} \cdot (l_1 - k_1) \\ &+ \frac{\int_{E_{\min}}^{E_{\max}} \Omega_i(E) \mu_2(E) \exp(-\mu_1(E) k_1 - \mu_2(E) k_2) \mathrm{d}E}{\int_{E_{\min}}^{E_{\max}} \Omega_i(E) \exp(-\mu_1(E) k_1 - \mu_2(E) k_2) \mathrm{d}E} (l_2 - k_2) \\ &+ R_n(l_1, l_2). \end{aligned} \tag{7}$$

MD acquires two measurements under different energy bins. We denote $\vec{p} = [p_1, p_2]^\top$ and $\vec{l} = [l_1, l_2]^\top$, and the material noise by $\delta l_i = l_i - \bar{l}_i$, where the mean value $\bar{l}_i$ represents the noise-free material. The noise term $\delta\vec{l}$ causes the change in the value of $\delta\vec{p}$. Since this paper investigates the correlation between $\vec{p}$ and $\vec{l}$, and such small shift can be proximate by first order proximation:

$$\delta\vec{p} = \mathbf{M} \delta\vec{l} \tag{8}$$

$$\mathbf{M} = \begin{bmatrix} \frac{\int_{E_{\min}}^{E_{\max}} \Omega_1(E) \mu_1(E) \mathbf{A} \mathrm{d}E}{\int_{E_{\min}}^{E_{\max}} \Omega_1(E) \mathbf{A} \mathrm{d}E} & \frac{\int_{E_{\min}}^{E_{\max}} \Omega_1(E) \mu_2(E) \mathbf{A} \mathrm{d}E}{\int_{E_{\min}}^{E_{\max}} \Omega_1(E) \mathbf{A} \mathrm{d}E} \\ \frac{\int_{E_{\min}}^{E_{\max}} \Omega_2(E) \mu_1(E) \mathbf{A} \mathrm{d}E}{\int_{E_{\min}}^{E_{\max}} \Omega_2(E) \mathbf{A} \mathrm{d}E} & \frac{\int_{E_{\min}}^{E_{\max}} \Omega_2(E) \mu_2(E) \mathbf{A} \mathrm{d}E}{\int_{E_{\min}}^{E_{\max}} \Omega_2(E) \mathbf{A} \mathrm{d}E} \end{bmatrix}, \tag{9}$$

where $\mathbf{A} = \exp(-\mu_1(E) k_1 - \mu_2(E) k_2)$. For convenience, we denote $\mathbf{M}$ as

$$\mathbf{M} = \begin{bmatrix} \mu_{11} & \mu_{21} \\ \mu_{12} & \mu_{22} \end{bmatrix} \tag{10}$$

where the two subscripts of μ indicate the basis material and the energy bin, respectively. According to Eq. 8, we also have

$$\delta\vec{l} = \mathbf{M}^{-1}\delta\vec{p}, \quad \mathbf{M}^{-1} = \frac{1}{|\mathbf{M}|}\begin{bmatrix}\mu_{22} & -\mu_{21}\\ -\mu_{12} & \mu_{11}\end{bmatrix}, \tag{11}$$

where $|\mathbf{M}|$ denotes the determinant of $\mathbf{M}$. Thus, the covariance matrix of $\vec{l}$ is

$$\begin{aligned}\mathrm{Cov}(\vec{l}) =& \mathbb{E}[\delta\vec{l}\,\delta\vec{l}^{\top}]\\ =& \mathbb{E}[(\mathbf{M}^{-1}\delta\vec{p})(\mathbf{M}^{-1}\delta\vec{p})^{\top}]\\ =& \mathbf{M}^{-1}\mathrm{Cov}(\vec{p})(\mathbf{M}^{-1})^{\top}\\ =& \mathbf{M}^{-1}\begin{bmatrix}\sigma_{p_1}^2 & \mathrm{Cov}(p_1,p_2)\\ \mathrm{Cov}(p_2,p_1) & \sigma_{p_2}^2\end{bmatrix}(\mathbf{M}^{-1})^{\top}.\end{aligned} \tag{12}$$

Eq. 12 demonstrates how the noise from the postlog-projection data is transformed into the noise of the basis materials. It is then clarified how to obtain the noise of the post-log projection data based on the number of photons received by the detector.

For a function $y = f(x)$ with a random variable $X = x$, its variance $\sigma_y^2 \approx (f'(\bar{X}))^2\sigma_x^2$, where $\bar{X}$ is the mean value of X. In our case, we have

$$\sigma_{p_i}^2 \approx \left(\frac{\mathrm{d}p_i}{\mathrm{d}N_i}\bigg|_{N_i=\bar{N}_i}\right)^2 \sigma_{N_i}^2 = \left(\frac{1}{\bar{N}_i}\right)^2 \sigma_{N_i}^2. \tag{13}$$

For a ground-truth quanta of photons λ, the measurement N as a random variable obeys the Poisson distribution. The mean value $\bar{N} = \lambda$ and variance $\sigma_N^2 = \lambda = \bar{N}$. Thus, $\sigma_{p_i}^2 \approx \frac{1}{\bar{N}_i}$. Then, by substituting Eq. 13 into Eq. 12, the covariance matrix of basis materials using the noise propagation model is

$$\begin{aligned}\mathrm{Cov}(\vec{l})_{poisson} &= \mathbf{M}^{-1}\begin{bmatrix}\frac{1}{\bar{N}_1} & 0\\ 0 & \frac{1}{\bar{N}_2}\end{bmatrix}(\mathbf{M}^{-1})^{\top}\\ &= \frac{1}{|\mathbf{M}|^2}\begin{bmatrix}\frac{\mu_{22}^2}{\bar{N}_1} + \frac{\mu_{21}^2}{\bar{N}_2} & k_p\\ k_p & \frac{\mu_{12}^2}{\bar{N}_1} + \frac{\mu_{11}^2}{\bar{N}_2}\end{bmatrix}\end{aligned} \tag{14}$$

where k_p represents the covariance of the materials. Here, we ignore the correlation term as the low flux leads to a low correlation between energy bins.

B. Calculation of the Noise Term

To calculate the noise term, it is necessary to know the four coefficients μ in Eq. 14, as well as the mean values $\bar{N}_1$ and $\bar{N}_2$. As can be seen from formulas Eq. 8, the noise terms of the two materials under the two energy bins can be expanded as:

$$\begin{bmatrix}\delta p_1\\ \delta p_2\end{bmatrix} = \begin{bmatrix}\mu_{11} & \mu_{21}\\ \mu_{12} & \mu_{22}\end{bmatrix}\begin{bmatrix}\delta l_1\\ \delta l_2\end{bmatrix}, \mathbf{M} = \begin{bmatrix}\mu_{11} & \mu_{21}\\ \mu_{12} & \mu_{22}\end{bmatrix} \tag{15}$$

Building upon our previous research on PCCT material decomposition, the basis materials can be estimated using polynomial fitting [9]:

$$\begin{cases}l_1 = \alpha_{1,0}p_1 + \alpha_{0,1}p_2 + \alpha_{2,0}(p_1)^2 + \alpha_{0,2}(p_2)^2 + \alpha_{1,1}p_1p_2 + \ldots\\ l_2 = \beta_{1,0}p_1 + \beta_{0,1}p_2 + \beta_{2,0}(p_1)^2 + \beta_{0,2}(p_2)^2 + \beta_{1,1}p_1p_2 + \ldots\end{cases} \tag{16}$$

where α and β represent the decomposition coefficients. The matrix form without higher-order terms is:

$$\begin{bmatrix}l_1\\ l_2\end{bmatrix} = \begin{bmatrix}\alpha_{1,0} & \alpha_{0,1}\\ \beta_{1,0} & \beta_{0,1}\end{bmatrix}\begin{bmatrix}p_1\\ p_2\end{bmatrix}, \mathbf{G} = \begin{bmatrix}\alpha_{1,0} & \alpha_{0,1}\\ \beta_{1,0} & \beta_{0,1}\end{bmatrix} \tag{17}$$

Therefore, the following conclusion can be easily drawn according to Eq. 11:

$$\mathbf{G} = \mathbf{M}^{-1} = \frac{1}{|\mathbf{M}|}\begin{bmatrix}\mu_{22} & -\mu_{21}\\ -\mu_{12} & \mu_{11}\end{bmatrix} \tag{18}$$

The decomposition coefficients $\mathbf{G}$ obtained via the polynomial fitting method [9] can be directly used to compute coefficients μ in Eq. 14, while the mean values $\bar{N}_1$ and $\bar{N}_2$ are derived from the local averages of the measured bin counts. Moreover, when the number of energy bins exceeds two, the non-invertibility issue of matrix $\mathbf{M}$ can be addressed using regularization or pseudo-inverse techniques to ensure the model's applicability.

