Review

Diffusion Models for MRI Reconstruction: A Systematic Review of Standard, Hybrid, Latent and Cold Diffusion Approaches

Marija Habijan[1], Zdravko Krpić[1], Juraj Perić[1], Irena Galić[1]

[1] J. J. Strossmayer University, Faculty of Electrical Engineering, Computer Science and Information Technology, 31000 Osijek, Croatia; marija.habijan@ferit.hr

* Correspondence: marija.habijan@ferit.hr

Abstract

Diffusion models have recently emerged as a new paradigm to solve inverse problems in magnetic resonance imaging (MRI). This paper systematically categorizes forty representative studies that use diffusion models for MRI reconstruction. The studies are systematically divided into three dimensional categories: (1) model type distribution that includes standard, score-based diffusion, hybrid and physics-informed diffusion, latent diffusion and cold diffusion, (2) application domain distribution that includes accelerated, motion-corrected, dynamic, quantitative and cross-modal MRI and (3) domain of reconstruction distribution that includes image-space, k-space, latent-space and hybrid approaches. In this context, diffusion-driven methods represent a significant methodological shift, evolving from purely data-driven denoising toward unified, physics-aligned reconstruction frameworks. Hybrid and cold diffusion approaches further integrate probabilistic generative priors with explicit MRI acquisition physics, which enables improved fidelity, artifact suppression and enhanced generalization across sampling regimes and anatomical domains. This study reviews current advances in the field of MRI reconstruction methods and gives future research directions, emphasizing the need for interpretable, data-efficient and multi-domain generative models that are capable of robustly addressing the diverse challenges of MRI reconstruction.

Keywords: Magnetic Resonance Imaging (MRI); Diffusion Models; Hybrid Diffusion Models; Latent Diffusion Models; Cold Diffusion Models; MRI Reconstruction

Received:
Revised:
Accepted:
Published:

Citation: Habijan, M.; Matić, K. ; Perić, M.; Galić, I. Diffusion Models for MRI Reconstruction: A Systematic Review of Standard, Hybrid, Latent and Cold Diffusion Approaches. *Electronics* **2025**, *1*, 0. https://doi.org/

1. Introduction

Magnetic Resonance Imaging (MRI) is one of the most powerful tools in modern medical diagnostics. It enables high soft-tissue contrast and non-invasive visualization of anatomical and functional processes [1]. MRI has the ability to differentiate fine structural and biochemical details, which is necessary for neurological, cardiovascular, oncological and musculoskeletal applications [2]. However, the greatest limitation of MRI is in its long acquisition time, a consequence of the need to sample k-space data sequentially. Extended scan durations not only reduce patient comfort and scanner throughput, but also increase the likelihood of motion-induced artifacts that compromise image quality and diagnostic reliability [3]. Overcoming this trade-off between scan time and image fidelity remains a big challenge in MRI reconstruction field.

Early efforts to accelerate MRI focused on acquisition-side techniques such as Parallel Imaging (PI), including SENSE and GRAPPA methods, which exploit multi-coil sensitivity to reduce k-space sampling [4,5]. Later, Compressed Sensing (CS) transformed MRI reconstruction by enabling image recovery from sparse data under the assumption of signal sparsity in a transform domain [3]. Combining PI and CS significantly shortened scan times and formed the basis for many clinical systems. However, both rely on handcrafted priors and iterative optimization, which can be slow and prone to artifacts at high acceleration rates.

Machine and deep learning have transformed MRI reconstruction by enabling models to learn data-driven priors. Convolutional neural networks (CNNs) and model-based architectures such as Generative Adversarial Networks (GANs), variational networks, unrolling optimization models and transformer-based approaches achieve higher reconstruction quality and efficiency than traditional methods [6,7]. Yet, most deep networks yield a single deterministic output which limits uncertainty representation when data are highly undersampled or corrupted. Recently, diffusion probabilistic models (DPMs) have shown remarkable success in solving computer vision tasks and are now rapidly advancing medical imaging [8,9]. By iteratively adding and removing noise, they model the full data distribution and produce realistic, uncertainty-aware reconstructions. Hybrid approaches combine diffusion priors with physics-based constraints [10], while cold diffusion models replace Gaussian noise with domain-specific degradations [11,12]. This aligns the model more closely with real acquisition physics. Previously mentioned advances have improved reconstruction fidelity, efficiency and physical realism, extending diffusion modeling to dynamic, motion-corrected and cross-modal imaging tasks [13,14].

