

EAGLE: an edge-aware gradient localization enhanced loss for CT image reconstruction

Yipeng Sun^{1,*,} Yixing Huang,² Zeyu Yang,^{3,4} Linda-Sophie Schneider^{1,5},
Mareike Thies^{1,5}, Mingxuan Gu^{1,5}, Siyuan Mei,¹ Siming Bayer^{1,5},
Frank G. Zöllner^{1,3,4} and Andreas Maier^{1,5}

¹University of Erlangen-Nuremberg, Pattern Recognition Lab, Erlangen, Germany

²Peking University, Institute of Medical Technology, Beijing, China

³Heidelberg University, Computer Assisted Clinical Medicine, Medical Faculty Mannheim, Heidelberg, Germany

⁴Heidelberg University, Mannheim Institute for Intelligent Systems in Medicine, Medical Faculty Mannheim, Heidelberg, Germany

ABSTRACT. **Purpose:** We aim to enhance deep learning-based computed tomography (CT) image reconstructions. Conventional loss functions such as mean squared error (MSE) yield blurry images, and alternative methods may introduce artifacts. To address these limitations, we propose Eagle-Loss, designed to improve sharpness and edge definition without increasing computational complexity. Eagle-Loss leverages spectral analysis of localized gradient variations to enhance visual quality and quantification in CT imaging.

Approach: Eagle-Loss enhances CT reconstructions by segmenting gradient maps into patches and calculating intra-block variance to create a variance map. This map is analyzed in the frequency domain to identify critical features. We evaluate Eagle-Loss on two public datasets for low-dose CT reconstruction and field-of-view (FOV) extension and on a private photon counting CT (PCCT) dataset for super-resolution. Eagle-Loss is integrated into various deep learning models and used as a regularizer in ART, demonstrating effectiveness across reconstruction methods.

Results: Our experiments show that Eagle-Loss outperforms existing methods across all evaluated tasks, consistently improving the visual quality of reconstructed CT images. It achieves better performance compared with current top-performing loss functions when used with different network architectures while maintaining similar speed without adding extra computational costs. For low-dose CT reconstruction, Eagle-Loss achieved the highest SSIM scores of 0.958 for the TF-FBP model and 0.972 for RED-CNN. In the CT FOV task, our method reached a best SSIM of 0.966. For PCCT super-resolution, Eagle-Loss attained a top SSIM score of 0.998.

Conclusions: Eagle-Loss effectively mitigates blurring and artifacts prevalent in current CT reconstruction methods by significantly improving image sharpness and edge definition. Our evaluation confirms that Eagle-Loss can be successfully integrated into various deep learning models and reconstruction techniques without incurring additional computational costs, underscoring its robustness and model-independent nature. This significant enhancement in visual quality is important for achieving more accurate diagnoses and improved clinical outcomes.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JMI.12.1.014001](https://doi.org/10.1117/1.JMI.12.1.014001)]

Keywords: deep learning; loss function; medical image reconstruction; computed tomography

*Address all correspondence to Yipeng Sun, yipeng.sun@fau.de

1 Introduction

1.1 Background and Motivation

Computed tomography (CT) imaging has become an important tool in modern healthcare, playing a critical role in quantitative diagnosis, evidence-based treatment planning, and patient-centric care across a wide spectrum of disease pathways.¹ The quality and speed of CT image reconstruction directly influence the accuracy of diagnosis, and the effectiveness of treatment plans, hence the overall efficiency of clinical workflows.² Traditional reconstruction methods such as filtered backprojection (FBP) and algebraic reconstruction technique (ART) often face challenges when noise or disturbance is present in the data. These approaches struggle to effectively balance image quality and computational efficiency, particularly in scenarios where such noise, artifacts, or disturbances exist.

In recent years, there has been a marked increase in the application of deep learning techniques to enhance various aspects of CT reconstruction.³⁻⁵ These approaches have shown promising results in addressing longstanding issues such as noise reduction and artifact removal.⁶⁻¹⁰ The potential of deep learning in CT reconstruction lies in its ability to learn complex, nonlinear mappings between input data and high-quality reconstructions, often outperforming traditional methods.^{11,12}

1.2 Loss Functions in Deep Learning-Based CT Reconstruction

At the core of deep learning approaches lies the optimization process, which relies on minimizing a specific objective function, known as the loss function, through backpropagation. The design of this loss function is critical in determining the quality and characteristics of the reconstructed CT images. It serves as the primary guide for the neural network during training, influencing what features are emphasized or suppressed in the final reconstruction.

The most commonly used loss function in CT image reconstruction is the pixel-wise mean squared error (MSE). Its popularity stems from its computational simplicity and its alignment with Gaussian noise models, which are often assumed in CT imaging.¹³ However, MSE has significant limitations when it comes to capturing the nuances of human visual perception. Research has shown that MSE may not accurately reflect human-perceived image quality,¹⁴ as it fails to account for the varying perception of noise based on image content, such as luminance and contrast.¹⁵

This discrepancy between MSE-based optimization and human visual assessment can lead to reconstructions that, while numerically accurate, may not be optimal for clinical interpretation. For instance, MSE tends to overly smooth images, potentially obscuring fine details that could be crucial for accurate diagnosis. Moreover, it does not adequately penalize structural distortions that might be visually significant but contribute little to the overall pixel-wise error.¹⁶

1.3 Alternative Loss Functions and Their Limitations

Recognizing the limitations of MSE, researchers have proposed various alternative loss functions.¹⁷⁻²⁶ These can be broadly categorized into three main types, each with its own strengths and limitations:

1. **Perception-driven loss functions:** These functions leverage pre-trained neural networks, typically convolutional neural networks (CNNs) trained on large-scale image classification tasks, to compare features between reconstructed and ground truth images. The underlying principle is that these pre-trained networks have learned to extract features relevant to human visual perception.²⁷ However, this approach faces significant challenges in medical imaging contexts:

Stylistic deviations and mitigation: Perceptual loss can bias reconstructed images towards the style of its training data,²⁸ which is particularly problematic for CT images where such deviations can obscure important diagnostic details. To address this issue, transfer learning

techniques can be employed. Pre-training the perceptual loss model on CT data can mitigate style biases,²⁹ helping to align the perceptual model's feature space with the characteristics of CT images.³⁰

Computational considerations: The use of perception-driven loss functions introduces significant computational challenges. The pre-trained network needs to process the images and calculate the loss for each iteration during training and validation, substantially increasing the time and resources required compared with simpler loss functions such as MSE. In addition, the complex nature of these networks reduces interpretability, which is crucial in medical applications.

These limitations underscore the need for more specialized and efficient loss functions that can capture the perceptual quality of CT images without introducing style biases or excessive computational burden. Future research could focus on developing lightweight perceptual models specifically tailored for CT imaging or exploring hybrid approaches that balance perceptual quality with computational efficiency.³¹

2. **Gradient-based methods:** These methods focus on preserving edge information by comparing gradient information between the reconstructed and ground truth images.^{18,32} They typically require fewer computations than perceptual loss functions, making them more efficient for training. However, current gradient-based techniques often examine gradients only in the spatial domain.²² This limitation can lead to globally blurred gradient maps, which in turn can hinder the reconstruction of sharp edges. Although these methods are generally better at preserving structural information than MSE, they may still struggle with fine detail preservation and can sometimes produce over-sharpened artifacts.
3. **Frequency domain approaches:** These methods analyze the differences between reconstructed and ground truth images in the frequency domain, typically using a Fourier transform.^{33,34} They excel at recovering high-frequency texture details, which can lead to visually sharp images. However, most current frequency domain methods focus solely on magnitude differences, omitting phase information. This omission is problematic because phase information is crucial for accurately defining edges and overall image structure. Consequently, although these methods can produce images with rich texture, they may struggle with accurate edge definition and overall structural coherence.

Each of these approaches offers certain advantages over MSE, but none fully addresses all the challenges inherent in CT image reconstruction.^{35,36} The ideal loss function would need to balance perceptual quality, structural accuracy, and computational efficiency while also being adaptable to the specific characteristics of CT images.

1.4 Contributions

Our work makes several contributions to the field of CT image reconstruction:

1. **Development of Eagle-Loss:** We introduce a loss function specifically designed for CT image reconstruction. Eagle-Loss combines gradient-based patch analysis, variance map generation, and frequency domain analysis to optimize both image sharpness and structural accuracy.
2. **Comprehensive evaluation across diverse CT scenarios:** We demonstrate the versatility and effectiveness of Eagle-Loss through extensive evaluations in multiple CT reconstruction contexts:
 - Low-dose CT reconstruction using two public datasets, addressing the challenge of producing high-quality images from reduced radiation exposure data.
 - CT field-of-view (FOV) extension, demonstrating the ability to reconstruct areas outside the scanner's primary field of view, effectively expanding the imaging area.
 - Photon counting CT (PCCT) super-resolution using a private dataset, showcasing Eagle-Loss's effectiveness in enhancing spatial resolution for this advanced CT modality.

This wide-ranging evaluation highlights Eagle-Loss's adaptability across diverse CT imaging paradigms and reconstruction challenges.

3. **Dataset contribution:** We modify and open-source a public dataset specifically for the CT FOV extension task, addressing the lack of standardized datasets in this area and facilitating further research in the field.
4. **Model-independent performance:** Our results demonstrate that Eagle-Loss consistently outperforms state-of-the-art loss functions in terms of visual quality and sharpness across various reconstruction models and CT modalities, offering a reliable solution independent of the specific reconstruction algorithm used.

By introducing Eagle-Loss and demonstrating its effectiveness across multiple scenarios, our work paves the way for improved CT image reconstruction methods that can potentially enhance diagnostic accuracy and patient care.

1.5 Paper Structure

The paper is structured as follows: Sec. 2 provides an in-depth discussion of the motivation behind Eagle-Loss as well as its mathematical foundations. Section 3 outlines the experimental setup and evaluation framework. Section 4 presents and analyzes the results of our experiments across various CT reconstruction scenarios and performs a comprehensive ablation study on the key components and hyperparameters of Eagle-Loss. Finally, Sec. 5 summarizes the main findings and contributions of our work, discussing implications and future directions.

