

# Medical Image Denoising with Task-Based Image-Adaptive Denoising CNN

Wentao Chen<sup>a</sup> and Weimin Zhou<sup>b,c</sup>

<sup>a</sup>Global College, Shanghai Jiao Tong University, Shanghai, 200240, China

<sup>b</sup>Wyant College of Optical Sciences, University of Arizona, Tucson, AZ 85721, USA

<sup>c</sup>Department of Radiology & Imaging Sciences, University of Arizona, Tucson, AZ 85721, USA

## ABSTRACT

Image denoising is a fundamental inverse problem in medical imaging. Traditional sparsity-promoting approaches, such as penalized least-squares (PLS) with total variation (TV), have been widely employed to address such inverse problems. However, the handcrafted regularizers used in these methods cannot fully model complex image statistics, often leading to suboptimal image quality (IQ) performance. Supervised learning-based methods employing convolutional neural networks (CNNs) have emerged as powerful denoising approaches that can capture complex image statistics and improve IQ. However, CNNs trained with loss functions based solely on traditional IQ measures may remove subtle image details that are critical for task performance (e.g., tumor detection accuracy), resulting in degraded task-based IQ measures. In this work, we propose a novel task-based image-adaptive strategy to fine-tune the pre-trained CNN denoiser for each individual image by further promoting data fidelity while penalizing deviations in test statistics of a model observer for classification tasks. Numerical studies involving a stochastic binary texture model are conducted, and it is demonstrated that the proposed method achieves substantial improvements compared to previous medical image denoising methods in both traditional and task-based IQ metrics.

**Keywords:** Medical image denoising, observer model, task-based measures of image quality, denoising CNN

## 1. INTRODUCTION

Medical imaging systems inherently produce noisy measurements, making image denoising a fundamental inverse problem in medical imaging. Model-based approaches, such as penalized least-squares (PLS) methods incorporating sparsity-promoting penalties, have been widely employed to solve inverse problems in imaging. However, such handcrafted penalties are typically designed to capture generic object characteristics and may fail to account for subtle features that represent task-specific information relevant to clinically important objectives. As such, for signal detection tasks (e.g., tumor detection), the detectability of signals in the resulting denoised images may be degraded. Building upon the conventional PLS framework with total variation penalty (PLS-TV), Chen *et al.*<sup>1</sup> introduced a task-based PLS-TV method that incorporates a task-based regularization to enable joint noise reduction and signal detectability enhancement.

However, handcrafted sparsity-promoting penalties used in model-based methods are often insufficient to fully capture complex image statistics associated with the underlying object distribution. For example, TV regularization promotes piecewise-constant solutions and may introduce patchy artifacts, resulting in degraded IQ. Deep learning-based methods that can learn complex image statistic have been widely explored for addressing imaging-related challenges. The powerful representational capacity of convolutional neural networks (CNNs) has led to many successes in image processing applications, as evaluated by traditional IQ metrics. However, traditional IQ metrics are often not correlated with task performance. As such, training CNNs with the objective of maximizing traditional IQ metrics alone may inadvertently lead to degraded task-based IQ metrics.<sup>2</sup>

To address this challenge, recent studies have explored task-based approaches to train CNNs by utilizing model observers. For example, Han *et al.*<sup>3</sup> proposed an observer-based perceptual loss for training CNN denoisers;

---

Further author information: (Send correspondence to Weimin Zhou.)

Weimin Zhou: E-mail: weiminzhou@arizona.edu

Li *et al.*<sup>4</sup> adopted a fine-tuning strategy in which the last several convolutional layers were updated using a task-specific loss function. Although promising, in practice, the fixed network weight parameters learned during training may not be optimal for every input image. As such, an image-adaptive refinement strategy that adjusts the model to each specific image is desirable.

In this work, we introduce a novel task-based image-adaptive denoising strategy that leverages a pretrained CNN and adaptively tailors it to each individual input image. Specifically, given a pretrained CNN denoiser, the network weights are further refined for each image by optimizing a task-based loss function that promotes data fidelity while preserving task-specific information. Numerical studies are conducted on a denoising problem involving images simulated from a stochastic binary texture model. The results demonstrate that the proposed task-based image-adaptive CNN denoiser achieves substantial improvements over prior model-based and deep learning-based methods in both traditional and task-based IQ metrics.

