MPI-Mamba: Latent Feature Fusion Mamba for Anisotropic Image Calibration and Deblurring in Magnetic Particle Imaging

Liwen Zhang[1,2*†], Zhaoji Miao[3*], Yusong Shen[3], Zechen Wei[2], Hui Hui[2], Jie Tian[2,4†]

[1]The Key Laboratory of Cognition and Decision Intelligence for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[2]CAS Key Laboratory of Molecular Imaging, Beijing Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China
[3]School of Computer Science and Engineering, Southeast University, Nanjing, China
[4]School of Engineering Medicine and the School of Biological Science and Medical Engineering, Beihang University, Beijing 100191, China,
{zhangliwen2018, weizechen2019, hui.hui}@ia.ac.cn,{230238543, shenyusong}@seu.edu.cn, tian@ieee.org

Abstract

Magnetic Particle Imaging (MPI) is an innovative medical modality, providing nanomolar-scale in vivo sensitivity and radiation-free dynamic real-time detection for precision medicine. However, MPI faces a challenging problem in accurately visualizing nanoparticle distributions, where the reconstructed images with unidirectional scanning exhibit anisotropy. The anisotropy in spatial resolution leads to distortion and blurred image boundaries. Existing deep learning methods for anisotropy calibration are only limited to simulation data due to lacking of real-world MPI datasets. To address the aforementioned problems, we spent over three years designing and constructing a real-world MPI anisotropic image datasets (20,156 images) with diverse phantoms (sensitivity, resolution, vessel, shape) and animal scanning. Then, we introduce a novel Mamba-based method, MPI-Mamba, for anisotropic image calibration. Specifically, we propose a latent feature fusion state space model (LFF-SSM) block for feature fusion and leverage conditional latent diffusion model (CL-DM) branch for feature extraction. The CL-DM is performed to extract latent features in a highly compressed latent space for guiding the calibration and deblurring process. Next, we exploit the LFF-SSM to fully fuse the extracted multi-scale features to capture contextual information from the image structure, enabling the model to learn the overall distribution of signal concentration. We evaluate our method and competing methods on simulation dataset and our constructed diverse real-world MPI datasets. The results show that our proposed approach outperforms competing methods for anisotropic image calibration and deblurring.

Code — https://github.com/dreamenwalker/MPI-Mamba.

Datasets — https://github.com/dreamenwalker/MPI-Mamba/tree/main/data.

[*]These authors contributed equally.
[†]Corresponding Authors

Introduction

Magnetic particle imaging (MPI) is an innovative tomographic technique with high-resolution and high-sensitivity imaging at the nanomolar-scale in vivo sensitivity and radiation-free dynamic real-time detection (Ilbey et al. 2019; Shi et al. 2023; Güngör et al. 2023; Billings et al. 2021). This modality is capable of directly detecting magnetic nanoparticles (MNPs) within living organisms in vivo. Additionally, MPI offers excellent temporal and spatial resolution (Güngör et al. 2022), without depth-of-field restrictions or the use of ionizing radiation for in vivo observation. These features make MPI particularly suitable for a broad array of preclinical applications, such as cell tracking (Panagiotopoulos et al. 2015), vascular imaging (Tong et al. 2023), and tumor visualization (Huang et al. 2023).

Currently, A critical challenge in MPI field is achieving fast precise localization of superparamagnetic iron oxide nanoparticle distributions based on X-space method (Goodwill and Conolly 2010). For fast visualize the distribution of nanoparticle, The X-space method directly reconstructs the measurement response function based on the MNPs and magnetic field using an analytical system model to reduce reconstruction time significantly (Mohn et al. 2022). However, the X-space method causes poor quality of reconstructed images due to the influence of the point spread function and the limitations of the analytical model in X-space theories (Shang et al. 2023; Goodwill and Conolly 2010). As shown in Fig. 1, one of representative issues is anisotropic resolution problem caused by uniaxial scanning in MPI (Shang et al. 2023, 2022). Especially for portable open-side MPI devices, the excitation and scanning directions are one-sided, resulting in attenuated magnetic fields and gradients along the depth direction. Consequently, the imaging resolution decreases, and the anisotropy issues that arise is hard to overcome at the hardware level. Therefore, many researchers pay attention to the existing techniques for anisotropic image calibration and deblurring.

In the past several years, deep learning technologies have shown promising performance in image restoration (Chen

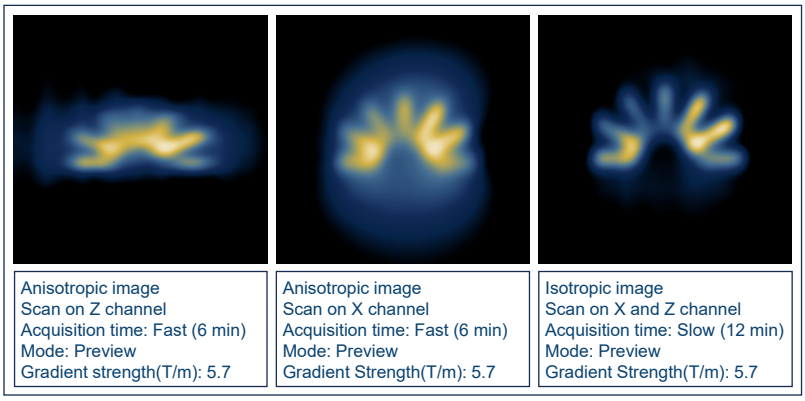

Figure 1: Visualization for anisotropic images and isotropic image with different scan modes. The defect of anisotropic images lies in their distortion of the true concentration distribution and the blurring of structural detail information.

et al. 2022; Zamir et al. 2022a; Liang et al. 2021; Zamir et al. 2022b) and deblurring (Shang et al. 2024; Tsai et al. 2022), primarily focusing on convolutional neural networks (CNNs), Transformer-based methods, and Diffusion Models(DM)-based approaches. CNNs excel at capturing local features and have been effective in tasks like denoising and deblurring (Chen et al. 2022; Zamir et al. 2022b), though they struggle with global feature extraction. Transformer-based methods (Shang et al. 2024; Liang et al. 2021; Kitaev, Kaiser, and Levskaya 2020) leverage global feature learning to outperform CNNs in tasks such as MPI image calibration and deblurring. However, these methods are evaluated based on simulation data for model training. The acquisition of diverse and high-quality real-world MPI image data faces significant challenges due to the high cost of commercial equipment, the time-consuming design of phantom and biological experiments, and the limited availability of suitable samples. Diffusion Models (Zamir et al. 2022a; Chen et al. 2024; Ren et al. 2023) have recently shown promise in natural and medical image analysis, particularly for denoising and super-resolution, but their application to MPI image resolution enhancement and deblurring remains unexplored. Notably, most studies evaluating these methods for MPI image analysis rely on simulation data, with a lack of comprehensive assessments on real-world datasets.

