






Article

Enabling Low-Dose In Vivo Benchtop X-ray Fluorescence Computed Tomography through Deep-Learning-Based Denoising

Naghmeh Mahmoodian ^{*}, Mohammad Rezapourian , Asim Abdulsamad Inamdar, Kunal Kumar ,
Melanie Facht  and Christoph Hoeschen 

Chair of Medical Systems Technology, Institute for Medical Technology, Faculty of Electrical Engineering and Information Technology, Otto von Guericke University, 39106 Magdeburg, Germany;

mohammad.rezapourianghahfarokhi@ovgu.de (M.R.); asim.inamdar@st.ovgu.de (A.A.I.);

kunal.kumar@ovgu.de (K.K.); melanie.facht@ovgu.de (M.F.); christoph.hoeschen@ovgu.de (C.H.)

* Correspondence: naghmeh.mahmoodian@ovgu.de

Abstract: X-ray Fluorescence Computed Tomography (XFCT) is an emerging non-invasive imaging technique providing high-resolution molecular-level data. However, increased sensitivity with current benchtop X-ray sources comes at the cost of high radiation exposure. Artificial Intelligence (AI), particularly deep learning (DL), has revolutionized medical imaging by delivering high-quality images in the presence of noise. In XFCT, traditional methods rely on complex algorithms for background noise reduction, but AI holds promise in addressing high-dose concerns. We present an optimized Swin-Conv-UNet (SCUNet) model for background noise reduction in X-ray fluorescence (XRF) images at low tracer concentrations. Our method's effectiveness is evaluated against higher-dose images, while various denoising techniques exist for X-ray and computed tomography (CT) techniques, only a few address XFCT. The DL model is trained and assessed using augmented data, focusing on background noise reduction. Image quality is measured using peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), comparing outcomes with 100% X-ray-dose images. Results demonstrate that the proposed algorithm yields high-quality images from low-dose inputs, with maximum PSNR of 39.05 and SSIM of 0.86. The model outperforms block-matching and 3D filtering (BM3D), block-matching and 4D filtering (BM4D), non-local means (NLM), denoising convolutional neural network (DnCNN), and SCUNet in both visual inspection and quantitative analysis, particularly in high-noise scenarios. This indicates the potential of AI, specifically the SCUNet model, in significantly improving XFCT imaging by mitigating the trade-off between sensitivity and radiation exposure.

Keywords: deep learning (DL); artificial intelligence (AI); X-ray fluorescence (XRF); XFCT; nanoparticles; cancer



Citation: Mahmoodian, N.; Rezapourian, M.; Inamdar, A.A.; Kumar, K.; Facht, M.; Hoeschen, C. Enabling Low-Dose In Vivo Benchtop X-ray Fluorescence Computed Tomography through Deep-Learning-Based Denoising. *J. Imaging* **2024**, *10*, 127. <https://doi.org/10.3390/jimaging10060127>

Academic Editor: Silvia Cipiccia

Received: 8 March 2024

Revised: 10 May 2024

Accepted: 18 May 2024

Published: 22 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent times, X-ray fluorescence (XRF) imaging has garnered significant interest as a promising tool for in vivo preclinical studies due to its ability to quantify and determine the biodistribution of labeled nanoparticle-based contrast agents at high resolution [1–4]. For this purpose, various imaging approaches have been proposed with different scanning methodologies and X-ray sources for the excitation of characteristic X-ray photons from specific metallic nanoparticles [1,5]. Some studies have measured the distribution of nanoparticles in mice using monochromatic X-rays from synchrotron sources [6–9]. Meanwhile, other studies have applied conventional X-ray tubes with polychromatic X-rays for specific biomedical applications at the laboratory scale [3,4,10–18]. To name a few, Cong et al. [16] effectively reconstructed gold nanoparticles by employing a fan-beam X-ray source and parallel single-hole collimation. Deng et al. [12] employed a conventional

X-ray tube to observe the distribution of gadolinium nanoparticles within mouse kidneys. Specifically, X-ray fluorescence computed tomography (XFCT), which leverages the principles of XRF imaging within a computed tomography (CT) framework, has been studied in several previous works [10,11,15,17,18].

In both approaches, the primary source of background noise is the appearance of Compton-scattered photons in the signal region, causing a low signal-to-noise ratio and reduced detection sensitivity. Contributions from Rayleigh scattering are, typically, minimal due to predominant scattering in forward directions, away from the detectors. Hence, enhancing image quality involves the essential task of minimizing Compton background noise to solve the mentioned problem. Approaches available for extracting XRF signals from Compton-scattered photons, including minimization of the latter, involve, e.g., special background-reduction schemes through spatial filtering algorithms and subtraction techniques, among others [3,9,15,17,18]. Specifically, subtraction techniques require consistency in the positioning and posture of the measured object, necessitating two scans (with and without contrast agents, such as nanoparticles) [13,14,19]. This may increase radiation dose, potentially leading to detrimental effects, such as radiation-induced cancer and metabolic abnormalities, among others [19]. Meanwhile, other methods can be cumbersome, may require highly specialized operational knowledge, or require significant user intervention. This motivates a shift from traditional noise suppression techniques to new AI-based approaches, like deep learning denoising, offering promising advantages and moving towards automation.

Recently, deep learning denoising methods have been widely used in the field of biomedical imaging and image processing [20–23]. The convolutional neural network (CNN) architecture can extract features from low to high levels, resulting in better identification and understanding of patterns [24]. The methods show a high performance of image-to-image translation issues [25,26]. For example, achieving a high-quality (high-dose) image from a low-dose image by removing image noise [25]. In contrast to the assumption made in non-blind image denoising, which presupposes knowledge of the image noise type and level, blind denoising addresses scenarios in which either the noise level, the specific noise type, or both are unknown. Zhang et al. [26] show that a deep model is capable of managing Gaussian denoising across different noise levels. Meanwhile, Chen et al. [27] suggest employing generative adversarial networks (GAN) for noise modeling from clean images. The generated noise is then used to create paired training data for subsequent training purposes. Guo et al. [28] introduce a convolutional blind denoising network, which incorporates a subnetwork for estimating noise. They further suggest training the model using both a practical noise model and pairs of real-world noisy–clean images. Krull et al. [29] introduce a method using variational inference for blind image denoising. This method unifies both noise estimation and image denoising within a Bayesian framework. Sun and Tappen [30] introduce a non-local deep learning approach that combines the benefits of block matching and 3D filtering (BM3D) and non-local means (NLM). Furthermore, the conventional BM3D technique is expanded into a four-dimensional space (BM4D), leading to enhanced preservation of image edge and texture intricacies Zhang et al. [31], Xu et al. [32]. Lefkimmiatis [33] creates an unrolled network capable of executing non-local processing, leading to enhanced denoising outcomes. Liang et al. [34] show improvement in peak signal-to-noise ratio (PSNR) performance achieved by a specific method, i.e., image restoration using the Swin transformer (SwinIR), compared to denoising convolutional neural network (DnCNN) on a benchmark dataset. Liu et al. [35] employ a transformer-based method, the shifted window transformer (Swin), as the primary building block. They demonstrate that the Swin model exhibits improved performance when handling images with repetitive structures, verifying the effectiveness of the transformer in enabling non-local modeling capabilities [35]. A further study proposed a new blind denoising network architecture named Swin-Conv-UNet (SCUNet), applied it to a real image dataset, and designed it for improved practical usage and enhanced local and non-local modeling abilities [26].