## 2. METHODS

### 2.1 Binary signal detection tasks and model observers

In this study, we consider a binary signal detection task that requires an observer to classify image data  $\mathbf{g}$  as satisfying either a signal-absent hypothesis  $H_0$  or a signal-present hypothesis  $H_1$ . The imaging processes under the two hypotheses can be described as:<sup>5</sup>

$$\begin{aligned} H_0 : \mathbf{g} &= \mathcal{H}\mathbf{f}_b + \mathbf{n}, \\ H_1 : \mathbf{g} &= \mathcal{H}(\mathbf{f}_b + \mathbf{f}_s) + \mathbf{n}. \end{aligned} \tag{1}$$

where  $\mathcal{H}$  denotes an imaging operator that maps objects to image data,  $\mathbf{f}_b$  and  $\mathbf{f}_s$  represent the background and signal object functions, respectively,  $\mathbf{n} \in \mathbb{R}^N$  denotes the random measurement noise, and  $\mathbf{g} \in \mathbb{R}^N$  denotes the measured imaging data acquired by a digital imaging system.

To perform a binary signal detection task, an observer computes a test statistic  $t(\mathbf{g})$  that maps the measured imaging data  $\mathbf{g}$  to a real-valued scalar, which is subsequently compared against to a predefined threshold  $\tau$  to determine whether  $\mathbf{g}$  satisfies the hypothesis  $H_0$  or  $H_1$ . Linear observers, such as matched filters and the Hotelling observer, have been frequently used for task-based image quality assessment. The general form of a linear test statistic can be written as:

$$t_{\text{lin}}(\mathbf{g}) = \mathbf{w}_{\text{lin}}^T \mathbf{g}, \tag{2}$$

where  $\mathbf{w}_{\text{lin}} \in \mathbb{R}^N$  is a detection template. In this work, we employ a linear observer to form a task-specific regularization term in our proposed task-based image-adaptive CNN, which will be described in the next subsection.

### 2.2 Task-based image-adaptive CNN denoiser

Supervised learning-based denoising approaches that employ CNNs have demonstrated substantial potential for medical imaging applications. Given a noisy image  $\mathbf{g} \in \mathbb{R}^N$ , the denoised image  $\hat{\mathbf{f}} \in \mathbb{R}^N$  can be generally expressed as:

$$\hat{\mathbf{f}} = \mathcal{F}(\mathbf{g}; \theta), \tag{3}$$

where  $\mathcal{F}$  denotes an image-to-image mapping implemented by a deep neural network (DNN) parameterized by the weight vector  $\theta$ . Various DNN architectures have been utilized to represent  $\mathcal{F}$ , among which convolutional neural networks (CNNs) are particularly popular. However, the training strategies that are commonly employed in general image processing applications typically do not incorporate information related to specific clinically relevant tasks. Recent studies have demonstrated that supervised learning-based CNN denoisers can result in the loss of task-specific information, leading to degraded task-based IQ.<sup>2</sup>

To address this challenge, in this work, we introduce a novel task-based image-adaptive strategy to fine-tune CNN-based denoisers for each input image. Specifically, given pre-trained CNN denoiser  $\mathcal{F}(\mathbf{g}; \theta_{\text{initial}})$ , with the pre-trained weight parameter  $\theta_{\text{initial}}$ , an objective function for individual input image is formed. The objective function includes a data fidelity term, a task-based regularization term, as well as a parameter drift loss term. The data fidelity term  $\|\mathbf{g} - \mathcal{F}(\mathbf{g}; \theta)\|_2^2$  is a squared distance between the data  $\mathbf{g}$  and the network-produced image