2 Methodology

To address the limitations of existing loss functions and the specific challenges of CT image reconstruction, we introduce a loss function termed “Eagle-Loss.” This approach is inspired by the observation that image blurring typically reduces the variance in gradient map patches.¹⁸ Eagle-Loss leverages this insight to develop a more effective method for assessing image quality during the reconstruction process.

The key innovations of Eagle-Loss are

- **Gradient-based patch analysis:** Eagle-Loss begins by computing gradient maps of the image, which capture edge details in both horizontal and vertical directions. These gradient maps are then divided into non-overlapping patches, allowing for a localized analysis of image features at different scales and positions.
- **Variance map generation:** Within each patch of the gradient maps, Eagle-Loss computes the variance. This step creates variance maps that effectively capture local contrast and edge information, providing a more nuanced representation of image structure compared with traditional gradient-based methods.
- **Frequency domain analysis:** The variance maps are then transformed into the frequency domain using a discrete Fourier transform (DFT). A Gaussian high-pass filter is applied to emphasize high-frequency components, which are associated with fine details and sharp edges in images.
- **Magnitude spectrum comparison:** Finally, Eagle-Loss quantifies the difference between the reconstructed image and the ground truth by calculating the L_1 loss of their respective magnitude spectra. This approach allows for a comprehensive comparison of image features across different scales and orientations.

To the best of our knowledge, this is the first application of frequency analysis to localized features within gradient-derived variance maps for CT image reconstruction. By combining spatial localization with frequency domain analysis, Eagle-Loss aims to capture both fine-grained texture information and larger-scale structural features, potentially overcoming the limitations of existing methods.

Our method builds upon and significantly extends the work of Abrahamyan et al.,¹⁸ who pioneered the use of gradient variance for image generation tasks. Figure 1 provides a comprehensive overview of the Eagle-Loss process, illustrating the steps involved in computing the magnitude spectrum from an input image. This approach is motivated by a key observation: blurring in reconstructed images typically manifests as reduced variance in the patches of their gradient maps.¹⁸ By focusing our analysis on the high-frequency components of these variance

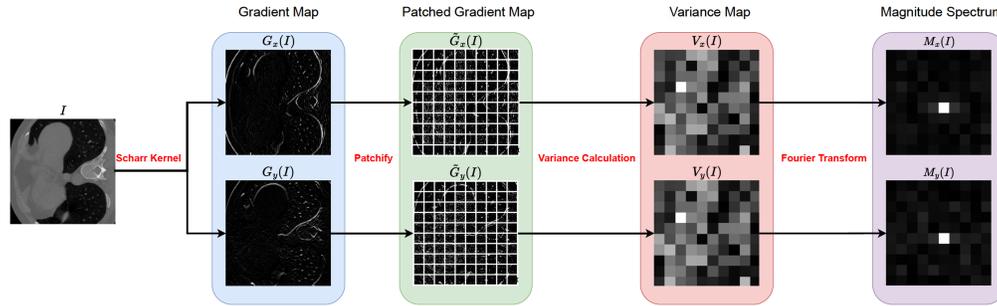


Fig. 1 Illustration of steps involved in computing the magnitude spectrum $M_x(I)$ and $M_y(I)$ from an image I . In this figure, larger patches are used for better visualization.

maps, we can more effectively capture and preserve fine details in reconstructed images, maintain sharp edges and intricate textures, and enhance overall image quality and fidelity.

Before explaining the specifics of Eagle-Loss, it is important to understand two key concepts: gradient maps and variance maps. A gradient map represents the directional change in intensity or color in an image. It highlights edges and textures, where pixel values change rapidly. Gradient maps are crucial for detecting features and understanding image structure, typically computed using convolution operations with specific kernels designed to detect changes in horizontal and vertical directions. A variance map quantifies the local contrast within an image. It is created by calculating the statistical variance of pixel values within small neighborhoods or patches of the image. High variance indicates areas with significant intensity changes (such as edges or textures), whereas low variance suggests more uniform regions. Variance maps can reveal subtle textures and structures that might not be immediately apparent in the original image.

Eagle-Loss leverages these concepts through a four-step process:

1. **Gradient map computation:** Let $I \in \mathbb{R}^{w \times h}$ be a grayscale image. We compute gradient maps of the input image $G_x(I)$ and $G_y(I)$ using Scharr kernels K_x and K_y ,³⁷ capturing edge details in both horizontal and vertical orientations. The process of computing Eagle-Loss is as follows:

$$G_x(I) = I * K_x, \quad G_y(I) = I * K_y, \quad (1)$$

where $*$ denotes convolution.

2. **Patch-wise variance calculation:** We then divide the gradient maps into non-overlapping patches and calculate the variance within each patch, creating variance maps that effectively capture local contrast and edge information. Here, we divide the gradient maps into non-overlapping patches of size $n \times n$:

$$\tilde{G}_x(I) = \mathcal{U}\{G_x(I), n\}, \quad \tilde{G}_y(I) = \mathcal{U}\{G_y(I), n\}, \quad (2)$$

where \mathcal{U} denotes the patching operation, whereas $\tilde{G}_x(I)$ and $\tilde{G}_y(I)$ represent the patchified gradient map.

$\mathcal{P}_{i,j}^x$ and $\mathcal{P}_{i,j}^y$ are the patches extracted from patchified gradient maps $\tilde{G}_x(I)$ and $\tilde{G}_y(I)$ along the x -axis and y -axis. We then calculate the variance:

$$v_{i,j}^x = \sigma^2(\mathcal{P}_{i,j}^x), \quad v_{i,j}^y = \sigma^2(\mathcal{P}_{i,j}^y), \quad i \in \left\{1, 2, \dots, \frac{w}{n}\right\}, \quad j \in \left\{1, 2, \dots, \frac{h}{n}\right\}. \quad (3)$$

The variance σ^2 for a patch \mathcal{P} is computed as

$$\sigma^2(\mathcal{P}) = \frac{1}{n^2} \sum_{k=1}^{n^2} (p_k - \mu)^2, \quad (4)$$

where p_k represents each element in the patch and μ is the mean value of the patch

$$\mu = \frac{1}{n^2} \sum_{k=1}^{n^2} p_k. \quad (5)$$

3. **Frequency domain transformation:** Next, we apply a DFT to the variance maps $V_x(I)$ and $V_y(I)$, where these maps denote the groups of variances from all patches, followed by a Gaussian high-pass filter,³⁸ allowing us to analyze the high-frequency components corresponding to fine details and sharp edges:

$$M_x(I) = W \odot |\mathcal{F}\{V_x(I)\}|, \quad M_y(I) = W \odot |\mathcal{F}\{V_y(I)\}|, \quad (6)$$

where \mathcal{F} is the DFT, \odot is element-wise multiplication, and W is the Gaussian high-pass filter:

$$W = 1 - e^{-\frac{(\sqrt{f_x^2 + f_y^2} - \kappa)^2}{2}}, \quad (7)$$

where f_x and f_y represent the frequency components in the x and y directions, respectively. These components correspond to the spatial frequencies in the image, with higher values representing finer details and lower values representing coarser structures. Specifically

- f_x ranges from $-\frac{w}{2}$ to $\frac{w}{2}$, where w is the width of the image.
- f_y ranges from $-\frac{h}{2}$ to $\frac{h}{2}$, where h is the height of the image.

The term $\sqrt{f_x^2 + f_y^2}$ calculates the radial distance from the origin in the frequency domain, effectively measuring how “high” the frequency is.

The parameter κ is the cutoff frequency, which determines the threshold between low and high frequencies. It acts as a control for the filter’s behavior:

- Frequencies below κ are attenuated (reduced in magnitude).
- Frequencies above κ are amplified or preserved.

A larger value of κ allows more low frequencies to pass through, resulting in less aggressive filtering, while a smaller value of κ creates a more stringent high-pass filter that emphasizes finer details and edges more strongly.

This Gaussian high-pass filter design ensures a smooth transition between attenuated and preserved frequencies, avoiding abrupt changes that could introduce artifacts in the filtered image.

4. **Magnitude spectrum comparison:** We compute the loss by comparing the magnitude spectra of the reconstructed image and the ground truth using the L_1 norm. For the reconstructed image I_{rec} and the ground truth image I_g , Eagle-Loss is computed as

$$\mathcal{L}_{\text{Eagle}} = \frac{1}{N} \|M_x(I_{\text{rec}}) - M_x(I_g)\|_1 + \frac{1}{N} \|M_y(I_{\text{rec}}) - M_y(I_g)\|_1, \quad (8)$$

where $N = \frac{wh}{n^2}$ is the number of pixels in the magnitude spectrum. The choice of the L_1 norm over the L_2 norm is supported by Parseval’s theorem,³⁹ which indicates that the Fourier transform is unitary. This property implies that the sum of the squares of function values remains invariant even after a Fourier transform is applied. Using the L_1 norm, we can effectively capture the differences in the magnitude spectrum while maintaining the energy conservation principle inherent in the Fourier transform.

This formulation of Eagle-Loss allows us to comprehensively assess the quality of reconstructed CT images by comparing both local and global features in the frequency domain, potentially leading to more accurate and visually pleasing reconstructions.