In this study, to address the limitations of existing studies, we firstly design and construct a diverse real-world anisotropic dataset, and we propose a novel Mamba-based method, named MPI-Mamba, for anisotropic image calibration. We summarize our contribution as: 1. We design and construct a diverse real-world anisotropic dataset comprised of 20,156 MPI images, and complete a full-chain workflow including phantom designs, breeding of eligible mice, and data scanning. To our knowledge, This is the first real-world dataset for improving the resolution for MPI images.

2. We propose a latent feature fusion state space model (LFF-SSM) blocks, specifically designed to enable effective multi-scale feature fusion for anisotropic image calibration in MPI. Additionally, we introduce a conditional latent diffusion model (CL-DM) branch that is tailored for extracting informative latent features within a highly compact latent space, thereby guiding the calibration and deblurring process.

3. Our method and competing methods are first assessed on our diversified real-world phantom and biological MPI image datasets. The results show that our method achieves the best performance among state-of-the-art methods.

Related Work

CNN-based Methods. The CNN is one of the earliest and most widely used deep learning approaches for improving image quality (Chen et al. 2022; Zamir et al. 2022b). Previous studies based on CNN also demonstrate that CNNs are powerful tools for MPI anisotropic image calibration and deblurring (Shang et al. 2023, 2022). Shang et al. (Shang et al. 2023, 2022) proposed CNN-based networks to improve the quality of MPI image based on simulation data, which demonstrate the effectiveness of the CNN-based methods. However, for feature extraction in global receptive field, CNNs show poor performance to capture long-range dependencies and global structure (Shang et al. 2024).

Transformer-based Methods. Transformer-based methods in image restoration have demonstrated impressive performance (Liang et al. 2021). For MPI image calibration, In previous studies (Shang et al. 2024; Yang et al. 2024), they proposed transformer-based network for deblurring task based on simulation data. Meanwhile, Shang et al. (Shang et al. 2024) constructed a multi-scale Transformer network for anisotropic image calibration based on simulation data. However, these transformer-based methods (Shang et al. 2024; Yang et al. 2024) are effective on large simulated datasets, whether they still perform well on real-world dataset of a certain size remains unknown.

Diffusion Models. Currently, Diffusion Models (DMs) draw much attention and show encouraging results in natural image analysis task, such as denoising (Liu et al. 2024a) and restoration (Zamir et al. 2022a; Ren et al. 2023). DMs also show good performance for medical image super-resolution (Mao et al. 2023), restoration (Chung and Ye 2022), and denoising tasks (Chung, Lee, and Ye 2022). However, to our

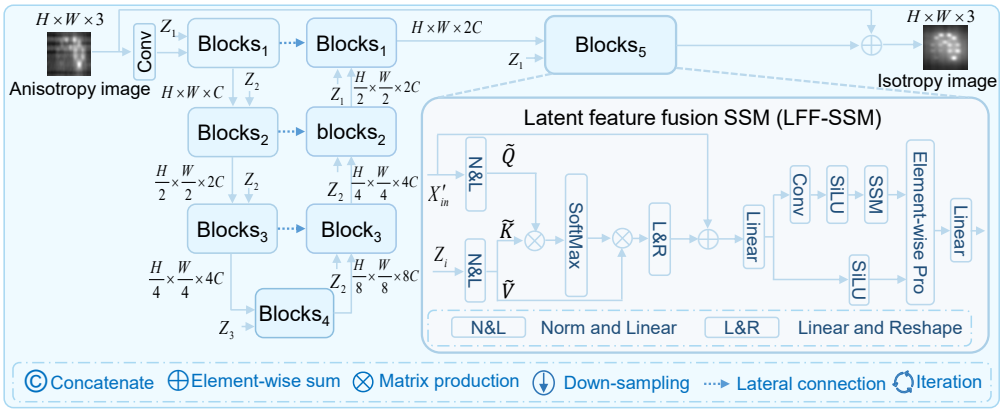

Figure 2: Architecture of MPI-Mamba. Each block is with same modules of LFF-SSM. SSM: state space model.

knowledge, diffusion models are unexplored for real-world MPI image analysis task of deblurring.

State Space Models. Recently, a new architecture of Mamba based on structured state-space sequence models (S4) is proposed and show good performance for sequence modeling (Gu and Dao 2023). Particularly the improved vision Mamba (Zhu et al. 2024), has gained significant attention as efficient backbones for different vison tasks, such as segmentation (Liu et al. 2024b), super-resolution (Qiao et al. 2024), restoration (Guo et al. 2024; Liu et al. 2025). Mamba offers a promising solution to the trade-off between global receptive field and computational efficiency in image restoration. By leveraging state-space equations from control systems, Mamba efficiently models global dependencies (Gu and Dao 2023). This allows Mamba-based networks to activate more pixels and improve restoration quality.

Methodology

Proposed MPI-Mamba

As shown in Fig. 2, we propose a novel framework, called MPI-Mamba, for anisotropic MPI image calibration and deblurring, which integrates two key components: 1) Latent feature fusion State Space Model (LFF-SSM), designed to model and learn the concentration distribution of MNPs; and 2) Conditional latent Diffusion Model (CL-DM), which focuses on removing blurring artifacts and reducing noise in MPI images, thereby enhancing the clarity of structural information.

