Sparse Bayesian Deep Learning for Cross Domain Medical Image Reconstruction

Jiaxin Huang[1], Qi Wu[1], Yazhou Ren[1,2], Fan Yang[2], Aodi Yang[1], Qianqian Yang[1], Xiaorong Pu[1,2]*

[1]School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China
[2]Shenzhen Institute for Advanced Study, University of Electronic Science and Technology of China, Shenzhen, China
hjx1134@gmail.com, qi.wu1@outlook.com, yazhou.ren@uestc.edu.cn, yf@std.uestc.edu.cn, aodi.yang@outlook.com, yangqq507@163.com, puxiaor@uestc.edu.cn

Abstract

Cross domain medical image reconstruction aims to address the issue that deep learning models trained solely on one source dataset might not generalize effectively to unseen target datasets from different hospitals. Some recent methods achieve satisfactory reconstruction performance, but often at the expense of extensive parameters and time consumption. To strike a balance between cross domain image reconstruction quality and model computational efficiency, we propose a lightweight sparse Bayesian deep learning method. Notably, we apply a fixed-form variational Bayes (FFVB) approach to quantify pixel-wise uncertainty priors derived from degradation distribution of the source domain. Furthermore, by integrating the uncertainty prior into the posterior sampled through stochastic gradient Langevin dynamics (SGLD), we develop a training strategy that dynamically generates and optimizes the prior distribution on the network weights for each unseen domain. This strategy enhances generalizability and ensures robust reconstruction performance. When evaluated on medical image reconstruction tasks, our proposed approach demonstrates impressive performance across various previously unseen domains.

Introduction

Image reconstruction is a crucial task in medical image processing that involves estimating the degradation distribution from the given input and removing it to obtain a clean image. Previous methods, such as deep neural networks (DNNs) and generative adversarial networks (GANs), have demonstrated remarkable performance in medical image denoising and artifact removal (Ronneberger, Fischer, and Brox 2015; Zhang et al. 2017; Chen et al. 2017; Zhu et al. 2017). Such deep learning (DL) methods typically operate under the assumption of consistent degradation features between training and testing datasets, a simplification that diverges from real-world application. Clinical datasets are inherently varied, affected by differences in scanning protocols, imaging equipment manufacturers, and patient demographics (Li et al. 2020), evident during simulations as well as actual patient assessments. Consider the example in Figure 1 which contrasts two distinct cross domain scenarios in medical image reconstruction: distribution shift and domain shift. The varying texture of smooth regions in low-dose abdominal and head computed tomography (CT) scans, along with the diversity in artifact expression across magnetic resonance image (MRI) datasets, highlight the substantial divergence between source (training data) and target (testing data) domains. This gap presents evident challenges for the scalability and real-world applicability of DL in a clinical setting, with models prone to overfit to training data and subsequently lose generalization capabilities when encountering new degradation types. Consequently, the ability to bridge source-target domain discrepancies is crucial for the deployment of DL in medical image reconstruction tasks.

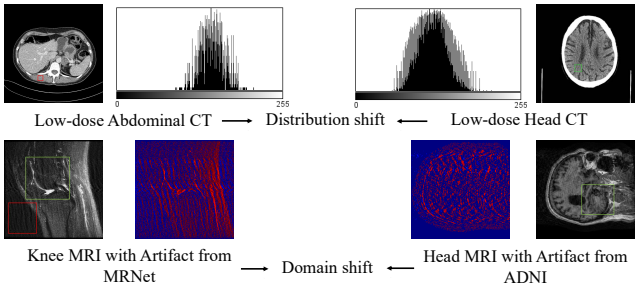


Figure 1: Examples of two kinds of cross domain scenarios in medical image reconstruction tasks: distribution shift and domain shift.

Recent progress in Bayesian deep learning offers promising solutions to the prevalence of overfitting in conventional reconstruction networks (Xu, Zhang, and Zhang 2020; Zhang et al. 2020a; El Helou and Süsstrunk 2020). Bayesian neural networks, in contrast to their traditional counterparts, infer a posterior distribution over model parameters, thus encapsulating the underlying uncertainty. This approach affords a more nuanced representation of parameter uncertainties, which is particularly advantageous in handling the erratic nature of real-world clinical data. To address the cross domain problem, some methods (Shankar et al. 2018; Manakov et al. 2019; Du, Chen, and Yang 2020; Huang et al. 2020) have been proposed to transfer the image from the noise domain to the clean domain, based on the similarity between the query image and support example. For instance, hybrid discriminator cycleGAN (Manakov et al.

*Corresponding authors.

2019) and CaGAN (Huang et al. 2020) have advanced the field of medical image reconstruction by focusing on dual-domain networks (Ganin et al. 2016). Nevertheless, these approaches tend to overlook the variations of degradation. In the pursuit of achieving both robustness and adaptability, certain hybrid approaches have emerged, aiming to integrate deep networks into traditional optimization algorithms. One such example is the utilization of deep plug-and-play (PNP) methods (Zhang et al. 2021; Huang et al. 2022), in which pretrained convolutional neural networks (CNN) are integrated as priors within iterative optimization frameworks designed for various image reconstruction tasks. Regrettably, these methods tend to face the drawback of time-consuming inference processes. Inspired by establishing a connection between model-based approaches and DL methodologies, Mou et al (Mou, Wang, and Zhang 2022) propose a deep generalized unfolding network (DGUNet) with good interpretability. While the proposed proximal gradient descent (PGD) algorithm enhances the deep unfolding networks (DUN), this method also requires numerous parameters. To rectify the above issues, we propose a sparse Bayesian deep learning method into a CNN-based reconstruction network, ensuring regularization and the retention of source domain knowledge to mitigate overfitting and maintain robustness.