Compared with conventional forward modeling [12], [33]–[35], our noise propagation model directly links photon count variations to basis material noise patterns and allows the network to learn noise characteristics directly from physical principles rather than purely data-driven fitting.

C. Deep Learning for Self-Supervised Denoising

This section mainly elaborates on the denoising process utilizing the noise model. First, the improved Ne2Ne network (SigmaNet) is employed for self-supervised denoising to generate preliminary denoising results while simultaneously predicting noise variance. Then, integrating the variance derived from the noise model, the negative log-likelihood loss function is utilized to further fine-tune the predicted values of SigmaNet. Finally, a robust training method is implemented to further improved the denoising results.

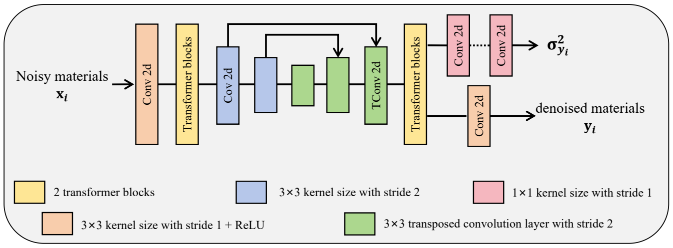


Fig. 2: The architectures of the proposed SigmaNet in Fig. 1. Not only does it employ the self-supervised learning strategy [36] to generate high-quality images, but it can also predict noise levels at specific locations via a series of 1×1 convolutional layers.

1) Improved Ne2Ne Denoising Method: The data-driven supervised learning generally explores the correlation between the low-quality sample x and the high-quality reference y. It aims to optimize a estimation network to predict the latent mapping space from x to y, which can be expressed as:

$$\arg\min_{\theta}\mathbb{E}_{x,y}\|g_\theta(x) - y\|_2^2 \tag{19}$$

where $g_\theta(x)$ is a deep neural network g_θ with a noisy input x and y is the corresponding clean reference (ground-truth). $\|\cdot\|_2^2$ denotes the L2 regularization.

To implement the training more effectively, we first utilized the self-supervised learning strategy Ne2Ne [36] to obtain the initial results. It relies on the assumption that adjacent pixels are weakly correlated, and uses random downsampling operators to form training labels and calculate reconstruction loss:

$$L_{rec} = \left\| g_\theta \left(s_1(x) \right) - s_2(x) \right\|_2^2 \tag{20}$$

Then, we utilized a down-sampling operator $s = (s_1, s_2)$ to obtain similar noisy pair $(s_1(x), s_2(x))$ from a noisy input x. Thus, the fidelity term of the self-supervised loss function can be expressed as. Although the ground truths of $s_1(x_1)$ and $s_2(x_2)$ are very close, they are not identical. Applying Eq. 20 directly may lead to over-smoothing. Therefore, the regularization is added to correct the gap between the similar ground truths:

$$L_{reg} = \left\| g_\theta \left(s_1 \left(x \right) \right) - s_2 \left(x \right) + \left[s_2 \left(g_\theta \left(x \right) \right) - s_1 \left(g_\theta \left(x \right) \right) \right] \right\|_2^2. \tag{21}$$

Therefore, the total loss function for optimizing network g_θ is:

$$L_{ne2ne} = L_{rec} + \beta L_{reg} \tag{22}$$

where β is a coefficient that balances the weight of the regularization term.

As illustrated in Fig. 2, the proposed SigmaNet is a transformer-based network that leverages a U-Net architecture [37] to estimate the parameters of the noise model. We assume that the noise is pixel-independent, signal-dependent, and image-specific. To achieve this, we employ multiple 1×1 convolutional layers to generate an image-specific noise level function for effective noise modeling. By utilizing all 1×1 convolutional layers, we ensure that the noise level at any given position depends solely on the input value at that same position. For a feature map y_i with n channels, the output of this module at one pixel can be represented as a $n \times n$ covariance matrix, where the variance terms on the main diagonal are $\{\sigma_{y_1}^2, \sigma_{y_2}^2, \ldots, \sigma_{y_n}^2\}$, thereby facilitating the modeling of material noise.

2) Integration of Noise Propagation Model: For a given material l_i, we have $x_i = \widetilde{y}_i + n_i$ with $n_i \sim \mathcal{N}(0, \sigma_{l_i}^2)$. Here, x_i and $\widetilde{y}_i$ are the noisy basis material and its underlying clean target, respectively.

Let y_i be the preliminary denoising result obtained by the neural network g_θ (the SigmaNet in Fig. 1). Assuming that $y_i = \widetilde{y}_i + n_i^*$ with $n_i^* \sim \mathcal{N}(0, \sigma_{y_i}^2)$ and i denotes the i_{th} material. Since y_i is the initial result, it should be closer to $\widetilde{y}$ than x. We denote $\alpha_i = x_i - y_i$ with $\alpha_i \sim \mathcal{N}(0, \sigma_{l_i}^2 + \sigma_{y_i}^2)$. Also, we denote a vector $x \in \mathbb{R}^{n \times 1}$ formed by concatenating l_1 and l_2 ($x = [l_1 \;\; l_2]^\top$) , which is fed into the SigmaNet as a multi-channel input. The negative log-likelihood of $x - y$ can be expressed as:

$$L_{\text{sigma}} = (x-y)^\top (\Sigma_l + \Sigma_y)^{-1} (x-y) + \log |\Sigma_l + \Sigma_y| \tag{23}$$

where $|\cdot|$ denotes the determinant of a matrix. $y \in \mathbb{R}^{n \times 1}$ is the denoising results of the proposed SigmaNet. Σ_l and Σ_y are $n \times n$ matrices. The elements on their main diagonals are $\{\sigma_{l_1}^2, \sigma_{l_2}^2, \ldots, \sigma_{l_n}^2\}$ and $\{\sigma_{y_1}^2, \sigma_{y_2}^2, \ldots, \sigma_{y_n}^2\}$, respectively, and the remaining elements are 0.

3) Robust Training: In general, in order to achieve optimal results, the neural network model will usually be trained according to clearly defined Settings. In the stage 1 in Fig. 1, all noise reduction methods need to be trained individually for different phantom and doses. In more complex clinical application scenarios, the use of a separate model for each task may lead to its inefficient use, hindering the deployment of deep learning-based CT imaging models from a single task to more general application scenarios. Recently, all-in-one image processing [38]–[41] has gained significant popularity in natural images, endeavoring to handle multiple distinct restoration tasks with a single universal model.