The LFF-SSM leverages the strengths of state space modeling to capture the complex spatial relationships and concentration variations inherent in MPI data, while the CL-DM employs generative diffusion techniques to restore fine details and suppress noise, ensuring that image structures are more distinct and interpretable. The following sections provide a comprehensive explanation of the architecture and methodologies employed in this study.

The training pipeline for our proposed method is divided into two phases (Chen and Liu 2024): 1) Latent feature encoder (LFE) training and LFF-SSM training. We jointly train the LFE and LFF-SSM for the extraction of latent features. 2) CL-DM training. Based on the pre-trained LFE from first phase, the latent features output by the trained LFE are used as supervisory information to train our CL-DM.

LFF-SSM for Feature Fusion

Latent Feature Extraction As shown in Fig. 2, in the first phase, taking an anisotropic image $\mathbf{x}_{\text{iniso}} \in \Phi^{H\times W\times 3}$ as input and its corresponding isotropic image (ground truth) $\mathbf{x}_{\text{label}} \in \Phi^{H\times W\times 3}$, we concatenate $\mathbf{x}_{\text{iniso}}$ and $\mathbf{x}_{\text{label}}$ along the channel dimension as $\mathbf{x}' \in \Phi^{H\times W\times 6}$. We take the $\mathbf{x}'$ as the input of our CL-DM. We use a CNN-based encoder of LFE to encode the anisotropic images and their corresponding labels. This structure extracts latent features (Z_i) in the latent space through several residual modules. The latent features are then used as a joint input and fed into our proposed blocks of LFF-SSM, which is based on a combination of cross-attention mechanisms and SSM attention mechanisms, for further processing. We should note that we compress the feature dimension of $H \times W$ of $\mathbf{x}'$ as N of Z_i with a ratio of $\frac{H\times W}{N}$.

Latent Feature Fusion To exploit the compressed latent information extracted from CL-DM, we integrate the latent feature and intermediate feature of the LFF-SSM block. To achieve this, we propose LFF-SSM block (Fig. 2) for feature fusion, and our architecture is inspired by the work of HI-Diff (Chen et al. 2024). Next, as shown in Fig. 2, given an intermediate input feature $X_{\text{inter}} \in \Phi^{H_1\times W_1\times C_1}$ of the LFF-SSM, we reshape it as an intermediate feature $\mathbf{x}_1 \in \Phi^{N_1\times C_1}$, where $N_1 = H_1 \times W_1$, and C_1 denotes the channel dimension.

SSM To better fuse latent features and the output $\mathbf{X}_{\text{out}} \in \Phi^{H_1\times W_1\times C_1}$ calculated with the cross-attention, we adopt SSM attention in our LFF-SSM blocks. They are inspired by a particular linear continuous system (Gu and Dao 2023) that maps a 1-dimensional function or sequence $z(t) \in \mathbb{R} \mapsto y(t) \in \mathbb{R}$ through an implicit latent state $x(t) \in \mathbb{R}^N$. N denotes state size in the space. Concretely, the SSM is defined with four parameters (Δ, A, B, C), which define a sequence-

to-sequence transformation in two stages:

$$\begin{cases} \dot{x}_t = Ax_t + Bu_t \\ y_t = Cx_t + Du(t) \end{cases}, \tag{1}$$

where $A \in \mathbb{R}^{N\times N}$, $B \in \mathbb{R}^{N\times 1}$, $C \in \mathbb{R}^{1\times N}$, and $D \in \mathbb{R}$ are variables that define the dynamic properties of the state space of the continuous system. To integrate this model to Transformer or CNN method, the state space can be discretized with the timescale operator of Δ for the continuous system (Eq.1). Here, we use zero-order hold (ZOH) discretization rule to transform the continuous parameters to discrete parameters (Gu and Dao 2023). The ZOH is defined as:

$$\begin{cases} \bar{A} = E(\Delta A) \\ \bar{B} = (\Delta A)^{-1}(E(\Delta A) - I) \cdot \Delta B' \end{cases}, \tag{2}$$

where $E(\cdot)$ denotes exponent operation. I denotes Identity Matrix. The corresponding discrete time system can be defined as:

$$\begin{cases} x(T) = \bar{A}x(T-1) + \bar{B}u(T) \\ y(T) = Cx(T) + Du(T) \end{cases}. \tag{3}$$

For the above Eq. 3, to facilitate use, the above formula is equivalent to a convolution operation, which can be implemented using convolution algorithms. Here, given a convolutional kernel K, as mentioned in (Gu and Dao 2023), the coefficients in above equation can be formulated as:

$$K \triangleq [C\bar{B},\ C\overline{AB},\ \ldots,\ C\bar{A}^n\bar{B},\ \ldots], \tag{4}$$

$$y = \text{Conv}(x, K), \tag{5}$$

where $\text{Conv}(\cdot)$ denotes convolution operation.

Proposed LFF-SSM Block In our study, we explore LFF-SSM for feature fusion, which contains mix mechanism of cross-attention and SSM (As shown in Fig. 2). The block lies in the design of a multi-scale feature fusion block, which integrates features from both denoised and structurally restored images at different scales, thereby enhancing detail preservation and global structural consistency. For each of our proposed LFF-SSM blocks, given an input of X''_{in}, the whole sequence computation process can be formulated as $X''_{\text{out}} = \text{LFF-SSM}(X''_{\text{in}})$, where LFF-SSM$(\cdot)$ is our module. We leverage the module to fuse latent features of CL-DM and intermediate features with cross-attention and SSM. Here, For the operation of cross-attention, it can be formulated as:

$$\begin{cases} \tilde{Q} = \text{Norm}((\text{Linear}(X''_{\text{in}}))) \\ \tilde{K}, \tilde{V} = \text{Norm}((\text{Linear}(Z_i))) \\ X'_{\text{out}} = \text{Softmax}(\dfrac{\tilde{Q} \times \tilde{K}^T}{\sqrt{C_0}}) \cdot \tilde{V} \end{cases}, \tag{6}$$

where the $\text{Norm}(\cdot)$ and $\text{Linear}(\cdot)$ is the regularization function and linear transformation, respectively. C_0 is denotes channel dimension. For the feature fusion with the SSM, as shown in Fig. 2, the process is formulated as:

$$\begin{cases} \bar{X}_{out} = \text{Linear}(X'_{\text{out}} + X'') \\ X_{11} = \text{Norm}(\text{SSM}(\text{SiLU}(\text{DwConv}(\bar{X}_{out})))) \\ X_{12} = \text{SiLU}(\bar{X}_{out}) \\ X''_{\text{out}} = \text{LinearEwPro}(X_{11}, X_{12}) \end{cases}, \tag{7}$$

where $\text{DwConv}(\cdot)$, $\text{SSM}(\cdot)$, $\text{SiLU}(\cdot)$, and $\text{EwPro}(\cdot)$ denote deep-wise convolution, our SSM module, SiLU activation, and element-wise product, respectively.