3 Experiments

3.1 Synthetic, Animal, and Clinical Data

Our experiments utilized three diverse datasets comprising synthetic, animal, and clinical data to evaluate the performance of Eagle-Loss across different CT reconstruction tasks comprehensively:

- **LoDoPaB-CT dataset:** We employed this public dataset⁴⁰ for low-dose CT reconstruction. From the original dataset comprising 35,802 training, 3522 validation, and 3553 testing samples, we randomly selected 1000 samples for training, 200 for validation, and 200 for testing. This dataset offers a rich collection of 362×362 phantom images and their corresponding 1000×513 sinograms, indicating that each CT scan comprises 1000 projections.
- **SMIR dataset:** The SMIR dataset⁴¹ containing head and neck CT data from 53 patients was used for CT FOV extension.⁴² We simulated a cone-beam computed tomography (CBCT) system with an FOV size of 32 cm to generate CBCT projections. The 3D Feldkamp-Davis-Kress Algorithm (FDK) reconstruction with water cylinder extrapolation was used to compute the reconstruction from truncated projection data. The reconstructed volumes have a size of $512 \times 512 \times 512$ with a voxel size of $1.27 \text{ mm} \times 1.27 \text{ mm} \times 1.27 \text{ mm}$. The original FOV diameter is 32 cm, and a U-Net⁴³ without pre-training was used to restore missing anatomical structures with a large FOV diameter of 65 cm. We used 2651 2D slices from 51 patients for training, 52 slices from one patient for validation, and 52 slices from one patient for testing.
- **Private PCCT dataset:** This study utilizes the PCCT dataset acquired with the Naetom Alpha CT scanner (Siemens Healthcare, Forchheim, Germany). Scans were performed using both tubes set to 120 kVp. The dataset includes scans of bones and lungs, conducted at a pitch of 1. The iterative reconstruction employed the “Qr40” quantitative reconstruction kernel, with the strength set to “QIR 3.” Reconstructions were carried out with a slice thickness of 0.60 mm, a field of view of 185 mm, and a matrix size of 512×512 . Consequently, PCCT images used in this study have a voxel size of $0.36 \text{ mm} \times 0.36 \text{ mm} \times 0.60 \text{ mm}$. SPP image files containing the spectral information are used for evaluation in this study. Our dataset includes scans of four different chicken bodies, sourced from a local supermarket, selected due to their small bone structures. This dataset is divided into 1080 slices for training, 136 slices for validation, and 135 slices for testing. The focus of the experiments is on $8 \times$ PCCT image super-resolution.

3.2 Experiment Setup

Our experiment framework was implemented based on Python 3.10 and PyTorch 2.0. For optimization, we employed the Adam optimizer ($\beta_1 = 0.9, \beta_2 = 0.99$), starting with an initial learning rate of 1×10^{-3} . A dynamic learning rate was implemented using a OneCycle learning rate scheduler, varying between 1×10^{-3} and 5×10^{-3} . All models were trained over 100 epochs on an NVIDIA RTX A6000 GPU.

A hybrid loss function, integrating MSE and our proposed Eagle-Loss [refer to Eq. (8)], was employed for training the models. This combination leverages the strengths of both components: MSE ensures global image fidelity, provides stability in training, and helps in noise reduction, whereas Eagle-Loss focuses on preserving high-frequency details and structural information critical in CT imaging. The loss function is formally defined as

$$\mathcal{L} = \mathcal{L}_{\text{MSE}} + \lambda \mathcal{L}_{\text{Eagle}}, \quad (9)$$

where $\lambda = 1 \times 10^{-3}$ denotes the weight coefficient for Eagle-Loss, empirically determined to balance the contributions of both components. This value ensures that Eagle-Loss significantly influences the optimization process without overshadowing the stabilizing effect of MSE. Based on our studies, we chose a patch size of $n = 3$ [see Eq. (2)] and a cutoff frequency of $\kappa = 0.5$ [see Eq. (7)] for Eagle-Loss. The small patch size allows for detailed analysis of local structures, while the selected κ value in the high-pass filter increases the model’s sensitivity to high-frequency components. This selection balances computational efficiency and the ability to capture important image features in CT reconstruction.

It is important to note that while these hyperparameters are not ideal for matrix calculations, they produced the sharpest images with well-defined edges and fine details, which aligns with our study’s goals. We provide a thorough analysis of how different hyperparameter choices affect the results in Sec. 4.4. Combining with MSE is important not only for Eagle-Loss but also for our competitive method. For instance, total variation (TV)⁴⁴ is a regularizer that cannot be used alone for training, and in our experiment, Gaussian edge-enhanced (GEE) also needs to be combined

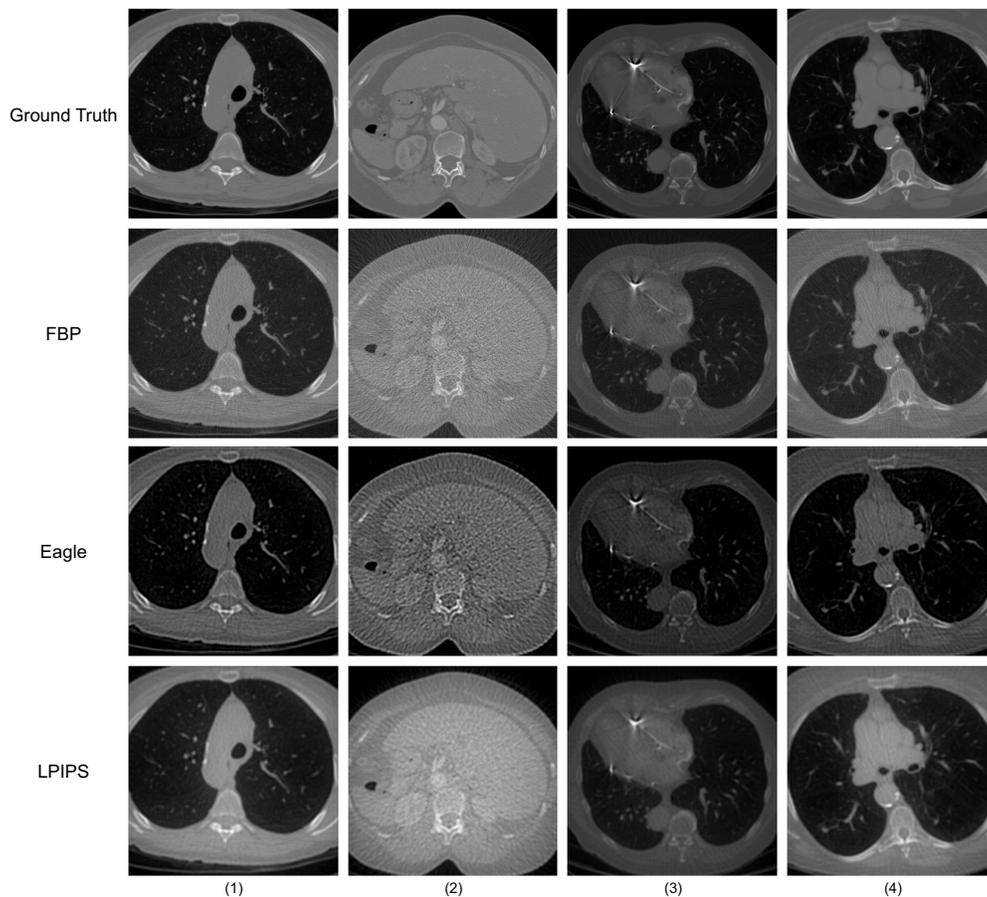


Fig. 2 Comparison of results from TF-FBP models trained with different loss functions without combining with MSE for low-dose CT reconstruction.

with MSE for stable training or it will induce a vanishing gradient problem. For the perceptual-driven approaches, because the backbone model is not trained on CT data, MSE serves as a strong constraint for optimizing the model with minimal style artifacts. Figure 2 presents the outcomes from TF-FBP models trained using individual loss functions. The results show that both Eagle-Loss and learned perceptual image patch similarity (LPIPS)⁴⁵ loss introduce some artifacts in the reconstructed images.

To ensure fair comparisons, all comparison methods use hybrid loss functions. We establish a systematic rule for selecting appropriate λ values across different loss functions to maintain fairness and optimize results. The rule works as follows: we evaluate both loss functions during the first training epoch, where λ always takes the form 1×10^n . We adjust n to ensure that the MSE and the weighted loss components (such as TV, perceptual loss, and others) are of the same order of magnitude. For example, if $L_{\text{MSE}} = 2.5 \times 10^{-2}$ and $L_{\text{TV}} = 3.7 \times 10^{-4}$, we would set $n = 2$ to obtain $\lambda = 100$, resulting in $\lambda \cdot L_{\text{TV}} = 3.7 \times 10^{-2}$, which matches the order of magnitude of L_{MSE} .

For each dataset, we employed different models:

- **Low-dose CT reconstruction:** We evaluated Eagle-Loss on two architectures: TF-FBP³³ and RED-CNN.⁴⁶ TF-FBP enhances the traditional FBP algorithm with a data-driven filter, utilizing trainable coefficients of the Fourier series. By contrast, RED-CNN improves FBP-reconstructed images through an encoder-decoder framework. The network structures for both models are depicted in Fig. 3. In addition, we explored the validity of Eagle-Loss as a regularizer in ART reconstructions.
- **CT FOV extension:** We used a U-Net⁴³ without pre-training to restore missing anatomical structures.

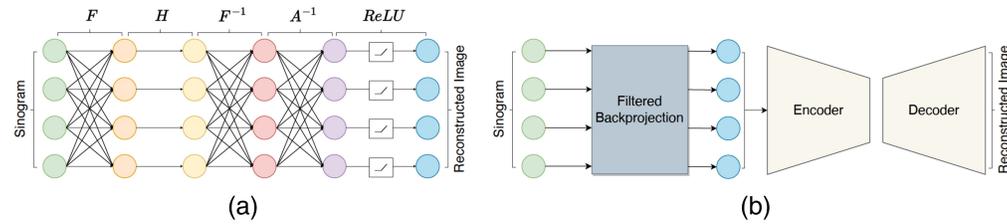


Fig. 3 Illustration of models used for low-dose CT reconstruction. (a) The structure of TF-FBP. Here, H represents the trainable filter and A^{-1} represents the differentiable backprojection operator, whereas F^{-1} and F denote the inverse and forward DFT. The differentiable backprojection operator was implemented using PYRO-NN.⁴⁷ (b) RED-CNN with its encoder-decoder structure. The input of the encoder is the reconstructed image using FBP.