This review aims to provide a systematic analysis of recent advances in diffusion-based models for MRI reconstruction. The primary objectives of this work are:

- To systematically classify and analyze the diffusion-based MRI reconstruction studies under (1) model type distribution, which includes standard, hybrid, latent and cold diffusion approaches, (2) application domain distribution, which includes accelerated, motion-corrected, dynamic, quantitative and multimodal, and cross-multimodal MRI approaches, (3) reconstruction domain, which includes the image-space, k-space, hybrid and latent-space approaches.
- To evaluate strengths, limitations and open challenges of diffusion based approaches and recommendations for future research directions.
- To highlight the potential of diffusion-based generative models for accurate and clinically trustworthy MRI reconstruction.

The remainder of this paper is organized as follows. Section 2 provides information regarding the research methodology. Section 3 classifies all included studies according to three dimensions: model type distribution, application domain distribution and reconstruction domain distribution. Section 3.1 reviews the model type distribution across standard diffusion approaches, hybrid and physics-informed diffusion approaches, latent diffusion approaches and cold diffusion approaches. Section 3.2 analyzes different application domains, including accelerated, motion-corrected, dynamic, quantitative and multidimensional as well as cross-modal MRI reconstructions. Section 3.3 focuses on the domain of reconstruction, including image-space, k-space, hybrid, and latent formulations. Section 4 discusses reviewed approaches, highlights challenges, future research directions and concludes the paper.

2. Review Methodology

In September 2025, we performed a comprehensive literature search across multiple databases in order to identify studies relevant to diffusion-based approaches for MRI reconstruction. The search was preformed using Web of Science platform. The databases included are listed in Table 1. The search strategy was defined using topic search (TS) terms designed to capture works focusing on diffusion models in the context of MRI reconstruction. The complete query was formulated as follows: TS = ((cold diffusion OR diffusion model OR denoising diffusion OR score-based generative model OR score-based diffusion OR cold diffusion model) AND (magnetic resonance imaging OR MRI OR k-space OR undersampled MRI OR accelerated MRI) AND (reconstruction OR image reconstruction OR inverse problem OR deep learning reconstruction OR compressed sensing OR MRI restoration OR MRI reconstruction)).

Table 1. Summary of databases and search terms used in the systematic review.

Database	Search terms
Science Citation Index Expanded	cold diffusion, diffusion model, denoising diffusion, score-based generative model, score-based diffusion, cold diffusion model, magnetic resonance imaging, MRI, k-space, undersampled MRI, accelerated MRI, reconstruction, image reconstruction, inverse problem, deep learning reconstruction, compressed sensing, MRI restoration, MRI reconstruction
Emerging Sources Citation Index	
MEDLINE	
Data Citation Index	
Conference Proceedings Citation Index	
KCI–Korean Journal Database	
BIOSIS Citation Index	
Derwent Innovations Index	
SciELO Citation Index	
Russian Science Citation Index	

After removing duplicates, 303 records were screened. Of these, 214 were excluded for being review papers, conference abstracts, or datasets rather than original research. A total of 89 records were used for retrieval and 89 were assessed for eligibility. Further, 49 works were excluded as not related to MRI reconstruction with diffusion models since they where mainly focused on image synthesis and generation tasks. Finally, 40 studies that met the inclusion criteria were analyzed in this review. The PRISMA flow diagram illustrating the study selection process is shown in Figure 1.