The contributions of this paper have three-folds:

- A sparse Bayesian deep learning method has been proposed to optimize the variational inference steps with less computational cost and higher reconstruction accuracy.
- We generalize fixed-form variational Bayes (FFVB) to a general form under Gaussian distribution, and prove the feasibility of sparse Bayesian Learning in Gaussian regression model. The model is adopted to quantify the pixel-wise uncertainty priors derived from the degradation distribution extracted from the source domain.
- By integrating uncertainty prior into the posterior sampled by stochastic gradient Langevin dynamics (SGLD), a training strategy is developed to dynamically optimize the prior distribution on the network weights under each unseen domain to improve generalizability.

Related Work

Bayesian Deep Learning

Bayesian learning techniques have received much attention in various image reconstruction tasks, which allows estimation of prediction uncertainty (Bishop and Nasrabadi 2006). The uncertainty is usually divided into random uncertainty and cognitive uncertainty (Kendall and Gal 2017), where the random uncertainty comes from the noise in the data. Bayesian neural network (BNN) is capable of incorporating uncertainty through priors on network weights, thus enabling variational inference about the loss function through posterior predictive distributions (MacKay 1992; Neal 2012). For instance, Chen et al. (Chen et al. 2023) propose a novel Bayesian noise uncertainty alignment (BNUA) method to conduct implicit noise distribution modeling and alignment in the latent space due to its impressive robustness. Bayesian deep learning methods have also been adeptly utilized in high-stakes applications such as enhancing the resolution of extremely low-dose CT (LDCT) images through Bayesian inversion coupled with conditional Wasserstein GANs (Adler and Öktem 2019), as well as in MRI super-resolution tasks that employ subpixel convolutional networks augmented with adaptive dropout rates (Tanno et al. 2017).

Cross Domain Medical Image Reconstruction

Cross domain framework is designed to improve the robustness in unseen domains (test datasets). Some proposed approaches, like Du et al. (Du, Chen, and Yang 2020), aim to learn an invariant representation devoid of noise through disentangled learning, aligning the features of reconstructed clean observations within a latent space for adaptation. Zhang et al. (Zhang et al. 2020b) propose a noise adaptation generative adversarial network (NAGAN) to generate identical noise patterns and preserve content by mapping data between domains. Despite their ability to retain certain semantic and background information, these methods depend on noise invariance. To address this constraint, Li et al. (Li et al. 2022) devise a Gaussian mixture model (GMM) to characterize noise distribution by classifying device parameters that influence noise. Although the GMM-based approach outperforms the parameter-dependent framework (PDF) introduced by Xia et al. (Xia et al. 2021) in modeling robust noise characteristics of LDCT images under various noise types, assuming a constant current during CT scanning, especially for the chest and abdomen, is untenable. On the other hand, Huang et al. (Huang et al. 2023) pioneer a novel cross domain denoising network (CDDnet) founded on a triplet loss to achieve local information alignment.

Proposed Method

Preliminary

In a cross domain scenario, domains are partitioned into distinct non-overlapping source domains denoted as S and target domains denoted as T. Throughout the training phase, solely the data from source domains is available, while data from the target domains remains unobserved. The objective is to construct a model exclusively using the data from source domains, which demonstrates effective generalization across the target domains. Given a source domain, there are n corrupted images with paired references, denoted respectively as $\mathbf{X}^S = \{\mathbf{x}_1^S, \cdots, \mathbf{x}_n^S\}$ and $\mathbf{Y}^S = \{\mathbf{y}_1^S, \cdots, \mathbf{y}_n^S\}$. In target domain, there are m corrupted images available, denoted as $\mathbf{X}^T = \{\mathbf{x}_1^T, \cdots, \mathbf{x}_m^T\}$.

Sparse Bayesian Formulation

Fixed-form Variational Bayes Bayesian Learning typically relies on Bayesian inference to estimate the parameters of a model and the associated uncertainties. In the context of image analysis where each pixel intensity is assumed to follow a Gaussian distribution, obtaining pixel-wise uncertainty involves the estimation of the variance for each pixel's Gaussian distribution. Variational Bayes (VB) is a method to approximate these posterior distributions without direct computation. To formalize this, we describe the generative

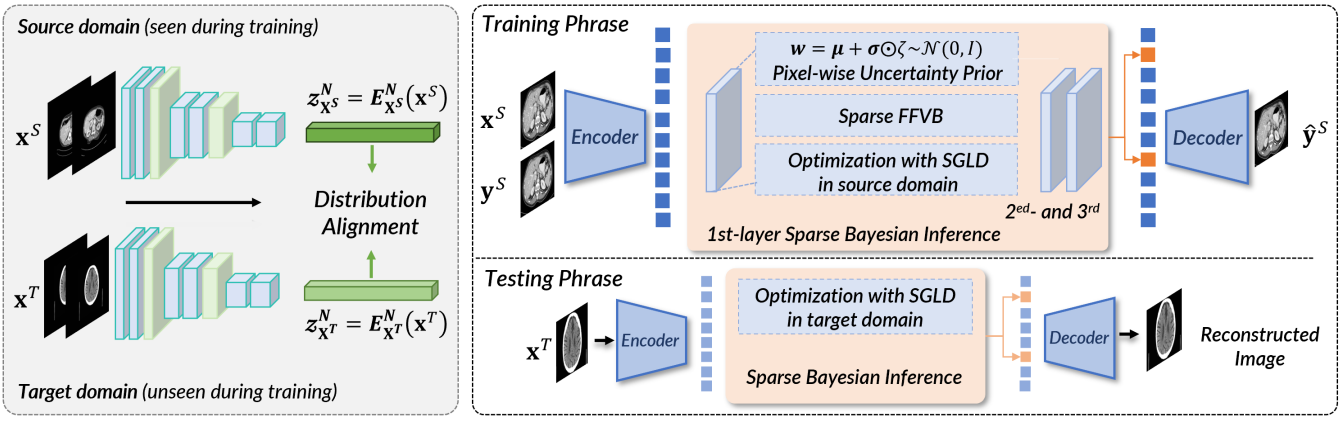


Figure 2: Overview of the proposed framework, using LDCT images as an example. The left side of the framework depicts the alignment process during both the training phase for source domain and the testing phase for target domain. On the right, two main phases are detailed: 1) Training Phase, showcasing a convolutional encoder-decoder network for reconstruction with a three-layer sparse Bayesian inference structure. 2) Testing Phase. As target domain data become available, dynamic updated prior accommodates for uncertainty and is fine-tuned to adapt through backpropagation, which also adjusts hyperparameters.