For convenience, we directly used the open source network [40] (the DenoisingNet in Fig. 1) for the second stage of training. Based on the first stage, we took the results of the SigmaNet as labels, the raw materials as the input of the DenoisingNet, and optimized it by Eq. 19. Consequently, there is no requirement for additional parameter settings during the inference phase, e.g., the type of phantom and the dose levels.

III. EXPERIMENT AND RESULTS

A. Experimental Datasets

The PCCT prototype system were built based on United Imaging's uCT 960+ system [44]. A large field of view (500mm in-plane) clinical prototype system equipped with photon counting detectors (PCDs) made of cadmium zinc telluride (CZT). The PCCT prototype is equipped with 38 photon-counting detector modules, each integrating 16 mini-modules. Each mini-module features a 12×8 array of macro pixels, with each macro pixel further divided into 9 ultra-high-resolution (UHR) sub-pixels. The **real head phantom** and **GAMMEX phantom data** were collected using the such prototype system. For rotational phantom scans, a 2-second acquisition time was used, with all experiments performed at a tube voltage of 120 kVp. Three exposure settings were employed: 100 mAs and 400 mAs for noisy data acquisition, and 800 mAs as high-quality ground truth (reference data) with reduced noise levels. The detector thresholds were set to 30 keV, 60 keV and above, thus the emitting spectrum was divided into two energy bins: [30, 60) keV and [60, 120] keV. The phantom sinogram dimensions are 912×2400×2, corresponding to 912 detector pixels per slice, 2400 projection views (evenly spaced over 360 degrees), and 2 energy bins. Each slice has a thickness of 0.625 cm, and reconstructed images are 512×512 pixels with a spatial resolution of 0.5 mm per pixel. For each energy bin, there are 32 slice sets in total. Each slice set contains projections acquired at different dose levels: 100 mAs and 400 mAs. Among these 32 sets per energy bin, 28 sets were allocated for training and 4 sets were reserved for testing.

B. Implementation Details

The proposed network was built based on the Pytorch platform with one NVIDIA RTX 4090 GPU. The parameters of the proposed network and all comparison methods were optimized using the Adam algorithm. The exponential decay rates β_1 and β_2 for Adam were 0.9, 0.999, respectively. In order to improve convergence efficiency, the initial learning rate was set to $2{\times}10^{-4}$ and then reduced by $1{\times}10^{-5}$ every 10 epochs until the final learning rate was $1{\times}10^{-5}$. The batch size was set to 32. All parameters were set to initial values

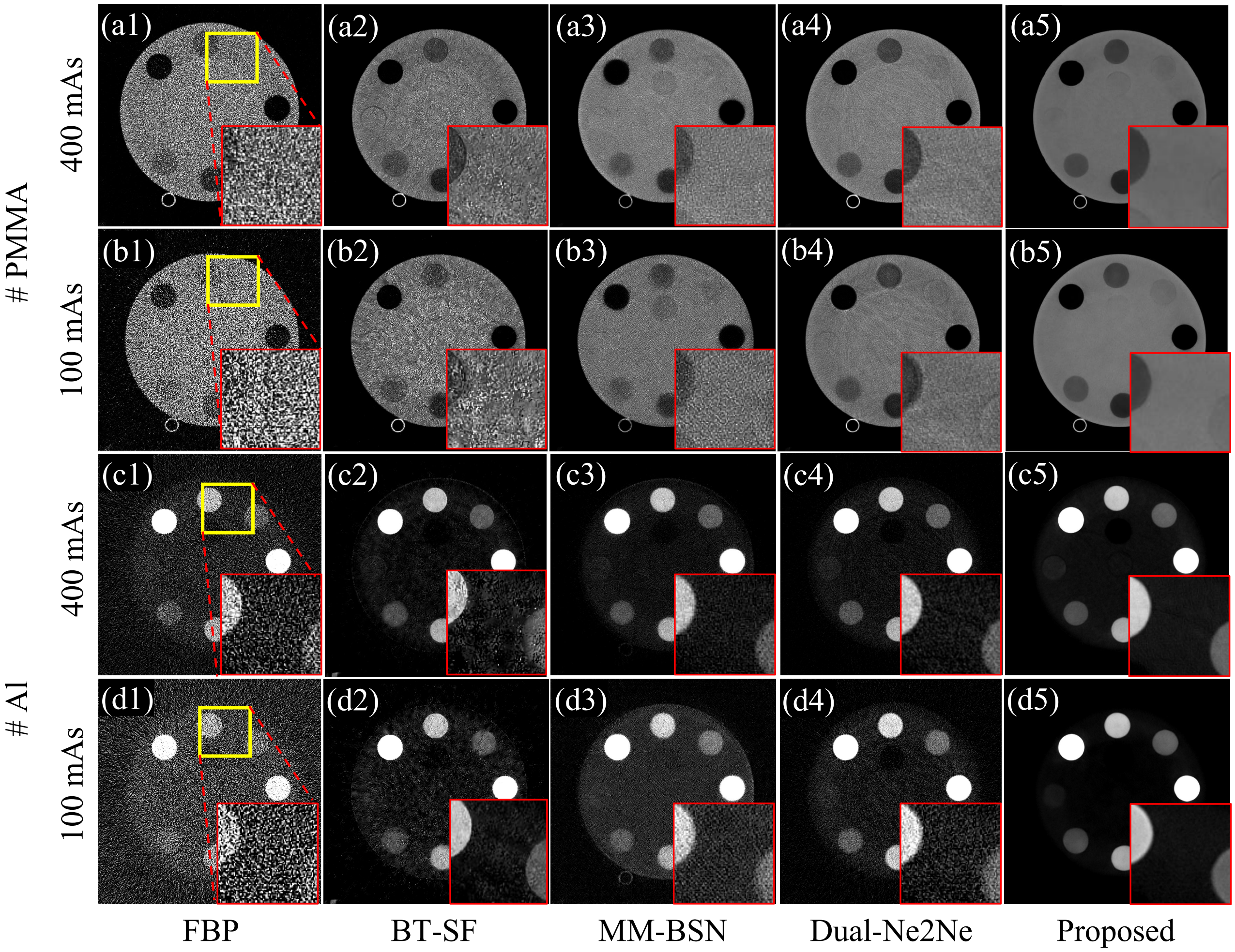


Fig. 3: Comparison of the reconstructed images for the PMMA and Al materials. Display range: [0,1500] mg/mL for PMMA basis images and [0,16] mg/mL for Al basis images.

based on experience and adjusted according to the specific experimental environment.

In training stage 1, we used data from a single phantom. Specifically, 32 slices scanned from the same phantom were split into 28 slices for training and 4 slices for testing. This setup ensured that the model first learned basic noise reduction capabilities under consistent imaging conditions. To enhance the model's generalizability, we then conducted a second training stage where data from different phantoms and varying dose conditions were mixed for robust training. The dataset splitting for robust training followed the same protocol as in training stage 1. Consequently, the training set for robust training comprised 112 slices (two phantoms and two dose levels), and the test set included 16 slices.

C. Experimental Results

In this section, the proposed method was compared with two well-established unsupervised methods, namely MM-BSN [43] and Dual-Ne2Ne [30], which have demonstrated reliable performance in the image domain and dual-domain noise reduction tasks respectively. Additionally, we utilized a classical noise reduction method BT-SF [42], to highlight the superiority of our proposed method. The conventional filtered back projection (FBP) was employed to obtain reconstructed images from projections. Moreover, we used the VMI obtained at a tube current of 400 mA as ground truth due to its lower noise levels. Nevertheless, due to the propensity of MD to amplify noise, the basis material images obtained at 400 mA exhibit considerable noise levels and are consequently inappropriate for use as reference images in assessing the reconstruction quality of the basis material.

