Diffusion Models with Awareness of Binary Signal Detection Tasks for Medical Image Denoising

Yaozhi Zhang[a], Wentao Chen[b], Weimin Zhou[c,d], and Yang Liu[a]

[a]Global Institute of Future Technology, Shanghai Jiao Tong University, Shanghai, 200240, China
[b]Global College, Shanghai Jiao Tong University, Shanghai, 200240, China
[c]Wyant College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA
[d]Department of Radiology & Imaging Sciences, University of Arizona, Tucson, AZ 85721, USA

ABSTRACT

It is widely recognized that the evaluation and optimization of medical imaging system performance should be guided by task-based measures of image quality (IQ). We propose a task-aware diffusion model (TADM), a sampling-based method that incorporates channelized model observers, to denoise medical images for binary signal detection tasks. Our study employs Partial Least Squares (PLS) channels and uses singular value decomposition (SVD) to decompose the image into task-dense and task-sparse components. During the denoising process, TADMs simultaneously reduce noise in both task-dense and task-sparse components via corresponding posterior sampling. It is demonstrated that the proposed TADMs outperform Denoising Diffusion Restoration Models (DDRMs) in both physical IQ metrics (e.g. PSNR and FID) and task-based IQ metrics across a variety of image noise levels and datasets.

Keywords: Image denoising, diffusion models, task-based measures of image quality, ideal observer, partial least squares

1. INTRODUCTION

It has been widely recognized that medical imaging systems should be assessed and optimized using task-based image quality (IQ) measures.[1,2] Task-based measures of IQ quantify the ability of an observer to perform specific tasks such as signal detection tasks. For binary signal classification tasks, the Bayesian ideal observer (IO) that models human visual system provides an upper bound of the AUC, measuring the classifier's ability to distinguish between signal-present and signal-absent cases (e.g., tumor detection task).

For medical image denoising problems, researchers have investigated various ways to establish task-based denoising methods to preserve task-specific information while denoising. By adjusting the network architecture and loss function of denoising autoencoders (DAEs),[3–5] models could learn task-specific information in the noisy images, ensuring that the denoised images maintain a similar AUC to that of the noisy input. However, DAEs rely on learning a direct mapping from noisy to clean images, which may not generalize well to the complex noise characteristics and structural variations present in real clinical data.

To solve this problem, denoising diffusion models (DDMs) such as Denoising Diffusion Restoration Models (DDRMs)[6] and Denoising Diffusion Null-space Models (DDNMs),[7] both based on pre-trained Denoising Diffusion Probabilistic Models (DDPMs),[8] performed well on traditional IQ metrics (e.g., PSNR and FID)[9] by progressively refining noisy inputs through an iterative denoising process. However, to meet the noise variance requirement of each step, these DDMs introduced additional noise during the sampling process to minimize the objective error to the pre-trained model[6] to improve the traditional IQ, which led to task-specific information loss and a degradation in task-based IQ.

To address the task-specific information loss caused by additional noise of DDMs, we propose task-aware diffusion models (TADMs), a task-based optimization and unsupervised posterior sampling method for binary

Further author information: (Send correspondence to Yang Liu.)
Yang Liu: E-mail: yang.liu1@sjtu.edu.cn

signal detection tasks based on pre-trained DDPMs. Specifically, TADMs introduce Partial Least Squares (PLS) channels,[10] which are widely used to extract features from images for signal detection tasks, and decompose the task vector, composed of the image vector and its channelized vector, into task-dense and task-sparse components by singular value decomposition (SVD). During the sampling process, our method ensures that less additional noise is introduced into the task-dense component while meeting the required noise variance of each step by adding the transformation form of the original noisy image vector. For task-sparse component, we add random noise to make the evidence lower bound (ELBO) objective of TADMs in the form of that of the DDPMs/DDIMs.[11] By incorporating task-based optimization method into the sampling process, TADMs enable high-quality image restorations with reduced task-specific information loss, making it well-suited for real medical images. Computer simulation studies on binary signal detection tasks involving a lumpy dataset and a multivariate normal lumpy background (MVNLumpy) dataset are conducted. For these two datasets, our method achieves higher AUC when noise standard deviation $\sigma_{\mathbf{g}} \geq 0.02$. On lumpy and MVNLumpy examples, our method outperforms DDRMs in PSNR and FID when $\sigma_{\mathbf{g}} \geq 0.02$. Experimental results demonstrated that the outputs of TADMs have better traditional and task-based IQ in most cases when noise standard deviation $\sigma_{\mathbf{g}} \geq 0.02$.