model for each pixel value. Assuming each pixel i is independent, the likelihood of an observed image $\mathbf{y}$ given the latent variables $\mathbf{x}$ and hyperparameters $\boldsymbol{\theta}$ is:

$$p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}) = \prod_{i=1}^{N} \mathcal{N}(y_i|f(x_i,\boldsymbol{\theta}),\sigma_i^2), \tag{1}$$

where N is the number of pixels per image, $\mathcal{N}$ represents the Gaussian distribution, y_i is the observed intensity of pixel i, $f(x_i,\boldsymbol{\theta})$ denotes the encoder of an autoencoder that maps the latent variable (x_i) to the mean of the Gaussian distribution with outputing a mean $\boldsymbol{\mu}_i$ and variance $\log \boldsymbol{\sigma}_i^2$ for each pixel's latent representation. σ_i^2 is the variance of the Gaussian for pixel i, representing the uncertainty in the intensity measurement.

The posterior distribution we want to estimate is:

$$p(\mathbf{x},\boldsymbol{\theta}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x},\boldsymbol{\theta}),p(\mathbf{x}),p(\boldsymbol{\theta}), \tag{2}$$

where $p(\mathbf{x})$ and $p(\boldsymbol{\theta})$ are priors on the latent variables and hyperparameters, respectively. Here, we impose a Laplace prior on both of these to induce sparsity. The Laplace distribution is chosen due to its heavier tails compared to the Gaussian distribution, leading to a higher probability of coefficients being close to zero, which encourages sparsity.

Assumption 1. *The variational posterior can be factorized as* $q_\phi(\boldsymbol{w}) = \prod_{i=1}^{L} \mathcal{N}(w_i|\mu_i,\sigma_i^2)$*, with the number of layers* L*. In each forward pass, all the weights are sampled using reparameterization* $\boldsymbol{w} = \boldsymbol{\mu} + \boldsymbol{\sigma} \odot \zeta \sim \mathcal{N}(\mathbf{0},\boldsymbol{I})$.

Followed the Assumption 1, FFVB approximates the true posterior distribution with a factorized distribution that can be optimized to be as close as possible to the true posterior by minimizing the Kullback-Leibler (KL) divergence from $q_\phi(\boldsymbol{w})$ to the true posterior.

Assumption 2. *A fixed parametric form for the VB approximation density* q*, e.g.,* $q = q_{\boldsymbol{v}}$ *belongs to some class of distributions* $\mathcal{Q}$ *indexed by a vector* $\boldsymbol{v}$ *called the variational parameter.*

Theorem 1. *Suppose the multiple satisfies Assumption (2),* $q_{\boldsymbol{v}}$ *can be a Gaussian distribution with mean* μ *and covariance matric* Λ.

The variational parameters $\boldsymbol{\phi} = \langle \boldsymbol{\mu}, \boldsymbol{\sigma} \rangle$ are optimized by minimizing the negative *log* evidence lower bound. This effectively doubles the number of trainable parameters and is known as Bayes by backdrop (Blundell et al. 2015).

Dynamic Optimization To update the prior distribution on network weights, we follow the Bayesian principle to define an appropriate objective based on Assumption 2 and Theorem 1. The target is the posterior distribution over the network weights, given data from both the source domain (seen during training) and target domains (unseen during training). With a focus on generalizability, we seek to update our prior knowledge as target domain data arrives, while integrating uncertainty estimation into the learning process. SGLD combines ideas from stochastic gradient descent (SGD) with elements from Langevin dynamics by injecting controlled noise that reflects uncertainty in the model. In this paper, this noise is directly related to the posterior distribution we are trying to sample from. The update rule for the weights using SGLD can be formularized as:

$$\Delta\theta_t = \epsilon_t \nabla \log p(\theta_{t-1}|D) + \sqrt{2\epsilon_t} N(0,I), \tag{3}$$

where D is the training data, θ_t are the parameters at iteration t, $\epsilon_t$is the learning rate, and $N(0,I)$ represents isotropic Gaussian noise.

As new data from the target domain becomes available, we update the prior distribution of the network weights to reflect the new information. The updated prior account for the uncertainty inherent to the unseen domain and dynamically adjust to fit it. This involves backpropagation to not only update the weights but also adjust the hyperparameters governing the prior distribution. Iterate between sampling (using SGLD) and adapting the prior as more data from the target domain is encountered. Regularize the learning pro-

Algorithm 1: Dynamic Prior Optimization with SGLD

Input: Source data D_{source}, target data D_{target}, Initial prior distribution $p(\theta)$, Initial network weights θ_0, Learning rate schedule $\{\epsilon_t\}$, Number of adaptation steps T_{adapt}, Number of SGLD iterations T_{sgld}.
Output: Adapted network weights θ.

```
Initialize network weights θ ~ p(θ).
for t = 1 to T_sgld do
    Sample mini-batch from D_source.
    Compute gradient of log-posterior: grad_logpost = ∇ log p(θ|mini-batch).
    Inject noise: noise_term = sqrt(2ε_t) N(0, I).
    Update weights by Eq. (3).
end for
for t = 1 to T_adapt do
    Sample mini-batch from D_target.
    Compute gradient of log-posterior with respect to the target domain.
    Update the prior p(θ) parameters based on the performance on the target domain.
    Inject noise: noise_term = sqrt(2ε_t) N(0, I).
    Update weights by Eq. (3).
end for
return θ.
```

cess to prevent overfitting to the target domain while still retaining the knowledge from the source domain.