$\mathcal{F}(\mathbf{g}; \theta)$ , and the task-specific regularization term  $\Psi(\theta)$  is designed to enforce consistency in test statistics for preserving task-relevant information:

$$\Psi(\theta) = \|t_{\text{lin}}(\mathbf{g}) - t_{\text{lin}}(\mathcal{F}(\mathbf{g}; \theta))\|_2^2 \equiv \|\mathbf{w}_{\text{lin}}^T \mathbf{g} - \mathbf{w}_{\text{lin}}^T \mathcal{F}(\mathbf{g}; \theta)\|_2^2, \quad (4)$$

where the observer template  $\mathbf{w}_{\text{lin}}$  can be computed from a training dataset and depends on the specific linear observer employed. Moreover, to leverage pre-trained networks while ensuring both data fidelity and consistency of the test statistic, we employ a parameter drift loss term. This loss term is defined as the squared distance between the weight parameters of the CNN being optimized, denoted as  $\theta$ , and those of the original pre-trained DNN denoiser,  $\theta_{\text{initial}}$ :

$$\Pi(\theta) = \|\theta_{\text{initial}} - \theta\|_2^2. \quad (5)$$

By constraining deviations of CNN weights during optimization, this term mitigates excessive divergence from the pre-trained model, thereby reducing the risk of distorting essential structures in the denoised images. Such a constraint serves as a form of regularization in the weight parameter space. Together with the data fidelity term, it balances task-specific adaptation with structural fidelity. The final optimization problem with the proposed loss function can be written as:

$$\hat{\mathbf{f}} = \arg \min_{\theta} \{\alpha \cdot \|\mathbf{g} - \mathcal{F}(\mathbf{g}; \theta)\|_2^2 + \beta \cdot \Psi(\theta) + \gamma \cdot \Pi(\theta)\}, \quad (6)$$

where  $\alpha$ ,  $\beta$ , and  $\gamma$  are tunable parameters. In this preliminary study, we employ a denoising convolutional neural network (DnCNN<sup>6</sup>-like architecture), and we employ the non-prewhitening matched filter (i.e., the signal template), denoted as  $\mathbf{s}$ , to construct the task-based regularization term  $\Psi(\theta)$ . Also, we choose to fine-tune all weight parameters of the pre-trained CNN denoiser by minimizing Eq. (6).

### 3. NUMERICAL STUDIES AND RESULTS

#### 3.1 Dataset and experimental setup

**Binary texture model:** In this study, we adopted a binary texture model,<sup>7</sup> which was designed to emulate breast CT textures. The background  $\mathbf{b} \equiv \mathcal{H}\mathbf{f}_b$  was generated by applying an element-wise thresholding operation to a Gaussian random field. In our implementation, we employed Gaussian  $1/f^\lambda$  noise with  $\lambda = 2.5$  in combination with a sigmoid thresholding function to produce binary texture backgrounds of size  $64 \times 64$ . Representative examples of the generated binary texture images are presented in Fig. 1. A circular disk signal with a radius of 2 pixels and an amplitude of 0.07 was added at the center of the images to simulate signal-present cases. To generate noisy measurement data, zero-mean Gaussian noise with a standard deviation of 0.1 was employed. Noisy realizations are also shown in Fig. 1.

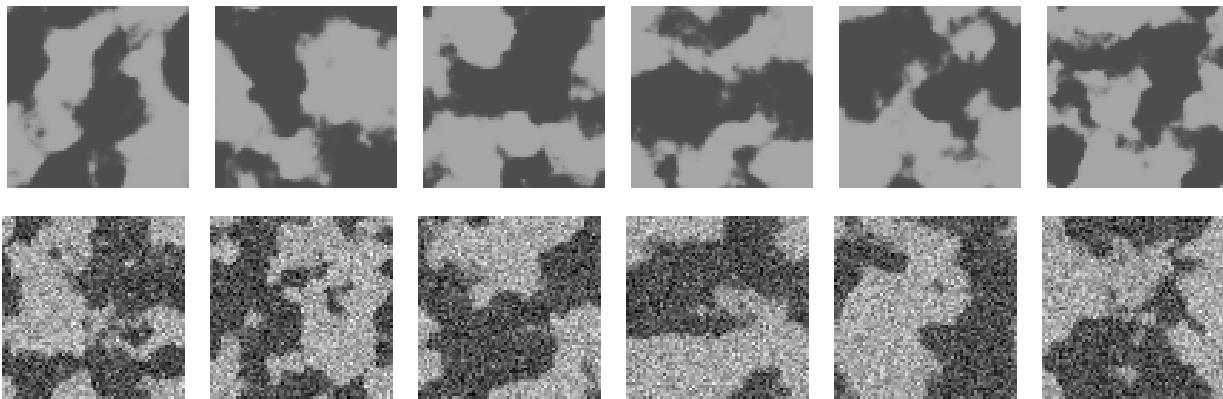


Figure 1. Top: examples of clean images; Bottom: examples of noisy images.