1) Results for GAMMEX and Head Phantom in Material Decomposition: Fig. 3 demonstrates visual comparisons of various methods applied to the GAMMEX Phantom data. The display ranges for PMMA and Al are [0,1500] mg/mL and [0,16] mg/mL, respectively. It is evident that the PMMA materials reconstructed using BT-SF and MM-BSN exhibit significant degradation, particularly at a tube current of 50 mA. In contrast, the proposed noise propagation model yields superior visual quality; key areas affected by noise are significantly restored, as illustrated in the zoomed region in Fig. 3(a5).

Quantitative comparisons are presented in Table I. The four regions of interest (ROIs) for assessing iodine concentration

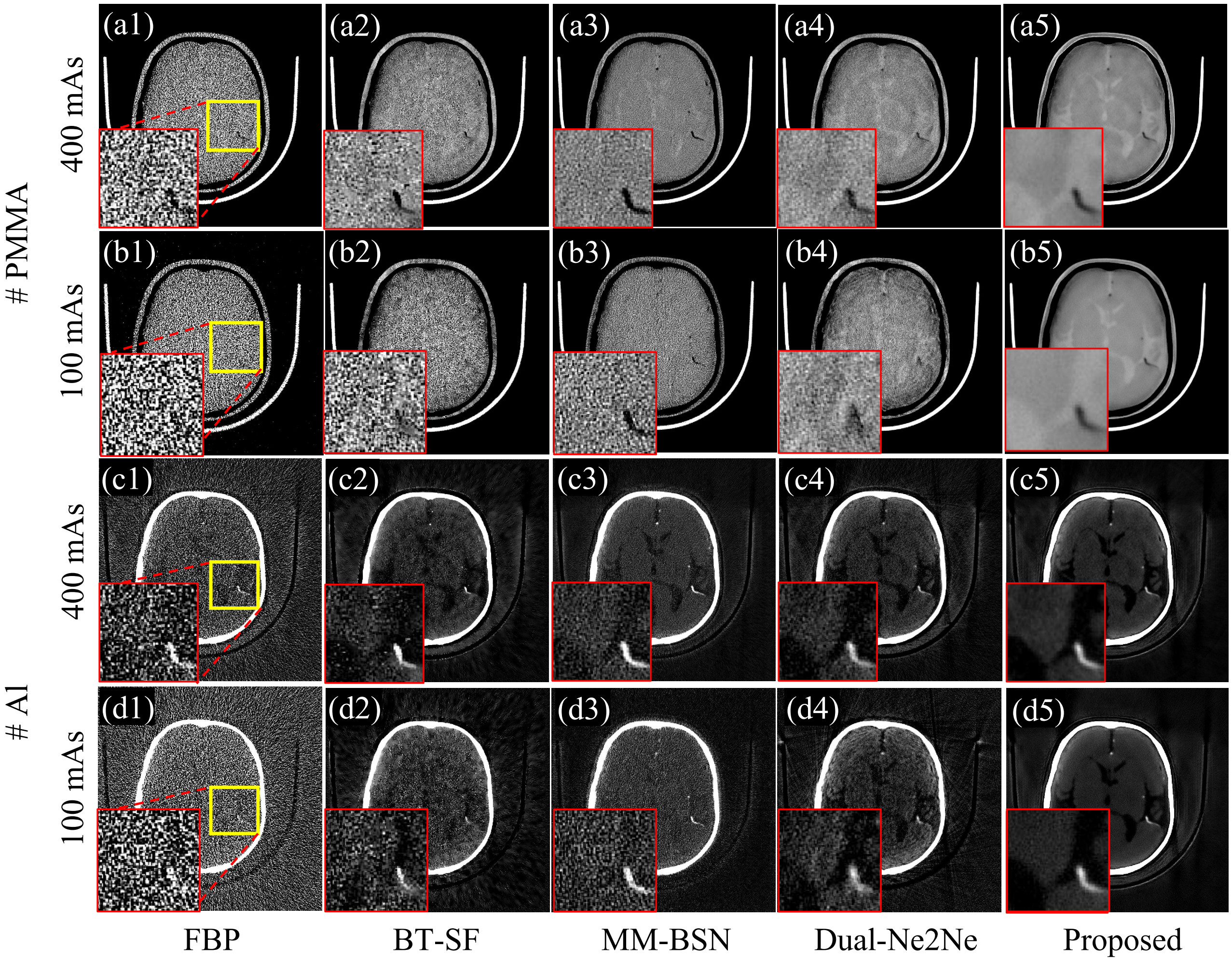


Fig. 4: Comparison of the reconstructed images for the PMMA and Al materials. Display range: [0,1500] mg/mL for PMMA basis images and [0,16] mg/mL for Al basis images.

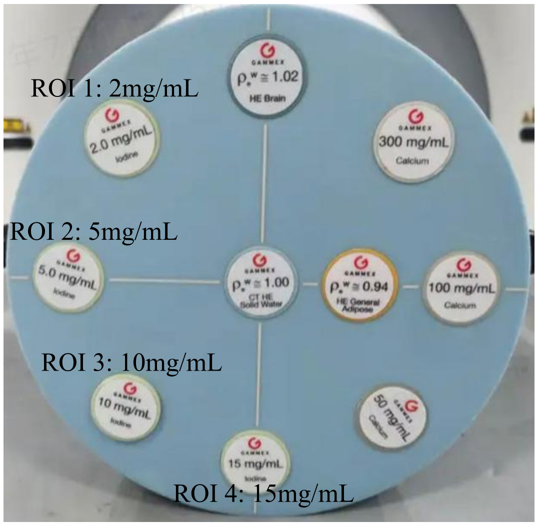


Fig. 5: Four regions of interest (ROIs) for the assessment of iodine concentration in the GAMMEX Phantom.

can be found in Fig. 5. The iodine concentrations obtained through all methods are very close to the ground truth; however, the standard deviations (STD) associated with the proposed method are markedly smaller than those observed with other techniques. For instance, at a tube current of 50 mA, the STD for images generated from our method is 3.955 lower than that of the images obtained by the FBP method in the four ROIs on average, indicating enhanced noise suppression by our proposed approach.

To further evaluate the effectiveness of our method, we also conducted experiments on head phantom data. As shown in Fig. 4(c5), the proposed method demonstrates superior visual performance compared to alternative approaches. Notably, edges within images processed through the learning-based noise propagation model exhibit greater clarity with minimal residual noise across identical display ranges.

2) Analysis of Virtual Monochromatic Imaging: Fig. 6 and Fig. 7 demonstrate the 60 keV VMI results obtained using different methods at 100 mAs and 400 mAs, respectively. The zoomed ROIs, indicated by yellow rectangles, are positioned below their corresponding images. It is evident that the proposed method exhibits commendable performance in both