CL-DM For Guiding Deblurring

As shown in Fig.3, we employ a conditional latent diffusion model (CL-DM) to learn and generate latent features of concentration distribution of MNPs. We adopt existing conditional denoising diffusion probabilistic models as our CL-DM (Chen et al. 2024). The model consists of two main processes: a forward diffusion process and a reverse denoising process. During the forward diffusion process, given an isotropic MPI image (ground truth), posterior $q(\mathbf{z}, t)$, and the latent feature $\mathbf{z} \in \mathbb{R}^{N\times C_1}$. the process can be defined as:

$$\begin{cases} q(\mathbf{z}_{1:T} \mid \mathbf{z}_0) = \prod_{t=1}^{T} q(\mathbf{z}_t \mid \mathbf{z}_{t-1}) \\ q(\mathbf{z}_t \mid \mathbf{z}_{t-1}) = \mathcal{N}(\mathbf{z}_t; \sqrt{1-\beta_t}\mathbf{z}_{t-1}, \beta_t\mathbf{I}) \end{cases}, \tag{8}$$

where $\mathbf{z}_t$ denotes the noisy features at the t-th step. T is the total number of steps. The parameters $\beta_{1:T} \in (0, 1)$ control the noise variance; $\mathcal{N}(\cdot)$ denotes the Gaussian distribution. By reparametrizing and deriving iteratively, the above equation can be reformulated as:

$$q(\mathbf{z}_t \mid \mathbf{z}_0) = \mathcal{N}(\mathbf{z}_t; \sqrt{\bar{\alpha}_t}\mathbf{z}_0, (1-\bar{\alpha}_t)\mathbf{I}), \tag{9}$$

$$\bar{\alpha}_t = \prod_{i=1}^{t} \bar{\alpha}_i, \tag{10}$$

where $\alpha_i = 1 - \beta_i$. We perform the reverse denoising process to generate the feature $\mathbf{z}_0$ from a conditional $\mathbf{z}_T$. Referring to (Chen et al. 2024), to estimate noise ϵ in the posterior distribution, we employ a neural network (denoted as Net_θ) as mentioned in the work. To perform DM in the latent space, we introduce another CNN-based latent feature encoder. This encoder compresses the anisotropic image $\mathbf{I}_{\text{aniso}}$ into the latent space, producing the conditional latent feature $\mathbf{c} \in \Phi^{N\times C'}$. The deblurring network $\text{Net}_\theta(\mathbf{z}_t, \mathbf{c}, t)$ predicts the blurry noise ϵ conditioned on $\mathbf{z}_t$, $\mathbf{c}$, and t. By predicting ϵ with the network Net_θ into the posterior mean and setting the variance to $(1-\alpha_t)$, the reverse process is rewritten as:

$$\mathbf{z}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{z}_t - \frac{1-\alpha_t}{\sqrt{1-\alpha_t}}\text{Net}_\theta(\mathbf{z}_t, \mathbf{c}, t)\right) + \mathbf{z}_{\epsilon_t}, \tag{11}$$

where $\mathbf{z}_{\epsilon_t} = \sqrt{1-\alpha_t}\epsilon_t$, $\epsilon_t \sim \mathcal{N}(0, \mathbf{I})$. Iteratively applying this equation for T steps, we can generate the predicted feature $\mathbf{Z}_i \in \Phi^{N\times C_1}$.

As shown in Fig. 2, the generated latent features are then fed into the different LFF-SSM blocks for MPI image calibration and deblurring. Notably, since the latent space $\Phi^{N\times C_1}$ has a simpler magnetic particle signal distribution compared to the image space $\Phi^{H\times W\times C}$ (Chen et al. 2024), generating latent features requires fewer iterations.

Training the diffusion model involves optimizing the denoising network Net_θ. Here, referring to previous work (Kawar et al. 2022), the model is trained by minimizing a weighted variational bound. The loss function is defined as:

$$\|\nabla_\theta\epsilon - \epsilon_\theta(\sqrt{\bar{\alpha}_t}\mathbf{z} + \sqrt{1-\alpha_t}\epsilon, \mathbf{c}, t)\|_2^2. \tag{12}$$

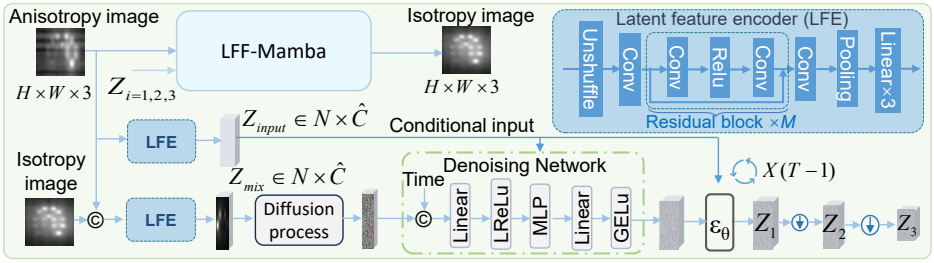

Figure 3: The Overview framework of conditional latent diffusion model.