In this study, our objective is to employ a blind hybrid deep learning model, SCUNet, and enhance its capability for denoising XFCT images. The model is a result of combining the U-net [36] and Swin [26] models, utilizing both local and non-local modeling. These two networks have individually shown promising denoising performance. The XRF images of a Gadolinium (Gd) contrast agent were generated from high-dose and low-dose measurements to estimate the effect of the radiation dose on background noise by applying the blind hybrid denoising method. The model is created to combine two specific abilities, i.e., the local modeling ability of a residual convolutional layer and the non-local modeling ability of a Swin transformer block. This combined block is suggested for use as a fundamental component within the image-to-image translation UNet architecture, potentially enhancing its performance or capabilities. However, this architecture alone does not suffice for attaining optimal results. Consequently, we incorporated a compound loss function to enhance the network's performance, specifically addressing the intricacies associated with denoising medical images.

2. Materials and Methods

2.1. Experimental Setup

The data used in this study consists of XRF images acquired in a previous study using small-animal-sized phantoms consisting of contrast-agent-filled target tubes [10]. The following briefly describes the experimental setup and image acquisition procedure for completeness.

Gadolinium (Gd) at varying concentrations, derived from Gadoteric acid ($C_{16}H_{25}GdN_4O_8$), was utilized as the contrast agent. The small-animal-sized surrogate phantoms (mimicking a size-scale on the order of, e.g., mice) consisted of water-filled hexagonal borosilicate glass containers with a maximum diameter of 50 mm and a height of 60 mm. To accommodate 3D breast cancer models immersed in the cell-culture growth medium, the phantoms were chosen with a diameter exceeding that of a typical mouse's torso (i.e., about 20–25 mm). The Gd-filled polypropylene microcentrifuge tubes, 8 mm in diameter, were embedded inside these water-filled phantoms. Gadolinium (Gd) was used as a contrast agent for various reasons. High atomic number (high-Z) elements, such as Gd (K-alpha emission around 43 keV), exhibit a relatively weaker energy-dependent photon attenuation compared to lower atomic number (low-Z) elements, e.g., iodine (K-alpha emission around 28.4 keV) at this size-scale. Additionally, cadmium telluride (CdTe) pixel detectors were used for the XFCT imaging that exhibits high sensor-intrinsic X-ray fluorescence (i.e., K-shell emission from Cd and Te), making the low energy range (roughly, 19–32 keV) unusable. Consequently, high-Z contrast agents become more favorable at this scale, making Gd a better choice over low-Z elements. These phantoms were used in the previous study to investigate the suitability and feasibility of cone-beam XFCT for both preclinical in vivo imaging of small animals and in vitro surrogate investigations with non-destructive analysis of biological samples [10]. The Gd tubes had concentrations ranging from 0 to 3.1 mg/mL (0, 0.031%, 0.1%, 0.2%, and 0.31% Gd by weight (wt. %), see Figure 1). The first row of Figure 1 shows a representative illustration of the varying Gd concentration, and the second row shows an example of the XFCT scan in the presence of different Gd concentrations.

Attenuation images, acquired using cone-beam computed tomography (CBCT), were used to correct for attenuation in the XRF images. The images were acquired using cone-beam polychromatic X-rays generated from a tungsten-target microfocus X-ray tube (Oxford Nova 600, Oxford Instruments X-ray Technology, Scotts Valley, CA, USA). The incident polychromatic X-rays with 90 kVp, a maximum beam current of 0.9 mA, 0.3 mm copper (Cu) filtration, and X-ray focal spot size of 14–20 μ m were used as an excitation source for both CBCT and XFCT imaging. The source-to-isocenter distance was 40.5 cm and the distance from source to transmission detector was 55.7 cm. Figure 2 shows a schematic illustration of the experimental benchtop imaging system, consisting of both CBCT and XFCT imaging configurations. For attenuation images, a total of 30 projections were captured with an

angular interval of 12° between each projection, and approximately 6 s exposure time per projection.

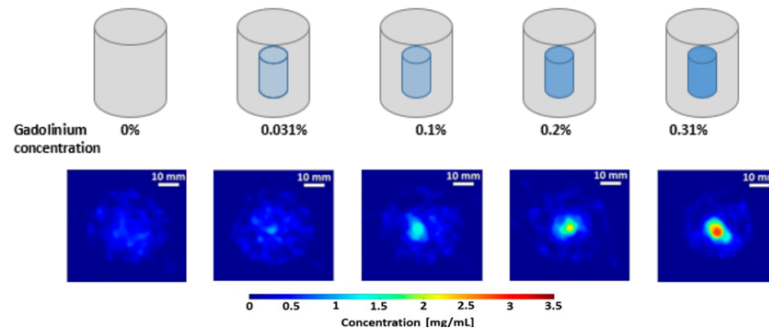


Figure 1. Gadolinium-contrast-filled small-animal phantoms with Gd concentrations ranging from 0 to 3.1 mg/mL.

The XFCT imaging system consisted of a Timepix3 HPCD (Minipix TPX3, Advacam, Prague, Czech Republic), with a pixel size of 55 μm, an active detection area of approximately 14 mm × 14 mm, 1 mm thick CdTe sensor, and an energy resolution of around 4–6 keV FWHM at Gd XRF energies. The detector was exposed to radiation using a circular-aperture single-pinhole arrangement. The pinhole collimator was constructed from lead, with an aperture diameter of 0.4 mm and a thickness of 1.5 mm. To conduct full-field scanning within a geometrically/mechanically constrained large CT detector arrangement, the distance between the X-ray source and the rotation isocenter was adjusted to 14 cm. Distances of 5.5 cm from the X-ray source to the isocenter were set, with a pinhole-to-detector distance of 1.5 cm. For XFCT imaging, a sparse-view image acquisition approach was employed to reduce radiation exposure. This strategy involved taking 10 angular projections within a 360° scan, each with an exposure time of 150 s per projection.

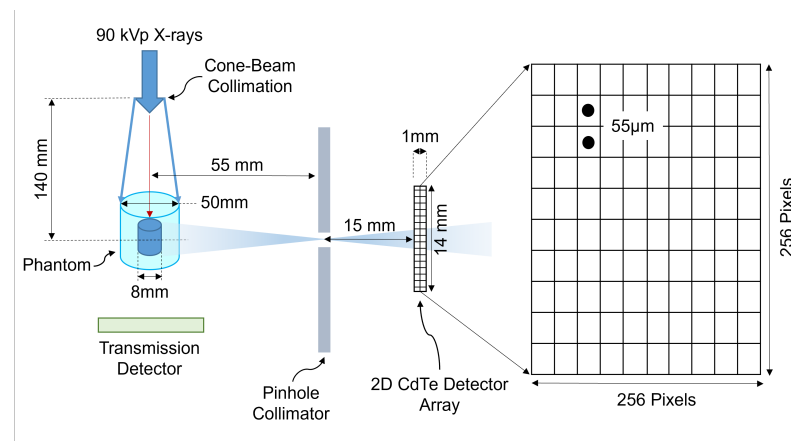


Figure 2. Schematic representation of the experimental benchtop XFCT system.