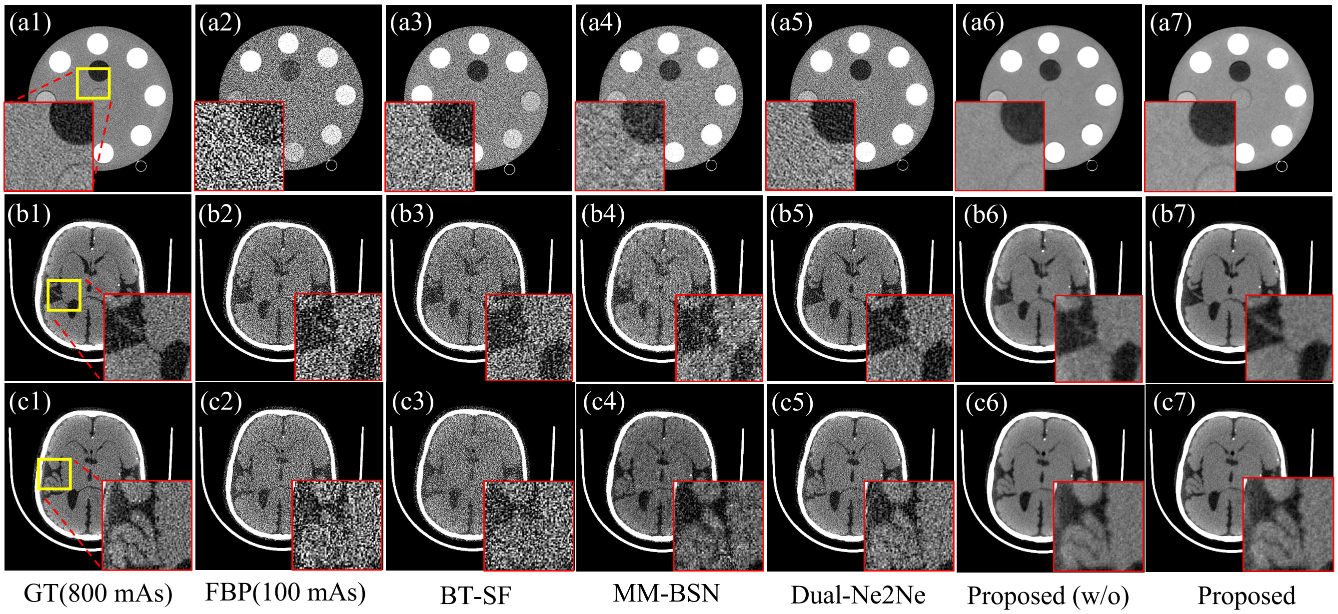


Fig. 6: VMI results of different methods at 60 keV (100 mAs). The notation "w/o" denotes the absence of the robust training strategy, i.e., without the all-in-one network.
Display range: [-100,100] HU for GAMMEX phantom and [-60,60] HU for head phantom.

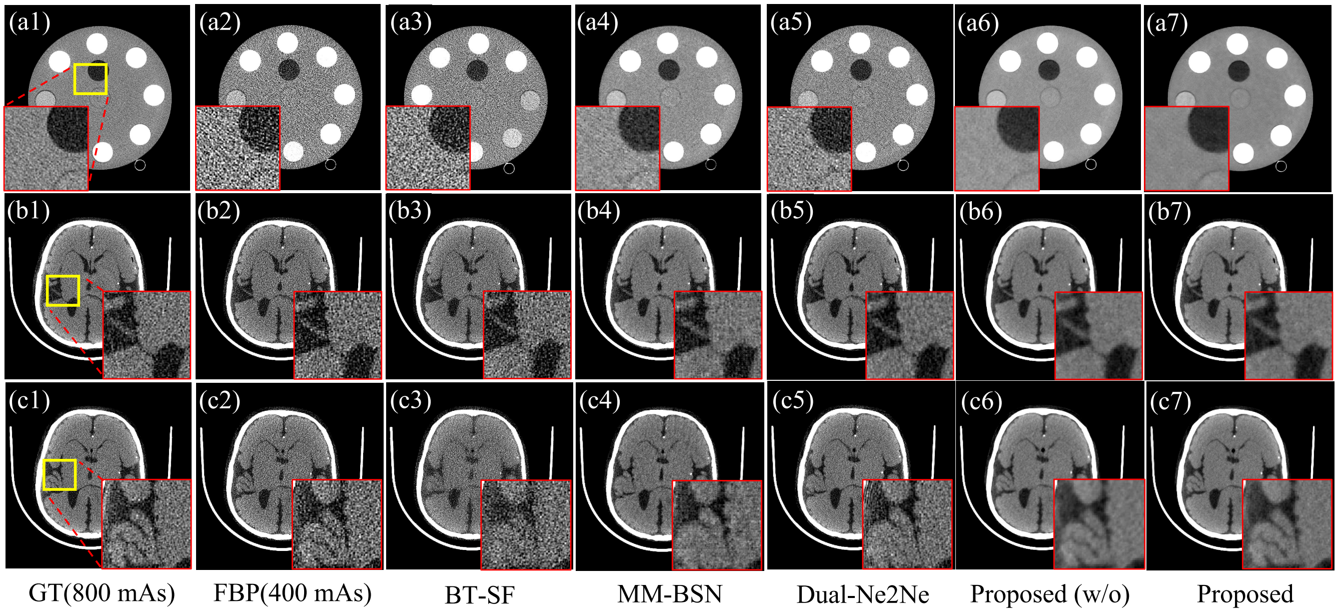


Fig. 7: VMI results of different methods at 60 keV (400 mAs). The notation "w/o" denotes the absence of the robust training strategy, i.e., without the all-in-one network.
Display range: [-100,100] HU for GAMMEX phantom and [-60,60] HU for head phantom.

TABLE I: Comparison of different methods for the GAMMEX iodine concentration accuracy. The four ROIs are shown in Fig. 5. The best results are highlighted in **bold**.

Exposure (mAs)	Method	Measured iodine concentration (mg/mL)			
	Ground truth	2	5	10	15
100	FBP	1.95 ± 4.37	5.08 ± 4.88	9.95 ± 5.50	14.99 ± 6.40
	BT-SF [42]	2.05 ± 1.05	5.04 ± 1.23	9.92 ± 1.68	15.04 ± 2.58
	MM-BSN [43]	2.07 ± 0.90	4.90 ± 1.24	9.99 ± 1.71	15.02 ± 2.41
	Dual-Ne2Ne [30]	1.95 ± 0.92	5.05 ± 1.30	9.98 ± 1.80	14.99 ± 2.71
	Proposed	1.94 ± **0.46**	5.05 ± **0.98**	10.05 ± **1.57**	14.95 ± **2.32**
400	FBP	1.95 ± 2.31	5.07 ± 2.73	9.96 ± 3.63	15.00 ± 4.70
	BT-SF [42]	2.06 ± 0.86	5.03 ± 1.11	9.92 ± 1.72	15.09 ± 2.62
	MM-BSN [43]	1.93 ± 0.61	5.04 ± 1.05	10.04 ± 1.61	14.96 ± 2.48
	Dual-Ne2Ne [30]	1.94 ± 0.65	5.07 ± 1.08	10.01 ± 1.63	14.97 ± 2.47
	Proposed	1.93 ± **0.42**	5.05 ± **0.96**	10.04 ± **1.57**	14.95 ± **2.35**

TABLE II: VMI results of different methods at 60 keV. The best results of the unsupervised methods are highlighted in **bold**.

Exposure (mAs)	Phantom	Metrics	Method				
			FBP	BT-SF [42]	MM-BSN [43]	Dual-Ne2Ne [30]	Proposed
100	GAMMEX	PSNR (dB)	15.89	21.80	21.24	23.09	**24.73**
		SSIM (%)	56.18	62.38	74.15	74.08	**80.05**
		MAE (HU)	75.81	49.95	37.29	30.95	**20.94**
	Brain	PSNR (dB)	23.73	24.00	25.74	28.36	**30.87**
		SSIM (%)	73.36	73.42	74.92	81.49	**87.61**
		MAE (HU)	32.83	31.91	27.96	19.94	**13.92**
400	GAMMEX	PSNR (dB)	21.11	21.74	27.30	24.98	**29.76**
		SSIM (%)	61.98	62.30	78.26	74.72	**80.95**
		MAE (HU)	52.35	38.16	29.90	25.79	**14.49**
	Brain	PSNR (dB)	28.11	28.23	30.09	30.50	**31.57**
		SSIM (%)	77.99	77.89	81.55	82.27	**89.98**
		MAE (HU)	19.60	19.43	16.28	**14.41**	17.38