Model Training

As shown in right side of Figure 2, the encoder denoted by E, extract pixel-wise features from the input $\mathbf{x}^S$. Subsequently, guided by the aforementioned Bayesian framework, the reconstruction network integrates a three-layer sparse Bayesian inference mechanism. This approach is pivotal for determining the joint distribution $P(\mathcal{X}, \mathcal{Y})$ and for parameterizing the posterior distribution $P(\mathcal{W}|\mathcal{X}, \mathcal{Y})$ through the application of SGLD within our conventional reconstruction network. Followed by this Bayesian regularization, the updated features are then transformed by a decoder G into the reconstructed image $\hat{\mathbf{y}}^S$. We also apply reconstruction loss $\mathcal{L}_{rec}$ and perceptual loss $\mathcal{L}_{per}$ to facilitate the training:

$$\mathcal{L}_{rec}(E, G) = \left\|\mathbf{y}^S - \hat{\mathbf{y}}^S\right\|_2^2, \tag{4}$$

$$\mathcal{L}_{per} = \left\|vgg(\mathbf{y}^S) - vgg(\hat{\mathbf{y}}^S)\right\|_2^2, \tag{5}$$

where $\mathcal{L}_{per}$ is applied to deal with potential issues like over-smoothing and distortion compared to per-pixel mean-square error (MSE) loss (Johnson, Alahi, and Fei-Fei 2016).

Furthermore, we introduce a KL divergence loss, D_{KL}, to enforce the alignment of the latent variable distribution within the network. This loss component ensures that the latent representation extracted from the autoencoder (as shown in the left side of Figure 2) adheres more closely to a predefined distribution (normal distribution) $p(\boldsymbol{z}_{\mathbf{X}^N} \sim \mathcal{N}(0, 1))$, thus assisting with the generalization of the model to target domains without compromising the quality of the reconstruction. In summary, our model can be trained by minimizing the following objectives as:

$$\mathcal{L}_{all} = \lambda_{align} D_{KL} + \lambda_{rec}\mathcal{L}_{rec}(E, G) + \lambda_{per}\mathcal{L}_{per}. \tag{6}$$

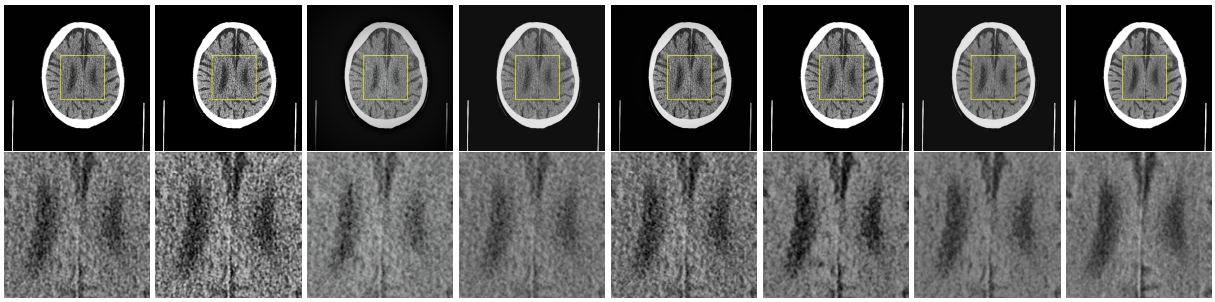
Experiments

Cross Domain LDCT Image Denoising

Dataset For this application, we train our proposed method on a phantom dataset ($Phantom$), in which 500 low-dose and normal-dose CT (NDCT) chest image pairs corrupted by random radiation levels with unknown distribution have been selected. The first testing dataset $AAPM$ is authorized by Mayo Clinic for the *2016 NIH-AAPM-Mayo Clinic Low Dose CT Grand Challenge* (McCollough 2016), among which 500 paired chest CT image slices and same amount of head CT image slices are randomly selected for testing. Another 500 pairs of chest and head CT image slices are randomly selected from the Low-Dose Parallel Beam (LoDoPaB)-CT dataset ($LoDoPaB$) (Leuschner et al. 2021) and the Low Dose CT Image and Projection Data (LDCT-and-Projection-data) (Moen et al. 2021), respectively. The experiments are divided into three groups: training on $Phantom$ dataset and testing on $AAPM$, $LoDoPaB$ and $LDCTPD$ (i.e, $Phantom \rightarrow AAPM$).

Experimental Settings The convolutional layers used in the encoder are of the kernel size 3, the padding size 1, and the stride 1, followed the protocol of the modularized CPCE (Shan et al. 2018). In all experiments, we randomly crop patches with a batch size of 16 for training. Hyper-parameters are set to λ_{align}=10, λ_{rec}=0.013 and λ_{per}=10. Number of SGLD iterations T_{sgld} for $Phantom$ data and number of adaptation steps T_{adapt} is set to 8. The learning rate was initially set as 0.001 and reduced to 0.0001 when the training errors held steady.

Analysis To visualize the reconstruction performance, we depict the results of head sets selected from LDCT-and-Projection-data ($LDCTPD$). To further evaluate the performance of the details, zoomed region-of-interests (ROIs) of the image are also given, marked by the yellow rectangles. We can observe in Figure 3 that our proposed method achieves relatively better noise suppression on both head scans from different domains compared with other methods and reserves better information about the background. From their enlarged ROIs shown in Figure 3, we see that they have good reconstruction performance but lose some fine structural objects compared to the ground truth (NDCT) image. We calculate four full-reference quality assessments, Peak Signal Noise Rate (PSNR), structural similarity (SSIM), information fidelity criterion (IFC) (Sheikh, Bovik, and De Veciana 2005), and visual information fidelity (VIF) (Sheikh and Bovik 2006). The values are calculated over all CT images selected from four datasets. Table 1 and Table 2 summarize the comparison results tested on the testing sets. Among the four metrics, PSNR and SSIM have more focus on pixel-level similarity, VIF and IFC have more focus on psychovisual features of the human visual system (HVS) by using natural statistics models. Our proposed method achieves the highest values of four metrics in



Corrupted DIP CycleGAN Du et al. DPIR DGUNet Ours Ground Truth

Figure 3: Performance comparison of different methods on the head slices from $LDCTPD$. All methods are trained on the chest sets from $Phantom$. The display window is [-10, 80]HU.