2. METHOD

A binary signal detection task is considered and the imaging processes under the signal-absent hypothesis (H_0) and the signal-present hypothesis (H_1) can be described as:

$$\begin{aligned} H_0 &: \mathbf{g} = \mathbf{b} + \mathbf{n} \equiv \mathbf{g}_0 + \mathbf{n}, \\ H_1 &: \mathbf{g} = \mathbf{b} + \mathbf{s} + \mathbf{n} \equiv \mathbf{g}_0 + \mathbf{n}. \end{aligned} \tag{1}$$

where $\mathbf{g} \in \mathbb{R}^N$ denotes the noisy image data, $\mathbf{b} \in \mathbb{R}^N$ denotes the image of background, $\mathbf{s} \in \mathbb{R}^N$ denotes the image of the to-be-detected signal, and $\mathbf{n} \in \mathbb{R}^N$ denotes the corrupting noise. To perform the considered signal detection task, an observer computes a real-valued test statistic $t(\mathbf{g})$ and calculates the AUC through $t(\mathbf{g})$.

2.1 Denoising Diffusion Probabilistic Models

Denoising Diffusion Probabilistic Models (DDPMs) provide a powerful approach to generate images from a random noise image $\mathbf{x}_T$. DDPMs use neural networks (typically U-Nets[12]) to predict the noise of step t, and for Gaussian noise, the image of step $t-1$ is:

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}} \epsilon_\theta(\mathbf{x}_t, t) \right) + \sigma_t \mathbf{z}, \tag{2}$$

where $\mathbf{x}_t$ and $\mathbf{x}_{t-1}$ are noisy images of step t and $t-1$, $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$, σ_t is the standard deviation of noise in step $t-1$, ϵ_θ is the predicted noise given by a pre-trained neural network, α_t and $\bar{\alpha}_t$ are given by hyper-parameters. By applying Eq. (2) from $t = T$ to $t = 1$, DDPMs finally generate noise-free output $\mathbf{x}_0$.

During the training process, noise sampled from the same probability distribution with that in medical images (e.g., Gaussian) is added to the noise-free medical images. In this way, DDPMs can be applied to the predictions of real noisy medical images.

2.2 Task-aware Diffusion Models

In this work, we propose a novel approach that integrates pre-trained DDPMs with task-based IQ optimization method to medical image denoising. Fig. 1 presents the overview of TADMs, comprising of three novel modules: task-based projection, task-based sampling and no-objective-error sampling. Noisy image $\mathbf{g}$ is projected onto the spectral space, where the vector is passed through task-based sampling module and no-objective-error sampling module to get the task-dense component and task-sparse component, and the vector combined of these two components is projected back to reconstruct the noise-free image.