- PCCT super-resolution:** The model employed for super-resolution is NinaSR.^{48,49} The head of the model starts with a rescale layer followed by a 3×3 convolution layer. The body of the model consists of a sequence of residual blocks, each containing two 3×3 convolutional layers and a channel attention mechanism. The attention mechanism uses local pooling, implemented with an average pooling layer, followed by 1×1 convolutions, ReLU and sigmoid activations, and upsampling with nearest-neighbor interpolation. The residual blocks also use scaling based on expected variance for stability. The tail of the model is responsible for upscaling the features to the target resolution. It includes a convolutional layer to increase the number of channels, followed by a pixel shuffle operation to achieve higher resolutions, and ends with a final rescaling layer. The largest pretrained NinaSR model was used for transfer learning in this study.⁴⁹ This configuration was chosen to leverage the capacity of the model for enhancing the resolution of the PCCT images by a factor of 8.

3.3 Evaluation Metrics

We used three primary metrics to evaluate the performance of various loss functions:

- Structural similarity index measure (SSIM):** SSIM is a perceptual metric that quantifies image quality degradation caused by processing such as data compression or transmission losses. It considers changes in structural information, luminance, and contrast. SSIM values range from -1 to 1 , where 1 indicates perfect structural similarity between the reconstructed and reference images. Higher SSIM values denote better reconstruction quality.¹³
- Peak signal-to-noise ratio (PSNR):** PSNR measures the ratio between the maximum possible value of a signal and the power of distorting noise that affects the fidelity of its representation. It is expressed in decibels (dB). Higher values indicate that the reconstructed image is closer to the original, implying better image quality. It is particularly useful for assessing the preservation of fine details in the reconstructed images.⁵⁰
- Training time (s):** Overall training time, measured in seconds (s), is a crucial metric for evaluating the computational efficiency of the model training process. It represents the total time taken to train the model for all epochs. Lower training times indicate more efficient training processes, which are essential for practical implementations,⁵¹ especially in clinical settings where time and computational resources are limited.

These metrics collectively provide a comprehensive assessment of reconstruction quality, structural preservation, and computational efficiency across our diverse experimental scenarios.

4 Results and Discussions

In this section, we present the results of comparing our proposed Eagle-Loss with other loss functions: GEE,³³ TV,⁴⁴ gradient variance (GV),¹⁸ SSIM, perceptual loss,¹⁹ and LPIPS.⁴⁵ GEE compares the high-frequency difference of images in the frequency domain using a high-pass filter. TV penalizes gradients to enhance smoothness, whereas GV compares the variance between image gradients. SSIM utilizes the structural similarity index as a loss function. The perceptual loss compares the MSE between different image features using an ImageNet⁵²

pre-trained VGG-19⁵³ model. Finally, LPIPS is a perceptual image similarity metric that leverages deep features extracted from ImageNet⁵² pre-trained AlexNet⁵⁴ and employs learned weights to compare images.

4.1 Low-Dose CT Reconstruction

Eagle-Loss demonstrates superior performance in low-dose CT reconstruction, as evidenced by both qualitative and quantitative evaluations. Figures 4 and 5 provide a visual comparison of two reconstruction samples using different loss functions.

The images reveal that MSE + Eagle preserves fine details and structural integrity more effectively than other methods, which tend to oversmooth the images. This is particularly noticeable in bone structures and soft tissue boundaries, where Eagle-Loss maintains sharpness while avoiding oversmoothing. The RED-CNN model generally produces better images compared to TF-FBP, regardless of the loss function used. The qualitative observations are corroborated by quantitative metrics presented in Table 1.

In terms of SSIM, the combination of MSE and Eagle-Loss outperforms other configurations, achieving the highest values for both TF-FBP (0.958) and RED-CNN (0.972) models. This represents a 0.24% and 1.66% improvement over the baseline MSE, respectively, indicating superior structural preservation. Regarding PSNR, MSE alone yields the highest value for TF-FBP (34.860 dB), whereas MSE + TV achieves the best for RED-CNN (37.296 dB). Interestingly, MSE + Eagle shows lower PSNR values (33.621 dB for TF-FBP and 36.233 dB for RED-CNN). This is attributed to its focus on preserving high-frequency details rather than oversmoothing the image, a trade-off often observed in image reconstruction tasks where methods that preserve more details may have lower PSNR due to the retention of some noise.

In terms of computational efficiency, the baseline MSE loss function exhibits the fastest training times for both models (8210.00 s for TF-FBP and 8306.20 s for RED-CNN). The MSE + Eagle combination is about 7.3% slower for TF-FBP and 6.4% slower for RED-CNN compared with MSE alone. However, it remains efficient compared with more complex losses

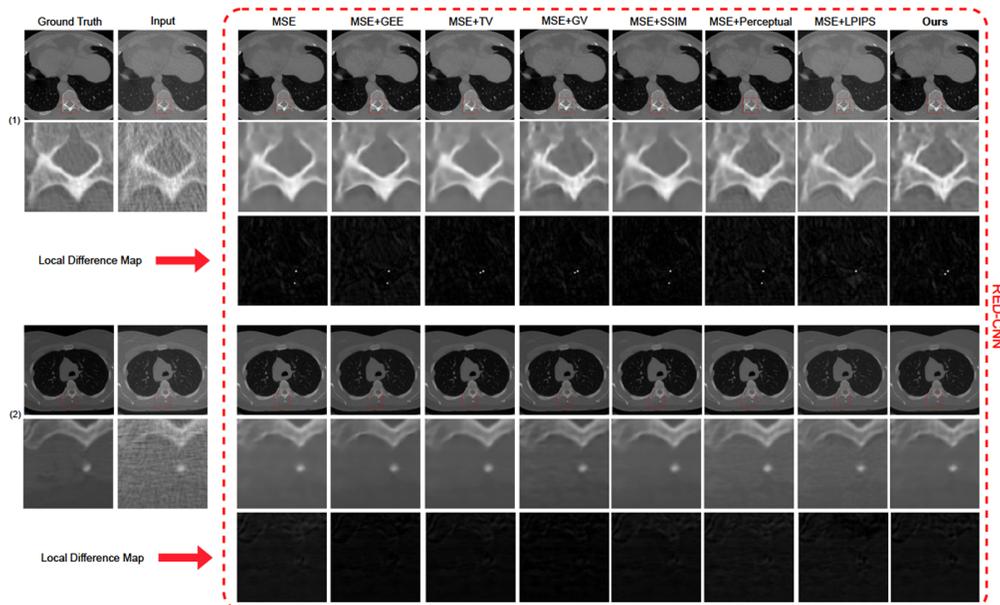


Fig. 4 Comparison of low-dose CT reconstruction results using different loss functions on the LoDoPaB-CT dataset, implemented with RED-CNN. Although the differences between methods are subtle, they are still discernible: (1) MSE, MSE+TV, and MSE + SSIM show a slight tendency to oversmooth, particularly in bone structures and soft tissue boundaries. (2) The proposed Eagle-Loss (rightmost column) achieves a marginally better balance between detail preservation and noise suppression. Note the subtle improvements in edge definition and texture preservation in the Eagle-Loss results compared with other methods, though the distinctions are less pronounced than in the TF-FBP implementation.

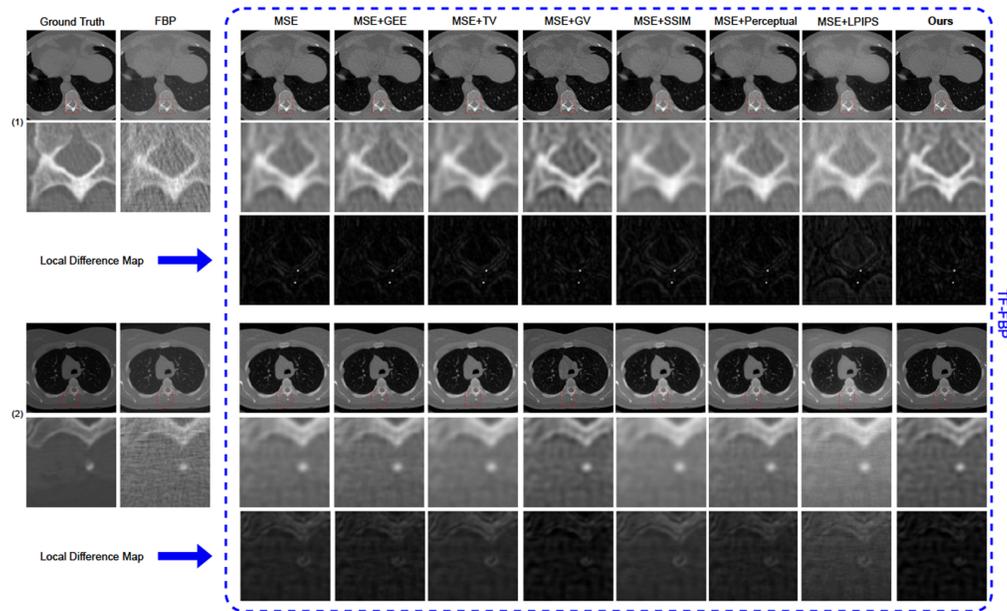


Fig. 5 Comparison of low-dose CT reconstruction results using different loss functions on the LoDoPaB-CT dataset, implemented with TF-FBP. The differences between methods are more pronounced in this case: (1) MSE, MSE + TV, and MSE + SSIM clearly oversmooth the image, resulting in significant loss of fine details, especially in bone structures and soft tissue boundaries. (2) The proposed Eagle-Loss (rightmost column) demonstrates a superior balance between detail preservation and noise suppression, maintaining sharpness in bone structures while avoiding oversmoothing in soft tissue areas. Note the markedly improved edge definition and texture preservation in the Eagle-Loss results. Below each zoomed-in image, its corresponding difference map is displayed, further highlighting the performance disparities between the methods.