For model training, we exploit a joint training strategy for the CL-DM and LFF-SSM. During each training iteration, the extracted feature $\mathbf{z}$ is used to generate the noisy sample $\mathbf{z}_t$ using $q(\mathbf{z}_t \mid \mathbf{z}_0) = \mathcal{N}(\mathbf{z}_t; \sqrt{\bar{\alpha}_t}\mathbf{z}_0, (1-\bar{\alpha}_t)\mathbf{I})$. Given the small number of iterations T in the latent space, we use the denoising network to predict the latent feature $\hat{\mathbf{z}}$. Then, this predicted latent distribution is used to guide the feature fusion via the LFF-SSM block. The total training loss is $\mathcal{L}_{\text{total}} = \mathcal{L}_{P1} + \mathcal{L}_{P2}$, where $\mathcal{L}_{P1} = \|\mathbf{I}_{\text{aniso}} - \mathbf{I}_{\text{label}}\|_1$ and $\mathcal{L}_{P2} = \|\mathbf{z} - \hat{\mathbf{z}}\|_1$.

Materials and Experiments

We formulate this part as: 1) Diverse dataset design and construction. 2) Experimental setup: Key implementation details and the computational environment are outlined to ensure reproducibility. 3) Competing methods: Several state-of-the-art methods are selected as benchmarks for performance evaluation.

Diverse Dataset Design and Construction

Simulation Dataset In this study, we simulate 10,000 anisotropic images for model evaluation. We adopt the Langevin equation to simulate the distribution of MNPs in MPI. Our simulated MPI anisotropic images are generated from digital binary images (Shang et al. 2023).

Real-world Phantom Dataset To construct a real-world data for evaluating our method, we spend more than three years for phantom design (Fig. 4) and data scanning with our MPI device (Fig. 5). We conducted experiments using around 1,000 customized phantoms and more than 200 mice. The phantoms are fabricated using 3D printing technology and were made from resin materials. To introduce the MNPs into the phantoms, we used commercially available tracers of MNPs of Perimag or Synomag. The MNPs were inserted into the phantoms with different concentration for different phantoms (1-10 mg/ml). The phantoms were then scanned using a commercial MPI scanner (MOMENTUM CT Magnetic Insight Inc., Alameda, CA, USA). The scanner utilized its proprietary software to generate the native images by scanning the prepared phantoms or mice.

In this study, as shown in Fig. 4, we design different diverse phantoms (sensitivity, vessel, resolution, shape) to construct first diverse real-world dataset. We scan and collect eligible 20,156 real-world MPI anisotropic images (we select eligible data that should scan with the mode of

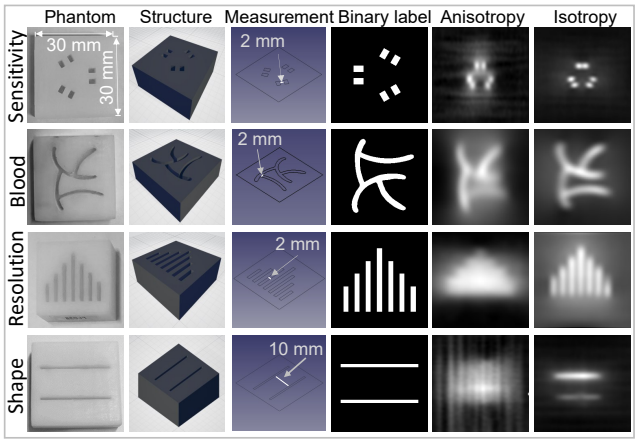

Figure 4: Real MPI data construction with diverse design of real phantoms. 1) Real phantom. It refers to custom-designed each phantom for diverse MPI datasets, focusing on resolution, sensitivity, vascular structures, and shapes. 2) Structure design. It shows each case 3D diagrams of these phantoms. 3) Measurement. Its details specific dimensions for algorithm validation. 4) Binary label. It visualizes the structure of each phantom in 2D. 5) Anisotropy image. This high sensitivity mode of MPI system offers quick, sensitive scans in about 6 minutes with low resolution. 6) Isotropy image. This high-resolution mode provides high-resolution images in about 12 minutes.

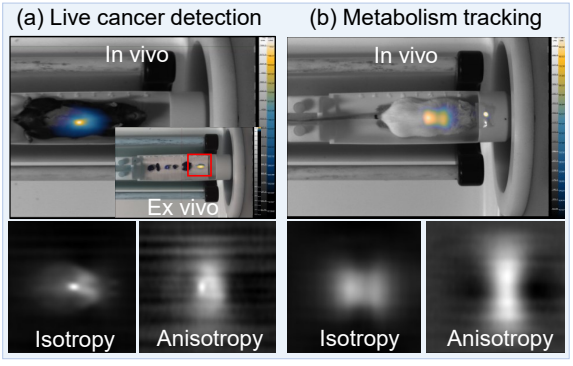

Figure 5: Biological MPI Images scanning with mice.

Isotropy image with anisotropic images in X and Z channel and their ground truth of isotropic images) using the commercial MPI scanner.

Real-world Biological Dataset We construct real-world iological MPI image database using synthetic imaging probes with superparamagnetic iron oxide nanoparticles (Fe_3O_4) injected into mice, such as tumor detection (Zhang et al. 2022), vessel plaque detection (Tong et al. 2021). The specific designs include: 1) Imaging diversity: our data is collected for different tasks, such as intravascular particle detection, tumor targeting, and metabolism tracking. As shown in Fig. 4, we show three examples for data design and acquisition. 2) Imaging application types: Examples include liver cancer detection and metabolic tracking. 3) Data structure: The database encompasses in vivo imaging, multi-timepoint metabolic tracking, and ex vivo assessments. 4) Data scanning parameters setting: We maintain consistency with the setting for phantom data scanning, obtaining anisotropic images through unidirectional scanning.