**Experimental setup:** To minimize the objective function in Eq. (6), Adam optimizer<sup>8</sup> was employed in the PyTorch framework.<sup>9</sup> A learning rate of  $1 \times 10^{-4}$  was employed to minimize the loss function by running 500 iterations. For the considered denoising task, we pre-trained a denoising CNN model on a dataset comprising 10,000 noisy/clean signal-absent image pairs and 10,000 noisy/clean signal-present image pairs. The denoising CNN model was trained by minimizing the mean squared error (MSE) loss function and the Adam optimizer was employed.

### 3.2 Results

An example of a signal-present clean image, its corresponding noisy image, and denoised images produced by the denoising CNN, PLS-TV, and task-based PLS-TV are shown in Fig. 2(a) to Fig. 2(e), while denoised images using the proposed task-based image-adaptive denoising CNN are shown in Fig. 2(f) through Fig. 2(h). It can be seen that, although the pre-trained CNN denoiser-produced image appears visually clean and retains much of the texture, it fails to preserve the signal to be detected. A similar limitation is observed for the PLS method employing only the TV penalty (PLS-TV). The denoised image generated by the task-based PLS-TV method<sup>1</sup> is displayed in Fig. 2(e), where the signal appears to be effectively restored. However, the result still suffers from patchy artifacts. In contrast, the proposed method can produce images with improved perceptual quality while preserving the signal to be detected.

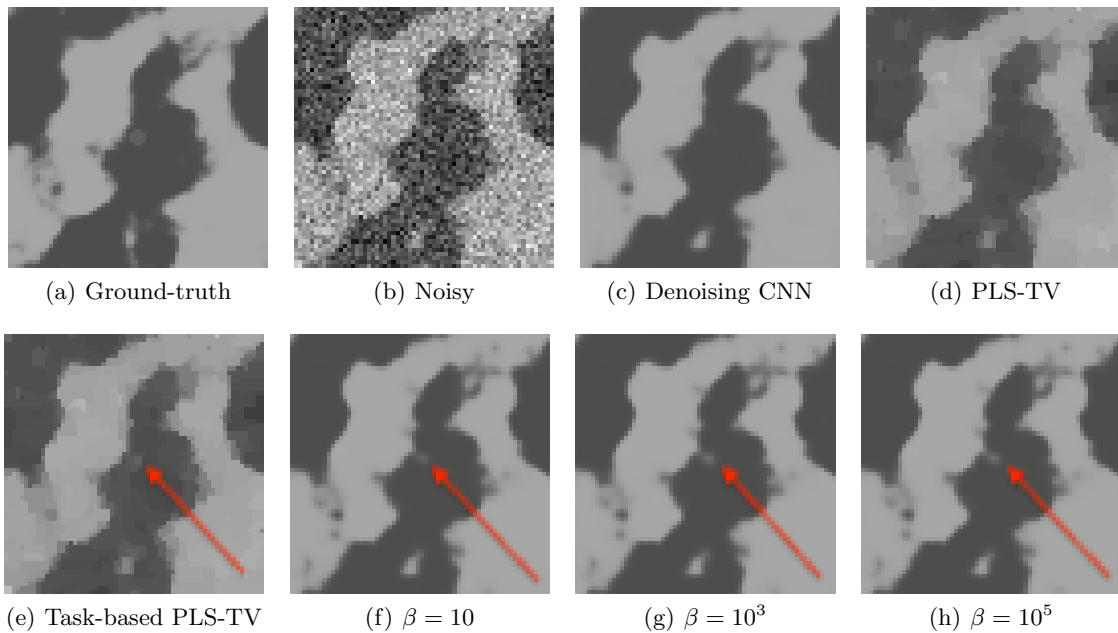


Figure 2. (a) ground-truth signal-present image; (b) noisy signal-present image; (c) denoised image using pre-trained CNN denoiser; (d) denoised image using PLS-TV; (e) denoised image using task-based PLS-TV; (f) - (h) denoised images using the proposed method with different  $\beta$  values when  $\alpha = 1$  and  $\gamma = 10^7$ .