2.2. Deep Blind Image Denoising Model

2.2.1. Dataset

To estimate the effect of radiation dose on background noise, images with various noise levels are generated (i.e., corresponding to radiation doses from high to low). The noisy datasets are generated by reducing the photon counts in raw data (i.e., where 100% correspond to high-dose images) by a factor of 25%, 50%, and 75% (corresponding to the low-dose images). Additionally, for each image, three bin widths corresponding to 0.05 keV, 0.1 keV, and 0.5 keV are used. Herein, smaller bin widths correspond to a finer sampling grid, however, with a reduced number of entries per XRF signal bin. Furthermore, to address the limitation posed by the small dataset, we employed a data augmentation technique known as rotation. Data augmentation involves creating additional training samples by applying various transformations to the existing data. While augmentation

techniques like rotation, flipping, or scaling increase the dataset size, the generated images primarily remain highly correlated with the original images. They have not been seen to introduce truly new anatomical variations or disease presentations. DL models generally perform better when they are trained on a diverse set of data that reflects the real-world variations that would be encountered during deployment. However, geometric augmentations may not fully provide this diversity. Rotating medical images can disrupt the correct anatomical orientation of organs and structures. This can make interpretation difficult or misleading for clinicians. Additionally, if the image labels are not adjusted after rotation, the model may learn incorrect associations between features and labels, leading to errors. Furthermore, certain imaging modalities might be more sensitive to rotations than others due to their specific acquisition and representation of data. However, it must be noted that these aspects cannot be fully investigated currently as the XFCT technique is only at a nascent stage. More data would be required to fully understand the impact of these augmentations.

In our case, rotation augmentation was chosen to enhance the diversity of the dataset. Rotation augmentation entails rotating the original images by certain angles to generate new perspectives and variations. We specifically chose rotation angles of 90° , 180° , and 270° for several reasons. Firstly, these angles represent orthogonal transformations, allowing the model to learn diverse features and patterns that may be present in different orientations. Secondly, these angles align with common geometric transformations, capturing potential variations that could be encountered in real-world scenarios. The SCUNet training pipeline involves utilizing denoised images as targets and noisy images as input for the model. Through this approach, the sample size was effectively increased by approximately tenfold. Within the training pipeline, a hold-out strategy was implemented, allocating 20% of the samples for testing in each category and utilizing the remaining 80% for training purposes. The DL model is trained using augmented data and evaluated for background noise reduction via the proposed denoising approach.

2.2.2. Proposed Model

One of the key challenges currently in X-ray fluorescence imaging, particularly when using benchtop/clinical-grade X-ray sources, is the high radiation exposure required to reach high sensitivity. This currently hinders the broader adoption of this technique for in vivo preclinical studies and its translation into future clinical applications. In this study, a blind denoising method was employed to remove noise present in the low-dose images. In blind image denoising, the process of obtaining an estimated clean image involves solving a specific Maximum A Posteriori (MAP) problem using an optimization algorithm. This allows us to estimate the original (clean) image X from noisy observations, a critical step in enhancing XFCT image quality, by employing the following optimization technique:

$$X = \underset{x}{\operatorname{argmin}} D(x, y) + \lambda P(x), \quad (1)$$

where $D(x, y)$ represents the data fidelity component, $P(x)$ represents the prior term, and λ is the trade-off parameter [26]. At this point, it becomes evident that the crux of addressing blind denoising is twofold: modeling the degradation process of a noisy image and designing the image prior to a clean image. Conceptually, by regarding the deep model as a condensed unrolled inference of Equation (1), the overarching objective of deep blind denoising typically involves tackling the following bi-level optimization problem [37,38]. The effectiveness of the deep blind learning model is largely dependent on its network architecture and the quality of the training data. The presence of noisy images in the training dataset significantly influences the model's understanding of the degradation process. Improvements in the network architecture and the inclusion of clean images within the training dataset play a crucial role in shaping this understanding. Improving the quality of clean data is achievable, but additional investigation is necessary to enhance and develop network architecture.

The network of the SCUNet model is shown in Figure 3. SCUNet (Spatial and Channel-wise Attention U-Net) specifically incorporates attention mechanisms to identify and preserve relevant image features while suppressing noise effectively. This is highly beneficial for the complex data generated by XFCT. The U-Net architecture, integrated within SCUNet, is particularly adept at medical image segmentation and denoising tasks. Its ability to analyze information at multiple scales ensures the preservation of both fine details and broader image context. By integrating the power of deep learning with the specific advantages of SCUNet, our study aims to provide a crucial solution for reducing radiation exposure in XFCT. This has the potential to enhance XFCT imaging, moving towards complete solutions that could enable its broader adoption into research and future medical applications. The core concept behind SCUNet involves merging the complementary network architecture designs from dilated-residual U-Net (DRUNet) and SwinIR. The central technical contribution of our study lies in the development of the optimized SCUNet model. This hybrid architecture uniquely combines the strengths of Convolutional Neural Networks (CNNs), and transformer-based models. Specifically, SCUNet incorporates novel swin-conv (SC) blocks into a UNet backbone [36]. The novel SC block is the heart of SCUNet. It integrates a Swin transformer for non-local modeling with a residual convolutional block for local modeling. This combination allows the model to effectively capture complex image features and noise patterns for superior denoising. The SC blocks are integrated into a multiscale UNet architecture. This enables the model to process information at varying resolutions, enhancing its ability to preserve fine details while suppressing noise across different scales. The hybrid design of SCUNet suggests that it may be more adaptable to the diverse noise characteristics present in XFCT data compared to models relying solely on convolutional or transformer-based architectures. The use of 1×1 convolutions within the SC block facilitates seamless information exchange between the transformer and convolutional components, likely improving denoising performance. In accordance with DRUNet [39], SCUNet's UNet backbone consists of four scales, each featuring a residual connection between 2×2 strided convolution (SConv)-based downscaling and 2×2 transposed convolution (TConv)-based upscaling. The channel counts in each layer vary from 64 in the first scale to 512 in the fourth scale. A key distinction between SCUNet and DRUNet lies in the adoption of four SC blocks, as opposed to four residual convolution blocks, in each scale of the downscaling and upscaling processes. The images of the XRF photons including Compton scattered photons were inputted into the network with a size of 256×256 pixels, and the output images were all the same size. The transformer layer SwinIR model consists of shallow feature extraction, deep feature extraction, and high-quality image reconstruction. We treat each patch as a token and 2×2 strided convolution with stride 2. The UNet backbone of our model has four scales, each of which has a residual connection between 2×2 SConv-based downscaling and TConv-based upscaling. The number of channels in each layer, from the first scale to the fourth scale, is 64, 128, 256, and 512, respectively.