Experimental Setup Our experiments are conducted using the PyTorch framework. We employ four NVIDIA V100 GPUs for training and validation. The learning rate is empirically set between 2×10^{-4} and 10^{-3}. We utilize the AdamW optimizer along with a cosine learning rate schedule to ensure convergence. Each training stage consists of 100,000 iterations with a batch size of 32. During training, images are cropped as 64×64 patches. We perform data augmentation (random rotations and flips) for training. We split our simulation dataset and real-world dataset as training, validation, and test dataset with a ratio of 8:1:1, respectively. The iteration is experimentally set as $T = 8$ for CL-DM. We adopted three widely used metrics to comprehensively assess the calibration performance: 1) Normalized Root Mean Square Error (nRMSE); 2) Peak Signal-to-Noise Ratio (PSNR); 3) Structural Similarity Index Measure (SSIM).

Competing Methods We evaluate our method against competing methods including:

1. AEPNet (Shang et al. 2023): This CNN-based method addresses the anisotropic resolution of MPI by constructing an asymmetric convolution network.

2. FDS (Shang et al. 2022) : A CNN-based approach that improves the spatial resolution of MPI by fusing a dual-sampling convolutional neural network.

3. SFM (Shang et al. 2024): This transformer-based method is based on simulation MPI image data, which restores both spatial and frequency domain features of native MPI images.

4. NAFNet (Chen et al. 2022): A simple and computationally efficient baseline for natural image restoration. It achieves SOTA results in tasks like image deblurring and denoising.

5. Restormer (Zamir et al. 2022a): An efficient Transformer model designed for natural image tasks. The Restormer achieves SOTA results on various image restoration tasks.

6. SwinIR (Liang et al. 2021): This is a transformer-based method for image restoration tasks. The SwinIR achieves SOTA results on various image restoration tasks.

7. DiffIR (Xia et al. 2023): A transformer-based Diffusion Model designed for image deblurring. The method achieves SOTA results on real-world image deblurring tasks.

Results

Results for Simulation Data

Firstly, we conducted comprehensive evaluations of our proposed method against state-of-the-art approaches for MPI image calibration with simulation data. As shown in Table 1, our method demonstrates superior performance across all evaluation metrics.

Method	Metric		
	PSNR	SSIM	nRMSE
AEPNet (Shang et al. 2023)	26.4488	0.9515	0.0664
FDS (Shang et al. 2022)	26.4364	0.9501	0.0653
SFM (Shang et al. 2024)	26.6631	0.9571	0.0654
NAFNet (Chen et al. 2022)	24.5609	0.8511	0.0948
Restormer (Zamir et al. 2022a)	26.5548	0.9546	0.0663
SwinIR (Liang et al. 2021)	26.5528	0.9503	0.0669
DiffIR (Xia et al. 2023)	26.6604	0.9542	0.0651
Ours	**29.4758**	**0.9684**	**0.0495**

Table 1: Performance evaluation on simulation data.

Results for Real-world Data

As shown in Table 2, our results demonstrate that our method outperforms existing state-of-the-art approaches. Compared to CNN-based methods, our method improves PSNR by 1.9-8.1 dB and SSIM by 3.0-13.0%. We also depict the error maps to visualize the differences between the calibrated images and corresponding labels for all methods. In terms of shape and resolution phantoms, the images restored using our method exhibit clear edge information and structural details. We have conducted 5-seed experiments showing robustness (PSNR: std=0.08 dB; SSIM: std=4e-3; nRMSE: std=3e-4).

Ablation Study

Impact of Different Proportion of Real-world Data We examined the effects of gradually integrating real-world MPI image data into simulations. As shown in Table 3, as the proportion of real data samples increases, the model's predictive performance also improves continuously. It was found that using just 10% of the real data could improve PSNR by 9.0 dB and increase SSIM by over 70%. These findings indicate that real-world data is significant for improving model accuracy.

Impact of Each Core Component As shown in Table 4, the backbone, without these modules, achieved a little poor performance. Incorporating the Mamba module, the model show better performance (better PSNR by 4.9 dB and SSIM by 33%). The addition of the DM module further enhanced performance (better PSNR by 4.9 dB and SSIM by 32%).

Method	Metric		
	PSNR	SSIM	nRMSE
AEPNet (Shang et al. 2023)	23.7242	0.8234	0.0910
FDS (Shang et al. 2022)	22.7843	0.8249	0.1320
SFM (Shang et al. 2024)	23.0343	0.8323	0.1074
NAFNet (Chen et al. 2022)	17.4842	0.7393	0.1683
Restormer (Zamir et al. 2022a)	21.6697	0.7656	0.1148
SwinIR (Liang et al. 2021)	24.3335	0.8425	0.0913
DiffIR (Xia et al. 2023)	24.2758	0.8495	0.0874
Ours	**25.6590**	**0.8700**	**0.0763**

Table 2: Performance evaluation on real-world MPI data.

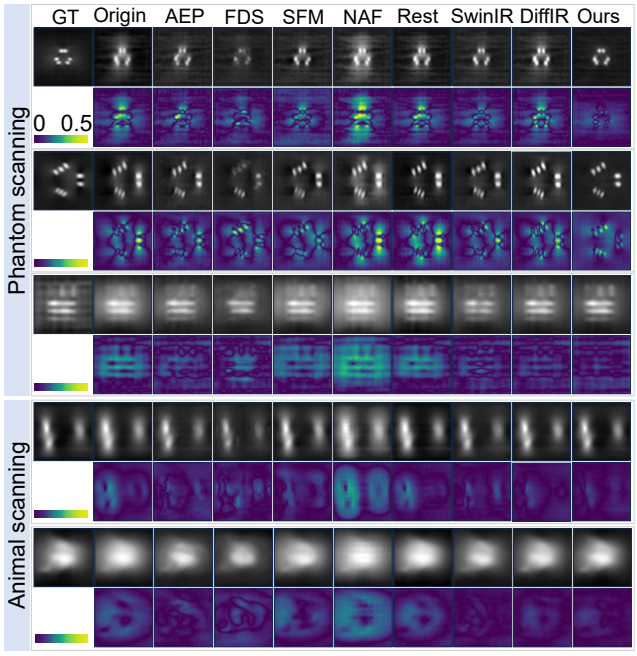

Figure 6: Performance evaluation of all methods for calibration of different anisotropic images with error maps. AEP: AEPNet, NAF: NAFNet, Rest: Restormer.