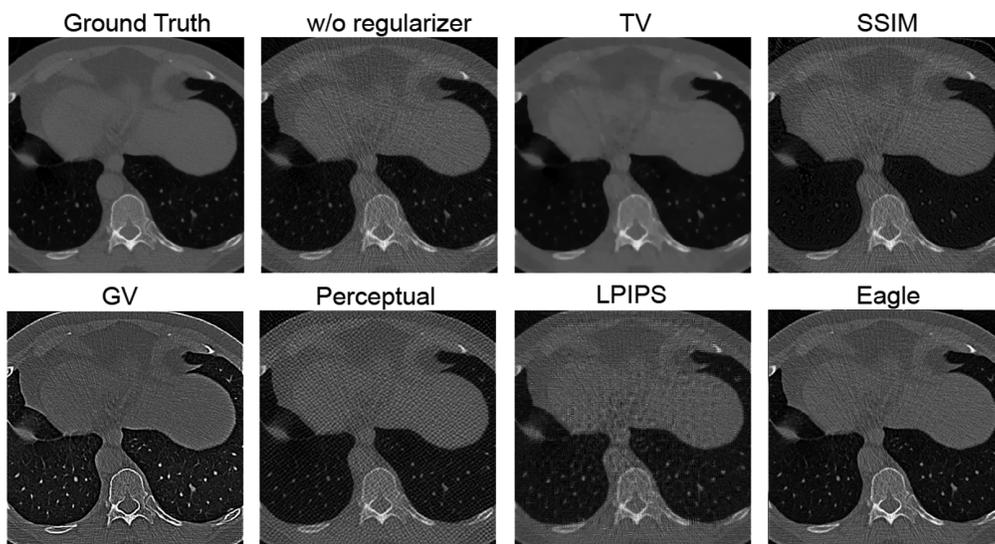
such as MSE + perceptual, which is $\sim 16.8\%$ slower for TF-FBP and 16.1% slower for RED-CNN. During the last training epoch, MSE accounted for 41.86% and 50.26% of our proposed hybrid loss in the two models, respectively. These percentages indicate that both MSE and Eagle-Loss make significant contributions to the overall loss function and final reconstructed image.

Table 1 Performance comparison of loss functions for low-dose CT reconstruction, evaluated on 200 samples. Metrics include SSIM, PSNR, and training speed for TF-FBP and RED-CNN models. Paired t tests are calculated⁵⁵ to compare each method with MSE + Eagle.

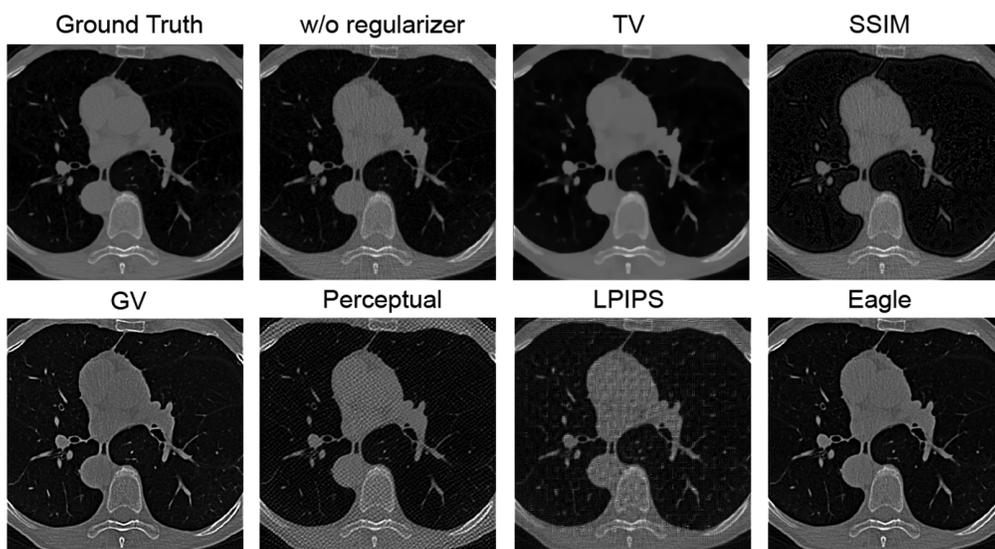
Loss function	SSIM \uparrow		PSNR (dB) \uparrow		Training time (s) \downarrow	
	TF-FBP	RED-CNN	TF-FBP	RED-CNN	TF-FBP	RED-CNN
MSE	0.956*	0.956	34.860	36.494*	8210.00	8306.20
MSE + GEE ³³	0.953	0.945	34.690	36.659	8518.08	9489.60
MSE + TV ⁴⁴	0.954*	0.971*	34.781	37.296	8378.40	8314.76
MSE + GV ¹⁸	0.952	0.952	33.190	35.558	8370.68	8337.00
MSE + SSIM	0.946	0.966	34.245	36.200*	8542.40	8402.48
MSE + Perceptual ¹⁹	0.953	0.961	34.489	36.639*	9589.40	9641.56
MSE + LPIPS ⁴⁵	0.947	0.955	33.074	35.996	11907.49	12,725.48
MSE + Eagle	0.958	0.972	33.621	36.233	8811.48	8840.88

*Indicates no statistically significant difference from Eagle-Loss ($p > 0.05$).

Note: Numbers shown in bold represent the best performance values within each matrix, making it easier for readers to identify the top results.



(a) Sample 1



(b) Sample 2

Fig. 6 Comparison of regularization techniques in ART reconstruction using two samples from the LoDoPaB-CT dataset. Perception-based regularizers (perceptual and LPIPS) introduce strong grid-like artifacts, whereas TV oversmooths the image. SSIM exhibits artifacts in high-contrast regions, and GV over-enhances edge contrast. By contrast, our proposed Eagle Loss achieves the best visual quality, successfully preserving fine details while avoiding noticeable artifacts, thus striking an optimal balance between noise reduction and detail preservation. GEE is not included in this comparison because of its vanishing gradient without a combination of MSE.

The effectiveness of Eagle-Loss extends beyond neural network-based approaches. Figure 6 illustrates the integration of Eagle-Loss as a regularization term into the ART algorithm, comparing it with TV regularization and no regularization.

The results demonstrate that Eagle-Loss significantly improves the sharpness and fidelity of CT reconstruction compared with no regularization and other regularization methods. Although TV regularization performs better than no regularization, it tends to oversmooth some details that Eagle-Loss successfully preserves. Other methods such as perception-based regularizers (perceptual and LPIPS) introduce strong grid-like artifacts, SSIM exhibits issues in high-contrast regions, and GV tends to over-enhance edge contrast.

These findings indicate that Eagle-Loss is a promising tool for low-dose CT reconstruction, offering enhanced image quality, particularly in terms of structural preservation and detail retention. Although there is a modest increase in computational cost, the improved image quality justifies this trade-off in many clinical scenarios where diagnostic accuracy is paramount. Eagle-Loss strikes an optimal balance between noise reduction and detail preservation, addressing the limitations observed in other regularization techniques and potentially enhancing the diagnostic value of low-dose CT images.

4.2 CT FOV Extension

Table 2 summarizes the performance metrics for various loss functions. The combination of MSE and Eagle-Loss (MSE + Eagle) demonstrates superior performance across key metrics. In terms of structural similarity, it achieves the highest SSIM of 0.966, a 1.09% improvement over baseline MSE. For image quality, it attains the best PSNR of 30.964 dB, surpassing MSE by 1.790 dB. Regarding computational efficiency, with a training time of 37602.32 s, it is only 1.15% slower than the fastest method (MSE + GEE) and 4.32% faster than MSE alone. These metrics indicate that Eagle-Loss enhances structural preservation and overall image quality without significantly increasing computational cost. During the last training epoch, MSE accounted for 48.13% of our proposed hybrid loss. This percentage indicates that both MSE and Eagle-Loss make important contributions to the overall loss function and final reconstructed image.

Figure 7 provides visual comparisons of FOV extension results. The images demonstrate that MSE + Eagle produces clearer structural boundaries compared with other methods. The transition between original and extended FOV appears more seamless with Eagle-Loss, and fine anatomical details are better preserved. By contrast, other methods, especially MSE alone, tend to generate blurrier extended regions.

Although Eagle-Loss shows overall superior performance, it exhibits some limitations in addressing streaky artifacts. This is possibly due to the high-pass filter interpreting these artifacts as low-frequency features. This observation points to potential areas for future improvements, such as incorporating additional terms specifically designed to mitigate such artifacts.

The ability of Eagle-Loss to maintain clear structural boundaries and seamlessly integrate the extended region is crucial for accurate interpretation and diagnosis in clinical settings. Its improved performance in both quantitative metrics and qualitative assessments suggests that Eagle-Loss could be a valuable tool for CT FOV extension tasks in practical applications.

In conclusion, Eagle-Loss demonstrates a compelling balance of improved image quality, structural preservation, and computational efficiency in CT FOV extension. Although there is

Table 2 Performance comparison of loss functions for CT FOV Extension, evaluated on 52 samples. Metrics include SSIM, PSNR, and training speed for U-Net. Paired t tests⁵⁵ are calculated to compare each method with MSE + Eagle.

Loss function	SSIM \uparrow	PSNR (dB) \uparrow	Training time (s) \downarrow
MSE	0.955	29.174	39,298.42
MSE + GEE	0.934	27.076	37,175.44*
MSE + TV	0.928	27.654	39,372.26
MSE + GV	0.955	28.947	39,046.87
MSE + SSIM	0.964*	30.241	40,115.87
MSE + Perceptual	0.951	28.900	77,336.19
MSE + LPIPS	0.964	30.048	80,114.13
MSE + Eagle	0.966	30.964	37,602.32

*Indicates no statistically significant difference from Eagle-Loss ($p > 0.05$).

Note: Numbers shown in bold represent the best performance values within each matrix, making it easier for readers to identify the top results.