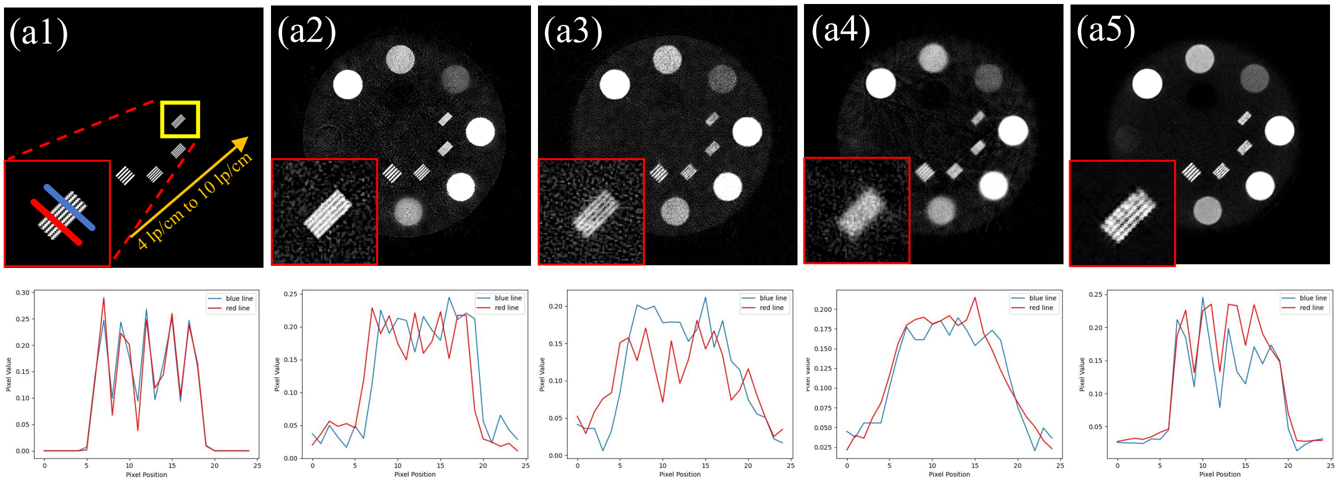
noise reduction and edge preservation.

The quantitative results presented in Table II further indicate that the proposed method significantly outperforms all other techniques based on PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity), and MAE (Mean Absolute Error) metrics. MAE quantifies the average absolute difference between the denoised images (predictions) and the ground-truth images. Mathematically, it is defined as MAE $= \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i|$, where y_i and $\hat{y}_i$ denote the pixel values of the ground-truth and denoised images, respectively, and N is the total number of pixels. From a visual assessment perspective, the effectiveness of the proposed method at 400 mAs appears superior to that at 100 mAs. For instance, Fig. 7 (b7) displays a lower noise level compared to Fig. 6 (b7); however, the metrics listed in Table II indicate a different conclusion. This inconsistency stems from our experiments using data at a tube current of 400 mA as ideal references, which also contain some noise. Consequently, when calculating metrics, images exhibiting better visual characteristics (such as low noise and high contrast) may yield lower numerical indicators.

3) Analysis of Spatial Resolution: To evaluate the impact of each method on the spatial resolution of the reconstructed image, we incorporated numerical iodine solutions of different sizes into the Gammex phantom (100 mAs). As illustrated in Fig. 8, the magnified region highlights one of the resolution patterns (10 lp/cm) within the phantom. It is evident that the proposed method effectively preserves spatial resolution while simultaneously reducing image noise.

4) Analysis of the Robust Training: Table III demonstrates the effectiveness of the robustness training strategy proposed in this study. As shown, after applying this strategy, most quantitative metrics have improved. More importantly, this strategy enables the model to be applicable to different phantoms and doses without requiring adjustments to hyper-parameter settings. Furthermore, we have compared the visual effects of VMI results obtained without the all-in-one network, as shown in Fig. 7 (b6) and Fig. 6 (b6). As can be seen, after removing the all-in-one network, the clarity of the texture has decreased.

5) Analysis of the weight β in Eq. 22: The specific value of β, which regulates the degree of denoising (with smaller values



GT (insertions) BT-SF MM-BSN Dual-Ne2Ne Proposed

Fig. 8: Comparison of the Al basis images from different methods for the digital quantification phantom (100 mAs). It is based on the GAMMEX phantom, and the inserts are iodine solutions of different sizes. The profile of **the blue and red lines in (a1)** are presented below each corresponding figure. Display range: [0 16] mg/mL.

TABLE III: Effectiveness of the robust training strategy (where Training stage 2 denotes the inclusion of the introduced all-in-one network).

Phantom/ Exposure	Metrics	Method	
		Training stage 1	Training stage 1+2
GAMMEX *100 mAs*	PSNR (dB)	23.99	24.73
	SSIM (%)	80.89	80.05
	MAE (HU)	21.74	20.94
Brain *100 mAs*	PSNR (dB)	28.99	30.87
	SSIM (%)	83.30	87.61
	MAE (HU)	17.45	13.92
GAMMEX *400 mAs*	PSNR (dB)	28.47	27.76
	SSIM (%)	82.36	80.95
	MAE (HU)	17.39	14.49
Brain *400 mAs*	PSNR (dB)	30.12	31.57
	SSIM (%)	88.28	89.98
	MAE (HU)	20.29	17.38

causing increased image blurring), was determined based on experimental results: Initially, we adopted the recommended β value of 0.1 from the original Ne2Ne denoising method. Subsequently, we conducted reconstruction experiments using β values within the range of 0.1 to 1, and selected the value that yielded the optimal performance. For instance, in the head phantom experiment, the best results for training stage 1 were achieved when $\beta = 0.3$, as illustrated in the Fig. 9.

IV. DISCUSSIONS

The proposed noise propagation model integrated with self-supervised learning demonstrates robust performance in PCCT MD and VMI. This section discusses the method's effectiveness, applications, and limitations addressing key considerations from experimental validation and clinical relevance.

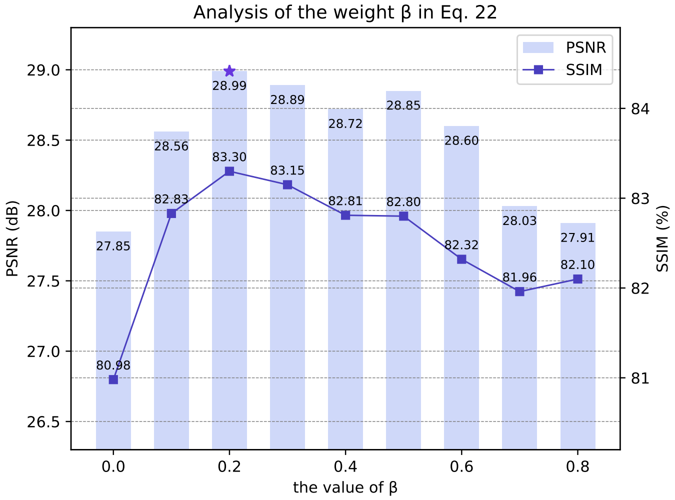


Fig. 9: Analysis of the weight β in Eq. 22

A. *Effectiveness of the Method*

The method achieves superior noise reduction while preserving material accuracy and structural details, as validated through quantitative metrics and visual assessments:

1) Noise-Resolution Trade-off: Compared to conventional and state-of-the-art self-supervised approaches, the proposed method balances noise suppression and edge preservation. Results confirm reduced noise without over-smoothing, with clearer vascular edges and anatomical details in visual comparisons.

2) Iodine Quantification Accuracy: In phantom experiments with varying iodine concentrations, the method exhibits minimal standard deviations, ensuring reliable quantitative measurements critical for clinical applications like contrast agent optimization.