For quantitative IQ evaluation, We first use traditional IQ metrics, as summarized in Table 1. Specifically, Structure Similarity Index Measure (SSIM), Peak Signal-to-Noise Ratio (PSNR), and Root Mean Squared Error (RMSE) were evaluated on an independent test dataset consisting of 2,000 signal-absent and 2,000 signal-present images. Compared with traditional model-based approaches and deep learning-based methods, the proposed image-adaptive denoising CNN demonstrated superior performance across all evaluated IQ metrics.

Moreover, we consider a binary signal detection task and employ the performance of the Hotelling observer (HO) for task-based IQ assessment. Specifically, the Hotelling template  $\mathbf{w}_{\text{HO}}$  was calculated using 190,000 noisy signal-absent images and 190,000 noisy signal-present images. This template was subsequently applied to the test dataset described above to assess the signal detection performance, which is quantified by the area under the

Method	PSNR ( $\uparrow$ )	SSIM ( $\uparrow$ )	RMSE ( $\downarrow$ )
Noisy	20.0058	0.4882	0.1
Denoising CNN	30.9179	0.9575	0.0286
PLS-TV	29.7211	0.9374	0.0328
Task-based PLS-TV	29.7157	0.9373	0.0328
Image-adaptive Denoising CNN ( $\beta = 10$ )	<b>31.3094</b>	<b>0.9589</b>	<b>0.0274</b>
Image-adaptive Denoising CNN ( $\beta = 10^3$ )	31.2214	0.9583	0.0276
Image-adaptive Denoising CNN ( $\beta = 10^5$ )	31.0413	0.9570	0.0282

Table 1. Performance comparison between our proposed task-based image-adaptive denoising CNN and other medical image denoising methods. Best results are **bolded**.

receiver operating characteristic (ROC) curve (AUC). The ROC curves for different choices of  $\beta$  are presented in Fig. 3(a), and Fig. 3(b) plots AUC values corresponding to ROC curves in Fig. 3(a). As expected, when  $\beta$  was increased, the corresponding AUC value also increased.

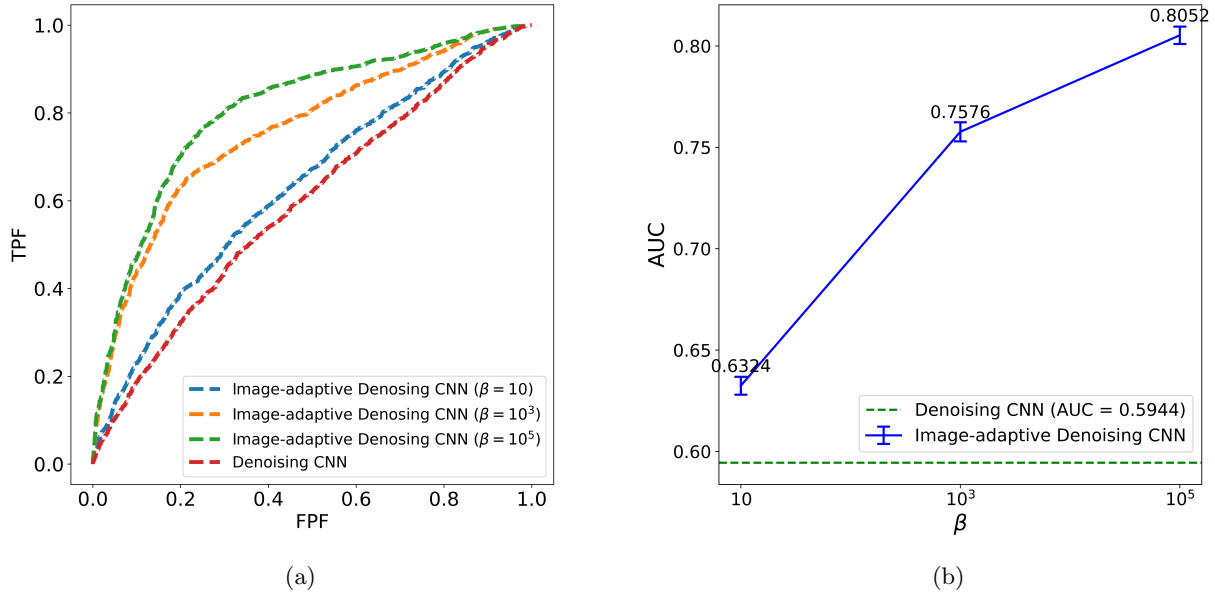


Figure 3. (a) The ROC curves produced by denoised images using different  $\beta$  values given  $\alpha = 1$  and  $\gamma = 10^7$ . (b) The AUC values corresponding to denoised images with different task-specific regularization parameters  $\beta$ , and the AUC value corresponding to the CNN denoised images is also plotted.