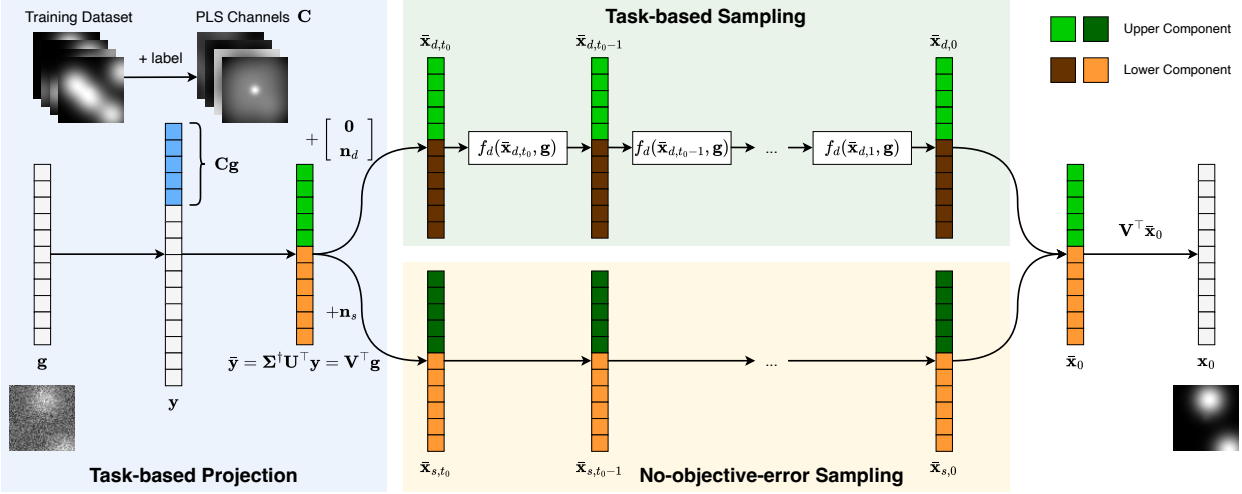


Figure 1. Process of task-aware diffusion models (TADMs). The image vector in spectral space $\bar{\mathbf{y}} = \mathbf{\Sigma}^\dagger \mathbf{U}^\top \mathbf{y} = \mathbf{V}^\top \mathbf{g}$ is passed to task-based sampling module and no-objective-error sampling module to recover task-dense component and task-sparse component of noise-free image in spectral space $\bar{\mathbf{x}}_0$. During the sampling process, the pre-trained model in the image space predicts the noise based on the noisy image reconstructed from the spectral space, i.e., left multiply vector in spectral space by $\mathbf{V}$, and the predicted noise is then used to perform denoising in the spectral space.

2.2.1 Task-based Projection

Various efficient channels are developed to extract features from images for signal detection tasks (e.g. PLS channels). For noisy image $\mathbf{g} = \mathbf{g}_0 + \mathbf{n}$, we use PLS channels $\mathbf{C} = \begin{bmatrix} \mathbf{C}_1^\top \\ \mathbf{C}_2^\top \\ \dots \\ \mathbf{C}_M^\top \end{bmatrix} \in \mathbb{R}^{M \times N}$ to extract task features from image $\mathbf{g}$. We define the task vector $\mathbf{y}$ composed of image and its channelized vector:

$$\mathbf{y} \equiv \begin{bmatrix} \mathbf{Cg} \\ \mathbf{g} \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{E} \end{bmatrix} \mathbf{g} \equiv \mathbf{Ag}, \tag{3}$$

where $\mathbf{y} \in \mathbb{R}^{M+N}$, and $\mathbf{A} \in \mathbb{R}^{(M+N) \times N}$ is the task matrix which extracts task features and also records full information.

Similar to DDRMs, we use the singular value decomposition (SVD) of $\mathbf{A}$ and perform the diffusion in the spectral space. Regardless of the type or number of channels used, applying SVD ensures that the transformed image vectors in the spectral space with noise that retains identical variances and exhibits zero covariance, i.e., the transformed components are uncorrelated. Moreover, SVD allows the transformation of the image vector into a spectral space where the contribution of each spectral component to the task-specific information is quantified by its associated singular value. For task matrix $\mathbf{A}$, its SVD is given as:

$$\mathbf{A} = \mathbf{U\Sigma V}^\top, \tag{4}$$

where $\mathbf{U} \in \mathbb{R}^{(M+N) \times (M+N)}$ and $\mathbf{V} \in \mathbb{R}^{N \times N}$ are orthogonal matrices, $\mathbf{\Sigma} \in \mathbb{R}^{(M+N) \times N}$ is a rectangular diagonal matrix containing N singular values of $\mathbf{A}$.