3) Basis Material Generalizability: The noise propagation model adapts to different basis materials by mapping detector Poisson noise to material-specific covariance matrices via coefficient matrix M. While specific materials were used in experiments, the framework's reliance on physics-driven noise modeling and self-supervised training enables extension to other material pairs, with performance dependent on decomposition algorithm accuracy.

B. Model Assumptions and Adaptability

1) Ignorance of Non-Diagonal Covariance Terms: Under low-flux conditions, photon count fluctuations in energy bins follow Poisson statistics, weakening inter-channel correlations. This indicates that ignoring the non-diagonal covariance term can simplify the derivation and keep the error within an acceptable range. Moreover, the neural network compensates for residual correlations via multi-channel convolutional layers, learning spatial dependencies from noisy data.

2) Basis Material Conditioning Impact: The method's performance is robust to basis material selection [45] due to: (1) the noise propagation model's generality, which converts energy-bin noise to basis material noise via coefficient matrix regardless of material type; (2) self-supervised training with SigmaNet, which adapts to noise distributions from different basis materials without high-quality labels. Performance depends on decomposition algorithm accuracy for basis material attenuation coefficients.

C. Clinical Implications and Limitations

1) Clinical Relevance: By suppressing noise amplification during MD, the method enhances contrast between iodine-containing tissues (e.g., blood vessels, tumor lesions) and background, aiding detection of subtle lesions [46]. Low noise in iodine regions allows reduced contrast agent dosage [47] or radiation exposure [48] while maintaining image quality. Also, the improved separation of vascular structures from soft tissues enhances diagnostic utility for conditions like stroke.

2) Limitations: Experiments relied on standard phantoms; validation with patient data is pending due to ethical constraints of the prototype PCCT system (non-clinical device). At ultra-low doses, high-contrast resolution shows marginal degradation compared to high-dose ground truth, though edge preservation remains superior to conventional denoising methods.

V. CONCLUSION

This study presents a self-supervised denoising framework integrated with a physics-driven noise propagation model, effectively mitigating noise amplification in PCCT material decomposition while preserving material accuracy and structural details. By leveraging Poisson noise modeling and robust training, the method achieves superior performance in low-dose scenarios and iodine quantification, which could have a clinical impact in future spectral CT applications.

VI. ACKNOWLEDGMENT

The authors would like to thank UIH (United Imaging Healthcare, China) CT RPA department for providing the real data. This research work was supported by the Big Data Computing Center of Southeast University.

REFERENCES

[1] E. Roessl and R. Proksa, "K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors," *Physics in Medicine and Biology*, vol. 52, no. 15, pp. 4679–96, 2007.

[2] R. Gutjahr, A. Halaweish, Z. Yu, S. Leng, L. Yu, Z. Li, S. Jorgensen, E. Ritman, S. Kappler, and C. McCollough, "Human imaging with photon counting-based computed tomography at clinical dose levels: Contrast-to-noise ratio and cadaver studies," *Investigative Radiology*, vol. 51, pp. 421–429, 2016.

[3] S. Leng, M. Bruesewitz, S. Tao, K. Rajendran, A. Halaweish, N. Campeau, J. Fletcher, and C. McCollough, "Photon-counting detector ct: System design and clinical applications of an emerging technology," *Radiographics*, vol. 39, no. 3, pp. 729–743, May 2019.

[4] Y. Ren, H. Xie, W. Long, X. Yang, and X. Tang, "On the conditioning of spectral channelization (energy binning) and its impact on multi-material decomposition based spectral imaging in photon-counting ct," *IEEE Transactions on Biomedical Engineering*, vol. 68, no. 9, pp. 2678–2688, 2021.

[5] A. N. Primak, J. G. Fletcher, T. J. Vrtiska, O. P. Dzyubak, J. C. Lieske, M. E. Jackson, J. C. Williams, and C. H. McCollough, "Noninvasive differentiation of uric acid versus non-uric acid kidney stones using dual-energy ct," *Academic Radiology*, vol. 14, no. 12, pp. 1441–1447, 2007.

[6] H. Chandarana, A. J. Megibow, B. A. Cohen, R. Srinivasan, D. Kim, C. Leidecker, and M. Macari, "Iodine quantification with dual-energy ct: phantom study and preliminary experience with renal masses," *AJR. American journal of roentgenology*, vol. 196, no. 6, pp. 693–700, June 2011.

[7] Y. Xue, Y. Jiang, C. Yang, Q. Lyu, J. Wang, C. Luo, L. Zhang, C. Desrosiers, K. Feng, X. Sun, X. Hu, K. Sheng, and T. Niu, "Accurate multi-material decomposition in dual-energy ct: A phantom study," *IEEE Transactions on Computational Imaging*, vol. 5, no. 4, pp. 515–529, 2019.

[8] X. Wang, J. Xiang, A. Mao, J. Xie, P. Jin, M. Ding, Y. Yuan, Y. Lu, L. Yu, H. Cai, B. Lei, and T. Niu, "Clip-driven universal model for multi-material decomposition in dual-energy ct," *IEEE Transactions on Computational Imaging*, vol. 11, pp. 349–361, 2025.

[9] M. Feng, X. Ji, R. Zhang, K. Treb, A. M. Dingle, and K. Li, "An experimental method to correct low-frequency concentric artifacts in photon counting ct," *Physics in Medicine and Biology*, vol. 66, no. 17, p. 175011, 2021.

[10] E. Y. Sidky, Y. Zou, and X. Pan, "Impact of polychromatic x-ray sources on helical, cone-beam computed tomography and dual-energy methods," *Physics in Medicine and Biology*, vol. 49, no. 11, p. 2293, may 2004.

[11] C. Maass, M. Baer, and M. Kachelriess, "Image-based dual energy ct using optimized precorrection functions: A practical new approach of material decomposition in image domain." *Medical physics*, vol. 368, pp. 3818–3829, 2009.

[12] Y. Jiang, Y. Xue, Q. Lyu, L. Xu, C. Luo, P. Yang, C. Yang, J. Wang, X. Hu, X. Zhang, K. Sheng, and T. Niu, "Noise suppression in image-domain multi-material decomposition for dual-energy ct," *IEEE Transactions on Biomedical Engineering*, vol. 67, no. 2, pp. 523–535, 2020.

[13] F. Kelcz, P. Joseph, and S. Hilal, "Noise considerations in dual energy ct scanning," *Medical physics*, vol. 6, no. 5, pp. 418–425, 1979.

[14] D. A. Hinshaw and J. T. Dobbins, "Recent progress in noise reduction and scatter correction in dual-energy imaging," in *Medical Imaging*, vol. 2432, 1995, pp. 134–142.

[15] S. Leng, L. Yu, J. G. Fletcher, and C. H. McCollough, "Maximizing iodine contrast-to-noise ratios in abdominal ct imaging through use of energy domain noise reduction and virtual monoenergetic dual-energy ct," *Radiology*, vol. 276, no. 2, pp. 562–570, Aug. 2015.

[16] R. J. Warp, "Quantitative evaluation of noise reduction strategies in dual-energy imaging," *Medical Physics*, vol. 30, no. 2, p. 190, Feb. 2003.

[17] M. Petrongolo and L. Zhu, "Noise suppression for dual-energy ct through entropy minimization," *IEEE Transactions on Medical Imaging*, vol. 34, no. 11, pp. 2286–2297, 2015.

[18] T. Niu, X. Dong, M. J. Petrongolo, and L. Zhu, “Iterative image-domain decomposition for dual-energy ct.” *Medical physics*, vol. 41 4, p. 041901, 2014.

[19] X. Dong, T. Niu, and L. Zhu, “Combined iterative reconstruction and image-domain decomposition for dual energy ct using total-variation regularization.” *Medical physics*, vol. 41 5, p. 051909, 2014.