Method	$Phantom \rightarrow AAPM$				$Phantom \rightarrow LoDoPaB$				$Phantom \rightarrow LDCTPD$			
	PSNR ↑	SSIM ↑	VIF ↑	IFC ↑	PSNR ↑	SSIM ↑	VIF ↑	IFC ↑	PSNR ↑	SSIM ↑	VIF ↑	IFC ↑
DIP	24.43	0.850	0.205	1.272	25.33	0.825	0.203	1.283	26.72	0.905	0.205	1.443
CycleGAN	25.98	0.857	0.212	1.321	26.64	0.842	0.201	1.395	27.92	0.918	0.208	1.452
Du et al.	26.24	0.852	0.193	1.393	27.10	0.892	0.210	1.475	28.97	0.921	0.218	1.462
DPIR	27.93	0.863	0.199	1.393	28.50	0.904	0.217	1.496	30.42	0.935	0.225	1.494
DGUNet	28.42	0.866	0.208	1.400	29.75	0.914	0.225	1.567	31.79	0.948	0.229	1.508
Ours	**29.46**	**0.899**	**0.234**	**1.495**	**30.45**	**0.926**	**0.248**	**1.592**	**32.81**	**0.955**	**0.231**	**1.516**

Table 1: Quantitative results of LDCT image denoising from chest to head. All the methods are trained on chest slices set from $Phantom$ and tested on each head set from $AAPM$, $LoDoPaB$ and $LDCTPD$. The best scores are highlighted in bold.

Method	$Phantom \rightarrow AAPM$				$Phantom \rightarrow LoDoPaB$				$Phantom \rightarrow LDCTPD$			
	PSNR ↑	SSIM ↑	VIF ↑	IFC ↑	PSNR ↑	SSIM ↑	VIF ↑	IFC ↑	PSNR ↑	SSIM ↑	VIF ↑	IFC ↑
DIP	25.62	0.840	0.209	1.478	27.03	0.853	0.207	1.458	28.93	0.911	0.214	1.448
CycleGAN	26.05	0.897	0.223	1.453	28.31	0.864	0.205	1.489	29.48	0.924	0.227	1.469
Du et al.	23.92	0.844	0.204	1.309	28.02	0.902	0.214	1.490	30.73	0.930	0.239	1.476
DPIR	25.84	0.892	0.213	1.396	29.03	0.910	0.219	1.509	31.16	0.935	0.240	1.486
DGUNet	27.24	0.934	0.225	1.425	29.35	0.913	0.220	1.510	33.00	0.946	0.249	1.490
Ours	**28.82**	**0.946**	**0.230**	**1.489**	**29.93**	**0.914**	**0.230**	**1.519**	**32.33**	**0.950**	**0.249**	**1.498**

Table 2: Quantitative results of LDCT image denoisng from head to chest. All the methods are trained on selected head slices set and tested on chest set. The best scores are highlighted in bold.

all experiments. The reconstructed images in Figure 3 also show that our proposed model achieves a better balance of reconstruction and structure preservation.

Cross Domain MRI Artifact Removal

Dataset We also conduct cross domain pseudo-artifact removal comparative experiments with mainstream models and the models proposed earlier on the MRNet dataset (Bien et al. 2018) and the ADNI dataset (Jack Jr et al. 2008). The experiments are divided into two groups: training on the ADNI dataset and testing on the MRNet dataset (i.e., $ADNI \rightarrow MRNet$), and training on the MRNet dataset and testing on the ADNI dataset (i.e., $MRNet \rightarrow ADNI$). The quantitative results of the comparative experiments are shown in Table 3, where the values in parentheses represent the results of the models trained and tested on the same-domain dataset (i.e., $ADNI \rightarrow ADNI$).

Experimental Settings In this experiment, we use 3D-CPCE (Shan et al. 2018). As the downsampling layers of the 3D-CPCE encoder do not perform on the depth, we can easily set kernel size as 3, the padding size as 1, and the stride as 1 in each 3D convolutional layers. Hyper-parameters are set to λ_{align}=7, λ_{rec}=0.02 and λ_{per}=5. Number of SGLD iterations T_{sgld} for $ADNI$ data and number of adaptation steps T_{adapt} for $MRNet$ is set to 8. Number of SGLD iterations T_{sgld} for $MRNet$ data and number of adaptation steps T_{adapt} for $ADNI$ is set to 10.

Analysis From Table 3, it can be observed that the notable disparities in data distribution between the training and testing sets lead to a significant deterioration in the model's artifact correction performance. Whether transitioning from the ADNI dataset to the MRNet dataset or vice versa, there is an average decrease of approximately 32.5% in PSNR and an average decrease of about 5.17% in SSIM. Additionally, the