For task matrix $\mathbf{A}$ containing PLS channels $\mathbf{C}$, the singular values can be analytically calculated. Row vectors of PLS channels matrix $\mathbf{C}$ constitute an orthonormal basis so the first M singular values are:

$$s_i = \sqrt{\lambda_i(\mathbf{A}^\top \mathbf{A})} = \sqrt{\lambda_i(\mathbf{C}^\top \mathbf{C} + \mathbf{E}^\top \mathbf{E})} = \sqrt{\lambda_i(\mathbf{CC}^\top + \mathbf{E})} = \sqrt{\lambda_i(2\mathbf{E})} = \sqrt{2}, \quad i \in [1, M] \tag{5}$$

and the remaining $N - M$ singular values are:

$$s_i = \sqrt{\lambda_i(\mathbf{E}^\top \mathbf{E})} = \sqrt{\lambda_i(\mathbf{E})} = 1, \quad i \in [M+1, N] \tag{6}$$

where λ_i is the eigenvalue. Unitary matrix $\mathbf{V}^\top$ can be decomposed into $\mathbf{V}_u$ and $\mathbf{V}_l$ according to different singular values:

$$\mathbf{V}^\top = \begin{bmatrix} \mathbf{V}_u \\ \mathbf{V}_l \end{bmatrix}, \tag{7}$$

where $\mathbf{V}_u \in \mathbb{R}^{M \times N}$ corresponding to M singular values of $\sqrt{2}$ extracts denser task-specific information and $\mathbf{V}_l \in \mathbb{R}^{(N-M) \times N}$ corresponding to $N - M$ singular values of 1 extracts sparser task-specific information.

The task vector in spectral space is:

$$\bar{\mathbf{y}} = \mathbf{\Sigma}^\dagger \mathbf{U}^\top \mathbf{y} = \mathbf{V}^\top \mathbf{g}. \tag{8}$$

Vectors in spectral space can be converted back to images by left multiplying $\mathbf{V}$:

$$\mathbf{g} = \mathbf{V}\bar{\mathbf{y}}. \tag{9}$$

For Gaussian noise $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_{\mathbf{g}}^2 \mathbf{I})$, the noise variance of $\bar{\mathbf{y}}$ is also $\sigma_{\mathbf{g}}^2$ because $\mathbf{V}^\top$ is a unitary matrix. In the following modules, vectors in spectral space are passed to denoise step by step.

2.2.2 Task-based Sampling

By adding noise to match the variance of a specific diffusion timestep, the image can be interpreted as an intermediate input within the sampling trajectory of pre-trained DDPMs. During the sampling process of DDPMs, the image $\mathbf{x}_t$ of step t has noise variance of σ_t^2 and $\bar{\mathbf{x}}_t$ has the same noise variance. We can find step t_0 whose noise variance is the closest to that of $\bar{\mathbf{y}}$ and t_0 meets:

$$\frac{\sqrt{1 - \bar{\alpha}_{t_0 - 1}}}{\sqrt{\bar{\alpha}_{t_0 - 1}}} < \sigma_{\mathbf{g}} \leq \frac{\sqrt{1 - \bar{\alpha}_{t_0}}}{\sqrt{\bar{\alpha}_{t_0}}}, \tag{10}$$

and the standard deviation difference $\Delta\sigma$ is defined as:

$$\Delta\sigma = \sqrt{\sigma_{t_0}^2 - \sigma_{\mathbf{g}}^2}. \tag{11}$$

To avoid introducing additional noise to vector $\mathbf{x}_t$ that causes task-specific information loss, we use the transformation form of $\bar{\mathbf{y}}$ and $\mathbf{x}_{t+1}$ to replace random noise. However, the replacement of random noise leads a weighted sum-of-squares error difference of ELBO objective of TADMs to DDPMs objective.[6]

In order to preserve task-specific information, we use the replacement of random noise for each step in task-based sampling module to get task-dense component, which is the upper component of output vector from task-based sampling module in spectral space; to reduce ELBO objective error, we still add random noise for each step in no-objective-error sampling module to get task-sparse component, the lower component of output vector from no-objective-error sampling module in spectral space. Finally, task-dense component and task-sparse component are concatenated in spectral space and projected back to image space as the final output.