The combined integration of both Mamba and DM modules yielded the best results (better PSNR by 5.4 dB and SSIM by 34%). As shown in Fig. 7, the visualized results also confirm our results.

Impact of patch size As shown in Table 5, we evaluated our model based on the different patch sizes in real-world data. The results show that our method exhibits good long-range learning capabilities.

Conclusion

We proposed a novel MPI-Mamba method to achieve anisotropic MPI image calibration and deblurring. Extensive experimental results confirm that our method outperforms state-of-the-art techniques on both simulation and real-world MPI datasets. Hence, our method is a potential tool to offer promising directions for future clinical applications and providing a available framework for MPI image calibration and resolution enhancement.

Metric	Proportion of real-world data size					
	0%	10%	30%	50%	80%	100%
PSNR	15.20	24.63	24.64	25.26	25.57	25.68
SSIM	0.105	0.846	0.859	0.864	0.873	0.870
nRMSE	0.277	0.090	0.082	0.082	0.078	0.076

Table 3: Impact of proportion of real-world MPI data size.

Backbone	Mamba	DM	SSIM	PSNR	nRMSE
√			0.531	20.202	0.126
√	√		0.860	25.120	0.081
√		√	0.855	25.129	0.078
√	√	√	0.870	25.659	0.076

Table 4: Impact of each core component.

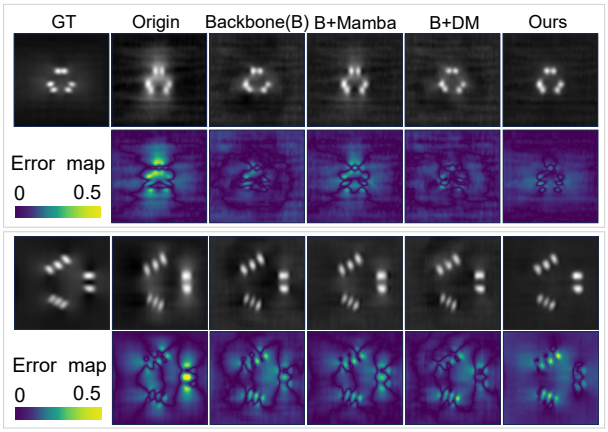

Figure 7: Performance evaluation for each component.

Method	Metric	Patch size		
		16	32	64
Ours	PSNR	20.212	23.437	25.659
	SSIM	0.822	0.830	0.870
	nRMSE	0.135	0.089	0.076

Table 5: Impact of different patch size.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 62576337, 82202269, 62027901, and in part by the National Key Research and Development Program for Young Scientists under Grant 2022YFC2505700.

References

Billings, C.; Langley, M.; Warrington, G.; Mashali, F.; and Johnson, J. A. 2021. Magnetic particle imaging: current and future applications, magnetic nanoparticle synthesis methods and safety measures. *International Journal of Molecular Sciences*, 22(14): 7651.

Chen, K.; and Liu, Y. 2024. Efficient image deblurring networks based on diffusion models. *arXiv preprint arXiv:2401.05907*.

Chen, L.; Chu, X.; Zhang, X.; and Sun, J. 2022. Simple baselines for image restoration. In *European conference on computer vision*, 17–33. Springer.

Chen, Z.; Zhang, Y.; Liu, D.; Gu, J.; Kong, L.; Yuan, X.; et al. 2024. Hierarchical integration diffusion model for realistic image deblurring. *Advances in neural information processing systems*, 36.

Chung, H.; Lee, E. S.; and Ye, J. C. 2022. MR image denoising and super-resolution using regularized reverse diffusion. *IEEE transactions on medical imaging*, 42(4): 922–934.

Chung, H.; and Ye, J. C. 2022. Score-based diffusion models for accelerated MRI. *Medical image analysis*, 80: 102479.

Goodwill, P. W.; and Conolly, S. M. 2010. The X-space formulation of the magnetic particle imaging process: 1-D signal, resolution, bandwidth, SNR, SAR, and magnetostimulation. *IEEE transactions on medical imaging*, 29(11): 1851–1859.

Gu, A.; and Dao, T. 2023. Mamba: Linear-time sequence modeling with selective state spaces. *arXiv preprint arXiv:2312.00752*.

Güngör, A.; Askin, B.; Soydan, D. A.; Saritas, E. U.; Top, C. B.; and Çukur, T. 2022. TranSMS: Transformers for super-resolution calibration in magnetic particle imaging. *IEEE Transactions on Medical Imaging*, 41(12): 3562–3574.

Güngör, A.; Askin, B.; Soydan, D. A.; Top, C. B.; Saritas, E. U.; and Çukur, T. 2023. DEQ-MPI: A deep equilibrium reconstruction with learned consistency for magnetic particle imaging. *IEEE Transactions on Medical Imaging*, 43(1): 321–334.

Guo, H.; Li, J.; Dai, T.; Ouyang, Z.; Ren, X.; and Xia, S.-T. 2024. Mambair: A simple baseline for image restoration with state-space model. In *European conference on computer vision*, 222–241. Springer.

Huang, X.; Hui, H.; Shang, W.; Gao, P.; Zhou, Y.; Pang, W.; Woo, C. M.; Tian, J.; and Lai, P. 2023. Deep penetrating and sensitive targeted magnetic particle imaging and photothermal therapy of early-stage glioblastoma based on a biomimetic nanoplatform. *Advanced Science*, 10(19): 2300854.

Ilbey, S.; Top, C. B.; Güngör, A.; Çukur, T.; Saritas, E. U.; and Güven, H. E. 2019. Fast system calibration with coded calibration scenes for magnetic particle imaging. *IEEE transactions on medical imaging*, 38(9): 2070–2080.

Kawar, B.; Elad, M.; Ermon, S.; and Song, J. 2022. Denoising diffusion restoration models. *Advances in Neural Information Processing Systems*, 35: 23593–23606.

Kitaev, N.; Kaiser, Ł.; and Levskaya, A. 2020. Reformer: The efficient transformer. *arXiv preprint arXiv:2001.04451*.