The second box in Figure 3 illustrates an SC block that combines a Swin transformer (SwinT) block [35] with a residual convolutional (RConv) block [40,41] through two 1×1 convolutions, split and concatenation operations, and a residual connection. Specifically, for an input feature tensor, it undergoes an initial 1×1 convolution. Following this, the tensor is evenly divided into two feature map groups, namely X_1 and X_2 . This entire process can be expressed as follows:

$$X_1, X_2 = \text{Split}(\text{Conv}_{1 \times 1}(X)) \quad (2)$$

Subsequently, X_1 and X_2 are individually inputted into a SwinT block and an RConv block, leading to:

$$Y_1, Y_2 = \text{SwinT}(X_1), \text{RConv}(X_2) \quad (3)$$

Finally, $Y1$ and $Y2$ are concatenated to form the input of a 1×1 convolution, which establishes a residual connection with the initial input tensor, X . Consequently, the ultimate output of the SC block is expressed as:

$$Z = \text{Conv}_{1 \times 1}(\text{Concat}(Y1, Y2)) + X \tag{4}$$

Our proposed SCUNet exhibits several advantages attributed to its innovative module designs. Firstly, the SC block integrates the local modeling capability of the RConv block with the non-local modeling ability of the SwinT block. Secondly, the local and non-local modeling capabilities of SCUNet are further enhanced through the incorporation of a multiscale UNet. Thirdly, the 1×1 convolution plays a crucial role in effectively and efficiently facilitating information fusion between the SwinT block and the RConv block. Fourthly, the split and concatenation operations serve as a form of group convolution with two groups, contributing to the reduction in computational complexity and the number of parameters. The parameters are optimized by minimizing the $L1$ loss with the Adam optimizer [42]. The learning rate starts from 1×10^{-4} and decays by a factor of 0.9 for each iteration for 40 epochs. The batch size is set to 2. We first train the model with a 25% noise level and then fine-tune the model for other noise levels. All experiments are implemented using PyTorch 2.0.1. It takes about 20 h to train a denoising model on an NVIDIA GTX 1080.

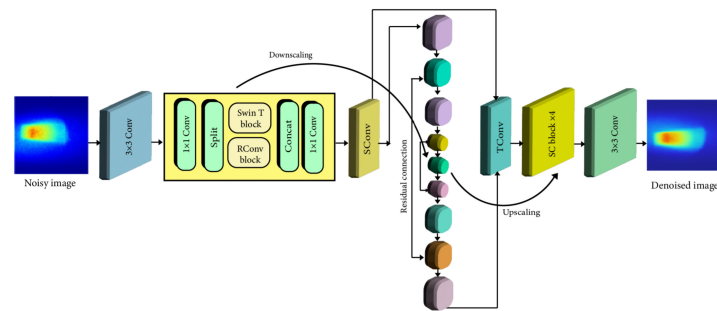


Figure 3. Architecture of the proposed denoising network method.

It is noteworthy that SCUNet operates as a hybrid CNN–Transformer network, integrating features from both architectures. Similar approaches exist in the literature, where researchers have explored the combination of CNNs and Transformers for effective network architecture design. It is significant to highlight the distinctions between our proposed SCUNet and two recent works, namely Uformer [43] and Swin-UNet [44].

Firstly, the motivation behind each approach differs significantly. SCUNet is inspired by the observation that state-of-the-art denoising methods, DRUNet [39] and SwinIR [34], leverage distinct network architecture designs. As a result, SCUNet seeks to integrate the complementary features of DRUNet and SwinIR. Conversely, Uformer and Swin-UNet aim to merge transformer variants with UNet, serving a different motivation.

Secondly, the primary building blocks employed in each model are distinct. SCUNet incorporates a novel swin-conv block, which integrates the local modeling capability of the residual convolutional layer [40] with the non-local modeling ability of the Swin transformer block [35] through 1×1 convolution, split, and concatenation operations. In contrast, Uformer adopts a novel transformer block by combining depth-wise convolution layers [45], while Swin-UNet utilizes the Swin transformer block as its primary building block.

In this research two evaluation metrics, i.e., peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), are utilized as they offer advantages and better performance over traditional metrics like Mean Squared Error (MSE) in perceptual quality [46]. Specifically, SSIM, for its feature and structural measures, has been used in several application areas such as denoising, pattern recognition, image restoration, image compression, and more [46]. Both PSNR and SSIM, have shown to perform well in predicting and reflecting the visual quality of images [46].

3. Results

As outlined in Section 2.2, enhancing the performance of the deep blind denoising model relies on the optimization of both the network architecture and the training data. In this section, we evaluate the performance of our model qualitatively and quantitatively and compare it with other deep and non-deep learning methods.

To achieve optimal denoising results, we initiated our model training by utilizing the official pre-trained weights of SCUNet. Subsequently, we fine-tuned the model on our specific dataset, enhancing its adaptability to our unique denoising requirements. During this training process, a compound loss function was employed to synergistically improve the overall performance of the model. Figure 4 shows examples of images in our dataset with varying noise levels and bin widths. In the first row, we present our original data with 100% of the initial number of photons. Subsequently, we reduce the number of photons by a factor of 25% in each following row, simultaneously increasing the sampling rate to 0.05, 0.1, and 0.5. In this figure, the highest quality image is the one with a noise level of 0% and maximum bin width (BW) of 0.5 (see first row and third column), and the lowest quality image is the one with a 75% noise level and bin width of 0.005. An enlarged image is shown in this figure for clarity, having an image size of 256×256 pixels.

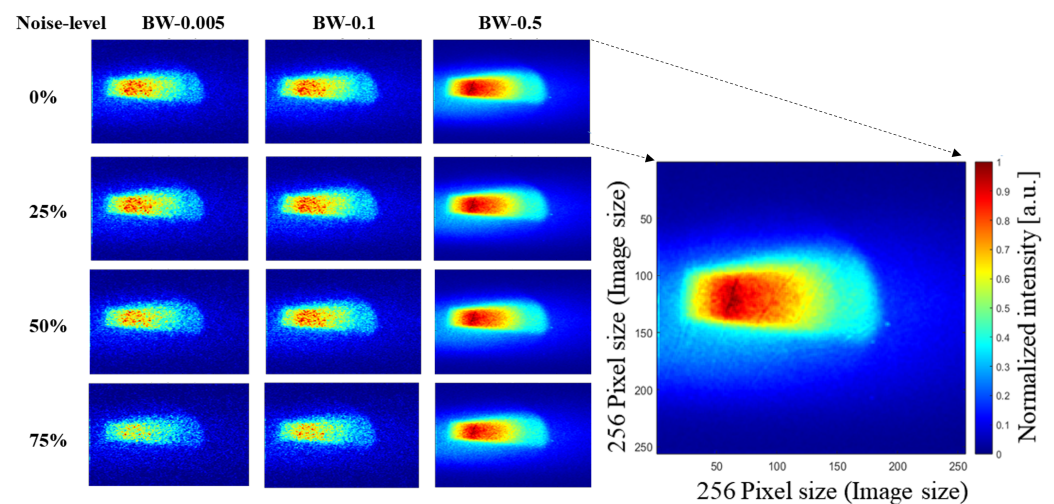


Figure 4. Examples from our dataset illustrating four different noise levels and three different bin widths (BW-0.005, BW-0.1, BW-0.5).