To get task-dense component, we conduct task-based sampling process from step t_0 in spectral space:

$$\bar{\mathbf{x}}_{d,t_0} = \bar{\mathbf{y}} + \begin{bmatrix} \mathbf{0} \\ \mathbf{n}_d \end{bmatrix} = \mathbf{V}^\top \mathbf{g} + \begin{bmatrix} \mathbf{0} \\ \mathbf{n}_d \end{bmatrix}, \tag{12}$$

$$\bar{\mathbf{x}}_{d,t} = f_d(\bar{\mathbf{x}}_{d,t+1}, \mathbf{g}) = \bar{\mathbf{x}}_{d,\theta,t} + \sigma_t \frac{\bar{\mathbf{y}} - \bar{\mathbf{x}}_{d,\theta,t}}{\sigma_{\mathbf{g}}} = \begin{bmatrix} \bar{\mathbf{x}}_{d,t,u} \\ \bar{\mathbf{x}}_{d,t,l} \end{bmatrix}, \quad t \in [0, t_0 - 1] \tag{13}$$

where $\mathbf{n}_d \sim \mathcal{N}(\mathbf{0}, \Delta\sigma \mathbf{I}_{N-M})$, $\bar{\mathbf{y}} = \mathbf{V}^\top \mathbf{g}$, $\bar{\mathbf{x}}_{d,t,u} \in \mathbb{R}^M$ is the upper component of $\bar{\mathbf{x}}_{d,t}$ representing the multiplying result of rows corresponding to singular values of $\sqrt{2}$ in $\mathbf{V}^\top$ and $\mathbf{x}_{d,t}$, $\bar{\mathbf{x}}_{d,t,l} \in \mathbb{R}^{N-M}$ is the lower component of $\bar{\mathbf{x}}_{d,t}$ representing the multiplying result of rows corresponding to singular values of 1 in $\mathbf{V}^\top$ and $\mathbf{x}_{d,t}$, the

predicted noise-free image $\bar{\mathbf{x}}_{d,\theta,t} = \mathbf{V}^\top \frac{\mathbf{V}\bar{\mathbf{x}}_{d,t+1} - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(\mathbf{V}\bar{\mathbf{x}}_{d,t+1}, t+1)}{\sqrt{\bar{\alpha}_t}}$ is in the spectral space. Eq. (12) is the boundary equation that convert $\bar{\mathbf{y}}$ into $\bar{\mathbf{x}}_{d,t_0}$ whose standard deviation σ_{t_0} is close to that of step t_0 of the pre-trained DDPMs and Eq. (13) samples image in spectral space from $t = t_0 - 1$ to step $t = 0$. $\frac{\bar{\mathbf{y}} - \bar{\mathbf{x}}_{d,\theta,t}}{\sigma_{\mathbf{g}}}$ provides a fitted standard normal distribution thus $\bar{\mathbf{x}}_{d,t}$ has a fitted normal noise whose variance is σ_t^2 that meets the requirement of step t. The output of task-based sampling module is $\bar{\mathbf{x}}_{d,0} = \begin{bmatrix} \bar{\mathbf{x}}_{d,0,u} \\ \bar{\mathbf{x}}_{d,0,l} \end{bmatrix}$.

The output of task-based sampling module has less task-specific information loss and better task-based IQ, meanwhile causing a degradation in the ELBO objective consistency with the pretrained model, which may lead to a degradation in traditional IQ, thus we reserve the upper component of this module as the task-dense component and propose the no-objective-error sampling module to get the task-sparse component.

2.2.3 No-objective-error Sampling

To get task-sparse component, we conduct no-objective-error sampling process from step t_0:

$$\bar{\mathbf{x}}_{s,t_0} = \bar{\mathbf{y}} + \mathbf{n}_s = \mathbf{V}^\top \mathbf{g} + \mathbf{n}_s, \tag{14}$$

$$\bar{\mathbf{x}}_{s,t} = \bar{\mathbf{x}}_{s,\theta,t} + \sigma_t \mathbf{I}_n = \begin{bmatrix} \bar{\mathbf{x}}_{s,t,u} \\ \bar{\mathbf{x}}_{s,t,l} \end{bmatrix}, \quad t \in [0, t_0 - 1] \tag{15}$$