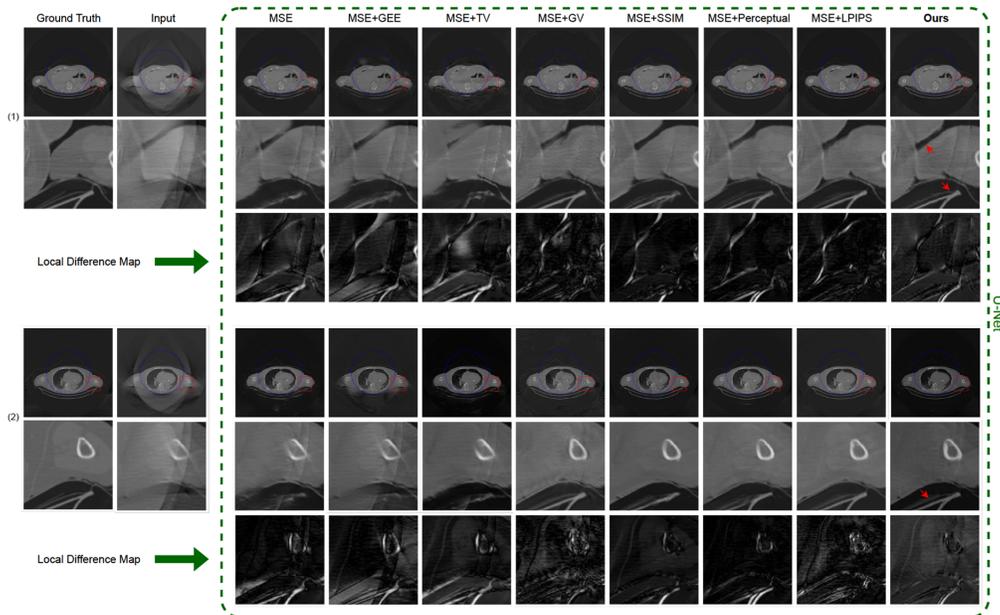


Fig. 7 Visualization of CT FOV extension results on the SMIR dataset using the U-Net model. The blue-dotted circle denotes the original FOV boundary. The input images show significant truncation artifacts outside the original FOV. Note the improved detail preservation and reduced artifacts in our method, particularly visible in the extended regions of the head and chest. The red arrows in our result highlight areas of notable improvement in structural definition and edge preservation compared with other methods. Below each zoomed-in picture, its corresponding difference map is displayed.

room for further refinement, particularly in artifact reduction, the current results indicate its strong potential for enhancing CT imaging capabilities in clinical practice.

4.3 PCCT Super-Resolution

Table 3 presents the quantitative results for the PCCT super-resolution task. In terms of SSIM, MSE + Eagle achieves the highest value (0.998), showing a 0.17% improvement over the baseline MSE. For PSNR, MSE + Eagle again achieves the highest value (45.105 dB), a substantial 3.217 dB improvement over MSE alone. These results demonstrate the exceptional performance

Table 3 Performance comparison of loss functions for PCCT super-resolution, evaluated on 135 samples. Metrics include SSIM, PSNR, and training speed for NinaSR. Paired t tests⁵⁵ are calculated to compare each method with MSE + Eagle.

Loss function	SSIM \uparrow	PSNR (dB) \uparrow	Training time (s) \downarrow
MSE	0.996	41.888	5567.85*
MSE + GEE	0.997*	42.697	5586.34*
MSE + TV	0.997*	42.951	5634.63*
MSE + GV	0.995	40.747	5558.84*
MSE + SSIM	0.997*	42.370	5968.55
MSE + perceptual	0.995	40.933	10,206.71
MSE + LPIPS	0.996	40.988	12,124.52
MSE + Eagle	0.998	45.105	5595.89

*Indicates no statistically significant difference from Eagle-Loss ($p > 0.05$)

Note: Numbers shown in bold represent the best performance values within each matrix, making it easier for readers to identify the top results.

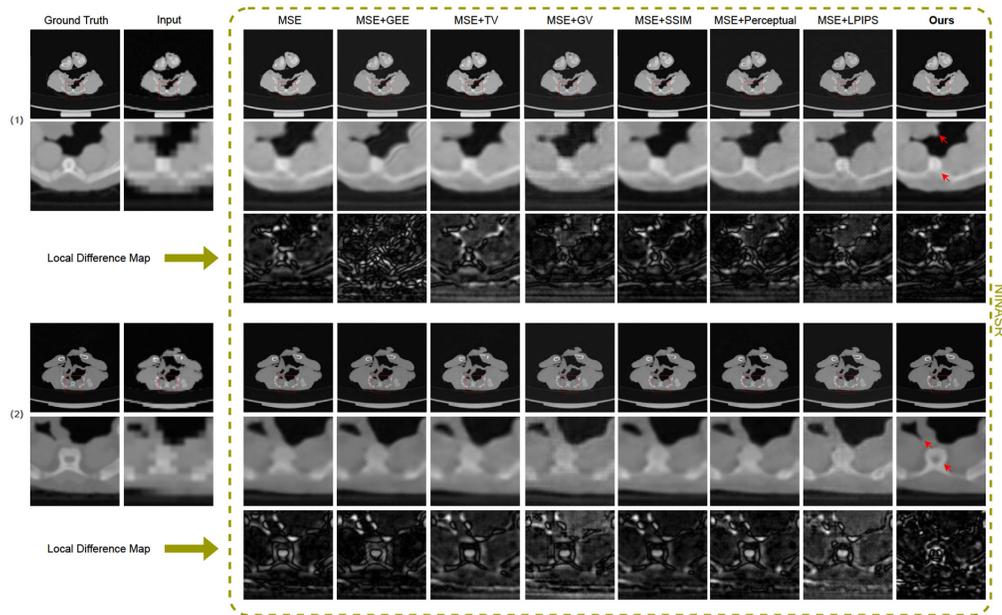


Fig. 8 Visualization of two PCCT super-resolution samples on the private PCCT dataset using the NinaSR model.⁴⁹ This figure compares the effectiveness of different loss functions on NinaSR in achieving 8 \times super-resolution. Note the improved detail preservation in our method, particularly visible in the fine structures of the small bones and surrounding soft tissues. Red arrows in our result highlight areas of notable improvement in edge definition and structural clarity compared with other methods. Below each zoomed-in picture, its corresponding difference map is displayed.

of Eagle-Loss in preserving structural details and enhancing overall image quality in the super-resolution task and during the last training epoch, MSE accounted for 40.57% of our proposed hybrid loss. This percentage indicates that both MSE and Eagle-Loss make important contributions to the overall loss function and final reconstructed image.

Regarding computational efficiency, MSE + GV has the fastest training time (5558.84 s), with MSE + Eagle being highly competitive (5595.89 s), only 0.67% slower than MSE + GV, and 0.50% slower than MSE alone. This demonstrates that the superior image quality achieved by Eagle-Loss comes with minimal additional computational cost.

Figure 8 provides a visual comparison of the PCCT super-resolution results using different loss functions. The images demonstrate that MSE + Eagle produces results with clearer and more defined structural boundaries. Fine details, particularly in areas of complex tissue structure, are better preserved with MSE + Eagle. By contrast, other loss functions, including MSE alone, tend to produce slightly blurrier results with less distinct edges.

The exceptional performance of Eagle-Loss in the PCCT super-resolution task is particularly noteworthy. The substantial improvement in both SSIM and PSNR, coupled with the visual enhancement of fine details and structural boundaries, suggests that Eagle-Loss could be a powerful tool for improving the resolution and quality of PCCT images. This could have significant implications for medical imaging, potentially enabling more accurate diagnoses and better visualization of subtle anatomical features.

The consistent superior performance of Eagle-Loss across all three tasks: low-dose CT reconstruction, CT FOV extension, and PCCT super-resolution, highlights its versatility and effectiveness in various medical imaging applications. Its ability to preserve structural integrity and fine details while maintaining computational efficiency makes it a promising tool for improving the quality of medical images across different modalities and tasks.

4.4 Ablation Study

An ablation study was conducted to analyze the effect of varying cutoff frequencies, denoted as κ , within the Gaussian high-pass filter [refer to Eq. (7)] on the quality of reconstructed images. The TF-FBP model was employed for this analysis due to its compact parameter space, where only

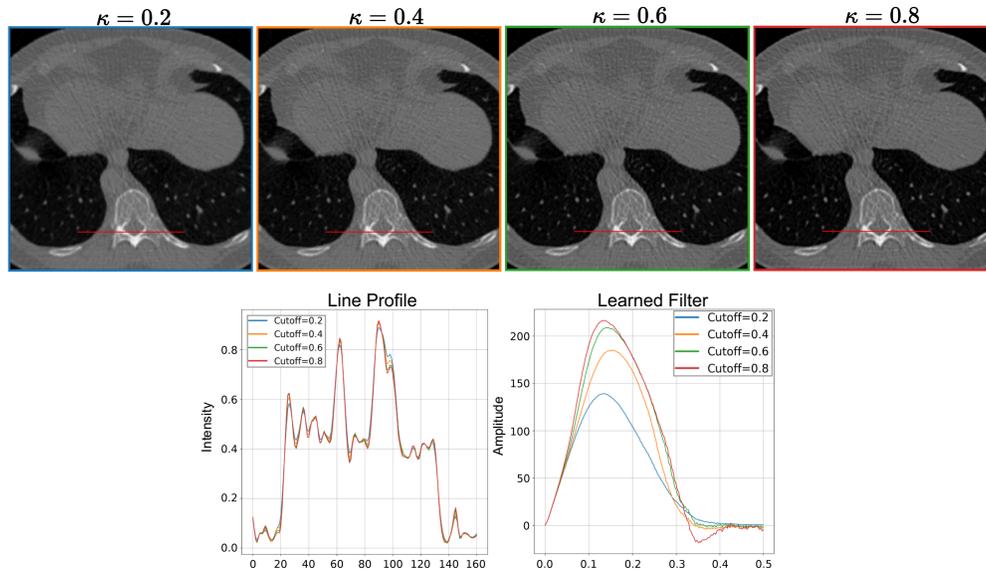


Fig. 9 Illustration of filters, line profiles, and reconstructed images for different κ values in the TF-FBP model. The red lines on the CT images indicate the locations of the line profiles. This figure demonstrates how the cutoff frequency directly affects the enhancement of high-frequency details during image reconstruction.