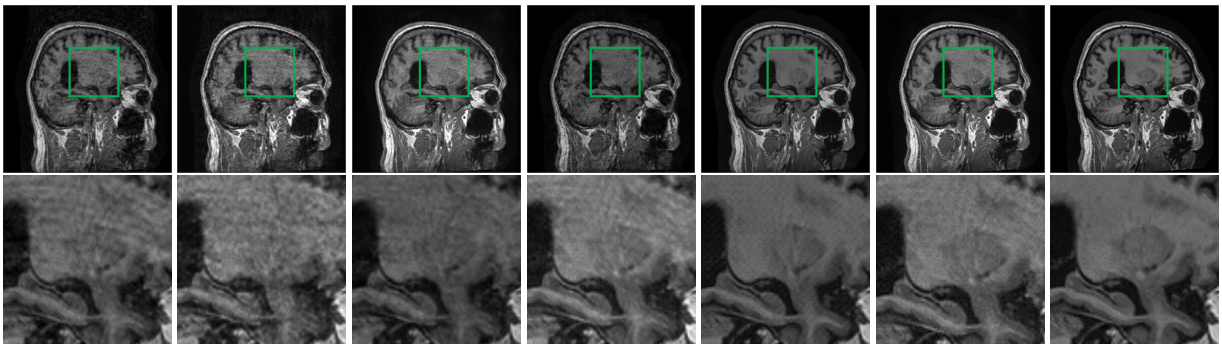


Figure 4: Visual comparison of MRI artifact removal. All the methods are trained on knee MRI data from $MRNet$ and tested on head data from $ADNI$, Our method performs better with the result are visually closer to the ground truth.

Model	$ADNI \rightarrow MRNet$		$MRNet \rightarrow ADNI$	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
CycleGAN	18.82(29.39)	0.904(0.958)	19.27(32.28)	0.737(0.894)
DeblurGAN	18.70(28.30)	0.895(0.950)	23.25(34.31)	0.722(0.958)
DPIR	24.38(32.55)	0.939(0.986)	25.17(37.42)	0.824(0.979)
DGUNet	26.42(33.49)	0.949(0.982)	26.49(37.90)	0.845(0.985)
Ours	**28.85**	**0.963**	**29.45**	**0.894**

Table 3: Quantitative results of cross domain MRI artifact removal between simulated datasets.

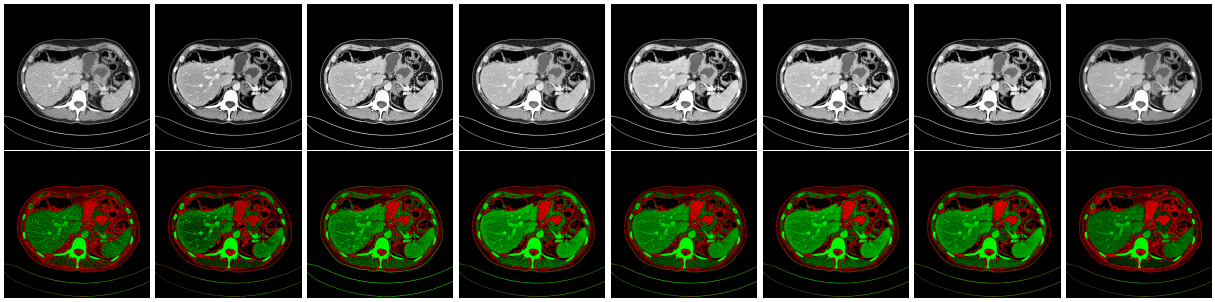


Figure 5: Ablation studies of different parts of Bayesian deep learning methods on $AAPM$.

transition from the MRNet dataset to the ADNI dataset experiences an average decrease of around 20.3% in PSNR and 6.52% in SSIM. This phenomenon might be attributed to the higher complexity of head structures compared to knee structures, a detailed analysis of which will be presented in subsequent sections in conjunction with visualized results. Furthermore, it can be observed in Figure 4, the proposed method achieves favorable generalization performance, with an average improvement of approximately 30.2% in PSNR and about 6.32% in SSIM. Compared with other methods, our proposed method achieves reconstructions that are closest to the ground truth image.

Ablation Study

To evaluate the effectiveness of our dynamic optimization approach, we benchmark it against five baseline Bayesian inference techniques. The reconstructions from Bayesian Deep image prior (DIP) and MAP exhibit noticeable artifacts; however, the results from SGLD, MCD, and FFVB combined with MCD (FFVB+MCD) are markedly smoother. Importantly, images reconstructed by our model (FFVB+SGLD) achieve an improved balance between detail retention and artifact reduction, as evident in Figure 5. Notably, Bayesian DIP tends to preserve more structural integrity than MAP, suggesting that MAP may discard subtler details. The application of Bayesian Optimization (BO) broadens the critical period before overfitting for both SGLD and MCD, underlining that early stopping is still a prerequisite for optimal reconstruction quality. Our FFVB-based method, however, successfully avoids overfitting to erroneous patterns and distinctly surpasses other methods in PSNR and SSIM across all test images, detailed in Table 4.

Method	$Phantom \rightarrow AAPM$		$Phantom \rightarrow LoDoPaB$		$Phantom \rightarrow LDCTPD$	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
Bayesian DIP	24.58	0.853	24.19	0.836	21.39	0.809
MAP	25.64	0.843	25.91	0.843	22.69	0.832
MCD	26.58	0.883	26.19	0.871	26.46	0.852
SGLD	26.28	0.888	26.98	0.872	25.36	0.855
FFVB+MCD	26.64	0.893	27.91	0.883	27.29	0.885
FFVB+SGLD (Ours)	**27.33**	**0.905**	**27.01**	**0.896**	**29.44**	**0.891**

Table 4: Quantitative results with different Bayesian inference methods.

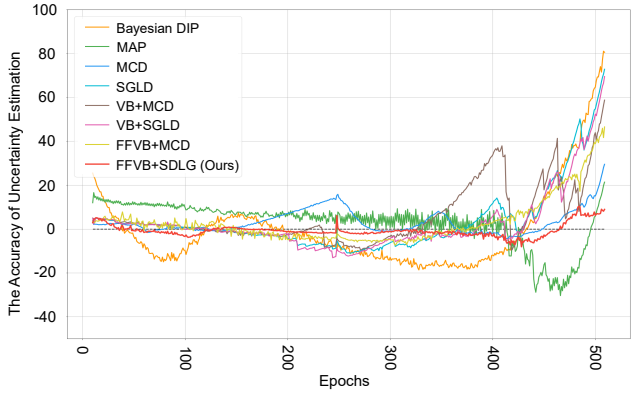


Figure 6: The accuracy of uncertainty estimation with different Bayesian inference methods during the training stage.