Figures 5, 6, and 7 show the predicted denoising qualitative results for various methods at noise levels of 25%, 50%, and 75%, respectively. These tables compare state-of-the-art techniques with our proposed model on our dataset. Results for the BM3D method are presented in the second row of each table. BM3D, a 3D block-matching algorithm primarily used for noise reduction in images, is an extension of the NLM methodology [47]. In the third column, the results for an enhanced version of MD3D are illustrated, which integrates a combination of BM4D within the 3D shearlet transform realm, alongside a generative adversarial network, for image denoising [31]. The fourth column displays the outcomes of the NLM method, which involves calculating the mean value for all pixels in the images, with weights assigned based on the similarity of each pixel to the target pixel [48]. Moving to the fifth column, we observe the results for the DnCNN [49]. This method aims to recover the clean image x from the noisy image $y = x + v$, assuming v is additive white Gaussian noise (AWGN). The network can handle Gaussian denoising with an unknown noise level. In general, image denoising methods can be categorized into two major groups: model-based methods, like BM3D, which are flexible in handling denoising problems with various noise levels but are time-consuming and require complex priors, and discriminative learning-based methods, like DnCNN, developed to overcome these drawbacks. In the sixth column, we present the results for blind image denoising via the SCUNet method [50],

while existing methods rely on simple noise assumptions, such as AWGN, the SCUNet method is designed to address all unknown noise types that remained unsolved in previous methods. The last column shows the results of our method, which extends the performance of the SCUNet method. The qualitative results of the proposed model consistently yield superior visual outcomes across all different bin widths, while BM3D exhibits better visual results, it tends to produce smoother images, consequently removing important details. In contrast, the proposed method generates images that closely resemble the original (0% noise level) image. Conversely, DnCNN produces results that closely match images at the same noise level.

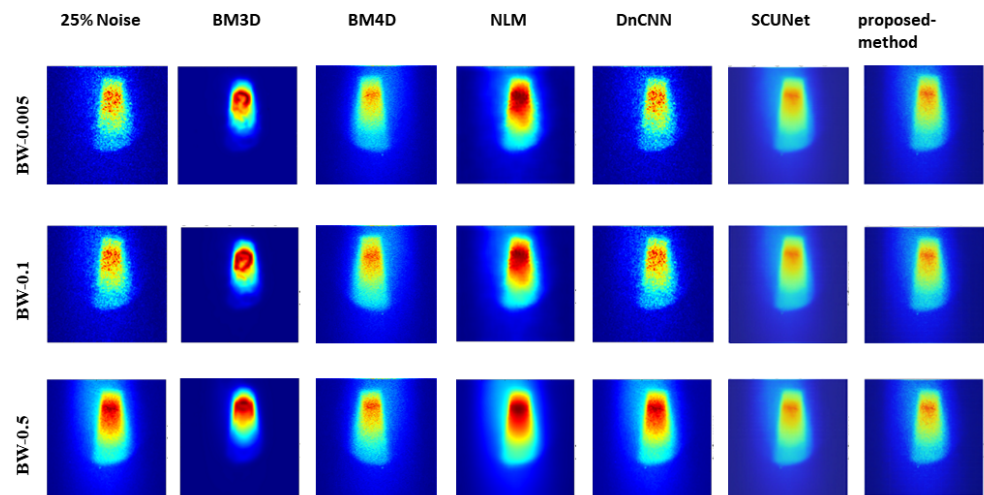


Figure 5. Visual comparison: predicted results of our method alongside those of other methods (noise level 25%).

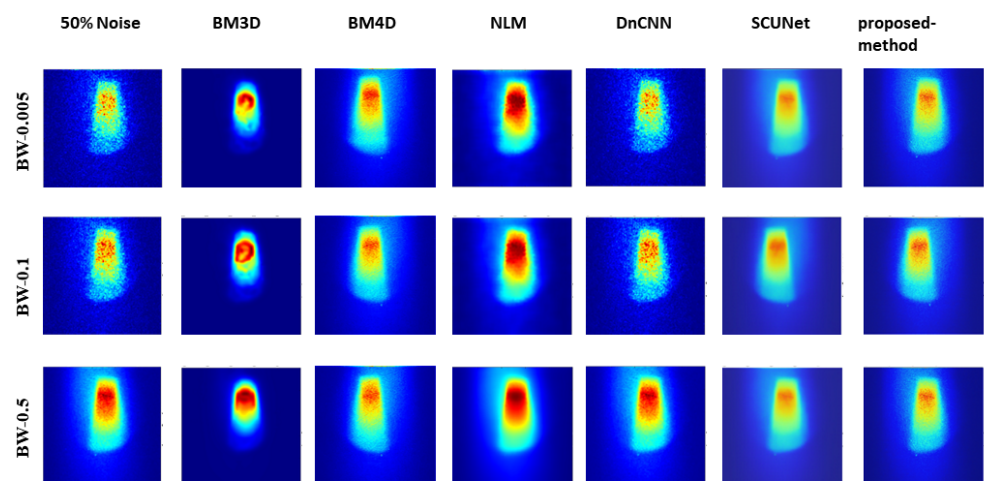


Figure 6. Visual comparison: predicted results of our method alongside those of other methods (noise level 50%).

Figure 8 presents the comparison results of predicted denoising for various methods at a 75% noise level with reference to the 0% noise condition. The findings indicate that our proposed model exhibits significantly improved visual results, closely resembling the original image even under high and low sampling rates.

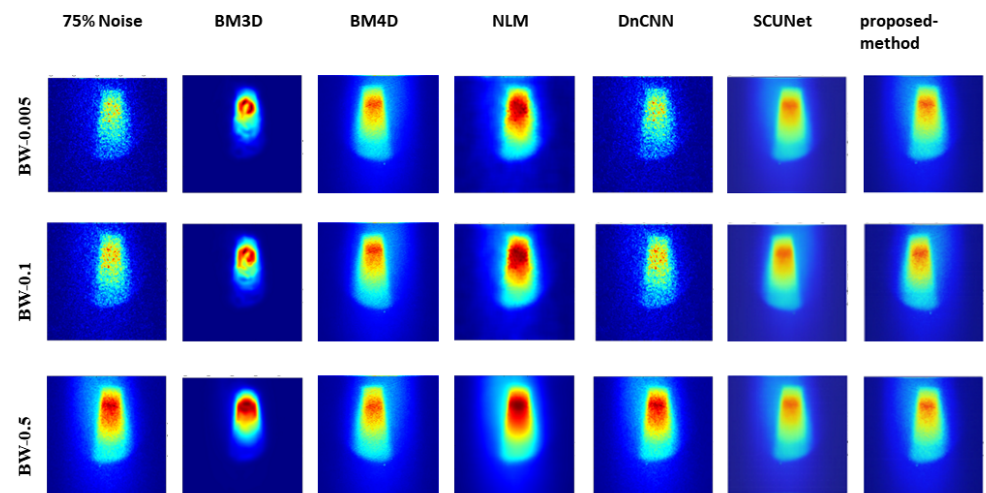


Figure 7. Visual comparison: predicted results of our method alongside those of other methods (noise level 75%).