Liang, J.; Cao, J.; Sun, G.; Zhang, K.; Van Gool, L.; and Timofte, R. 2021. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF international conference on computer vision*, 1833–1844.

Liu, H.; Liu, C.; Xu, J.; Jiang, P.; and Lu, M. 2025. XYScanNet: A State Space Model for Single Image Deblurring. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, 779–789.

Liu, J.; Wang, Q.; Fan, H.; Wang, Y.; Tang, Y.; and Qu, L. 2024a. Residual denoising diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2773–2783.

Liu, J.; Yang, H.; Zhou, H.-Y.; Yu, L.; Liang, Y.; Yu, Y.; Zhang, S.; Zheng, H.; and Wang, S. 2024b. Swin-UMamba†: Adapting Mamba-based vision foundation models for medical image segmentation. *IEEE Transactions on Medical Imaging*.

Mao, Y.; Jiang, L.; Chen, X.; and Li, C. 2023. Disc-diff: Disentangled conditional diffusion model for multi-contrast mri super-resolution. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 387–397. Springer.

Mohn, F.; Knopp, T.; Boberg, M.; Thieben, F.; Szwargulski, P.; and Graeser, M. 2022. System matrix based reconstruction for pulsed sequences in magnetic particle imaging. *IEEE transactions on medical imaging*, 41(7): 1862–1873.

Panagiotopoulos, N.; Duschka, R. L.; Ahlborg, M.; Bringout, G.; Debbeler, C.; Graeser, M.; Kaethner, C.; Lüdtke-Buzug, K.; Medimagh, H.; Stelzner, J.; et al. 2015. Magnetic particle imaging: current developments and future directions. *International journal of nanomedicine*, 3097–3114.

Qiao, J.; Liao, J.; Li, W.; Zhang, Y.; Guo, Y.; Wen, Y.; Qiu, Z.; Xie, J.; Hu, J.; and Lin, S. 2024. Hi-mamba: Hierarchical mamba for efficient image super-resolution. *arXiv preprint arXiv:2410.10140*.

Ren, M.; Delbracio, M.; Talebi, H.; Gerig, G.; and Milanfar, P. 2023. Multiscale structure guided diffusion for image deblurring. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 10721–10733.

Shang, Y.; Liu, J.; Liu, Y.; Wang, Y.; Shen, Y.; Wu, X.; Zhang, L.; Hui, H.; and Tian, J. 2024. Spatial-frequency multi-scale transformer for deblurring and shape-preserving reconstruction in magnetic particle imaging. *IEEE Transactions on Computational Imaging*, 10: 196–207.

Shang, Y.; Liu, J.; Liu, Y.; Zhang, B.; Wu, X.; Zhang, L.; Tong, W.; Hui, H.; and Tian, J. 2023. Anisotropic edge-preserving network for resolution enhancement in unidirectional Cartesian magnetic particle imaging. *Physics in Medicine & Biology*, 68(4): 045014.

Shang, Y.; Liu, J.; Zhang, L.; Wu, X.; Zhang, P.; Yin, L.; Hui, H.; and Tian, J. 2022. Deep learning for improving the spatial resolution of magnetic particle imaging. *Physics in Medicine & Biology*, 67(12): 125012.

Shi, G.; Yin, L.; An, Y.; Li, G.; Zhang, L.; Bian, Z.; Chen, Z.; Zhang, H.; Hui, H.; and Tian, J. 2023. Progressive pretraining network for 3D system matrix calibration in magnetic particle imaging. *IEEE Transactions on Medical Imaging*, 42(12): 3639–3650.

Tong, W.; Hui, H.; Shang, W.; Zhang, Y.; Tian, F.; Ma, Q.; Yang, X.; Tian, J.; and Chen, Y. 2021. Highly sensitive magnetic particle imaging of vulnerable atherosclerotic plaque with active myeloperoxidase-targeted nanoparticles. *Theranostics*, 11(2): 506.

Tong, W.; Zhang, Y.; Hui, H.; Feng, X.; Ning, B.; Yu, T.; Wang, W.; Shang, Y.; Zhang, G.; Zhang, S.; et al. 2023. Sensitive magnetic particle imaging of haemoglobin degradation for the detection and monitoring of intraplaque haemorrhage in atherosclerosis. *EBioMedicine*, 90.

Tsai, F.-J.; Peng, Y.-T.; Lin, Y.-Y.; Tsai, C.-C.; and Lin, C.-W. 2022. Stripformer: Strip transformer for fast image deblurring. In *European conference on computer vision*, 146–162. Springer.

Xia, B.; Zhang, Y.; Wang, S.; Wang, Y.; Wu, X.; Tian, Y.; Yang, W.; and Van Gool, L. 2023. Diffir: Efficient diffusion model for image restoration. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 13095–13105.

Yang, F.; Zhang, L.; Zhang, J.; Wei, Z.; Yang, X.; Tian, J.; and Hui, H. 2024. Multi-Scale Adaptive Transformer for Image Deblurring in Magnetic Particle Imaging. In *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, 1–5. IEEE.

Zamir, S. W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F. S.; and Yang, M.-H. 2022a. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 5728–5739.

Zamir, S. W.; Arora, A.; Khan, S.; Hayat, M.; Khan, F. S.; Yang, M.-H.; and Shao, L. 2022b. Learning enriched features for fast image restoration and enhancement. *IEEE transactions on pattern analysis and machine intelligence*, 45(2): 1934–1948.

Zhang, W.; Liang, X.; Zhu, L.; Zhang, X.; Jin, Z.; Du, Y.; Tian, J.; and Xue, H. 2022. Optical magnetic multimodality imaging of plectin-1-targeted imaging agent for the precise detection of orthotopic pancreatic ductal adenocarcinoma in mice. *EBioMedicine*, 80.

Zhu, L.; Liao, B.; Zhang, Q.; Wang, X.; Liu, W.; and Wang, X. 2024. Vision mamba: efficient visual representation learning with bidirectional state space model. In *Proceedings of the 41st International Conference on Machine Learning*, ICML'24. JMLR.org.