[20] Z. Su, M. Zhou, M. Li, Z. Zhang, D. Han, and G. Li, “Revisiting the application of twin connected parallel networks and regression loss functions in industrial defect detection,” *Adv. Eng. Informatics*, vol. 62, p. 102611, 2024.

[21] M. Zhou, Z. Su, M. Li, Y. Wang, and G. Li, “Csdd-net: A cross semi-supervised dual-feature distillation network for industrial defect detection,” *Knowl. Based Syst.*, vol. 306, p. 112751, 2024.

[22] D. Han, L. Xu, M. Zhou, J. Wan, M. Li, and G. Li, “Reconsidering learnable fine-grained text prompts for few-shot anomaly detection in visual-language models,” *Neural networks : the official journal of the International Neural Network Society*, vol. 182, p. 106906, 2024.

[23] Y. Xu, B. Yan, J. Zhang, J. Chen, L. Zeng, and L. Wang, “Image decomposition algorithm for dual-energy computed tomography via fully convolutional network,” *Computational and Mathematical Methods in Medicine*, vol. 2018, pp. 1–9, 2018.

[24] Z. Chen and L. Li, “Robust multimaterial decomposition of spectral CT using convolutional neural networks,” *Optical Engineering*, vol. 58, no. 1, p. 013104, 2019.

[25] M. Chen, Z. Chen, Y. Xi, X. Qiao, X. Chen, and Q. Huang, “Multimodal fusion network for detecting hyperplastic parathyroid glands in spect/ct images,” *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 3, pp. 1524–1534, 2023.

[26] H. Li, X. Yang, S. Yang, D. Wang, and G. Jeon, “Transformer with double enhancement for low-dose ct denoising,” *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 10, pp. 4660–4671, 2023.

[27] M. Meng, Y. Wang, M. Zhu, X. Tao, Z. Mao, J. Liao, Z. Bian, D. Zeng, and J. Ma, “Ddt-net: Dose-agnostic dual-task transfer network for simultaneous low-dose ct denoising and simulation,” *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 6, pp. 3613–3625, 2024.

[28] Y. Ma, J. Wang, C. Xu, Y. Huang, M. Chu, Z. Fan, Y. Xu, and D. Wu, “Cdaf-net: A contextual contrast detail attention feature fusion network for low-dose ct denoising,” *IEEE Journal of Biomedical and Health Informatics*, vol. 29, no. 3, pp. 2048–2060, 2025.

[29] D. Hu, W. Wu, M. Xu, Y. Zhang, J. Liu, R. Ge, Y. Chen, L. Luo, and G. Coatrieux, “Sister: Spectral-image similarity-based tensor with enhanced-sparsity reconstruction for sparse-view multi-energy ct,” *IEEE Transactions on Computational Imaging*, vol. 6, pp. 477–490, 2020.

[30] C. Niu, M. Li, X. Guo, and G. Wang, “Self-supervised dual-domain network for low-dose ct denoising,” *SPIE Proceedings*, p. 15, 10 2022.

[31] Y. Wu, S. Qiu, M. L. Groot, and Z. Zhang, “Self-supervised image denoising of third harmonic generation microscopic images of human glioma tissue by transformer-based blind spot (tbs) network,” *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 8, pp. 4688–4700, 2024.

[32] Z. Wang, Z. Cao, J. Xie, H. Li, and Z. He, “Towards applicable unsupervised signal denoising via subsequence splitting and blind spot network,” *IEEE Transactions on Signal Processing*, vol. 72, pp. 4967–4982, 2024.

[33] R. E. Alvarez and A. Macovski, “Energy-selective reconstructions in x-ray computerised tomography,” *Physics in Medicine and Biology*, vol. 21, no. 5, p. 733, sep 1976. [Online]. Available: https://dx.doi.org/10.1088/0031-9155/21/5/002

[34] E. Roessl, A. Ziegler, and P. R, “On the influence of noise correlations in measurement data on basis image noise in dual-energylike x-ray imaging,” *Medical physics*, vol. 34, no. 3, pp. 959–966, 2007.

[35] X. Tang, Y. Ren, and H. Xie, “Noise correlation and its impact on the performance of multi-material decomposition-based spectral imaging in photon-counting ct,” *Medical physics*, vol. 24, no. 1, p. e13830, 2023.

[36] T. Huang, S. Li, X. Jia, H. Lu, and J. Liu, “Neighbor2neighbor: Self-supervised denoising from single noisy images,” in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 14 776–14 785.

[37] S. Laine, T. Karras, J. Lehtinen, and T. Aila, “High-quality self-supervised deep image denoising,” in *Neural Information Processing Systems*, 2019, pp. 6968–6978.

[38] M. Yao, R. Xu, Y. Guan, J. Huang, and Z. Xiong, “Neural degradation representation learning for all-in-one image restoration,” *IEEE Transactions on Image Processing*, vol. 33, pp. 5408–5423, 2024.

[39] D. Cheng, Y. Ji, D. Gong, Y. Li, N. Wang, J. Han, and D. Zhang, “Continual all-in-one adverse weather removal with knowledge replay on a unified network structure,” *IEEE Transactions on Multimedia*, vol. 26, pp. 8184–8196, 2024.

[40] H. Gao, J. Yang, Y. Zhang, N. Wang, J. Yang, and D. Dang, “Prompt-based ingredient-oriented all-in-one image restoration,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 10, pp. 9458–9471, 2024.

[41] Z. Feng, W. Wang, W. Li, G. Li, M. Li, and M. Zhou, “Mfur-net: Multimodal feature fusion and unimodal feature refinement for rgb-d salient object detection,” *Knowl. Based Syst.*, vol. 299, p. 112022, 2024.

[42] X. Ji, X. Zhuo, Y. Lu, W. Mao, S. Zhu, G. Quan, Y. Xi, T. Lyu, and Y. Chen, “Image domain multi-material decomposition noise suppression through basis transformation and selective filtering,” *IEEE Journal of Biomedical and Health Informatics*, vol. 28, no. 5, pp. 2891–2903, 2024.

[43] D. Zhang, F. Zhou, Y. Jiang, and Z. Fu, “Mm-bsn: Self-supervised image denoising for real-world with multi-mask based on blind-spot network,” *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 4189–4198, 2023.

[44] T. Fan, Z. Zhang, M. Yang, G. Quan, and X. Lai, “Image-domain material decomposition and beam hardening correction for photon-counting ct,” in *2023 IEEE Nuclear Science Symposium, Medical Imaging Conference and International Symposium on Room-Temperature Semiconductor Detectors (NSS MIC RTSD)*, 2023, pp. 1–1.

[45] X. Tang and Y. Ren, “On the conditioning of basis materials and its impact on multimaterial decomposition-based spectral imaging in photon-counting ct,” *Medical physics*, vol. 3, no. 48, pp. 1100–1116, 2021.

[46] Q. Wang, G. Shi, X. Qi, X. Fan, and L. Wang, “Quantitative analysis of the dual-energy ct virtual spectral curve for focal liver lesions characterization,” *European Journal of Radiology*, vol. 83, no. 10, pp. 1759–1764, 2014.

[47] W. Cai, J.-G. Lee, D. Zhang, S. H. Kim, M. Zalis, and H. Yoshida, “Electronic cleansing in fecal-tagging dual-energy ct colonography based on material decomposition and virtual colon tagging,” *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 2, pp. 754–765, 2015.

[48] X. Jiang, G. J. Gang, and J. W. Stayman, “Multi-material decomposition using spectral diffusion posterior sampling,” *IEEE Transactions on Biomedical Engineering*, pp. 1–15, 2025.