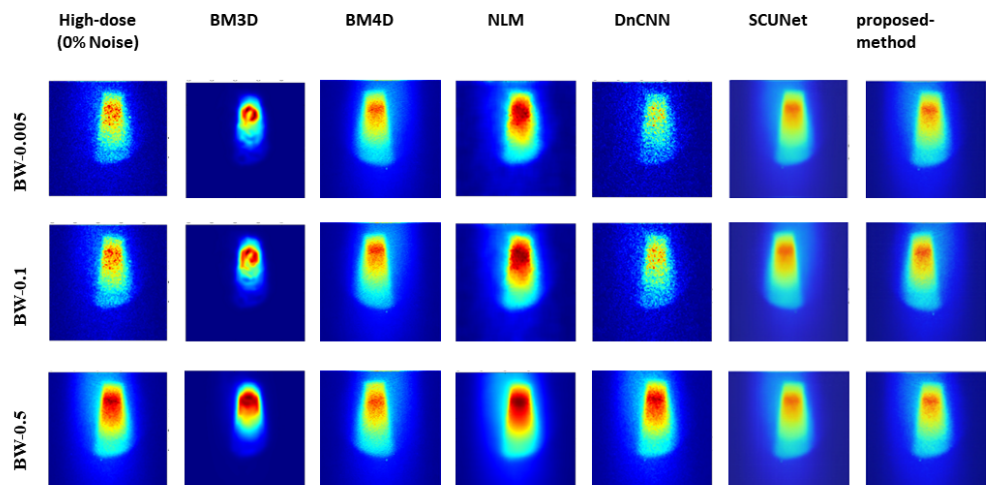


Figure 8. Visual comparison: predicted results of our method alongside those of other methods with the original image (noise level 0%).

In addition to the qualitative comparison of different methods, we present the quantitative results for noise levels of 25%, 50%, and 75%, based on two measurement factors, PSNR and SSIM, respectively, in Tables 1 and 2. The compared methods are BM3D, BM4D, NLM, DnCNN, and SCUNet. It can be observed that our proposed method achieved significantly better PSNR and SSIM results than other methods for the highest noise levels. Except in comparison with BM3D and DnCNN, which have in some cases higher PSNR and SSIM for the two noise levels of 25% and 50%, our method outperformed the others in terms of PSNR and SSIM. However, possible reasons for the better performance of DnCNN could be because of its match with the same noise level. We can see from the results of both tables that the SCUNet method achieves lower quantitative metrics (i.e., PSNR and SSIM) at all noise levels compared to our proposed model. The computational time performance for other methods, such as the SCUNet model, is around 0.072 s, whereas the computational time for BM4D inference is around 0.064 s, indicating a negligible increase in performance. The proposed model’s computational time is at around 0.059 s. Thus, the time performances of these methods are roughly on the same order magnitude.

Table 1. Quantitative results of the various denoising methods based on PSNR.

Bin Widths	Noise Level	BM3D	BM4D	NLM	DnCNN	SCUNet	Proposed Model
BW-0.05	25%	13.93	31.44	27.92	36.82	22.77	29.68
	50%	13.91	28.61	25.16	31.07	24.51	26.97
	75%	13.89	29.18	24.58	24.75	25.29	31.44
BW-0.1	25%	13.70	29.78	32.42	38.88	22.87	35.48
	50%	13.70	31.20	28.19	34.50	27.82	29.06
	75%	13.67	27.02	27.71	27.75	26.77	29.33
BW-0.5	25%	11.76	28.18	39.94	49.35	22.54	31.81
	50%	11.75	33.27	38.77	43.66	22.74	32.67
	75%	11.75	31.22	36.75	38.40	25.58	39.05

Table 2. Quantitative results of the various denoising methods based on SSIM.

Bin Widths	Noise Level	BM3D	BM4D	NLM	DnCNN	SCUNet	Proposed Model
BW-0.05	25%	0.073	0.7902	0.7435	0.9430	0.4661	0.7868
	50%	0.0720	0.7710	0.739	0.872	0.5177	0.8369
	75%	0.071	0.7917	0.7156	0.7431	0.6472	0.8654
BW-0.1	25%	0.0716	0.7425	0.7985	0.8594	0.6153	0.8284
	50%	0.0715	0.7154	0.7984	0.8023	0.6206	0.8658
	75%	0.0712	0.7556	0.7801	0.7867	0.6453	0.8218
BW-0.5	25%	0.0617	0.8349	0.8029	0.9139	0.5773	0.8786
	50%	0.0616	0.8314	0.7626	0.9014	0.4972	0.8466
	75%	0.0615	0.8424	0.7611	0.7582	0.5290	0.8638

4. Discussion and Conclusions

In this study, we refined the SCUNet deep learning model to enhance its applicability for background noise reduction in XFCT images. To optimize the application to our image data, we implemented a compound loss function [51] aimed at capturing shape-aware weight maps, addressing specific challenges posed by medical images within the SCUNet architecture. Our approach, as indicated in Tables 1 and 2, demonstrated improved outcomes, showcasing the effectiveness of the compound loss function [51].

To assess the efficacy of our approach, we conducted comparative analyses against five existing models BM3D, BM4D, NLM, DnCNN, and the original SCUNet. These comparisons demonstrate the efficiency of our blind deep learning model in eliminating all types of noise, including unknown noise, alongside known image noise such as AGN. Our overarching goal is to produce clean XFCT images from concurrent noisy images, corresponding to low-dose image acquisition, in order to enable low-dose XFCT imaging.

To facilitate our research, we curated a dataset comprising reconstructed images derived from sparse-view benchtop XFCT images. The deep learning model underwent training and evaluation, utilizing two key metrics, PSNR and SSIM. The results on experimental data provide insights into the effectiveness of our proposed method for denoising Gaussian noise and suggest the potential practicality of the trained deep blind model in handling real noisy images. In this study, we only present the preliminary results of our proposed method due to the constraints imposed by the limited availability of training data.

Analysis of the results, as illustrated in Figure 8, shows the robust performance of our model across varying noise levels and low-dose scenarios, transcending three different bin widths. Conversely, other methods exhibited performance dependency on bin width values and noise levels. Further insights from Figures 5, 6, and 7 showcased the output of all models under different noise levels of 25%, 50%, and 75%, respectively.

Quantitative assessments of noise removal demonstrated better performance of our proposed model compared to BM3D, NLM, and the original SCUNet. Our method demonstrated notable proficiency in handling images with high noise levels and sparse sampling, surpassing other studied methods, including SCUNet. Finally, our method achieved a maximum PSNR of 39.05 and SSIM of 0.86, indicating the improved performance of our model. We believe this framework offers a strong foundation for further optimization and tailoring of the SCUNet model specifically for the complexities of XFCT image denoising. However, our study has several limitations, primarily arising from the limited XFCT data available for model training. The presented results consider data and augmented samples from only a single phantom shape and size, corresponding to the small-animal-size scale. These phantoms are homogeneous water-filled borosilicate containers with hexagonal outer shapes, embedding contrast-filled targets. Additionally, we only have data corresponding to a single contrast agent (i.e., gadolinium at varying concentrations) and target size (i.e., around 8 mm diameter polypropylene microcentrifuge tubes filled with Gd). Correspondingly, due to the limited availability of training and reference data, unlike established anatomical imaging methods like CT and MRI, our present study has limits in comprehensively identifying potential artifacts and describing the detailed impact of our denoising model on the shape, size, and other properties of both target and background features. To improve upon current limitations and extend the model's capabilities, our future work could investigate the proposed method on a wider range of X-ray fluorescence data generated using Monte Carlo simulations. This data could encompass a variety of target and background shapes, sizes, and features, as well as different nanoparticle-based contrast agents. This would also enable us to investigate the various influencing factors, physical or algorithmic, affecting image quality and overall noise removal performance.