where $\mathbf{n}_s \sim \mathcal{N}(\mathbf{0}, \Delta\sigma \mathbf{I}_n)$, $\bar{\mathbf{x}}_{s,t,u} \in \mathbb{R}^M$ is the upper component of $\bar{\mathbf{x}}_{s,t}$ representing the multiplying result of rows corresponding to singular values of $\sqrt{2}$ in $\mathbf{V}^\top$ and $\mathbf{x}_{s,t}$, $\bar{\mathbf{x}}_{s,t,l} \in \mathbb{R}^{N-M}$ is the lower component of $\bar{\mathbf{x}}_{s,t}$ representing the multiplying result of rows corresponding to singular values of 1 in $\mathbf{V}^\top$ and $\mathbf{x}_{s,t}$, the predicted noise-free image $\bar{\mathbf{x}}_{s,\theta,t} = \mathbf{V}^\top \frac{\mathbf{V}\bar{\mathbf{x}}_{s,t+1} - \sqrt{1-\bar{\alpha}_t}\epsilon_\theta(\mathbf{V}\bar{\mathbf{x}}_{s,t+1}, t+1)}{\sqrt{\bar{\alpha}_t}}$ is in spectral space. In this way, the sampling process is equivalent to that of DDRMs with no objective error.[6] Eq. (14) is the boundary equation that convert $\bar{\mathbf{y}}$ into $\bar{\mathbf{x}}_{s,t_0}$ whose standard deviation σ_{t_0} is that of step t_0 of the pre-trained DDPMs and Eq. (15) samples image in spectral space $\bar{\mathbf{x}}_{s,t}$ with standard normal distribution noise with variance of σ_t^2 from $t = t_0 - 1$ to step $t = 0$. The output of no-objective-error sampling module is $\bar{\mathbf{x}}_{s,0} = \begin{bmatrix} \bar{\mathbf{x}}_{s,0,u} \\ \bar{\mathbf{x}}_{s,0,l} \end{bmatrix}$.

The output of no-objective-error sampling module has better traditional IQ. However, the introducing of additional noise during the sampling process causes a degradation in task-based IQ, thus we reserve the lower component of this module as the task-sparse component.

The final output $\bar{\mathbf{x}}_0$ in spectral space composed of task-dense component and task-sparse component is:

$$\bar{\mathbf{x}}_0 = \begin{bmatrix} \bar{\mathbf{x}}_{d,0,u} \\ \bar{\mathbf{x}}_{s,0,l} \end{bmatrix}, \tag{16}$$

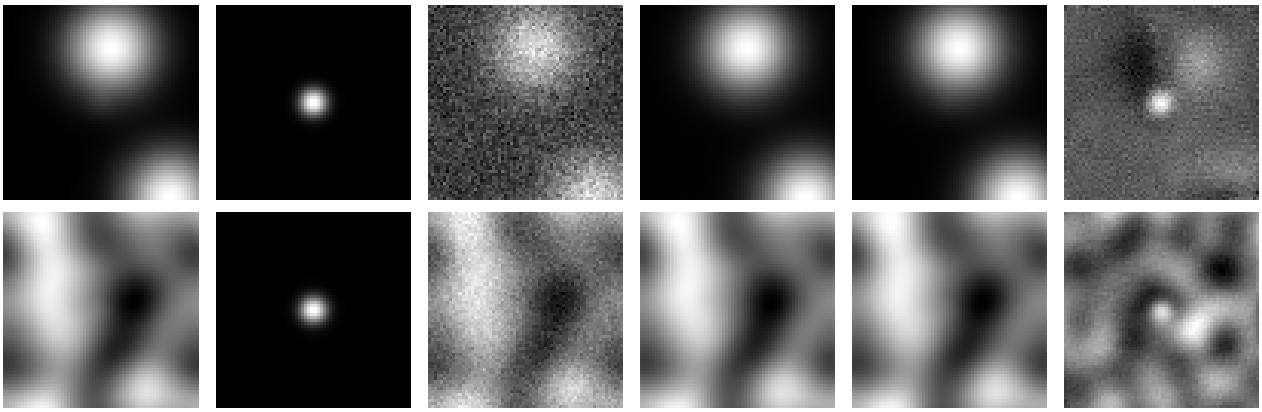
and noise-free image $\mathbf{x}_0$ is:

$$\mathbf{x}_0 = \mathbf{V}\bar{\mathbf{x}}_0. \tag{17}$$

In this way, we get noise-free image $\mathbf{x}_0$ by combining task-dense component and task-sparse component together. The task-dense component has less additional noise introduced during the sampling process so that there is lower task-specific information loss. The task-sparse has the same ELBO objective with DDPMs so that pre-trained DDPMs are good approximations to our method, leading to better traditional IQ at the same time.

3. NUMERICAL STUDIES

In this work, simulations on lumpy and MVNLumpy datasets were conducted to investigate the performance of DDRMs and our method. A fixed Gaussian signal was employed to images to simulate the tumor signal and the noise was sampled from Gaussian distribution with the mean of 0 and the standard deviation $\sigma_{\mathbf{g}}$ from 0.01 to 0.05. Examples of the background image, signal image, ground-truth signal-present image, and noisy signal-present image are shown in Fig. 2.



(a) (b) (c) (d) (e) (f)

Figure 2. First row: Lumpy dataset; second row: MVNLumpy dataset. The standard variance of noise is 0.02. (a) Original signal-present image. (b) The signal contained in image (a). (c) Noisy image. (d) Output of DDRMs. (e) Output of TADMs. (f) Subtracting (d) from (e). The difference of outputs between TADMs and DDRMs is mainly reflected in the vicinity of the signal position.

$\sigma_{\mathbf{g}}$	Lumpy			MVNLumpy		
	Input	DDRMs	TADMs	Input	DDRMs	TADMs
0.01	0.9488	**0.9311**	0.9302	0.9924	0.9888	**0.9889**
0.02	0.8613	0.7261	**0.8365**	0.9244	0.8244	**0.9171**
0.03	0.7763	0.6862	**0.7535**	0.8463	0.7896	**0.8427**
0.04	0.7203	0.6648	**0.7028**	0.8198	0.7847	**0.8120**
0.05	0.6817	0.6249	**0.6715**	0.7995	0.7440	**0.7845**

Table 1. AUC values of input, DDRMs output and TADMs output on lumpy and MVNLumpy datasets with different standard deviation $\sigma_{\mathbf{g}}$ of noise added to the input images.

The pre-trained DDPMs used self-attention U-Nets to predict the noise of each step during the sampling process. To train the U-Nets, 50,000 signal-present and 50,000 signal-absent images were used and the semi-online learning in which the image noise was generated and added on-the-fly for each dataset. The PLS channels were calculated using these noise-free images. After the U-Nets were trained, DDRMs and our method were applied to the test dataset, which contained 10,000 signal-present and 10,000 signal-absent images with noise with the mean of 0 and the variance from ranging from 0.01 to 0.05. After denoising through DDRMs and our method, the AUC of noisy input and output noise-free images calculated by the test statistics given by CNN observer[13] are shown in Table 1. The AUC value of input and output images was decreased when the standard deviation $\sigma_{\mathbf{g}}$ was increased. For the same input images, the AUC value of our method was closer to that of input, meaning that the task-specific information loss during our sampling process is lower.

Dataset	Method	$\sigma_{\mathbf{g}} = 0.01$		$\sigma_{\mathbf{g}} = 0.02$		$\sigma_{\mathbf{g}} = 0.03$		$\sigma_{\mathbf{g}} = 0.04$		$\sigma_{\mathbf{g}} = 0.05$	
		PSNR↑	FID↓	PSNR↑	FID↓	PSNR↑	FID↓	PSNR↑	FID↓	PSNR↑	FID↓
Lumpy	DDRMs	**62.92**	**2.248**	53.17	2.222	52.20	2.506	51.67	3.024	48.65	2.438
	TADMs	62.17	2.692	**57.61**	**2.208**	**54.72**	**2.196**	**52.63**	**2.348**	**50.97**	**2.200**
MVNLumpy	DDRMs	**56.57**	**1.692**	47.14	3.016	46.13	3.794	45.58	4.782	42.71	4.082
	TADMs	55.03	3.020	**50.92**	**2.952**	**48.18**	**3.154**	**46.15**	**3.308**	**44.54**	**3.382**

Table 2. PSNR and FID comparison on lumpy and MVNLumpy datasets with different standard deviation $\sigma_{\mathbf{g}}$ of noise added to the input images.