255 parameters define the outcome. In TF-FBP, the quality of the reconstructed images is directly linked to the characteristics of the learned filter.

Figure 9 illustrates the resultant filters and corresponding reconstructed images for different values of κ . A notable observation from this study is that higher values of κ enhance the sharpness of the reconstructed image while retaining more of the high-frequency noise. This demonstrates the direct relationship between the cutoff frequency and the enhancement of high-frequency components during image reconstruction, providing insights into the mechanism by which Eagle-Loss achieves its superior performance in preserving fine structural details.

We also investigated the impact of the weighting coefficient λ in our hybrid loss function [Eq. (9)] using the TF-FBP model with a fixed patch size of $n = 3$. Figure 10 shows visual results, whereas Table 4 provides quantitative analysis.

Figure 10 demonstrates improved edge definition and image sharpness as λ increases, with $\lambda = 2 \times 10^{-3}$ showing the sharpest image. Our analysis reveals that SSIM is the highest at $\lambda = 1 \times 10^{-3}$, but varies minimally across all values. The discrepancy between optimal λ for SSIM versus PSNR and LPIPS highlights the complexity of image quality assessment.

In our investigation of Eagle-Loss, our primary objective is to enhance image sharpness rather than merely optimizing evaluation metrics such as SSIM or PSNR. To explore the influence of patch size n on Eagle-Loss, we conducted experiments using the TF-FBP model, fixing λ at 8×10^{-4} [see Eq. (9)] while varying n from 3 to 15. Figure 11 illustrates that as n decreases, there are notable improvements in edge definition and fine detail preservation. This visual

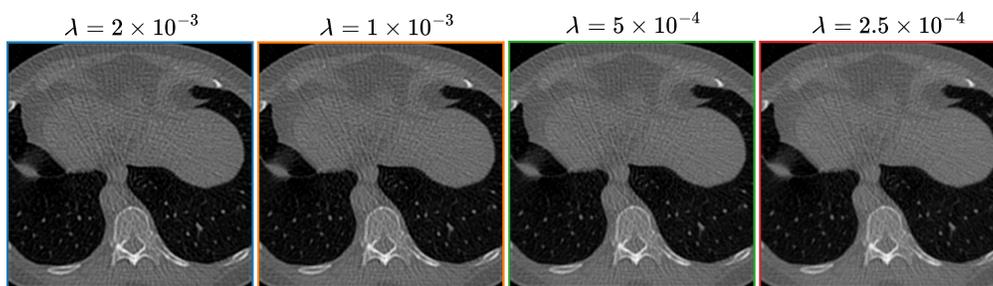
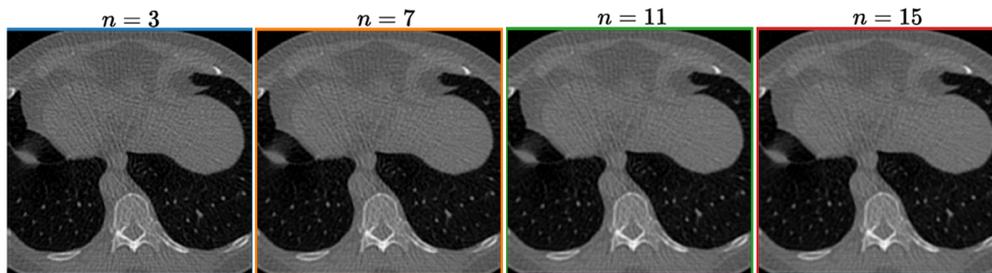


Fig. 10 Illustration of reconstructed images for different λ values in the TF-FBP model with a fixed patch size of $n = 3$ [refer to Eq. (2)].

Table 4 Performance of different λ values in SSIM, PSNR, and LPIPS. Here, the model is TFFBP with a fixed patch size of $n = 3$ [refer to Eq. (2)].

λ	SSIM \uparrow	PSNR (dB) \uparrow	LPIPS \downarrow
2×10^{-3}	0.944	32.804	0.248
1×10^{-3}	0.958	33.621	0.236
5×10^{-4}	0.953	33.613	0.247
2.5×10^{-4}	0.957	34.340	0.223

Note: Numbers shown in bold represent the best performance values within each matrix, making it easier for readers to identify the top results.

**Fig. 11** Illustration of reconstructed images for different n values in the TF-FBP model with a fixed $\lambda = 8 \times 10^{-4}$ [refer to Eq. (9)].**Table 5** Performance of different patch sizes n in SSIM, PSNR, and LPIPS. Here, the model is TF-FBP with a fixed $\lambda = 8 \times 10^{-4}$ [refer to Eq. (9)].

n	SSIM \uparrow	PSNR (dB) \uparrow	LPIPS \downarrow
3	0.950	33.454	0.243
7	0.956	34.050	0.227
11	0.958	34.332	0.219
15	0.960	34.506	0.214

Note: Numbers shown in bold represent the best performance values within each matrix, making it easier for readers to identify the top results.

enhancement aligns with our goal of improving image sharpness. Although Table 5 presents quantitative results showing that traditional metrics like SSIM and PSNR improve with larger patch sizes, these metrics do not fully capture the enhancements in image sharpness achieved with smaller patches. Our comprehensive ablation studies on key parameters—weighting parameter λ , cutoff frequency κ , and patch size n —revealed a consistent trend: image sharpness increased with larger λ values, higher cutoff frequencies, and smaller patch sizes. Specifically, larger λ values amplified the influence of Eagle-Loss, preserving more high-frequency details, higher cutoff frequencies allowed more high-frequency components to pass through the Gaussian high-pass filter, and smaller patch sizes enabled more localized feature analysis, effectively capturing fine structures. Although these parameter choices maximized sharpness, they did not always optimize traditional metrics, highlighting a trade-off between image clarity and metric-based quality. These findings underscore Eagle-Loss's flexibility and its potential to significantly enhance CT image reconstruction, particularly in scenarios that prioritize image sharpness and detail preservation over conventional quantitative metrics. They also emphasize the importance of careful parameter tuning based on specific application requirements, paving the way for future research into adaptive parameter selection methods and variant forms of Eagle-Loss tailored to different imaging modalities or reconstruction challenges.

5 Conclusion

In this paper, we propose Eagle-Loss, a new loss function specifically designed to improve the quality of reconstructed images by emphasizing high-frequency details crucial for accurate edge and texture representation. Eagle-Loss leverages localized feature analysis within gradient maps and applies frequency analysis to capture and preserve essential image features. Our experimental results demonstrate that Eagle-Loss effectively reduces image blur and enhances edge sharpness, leading to reconstructed images that exhibit superior fidelity to the ground truth. In addition, Eagle-Loss maintains computational efficiency in both speed and quality across all tasks. However, Eagle-Loss is sensitive to hyperparameter settings, and the optimal selection of these parameters varies across reconstruction tasks. Future efforts will focus on devising an automated hyperparameter optimization strategy to enhance the versatility and efficacy of Eagle-Loss in diverse imaging contexts. Our findings suggest that Eagle-Loss holds significant promise for CT image reconstruction and has the potential for broader applications in other fields.

Disclosures

The authors declare no conflicts of interest related to this work. No financial, personal, or professional affiliations influenced the research presented in this paper. All funding sources supporting this study are acknowledged within the paper. Any affiliations, financial involvement, or collaborations that could potentially influence the interpretation or outcome of this research have been disclosed. The development of Eagle-Loss was conducted with strict adherence to ethical guidelines and standards to ensure unbiased and objective results.

Code and Data Availability

The code and data for this study are available at https://github.com/sypsyp97/Eagle_Loss.

Acknowledgments

This research was financed by the “Verbundprojekt 05D2022 - KI4D4E: Ein KI-basiertes Framework für die Visualisierung und Auswertung der massiven Datenmengen der 4D-Tomographie für Endanwender von Beamlines. Teilprojekt 5” (Grant No. 05D23WE1). Part of this work was carried out at the Research Campus M2OLIE and funded by the German Federal Ministry of Education and Research (BMBF) within the framework “Forschungscampus: Public–Private Partnership for Innovations” under the funding code 13GW0388A. We also acknowledge the use of Claude 3.5 Sonnet for language and grammar clean-up of this paper.

References

1. W. A. Kalender, *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications*, John Wiley & Sons (2011).
2. T. M. Buzug, “Computed tomography,” in *Springer Handbook of Medical Technology*, R. Kramme, K.-P. Hoffmann, and R. S. Pozos, Eds., pp. 311–342, Springer Berlin Heidelberg, Berlin, Heidelberg (2011).
3. L. Wang, J. A. Sahel, and S. Pi, “Sub2Full: split spectrum to boost optical coherence tomography despeckling without clean data,” in *Opt. Lett.* **49**(11), 3062–3065 (2024).
4. Y. Chen et al., “Self-supervised neuron segmentation with multi-agent reinforcement learning,” in *Proc. Thirty-Second Int. Joint Conf. Artif. Intell.*, pp. 609–617 (2023).
5. G. Wang, J. C. Ye, and B. De Man, “Deep learning for tomographic image reconstruction,” *Nat. Mach. Intell.* **2**(12), 737–748 (2020).
6. B. Yu et al., “Ea-GANs: edge-aware generative adversarial networks for cross-modality MR image synthesis,” *IEEE Trans. Med. Imaging* **38**(7), 1750–1762 (2019).
7. W. Weimin et al., “Enhancing liver segmentation: a deep learning approach with eas feature extraction and multi-scale fusion,” *Int. J. Innov. Res. Comput. Sci. Technol.* **12**(1), 26–34 (2024).
8. J. Chi et al., “Low-dose CT image super-resolution network with dual-guidance feature distillation and dual-path content communication,” *Lect. Notes Comput. Sci.* **14229**, 98–108 (2023).
9. J. Pan et al., “Motion-compensated MR CINE reconstruction with reconstruction-driven motion estimation,” *IEEE Trans. Med. Imaging* **43**(7), 2420–2433 (2024).
10. Z. A. Balogh, Z. Barna, and E. Majoros, “Comparison of iterative reconstruction implementations for multi-slice helical CT,” *Zeitschr. Med. Phys.* (2024).