Consumption	$Time$	$Params$	$MACs$	$Memory$
DIP	11.46	1.324	2.100	9.160
CycleGAN	22.24	7.500	5.000	19.34
Du et al.	26.24	10.24	13.00	18.00
DPIR	30.05	15.93	1.320	43.00
DGUNet	46.44	22.30	12.07	63.70
Ours	27.19	13.32	2.650	11.26

Table 5: Comparison of computational cost.

Feature	Ours	DIP	MAP
Laplace and Guassian (LOG)	1	5	3
Scale-Invariant Feature	3	4	2
Fourier Energy Spectrum	2	3	1
Laws Texture Energy Metric	5	1	5
Normal Vector Flows (NVFs)	4	2	4

Table 6: Feature importance ranking for reconstruction.

Furthermore, as illustrated in Figure 6, our proposed approach demonstrates the smallest absolute error in comparison to all evaluated Bayesian methods.

Evaluation

Computational Cost We compare the computational cost of each model in the four dimensions of $Time$, $Paras$, $MACs$, and $Memory$. Among them, $Time$ refers to the average training time of one epoch. $Params$ refers to the total number of parameters that a layer model needs to train. $MACs$ stands for multiply–accumulate operations when a layer model is trained. $Memory$ refers to the maximum memory required by the entire model during training. Both $Params$ and $MACs$ are calculated using PyTorch-OpCounter, and $Memory$ is calculated using Pytorch-Memory-Utils. Table 5 shows that the training time of our proposed method is shorter. And $Params$, $MACs$ and $Memory$ required for training are far within the range that the server can afford. Therefore, the proposed method can obtain excellent results at a small computational cost.

Interpretability Our proposed method, like DIP and MAP, ensures that the features and the target results maintain the same development trend by imposing partial monotonicity constraints, thus generating quantifiable feature importance. Table 6 shows the feature importance ranking of the models from high to low, and the correlation between different models' rankings is shown by Kendall's tau coefficient (Van Doorn et al. 2018). Due to the different degrees of attention to features among models, the correlation between our method and other models also varies. Our method can achieve a correlation of 0.733 and pays more attention to Laplace and Guassian (LOG) and Fourier energy spectrum, two features that describe the image edge and noise characteristic, than other models, and these two features rank first and second in the ranking, respectively.

Conclusion

In addressing the challenge of medical image reconstruction across diverse domains, we present a lightweight sparse Bayesian learning approach, tailored for cross domain adaptability and efficiency. Our method utilizes FFVB to capture pixel-wise uncertainty and leverages SGLD to dynamically optimize the prior distribution for network weights. Through this approach, we enhance the model's generalizability and robustness in reconstructing medical images from unseen domains. Comparative evaluations on LDCT and MRI datasets confirm the effectiveness of our methodology over existing techniques, achieving better performance in cross domain reconstruction tasks.

Acknowledgements

We'd like to thank Kecheng Chen for his valuable suggestions and guidance related to this manuscript, and to Junchi Liu for his insightful feedback on the MRI Artifact Removal experiments conducted in this study. This work was supported in part by NHC Key Laboratory of Nuclear Technology Medical Transformation (Mianyang Central Hospital) (No. 2021HYX017), Sichuan Science and Technology Program (No. 2022JDZH0027), Clinical Research Incubation Project, West China Hospital, Sichuan University (No. 2021HXFH004), and Shenzhen Science and Technology Program (Nos. JCYJ20230807120010021 and JCYJ20230807115959041).

References

Adler, J.; and Öktem, O. 2019. Deep posterior sampling: Uncertainty quantification for large scale inverse problems. In *MIDL*.

Bien, N.; Rajpurkar, P.; Ball, R. L.; Irvin, J.; Park, A.; Jones, E.; Bereket, M.; Patel, B. N.; Yeom, K. W.; Shpanskaya, K.; et al. 2018. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. *PLoS medicine*, 15(11): e1002699.

Bishop, C. M.; and Nasrabadi, N. M., eds. 2006. *Pattern recognition and machine learning*, volume 4. Springer.

Blundell, C.; Cornebise, J.; Kavukcuoglu, K.; and Wierstra, D. 2015. Weight uncertainty in neural network. In *ICML*, 1613–1622. PMLR.

Chen, H.; Zhang, Y.; Kalra, M. K.; Lin, F.; Chen, Y.; Liao, P.; Zhou, J.; and Wang, G. 2017. Low-dose CT with a residual encoder-decoder convolutional neural network. *TMI*, 36(12): 2524–2535.

Chen, K.; Li, H.; Wan, R.; and Yan, H. 2023. Robust Cross-domain CT Image Reconstruction via Bayesian Noise Uncertainty Alignment.

Du, W.; Chen, H.; and Yang, H. 2020. Learning invariant representation for unsupervised image restoration. In *CVPR*, 14483–14492. IEEE.

El Helou, M.; and Süsstrunk, S. 2020. Blind universal Bayesian image denoising with Gaussian noise level learning. *TIP*, 29: 4885–4897.

Ganin, Y.; Ustinova, E.; Ajakan, H.; Germain, P.; Larochelle, H.; Laviolette, F.; Marchand, M.; and Lempitsky, V. 2016. Domain-adversarial training of neural networks. *JMLR*, 17(1): 2096–2030.

Huang, J.; Chen, K.; Ren, Y.; Sun, J.; Wang, Y.; Tao, T.; and Pu, X. 2023. CDDnet: Cross-domain denoising network for low-dose CT image via local and global information alignment. *Computers in Biology and Medicine*, 163: 107219.

Huang, J.; Chen, K.; Sun, J.; Pu, X.; and Ren, Y. 2022. Cross Domain Low-Dose CT Image Denoising With Semantic Information Alignment. In *ICIP*, 4228–4232.