We also measured the performance of DDRMs and our method through traditional IQ assessment indicators

including Peak Signal-to-Noise Ratio (PSNR) and Fréchet Inception Distance (FID). The corresponding results are summarized in the Table 2. The higher PSNR scores along with lower FID values, indicate that our method achieved superior reconstruction accuracy both at the pixel level and in terms of perceptual quality, producing images that were both quantitatively precise and visually natural.

4. CONCLUSIONS

In this work, a novel approach for medical image denoising based on pre-trained DDPMs was proposed. By combining image and its channelized vector given by PLS channels and decomposing the combined vector to task-dense and task-sparse components by SVD, our method used the transformation form of the original noisy image to meet the noise requirement for task-dense component and random noise for task-sparse component during the sampling process to preserve task-specific information and keep the ELBO objective close to that of DDPMs/DDIMs. Through our method, the output noise-free images kept signal information close to that of original noisy images, leading to a high signal classification ability.

REFERENCES

[1] Barrett, H. H. and Myers, K. J., [*Foundations of image science*], John Wiley & Sons (2013).

[2] Chen, W., Xu, T., and Zhou, W., "Task-based regularization in penalized least squares for binary signal detection tasks in medical image denoising," in [*Medical Imaging 2025: Image Perception, Observer Performance, and Technology Assessment*], **13409**, 130–136, SPIE (2025).

[3] Vincent, P., Larochelle, H., Bengio, Y., and Manzagol, P.-A., "Extracting and composing robust features with denoising autoencoders," in [*Proceedings of the 25th international conference on Machine learning*], 1096–1103 (2008).

[4] Gondara, L., "Medical image denoising using convolutional denoising autoencoders," in [*2016 IEEE 16th international conference on data mining workshops (ICDMW)*], 241–246, IEEE (2016).

[5] Ongie, G., Sidky, E. Y., Reiser, I. S., and Pan, X., "Optimizing model observer performance in learning-based ct reconstruction," in [*Medical Imaging 2022: Image Perception, Observer Performance, and Technology Assessment*], **12035**, 55–59, SPIE (2022).

[6] Kawar, B., Elad, M., Ermon, S., and Song, J., "Denoising diffusion restoration models," *Advances in Neural Information Processing Systems* **35**, 23593–23606 (2022).

[7] Wang, Y., Yu, J., and Zhang, J., "Zero-shot image restoration using denoising diffusion null-space model," *arXiv preprint arXiv:2212.00490* (2022).

[8] Ho, J., Jain, A., and Abbeel, P., "Denoising diffusion probabilistic models," *Advances in neural information processing systems* **33**, 6840–6851 (2020).

[9] Wang, Z., Bovik, A. C., Sheikh, H. R., and Simoncelli, E. P., "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing* **13**(4), 600–612 (2004).

[10] Witten, J. M., Park, S., and Myers, K. J., "Partial least squares: a method to estimate efficient channels for the ideal observers," *IEEE Transactions on medical imaging* **29**(4), 1050–1058 (2010).

[11] Song, J., Meng, C., and Ermon, S., "Denoising diffusion implicit models," *arXiv preprint arXiv:2010.02502* (2020).

[12] Ronneberger, O., Fischer, P., and Brox, T., "U-net: Convolutional networks for biomedical image segmentation," in [*Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*], 234–241, Springer (2015).

[13] Zhou, W., Li, H., and Anastasio, M. A., "Approximating the ideal observer and hotelling observer for binary signal detection tasks by use of supervised learning methods," *IEEE transactions on medical imaging* **38**(10), 2456–2468 (2019).