11. A. Maier et al., “A gentle introduction to deep learning in medical image processing,” *Zeitschr. Med. Phys.* **29**(2), 86–101 (2019).
12. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature* **521**(7553), 436–444 (2015).
13. Z. Wang et al., “Image quality assessment: from error visibility to structural similarity,” *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
14. L. Zhang et al., “A comprehensive evaluation of full reference image quality assessment algorithms,” in *19th IEEE Int. Conf. Image Process.*, IEEE, pp. 1477–1480 (2012).
15. H. Zhao et al., “Loss functions for image restoration with neural networks,” *IEEE Trans. Comput. Imaging* **3**(1), 47–57 (2016).
16. Y. Song et al., “Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images,” *IEEE/ACM Trans. Comput. Biol. Bioinf.* **18**(6), 2775–2780 (2021).
17. B. Stimpel et al., “Projection-to-projection translation for hybrid X-ray and magnetic resonance imaging,” *Sci. Rep.* **9**(1), 18814 (2019).
18. L. Abrahamyan et al., “Gradient variance loss for structure-enhanced image super-resolution,” in *ICASSP 2022-2022 IEEE Int. Conf. Acoust. Speech and Signal Process. (ICASSP)*, IEEE, pp. 3219–3223 (2022).
19. J. Johnson, A. Alahi, and L. Fei-Fei, “Perceptual losses for real-time style transfer and super-resolution,” *Lect. Notes Comput. Sci.* **9906**, 694–711 (2016).
20. G. Seif and D. Androustos, “Edge-based loss function for single image super-resolution,” in *IEEE Int. Conf. Acoust. Speech and Signal Process. (ICASSP)*, IEEE, pp. 1468–1472 (2018).
21. L. Jiang et al., “Focal frequency loss for image reconstruction and synthesis,” in *Proc. IEEE/CVF Int. Conf. Comput. Vision*, pp. 13919–13929 (2021).
22. L. Ge and L. Dou, “G-Loss: a loss function with gradient information for super-resolution,” *Optik* **280**, 170750 (2023).
23. C. Ledig et al., “Photo-realistic single image super-resolution using a generative adversarial network,” in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 4681–4690 (2017).
24. B. Benjdირaa, A. M. Alia, and A. Koubaa, “Guided frequency loss for image restoration,” arXiv:2309.15563 (2023).
25. Z. Fu et al., “Edge-aware deep image deblurring,” *Neurocomputing* **502**, 37–47 (2022).
26. Z. Wang et al., “Depth upsampling based on deep edge-aware learning,” *Pattern Recognit.* **103**, 107274 (2020).
27. S. Ramanathan and M. Ramasundaram, “Low-dose CT image reconstruction using vector quantized convolutional autoencoder with perceptual loss,” *Sādhanā* **48**, 43 (2023).
28. C. Ma et al., “Structure-preserving super resolution with gradient guidance,” in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, pp. 7769–7778 (2020).
29. M. Han, H. Shim, and J. Baek, “Perceptual CT loss: implementing ct image specific perceptual loss for cnn-based low-dose ct denoiser,” *IEEE Access* **10**, 62412–62422 (2022).
30. F. Piri, N. Karimi, and S. Samavi, “Enhanced segmentation in abdominal CT images: leveraging hybrid CNN-transformer architectures and compound loss function,” in *IEEE World AI IoT Congr. (AIoT)*, pp. 363–369 (2024).
31. V. K. Tanwar, V. Anand, and B. Sharma, “A deep-learning model based on channel and spatial attention for cervical fractures prediction using CT images,” in *5th Int. Conf. Emerg. Technol. (INCET)*, pp. 1–5 (2024).
32. J. Si and S. Kim, “Defect detection of wood grain images with image gradient applied to ae-based generation,” in *IEEE 6th Int. Conf. Knowl. Innov. and Invent. (ICKII)*, pp. 335–337 (2023).
33. Y. Sun et al., “Data-driven filter design in FBP: transforming CT reconstruction with trainable Fourier series,” arXiv:2401.16039 (2024).
34. Q. Zhang et al., “Wavefront coding image reconstruction via physical prior and frequency attention,” *Opt. Express* **31**(20), 32875–32886 (2023).
35. Z. Li et al., “Low-dose CT image denoising with improving wgan and hybrid loss function,” *Comput. Math. Methods Med.* **2021**(1), 2973108 (2021).
36. Y. Kim and H. Kudo, “Nonlocal total variation using the first and second order derivatives and its application to ct image reconstruction,” *Sensors* **20**(12), 3494 (2020).
37. H. Scharr, “Optimal operators in digital image processing dissertation,” (2000).
38. A. Dogra and P. Bhalla, “Image sharpening by Gaussian and butterworth high pass filter,” *Biomed. Pharmacol. J.* **7**(2), 707–713 (2014).
39. M.-A. Parseval, “Mémoire sur les séries et sur l’intégration complète d’une équation aux différences partielles linéaires du second ordre, à coefficients constants,” *Mém. prés. par divers savants, Acad. des Sci. Paris (1)* **1**, 638–648 (1806).
40. J. Leuschner et al., “Lodopab-CT, a benchmark dataset for low-dose computed tomography reconstruction,” *Sci. Data* **8**(1), 109 (2021).
41. M. Kistler et al., “The virtual skeleton database: an open access repository for biomedical research and collaboration,” *J. Med. Internet Res.* **15**(11), e245 (2013).

42. Y. Huang et al., "Data extrapolation from learned prior images for truncation correction in computed tomography," *IEEE Trans. Med. Imaging* **40**(11), 3042–3053 (2021).
43. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).
44. L. A. Gatys, A. S. Ecker, and M. Bethge, "Image style transfer using convolutional neural networks," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 2414–2423 (2016).
45. R. Zhang et al., "The unreasonable effectiveness of deep features as a perceptual metric," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 586–595 (2018).
46. H. Chen et al., "Low-dose CT with a residual encoder-decoder convolutional neural network," *IEEE Trans. Med. Imaging* **36**(12), 2524–2535 (2017).
47. C. Syben et al., "PYRO-NN: python reconstruction operators in neural networks," *Med. Phys.* **46**(11), 5110–5115 (2019).
48. G. Gouvine, "Super-resolution networks for PyTorch," 2021, <https://github.com/Coloquinte/torchSR>.
49. G. Gouvine, "NinaSR: efficient small and large ConvNets for super-resolution," 2021, <https://github.com/Coloquinte/torchSR/blob/main/doc/NinaSR.md>.
50. A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *20th Int. Conf. Pattern Recognit.*, IEEE, pp. 2366–2369 (2010).
51. I. Goodfellow, Y. Bengio, and A. Courville, "Deep feedforward networks," in *Deep Learn.*, MIT press, Cambridge, Massachusetts (2016).
52. J. Deng et al., "ImageNet: a large-scale hierarchical image database," in *IEEE Conf. Comput. Vision and Pattern Recognit.*, IEEE, pp. 248–255 (2009).
53. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556 (2014).
54. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Adv. Neural Inf. Process. Syst.* **25** (2012).
55. H. Hsu and P. A. Lachenbruch, "Paired *t* test," Wiley StatsRef: Statistics Reference Online (2014).

Yipeng Sun is a PhD student at the Pattern Recognition Lab, University of Erlangen-Nuremberg, focusing on deep learning techniques for image reconstruction. He began his PhD in July 2023, following an MSc in medical engineering from FAU (2020–2023) and his BEng in measurement and control technology and instruments from Nanjing University of Science and Technology, China (2015–2019).

Yixing Huang received his BE degree in 2013 from the Department of Biomedical Engineering, Peking University, China. He received his MS degree and PhD in 2016 and 2020, respectively, from Pattern Recognition Lab, Department of Computer Science, University of Erlangen-Nuremberg. In April 2021, he became a leading computer science researcher with the Department of Radiation Oncology, University Hospital Erlangen, Germany. Since September 2024, he has become an assistant professor at the Institute of Medical Technology, Peking University, China. His current research interests include machine learning in medical imaging and radiation oncology.

Zeyu Yang is a PhD student at the Computer-Assisted Clinical Medicine, Medical Faculty Mannheim, Heidelberg University, focusing on deep learning techniques for photon counting CT reconstruction and postprocessing. He began his PhD in September 2022, following his MSc in electrical engineering from KIT (2019–2022) and his BEng in telecommunication technology from Beijing University of Post- and Telecommunications (2014–2018).

Linda-Sophie Schneider is a PhD student at the Pattern Recognition Lab, University of Erlangen-Nuremberg, focusing on deep learning and optimization-based techniques for CT trajectory optimization and CT reconstruction. She began her PhD in August 2021, following her MSc in applied mathematics from FAU (2018–2021). Her current research interests include machine learning in CT imaging and combinatorial optimization.

Siyuan Mei is a PhD student at the Pattern Recognition Lab, University of Erlangen-Nuremberg, focusing on deep learning techniques for image reconstruction. He began his PhD in September 2023, following his MSc in medical engineering from FAU (2020–2023) and his BEng in automation from Zhejiang University of Science and Technology (2016–2020).

Andreas Maier received his bachelor's degree and PhD in computer science from the University of Erlangen-Nuremberg, Erlangen, in 2005 and 2009, respectively. From 2005 to 2009, he

was with the Pattern Recognition Laboratory, Computer Science Department, University of Erlangen-Nuremberg. From 2009 to 2010, he worked on flat-panel C-arm CT as a postdoctoral fellow with the Radiological Sciences Laboratory, Department of Radiology, Stanford University, Stanford, CA, United States. From 2011 to 2012, he joined Siemens Healthcare, Erlangen, Germany, as an Innovation Project Manager. In 2012, he returned to the University of Erlangen-Nuremberg as the Leader of the Medical Reconstruction Group, Pattern Recognition Laboratory, where he became a professor and the head of the Pattern Recognition Laboratory, in 2015.

Biographies of the other authors are not available.