## 4. CONCLUSIONS

In this work, we introduced a novel task-based image-adaptive denoising framework for fine-tuning pre-trained CNN denoisers during inference (test-time optimization). The framework incorporates a task-based regularization term to ensure that task-relevant information is preserved during the denoising process. Numerical studies involving a stochastic binary texture model were conducted, and the results demonstrated that the proposed method can produce noise-reduced images with substantially improved performance in both task-based IQ and traditional IQ metrics. This preliminary study focused on utilizing linear observers to construct the task-specific regularization term. Future work will explore the application of deep learning-based non-linear ideal observers (IO)<sup>10–12</sup> and task-relevant feature vectors extracted from images<sup>13,14</sup> for regularizing inverse problems. This

study addressed the binary signal detection task. Future investigations will consider more complex inference tasks, such as joint signal detection and estimation tasks.<sup>15,16</sup>

## REFERENCES

- [1] 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).
- [2] Li, K., Zhou, W., Li, H., and Anastasio, M. A., “Assessing the impact of deep neural network-based image denoising on binary signal detection tasks,” *IEEE transactions on medical imaging* **40**(9), 2295–2305 (2021).
- [3] Han, M., Shim, H., and Baek, J., “Low-dose ct denoising via convolutional neural network with an observer loss function,” *Medical physics* **48**(10), 5727–5742 (2021).
- [4] Li, K., Li, H., and Anastasio, M. A., “Investigating the use of signal detection information in supervised learning-based image denoising with consideration of task-shift,” *Journal of Medical Imaging* **11**(5), 055501–055501 (2024).
- [5] Barrett, H. H. and Myers, K. J., [*Foundations of image science*], John Wiley & Sons (2013).
- [6] Zhang, K., Zuo, W., Chen, Y., Meng, D., and Zhang, L., “Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising,” *IEEE transactions on image processing* **26**(7), 3142–3155 (2017).
- [7] Abbey, C. K. and Boone, J. M., “An ideal observer for a model of x-ray imaging in breast parenchymal tissue,” in [*Digital Mammography: 9th International Workshop, IWDM 2008 Tucson, AZ, USA, July 20-23, 2008 Proceedings 9*], 393–400, Springer (2008).
- [8] Kingma, D. P. and Ba, J., “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980* (2014).
- [9] Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al., “Pytorch: An imperative style, high-performance deep learning library,” *Advances in neural information processing systems* **32** (2019).
- [10] 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).
- [11] Zhou, W., Li, H., and Anastasio, M. A., “Approximating the ideal observer for joint signal detection and localization tasks by use of supervised learning methods,” *IEEE transactions on medical imaging* **39**(12), 3992–4000 (2020).
- [12] Zhou, W., Villa, U., and Anastasio, M. A., “Ideal observer computation by use of markov-chain monte carlo with generative adversarial networks,” *IEEE Transactions on Medical Imaging* **42**(12), 3715–3724 (2023).
- [13] 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).
- [14] Zhou, W., “Using gradient of lagrangian function to compute efficient channels for the ideal observer,” in [*Medical Imaging 2025: Image Perception, Observer Performance, and Technology Assessment*], **13409**, 61–67, SPIE (2025).
- [15] Li, K., Zhou, W., Li, H., and Anastasio, M. A., “A hybrid approach for approximating the ideal observer for joint signal detection and estimation tasks by use of supervised learning and markov-chain monte carlo methods,” *IEEE transactions on medical imaging* **41**(5), 1114–1124 (2021).
- [16] Li, D., Li, K., Zhou, W., and Anastasio, M. A., “Approximating the ideal observer for joint signal detection and estimation tasks by the use of markov-chain monte carlo with generative adversarial networks,” *Journal of Medical Imaging* **12**(5), 051810–051810 (2025).