Huang, Z.; Chen, Z.; Zhang, Q.; Quan, G.; Ji, M.; Zhang, C.; Yang, Y.; Liu, X.; Liang, D.; Zheng, H.; et al. 2020. CaGAN: A cycle-consistent generative adversarial network with attention for low-dose CT imaging. *IEEE TCI*, 6: 1203–1218.

Jack Jr, C. R.; Bernstein, M. A.; Fox, N. C.; Thompson, P.; Alexander, G.; Harvey, D.; Borowski, B.; Britson, P. J.; L. Whitwell, J.; Ward, C.; et al. 2008. The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 27(4): 685–691.

Johnson, J.; Alahi, A.; and Fei-Fei, L. 2016. Perceptual losses for real-time style transfer and super-resolution. In *ECCV*, 694–711. Springer.

Kendall, A.; and Gal, Y. 2017. What uncertainties do we need in bayesian deep learning for computer vision? *NIPS*, 30.

Leuschner, J.; Schmidt, M.; Baguer, D. O.; and Maass, P. 2021. LoDoPaB-CT, a benchmark dataset for low-dose computed tomography reconstruction. *SCI DATA*, 8(1): 109.

Li, D.; Bian, Z.; Li, S.; He, J.; Zeng, D.; and Ma, J. 2022. Noise Characteristics Modeled Unsupervised Network for Robust CT Image Reconstruction. *TMI*, 41(12): 3849–3861.

Li, H.; Wang, Y.; Wan, R.; Wang, S.; Li, T.-Q.; and Kot, A. 2020. Domain generalization for medical imaging classification with linear-dependency regularization. In *NIPS*, volume 33, 3118–3129.

MacKay, D. J. 1992. A practical Bayesian framework for backpropagation networks. *Neural computation*, 4(3): 448–472.

Manakov, I.; Rohm, M.; Kern, C.; Schworm, B.; Kortuem, K.; and Tresp, V. 2019. Noise as domain shift: Denoising medical images by unpaired image translation. In *MICCAI*, 3–10. Springer.

McCollough, C. 2016. TU-FG-207A-04: overview of the low dose CT grand challenge. *MED PHYS*, 43(6Part35): 3759–3760.

Moen, T. R.; Chen, B.; Holmes III, D. R.; Duan, X.; Yu, Z.; Yu, L.; Leng, S.; Fletcher, J. G.; and McCollough, C. H. 2021. Low-dose CT image and projection dataset. *MED PHYS*, 48(2): 902–911.

Mou, C.; Wang, Q.; and Zhang, J. 2022. Deep Generalized Unfolding Networks for Image Restoration. In *CVPR*, 17399–17410.

Neal, R. M. 2012. *Bayesian learning for neural networks*, volume 118. Springer Science & Business Media.

Ronneberger, O.; Fischer, P.; and Brox, T. 2015. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, 234–241. Springer.

Shan, H.; Zhang, Y.; Yang, Q.; Kruger, U.; Kalra, M. K.; Sun, L.; Cong, W.; and Wang, G. 2018. 3-D convolutional encoder-decoder network for low-dose CT via transfer learning from a 2-D trained network. *TMI*, 37(6): 1522–1534.

Shankar, S.; Piratla, V.; Chakrabarti, S.; Chaudhuri, S.; Jyothi, P.; and Sarawagi, S. 2018. Generalizing across domains via cross-gradient training.

Sheikh, H. R.; and Bovik, A. C. 2006. Image information and visual quality. *TIP*, 15(2): 430–444.

Sheikh, H. R.; Bovik, A. C.; and De Veciana, G. 2005. An information fidelity criterion for image quality assessment using natural scene statistics. *TIP*, 14(12): 2117–2128.

Tanno, R.; Worrall, D. E.; Ghosh, A.; Kaden, E.; Sotiropoulos, S. N.; Criminisi, A.; and Alexander, D. C. 2017. Bayesian image quality transfer with CNNs: exploring uncertainty in dMRI super-resolution. In *MICCAI*, 611–619. Springer.

Van Doorn, J.; Ly, A.; Marsman, M.; and Wagenmakers, E.-J. 2018. Bayesian inference for Kendall's rank correlation coefficient. *The American Statistician*, 72(4): 303–308.

Xia, W.; Lu, Z.; Huang, Y.; Liu, Y.; Chen, H.; Zhou, J.; and Zhang, Y. 2021. CT reconstruction with PDF: Parameter-dependent framework for data from multiple geometries and dose levels. *IEEE TMI*, 40(11): 3065–3076.

Xu, S.; Zhang, C.; and Zhang, J. 2020. Bayesian deep matrix factorization network for multiple images denoising. *Neural Networks*, 123: 420–428.

Zhang, K.; Li, Y.; Zuo, W.; Zhang, L.; Van Gool, L.; and Timofte, R. 2021. Plug-and-play image restoration with deep denoiser prior. *IEEE TPAMI*, 44(10): 6360–6376.

Zhang, K.; Zuo, W.; Chen, Y.; Meng, D.; and Zhang, L. 2017. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *TIP*, 26(7): 3142–3155.

Zhang, Q.; Yuan, Q.; Li, J.; Sun, F.; and Zhang, L. 2020a. Deep spatio-spectral Bayesian posterior for hyperspectral image non-iid noise removal. *ISPRS*, 164: 125–137.

Zhang, T.; Cheng, J.; Fu, H.; Gu, Z.; Xiao, Y.; Zhou, K.; Gao, S.; Zheng, R.; and Liu, J. 2020b. Noise Adaptation Generative Adversarial Network for Medical Image Analysis. *IEEE TMI*, 39(4): 1149–1159.

Zhu, J.-Y.; Park, T.; Isola, P.; and Efros, A. A. 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *ICCV*, 2223–2232. IEEE.