Author Contributions: N.M. and M.R. developed the methodology. M.F., K.K. and C.H. conceived the project and designed experiments. M.F. and K.K. acquired data. N.M. analyzed and interpreted data. N.M., M.R., K.K. and M.F. wrote, reviewed, or revised the manuscript. A.A.I. provided the results of the literature methods. C.H. supervised the study. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used within this study can be obtained here: <https://zenodo.org/records/11243130> (accessed on 7 March 2024). All other related raw data and datasets supporting the results and conclusions of this article will be made available by the authors without undue reservations.

Conflicts of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

XFCT	X-ray fluorescence computed tomography
XRF	X-ray fluorescence
AI	Artificial Intelligence
DL	Deep Learning
SCUNet	Swin-Conv-UNet
CT	Computed Tomography
PSNR	Peak signal-to-noise ratio
SSIM	Structural Similarity Index

BM3D	Block-Matching and 3D filtering
BM4D	Block-Matching and 3D filtering
MSE	Mean Squared Error
NLM	Non-local means
DnCNN	Denoising Convolutional Neural Network
CNN	Convolutional Neural Network
GAN	Generative Adversarial Networks
Gd	Gadolinium
wt	Weight
Swin	Shifted window Transformer
keV	Kilo electron volt
mA	milliampere
CBCT	Cone-Beam Computed Tomography
MAP	Maximum A Posteriori
DRUNet	Dilated-Residual U-Net
SwinIR	Image Restoration Using Swin Transformer
SConv	Strided Convolution
TConv	Transposed Convolution
RConv	Residual Convolutional
SwinT	Swin Transformer
BW	Bin width
AWGN	Additive White Gaussian Noise

References

1. Staufer, T.; Grüner, F. Review of development and recent advances in biomedical X-ray fluorescence imaging. *Int. J. Mol. Sci.* **2023**, *24*, 10990. [[CrossRef](#)] [[PubMed](#)]
2. Shaker, K.; Vogt, C.; Katsu-Jiménez, Y.; Kuiper, R.V.; Andersson, K.; Li, Y.; Larsson, J.C.; Rodriguez-Garcia, A.; Toprak, M.S.; Arsenian-Henriksson, M.; et al. Longitudinal in vivo X-ray fluorescence computed tomography with molybdenum nanoparticles. *IEEE Trans. Med. Imaging* **2020**, *39*, 3910–3919. [[CrossRef](#)] [[PubMed](#)]
3. Manohar, N.; Reynoso, F.J.; Diagaradjane, P.; Krishnan, S.; Cho, S.H. Quantitative imaging of gold nanoparticle distribution in a tumor-bearing mouse using benchtop X-ray fluorescence computed tomography. *Sci. Rep.* **2016**, *6*, 22079. [[CrossRef](#)] [[PubMed](#)]
4. Larsson, J.C.; Vogt, C.; Vågberg, W.; Toprak, M.S.; Dzieran, J.; Arsenian-Henriksson, M.; Hertz, H.M. High-spatial-resolution X-ray fluorescence tomography with spectrally matched nanoparticles. *Phys. Med. Biol.* **2018**, *63*, 164001. [[CrossRef](#)] [[PubMed](#)]
5. Dao, A.T.N.; Mott, D.M.; Maenosono, S. Characterization of metallic nanoparticles based on the abundant usages of X-ray techniques. In *Handbook of Nanoparticles*; Springer International Publishing: Cham, Switzerland, 2015; pp. 217–244. [[CrossRef](#)]
6. Takeda, T.; Yu, Q.; Yashiro, T.; Zeniya, T.; Wu, J.; Hasegawa, Y.; Hyodo, K.; Yuasa, T.; Dilmanian, F.; Akatsuka, T.; et al. Iodine imaging in thyroid by fluorescent X-ray CT with 0.05 mm spatial resolution. *Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrometers Detect. Assoc. Equip.* **2001**, *467*, 1318–1321. [[CrossRef](#)]
7. Körnig, C.; Staufer, T.; Schmutzler, O.; Bedke, T.; Machicote, A.; Liu, B.; Liu, Y.; Gargioni, E.; Feliu, N.; Parak, W.J.; et al. In-situ X-ray fluorescence imaging of the endogenous iodine distribution in murine thyroids. *Sci. Rep.* **2022**, *12*, 2903. [[CrossRef](#)] [[PubMed](#)]
8. Staufer, T.; Körnig, C.; Liu, B.; Liu, Y.; Lanzloth, C.; Schmutzler, O.; Bedke, T.; Machicote, A.; Parak, W.J.; Feliu, N.; et al. Enabling X-ray fluorescence imaging for in vivo immune cell tracking. *Sci. Rep.* **2023**, *13*, 11505. [[CrossRef](#)] [[PubMed](#)]
9. Grüner, F.; Blumendorf, F.; Schmutzler, O.; Staufer, T.; Bradbury, M.; Wiesner, U.; Rosentreter, T.; Loers, G.; Lutz, D.; Richter, B.; et al. Localising functionalised gold-nanoparticles in murine spinal cords by X-ray fluorescence imaging and background-reduction through spatial filtering for human-sized objects. *Sci. Rep.* **2018**, *8*, 16561. [[CrossRef](#)]
10. Kumar, K.; Facht, M.; Al-Maatoq, M.; Chakraborty, A.; Khismatrao, R.S.; Oka, S.V.; Staufer, T.; Grüner, F.; Michel, T.; Walles, H.; et al. Characterization of a polychromatic microfocus X-ray fluorescence imaging setup with metallic contrast agents in a microphysiological tumor model. *Front. Phys.* **2023**, *11*, 1125143. [[CrossRef](#)]
11. Zhang, S.; Li, L.; Chen, J.; Chen, Z.; Zhang, W.; Lu, H. Quantitative imaging of Gd nanoparticles in mice using benchtop cone-beam X-ray fluorescence computed tomography system. *Int. J. Mol. Sci.* **2019**, *20*, 2315. [[CrossRef](#)]
12. Deng, L.; Ahmed, M.F.; Jayarathna, S.; Feng, P.; Wei, B.; Cho, S.H. A detector's eye view (DEV)-based OSEM algorithm for benchtop X-ray fluorescence computed tomography (XFCT) image reconstruction. *Phys. Med. Biol.* **2019**, *64*, 08NT02. [[CrossRef](#)] [[PubMed](#)]
13. Jung, S.; Kim, T.; Lee, W.; Kim, H.; Kim, H.S.; Im, H.J.; Ye, S.J. Dynamic in vivo X-ray fluorescence imaging of gold in living mice exposed to gold nanoparticles. *IEEE Trans. Med. Imaging* **2019**, *39*, 526–533. [[CrossRef](#)]
14. Jung, S.; Sung, W.; Ye, S.J. Pinhole X-ray fluorescence imaging of gadolinium and gold nanoparticles using polychromatic X-rays: A Monte Carlo study. *Int. J. Nanomed.* **2017**, *12*, 5805–5817. [[CrossRef](#)]

15. Ahmad, M.; Bazalova-Carter, M.; Fahrig, R.; Xing, L. Optimized detector angular configuration increases the sensitivity of X-ray fluorescence computed tomography (XFCT). *IEEE Trans. Med. Imaging* **2014**, *34*, 1140–1147. [[CrossRef](#)] [[PubMed](#)]
16. Cong, W.; Shen, H.; Cao, G.; Liu, H.; Wang, G. X-ray fluorescence tomographic system design and image reconstruction. *J. X-ray Sci. Technol.* **2013**, *21*, 1–8. [[CrossRef](#)]
17. Jones, B.L.; Manohar, N.; Reynoso, F.; Karellas, A.; Cho, S.H. Experimental demonstration of benchtop X-ray fluorescence computed tomography (XFCT) of gold nanoparticle-loaded objects using lead-and tin-filtered polychromatic cone-beams. *Phys. Med. Biol.* **2012**, *57*, N457. [[CrossRef](#)] [[PubMed](#)]
18. Cheong, S.K.; Jones, B.L.; Siddiqi, A.K.; Liu, F.; Manohar, N.; Cho, S.H. X-ray fluorescence computed tomography (XFCT) imaging of gold nanoparticle-loaded objects using 110 kVp X-rays. *Phys. Med. Biol.* **2010**, *55*, 647. [[CrossRef](#)] [[PubMed](#)]
19. Feng, P.; Luo, Y.; Zhao, R.; Huang, P.; Li, Y.; He, P.; Tang, B.; Zhao, X. Reduction of Compton background noise for X-ray fluorescence computed tomography with deep learning. *Photonics* **2022**, *9*, 108. [[CrossRef](#)]
20. Li, W.; Liu, H.; Wang, J. A deep learning method for denoising based on a fast and flexible convolutional neural network. *IEEE Trans. Geosci. Remote. Sens.* **2021**, *60*, 1–13. [[CrossRef](#)]
21. Tian, C.; Fei, L.; Zheng, W.; Xu, Y.; Zuo, W.; Lin, C.W. Deep learning on image denoising: An overview. *Neural Netw.* **2020**, *131*, 251–275. [[CrossRef](#)]
22. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
23. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015, pp. 1–9. [[CrossRef](#)]
24. Rahim, T.; Usman, M.A.; Shin, S.Y. A survey on contemporary computer-aided tumor, polyp, and ulcer detection methods in wireless capsule endoscopy imaging. *Comput. Med. Imaging Graph.* **2020**, *85*, 101767. [[CrossRef](#)] [[PubMed](#)]
25. Chen, H.; Zhang, Y.; Kalra, M.K.; Lin, F.; Chen, Y.; Liao, P.; Zhou, J.; Wang, G. Low-dose CT with a residual encoder-decoder convolutional neural network. *IEEE Trans. Med. Imaging* **2017**, *36*, 2524–2535. [[CrossRef](#)] [[PubMed](#)]
26. Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; Zhang, L. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Trans. Image Process.* **2017**, *26*, 3142–3155. [[CrossRef](#)] [[PubMed](#)]
27. Chen, J.; Chen, J.; Chao, H.; Yang, M. Image blind denoising with generative adversarial network based noise modeling. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 3155–3164. [[CrossRef](#)]
28. Guo, S.; Yan, Z.; Zhang, K.; Zuo, W.; Zhang, L. Toward convolutional blind denoising of real photographs. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 1712–1722.
29. Krull, A.; Buchholz, T.O.; Jug, F. Noise2void-learning denoising from single noisy images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 2129–2137.
30. Sun, J.; Tappen, M.F. Learning non-local range Markov random field for image restoration. In Proceedings of the CVPR 2011, Colorado Springs, CO, USA, 20–25 June 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 2745–2752. [[CrossRef](#)]
31. Zhang, S.; Wang, L.; Chang, C.; Liu, C.; Zhang, L.; Cui, H. An image denoising method based on BM4D and GAN in 3D shearlet domain. *Math. Probl. Eng.* **2020**, *2020*, 1–11. [[CrossRef](#)]
32. Xu, P.; Chen, B.; Xue, L.; Zhang, J.; Zhu, L.; Duan, H. A new MNF–BM4D denoising algorithm based on guided filtering for hyperspectral images. *ISA Trans.* **2019**, *92*, 315–324. [[CrossRef](#)] [[PubMed](#)]
33. Lefkimmatis, S. Non-local color image denoising with convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 3587–3596.
34. Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; Timofte, R. Swinir: Image restoration using swin transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 1833–1844.
35. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 10012–10022. [[CrossRef](#)]
36. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241. [[CrossRef](#)]
37. Schmidt, U.; Roth, S. Shrinkage fields for effective image restoration. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 23–28 June 2014; pp. 2774–2781.
38. Chen, Y.; Pock, T. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 1256–1272. [[CrossRef](#)] [[PubMed](#)]
39. Zhang, K.; Li, Y.; Zuo, W.; Zhang, L.; Van Gool, L.; Timofte, R. Plug-and-play image restoration with deep denoiser prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 6360–6376. [[CrossRef](#)] [[PubMed](#)]
40. He, K.; Zhang, X.; Ren, S.; Sun, J. Identity mappings in deep residual networks. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Proceedings, Part IV 14; Springer: Berlin/Heidelberg, Germany, 2016; pp. 630–645. [[CrossRef](#)]

41. Lim, B.; Son, S.; Kim, H.; Nah, S.; Mu Lee, K. Enhanced deep residual networks for single image super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 136–144. [[CrossRef](#)]
42. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
43. Wang, Z.; Cun, X.; Bao, J.; Zhou, W.; Liu, J.; Li, H.U. A General U-Shaped Transformer for Image Restoration. *arXiv* **2021**, arXiv:2106.03106.
44. Cao, H.; Wang, Y.; Chen, J.; Jiang, D.; Zhang, X.; Tian, Q.; Wang, M. Swin-unet: Unet-like pure transformer for medical image segmentation. In Proceedings of the European Conference on Computer Vision, Tel Aviv, Israel, 23–27 October 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 205–218.
45. Li, Y.; Zhang, K.; Cao, J.; Timofte, R.; Van Gool, L.L. Bringing locality to vision transformers. *arXiv* **2021**, arXiv:2104.05707.
46. Rehman, A.; Wang, Z. SSIM-based non-local means image denoising. In Proceedings of the 2011 18th IEEE International Conference on Image Processing, Brussels, Belgium, 11–14 September 2011; IEEE: Piscataway, NJ, USA, 2011; pp. 217–220. [[CrossRef](#)]
47. Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095. [[CrossRef](#)] [[PubMed](#)]
48. Heo, Y.C.; Kim, K.; Lee, Y. Image denoising using non-local means (NLM) approach in magnetic resonance (MR) imaging: A systematic review. *Appl. Sci.* **2020**, *10*, 7028. [[CrossRef](#)]
49. Babu, D.; K Jose, S. Review on CNN based image denoising. In Proceedings of the International Conference on Systems, Energy & Environment (ICSEE), Singapore, 3–5 February 2021. [[CrossRef](#)]
50. Zhang, K.; Li, Y.; Liang, J.; Cao, J.; Zhang, Y.; Tang, H.; Fan, D.P.; Timofte, R.; Gool, L.V. Practical blind image denoising via Swin-Conv-UNet and data synthesis. *Mach. Intell. Res.* **2023**, *20*, 822–836. [[CrossRef](#)]
51. Liang, T.; Jin, Y.; Li, Y.; Wang, T. Edcnn: Edge enhancement-based densely connected network with compound loss for low-dose ct denoising. In Proceedings of the 2020 15th IEEE International Conference on Signal Processing (ICSP), Beijing, China, 6–9 December 2020; IEEE: Piscataway, NJ, USA, 2020; Volume 1, pp. 193–198. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.