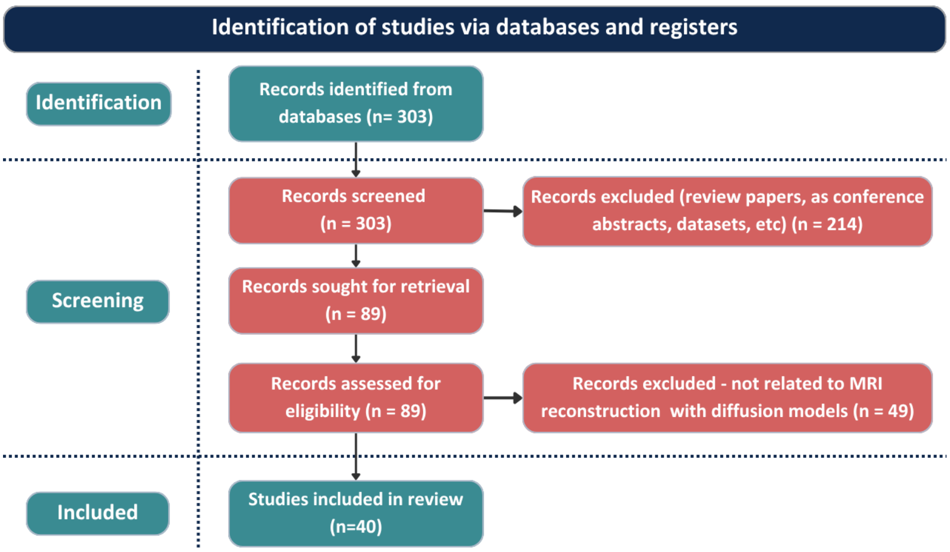

Figure 1. PRISMA flow diagram illustrating the identification, screening and inclusion of studies.

Early work in accelerated MRI reconstruction, before 2010, was largely based on traditional optimization and CS methods that relied on hand-crafted priors, such as sparsity and total variation. Between 2010 and 2015, model-driven techniques incorporating MRI physics, such as PI and low-rank modeling, became the standard. From 2015 to 2020, deep learning approaches, including convolutional and unrolling networks, brought major improvements in reconstruction accuracy, robustness and computational efficiency. More recently, diffusion and cold diffusion models have emerged as a new generation of reconstruction techniques. Studies published between 2022 and 2025 demonstrate a clear shift toward physics-aligned, diffusion-based approaches. The temporal evolution and methodological trends of these studies are summarized in Figure 2.

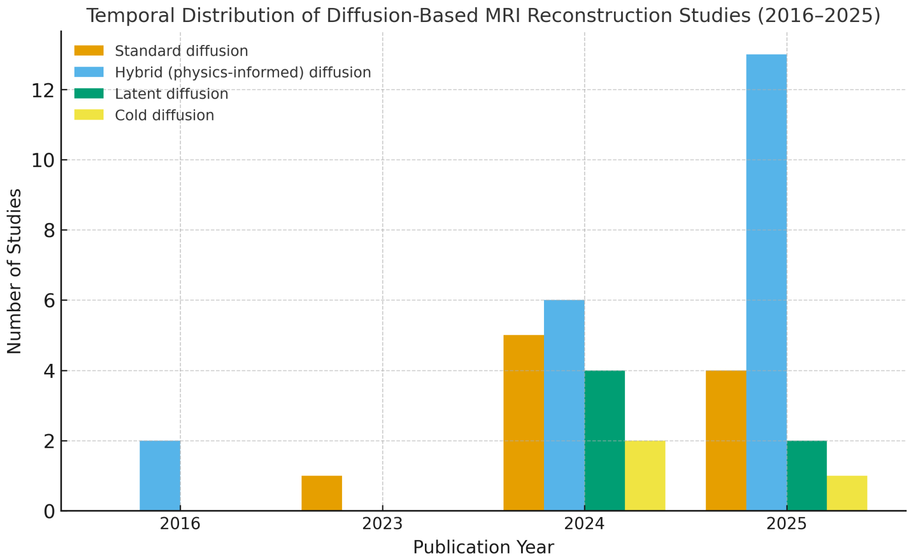

Figure 2. The temporal distribution of 40 reviewed studies categorized by model type. A clear increase in publications after 2023 highlights the rapid expansion of the field.

3. MRI Reconstruction Methods

The reviewed literature shows that diffusion-based MRI reconstruction has rapidly evolved in recent years. As generative modeling techniques have matured, different methodological directions have emerged, each addressing specific reconstruction challenges and levels of physical modeling. To provide a structured synthesis of these advances, this review systematically classifies all included studies according to three complementary dimensions: model type distribution, application domain distribution and reconstruction domain distribution as shown in Figure 3. Such categorization enables a clear comparison of methodological trends and highlights how diffusion approaches differ in mathematics and theoretical formulation, practical application and integration. In the following sections, we discuss the distribution of MRI reconstruction approaches across the reviewed literature and outline recent and dominant design paradigms in this research area.

3.1. Model Type Distribution

Diffusion-based generative modeling is grounded in non-equilibrium thermodynamics and stochastic differential equations (SDEs) that describe how a data distribution gradually transforms into a tractable noise distribution and back. In the forward, i.e. diffusion process, clean data $\mathbf{x}_0$ are progressively corrupted through a Markov chain:

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\sqrt{1-\beta_t}\mathbf{x}_{t-1}, \beta_t \mathbf{I}),$$

where β_t controls the noise variance schedule. In the reverse process, a neural network parameterized as a score function $s_\theta(\mathbf{x}_t, t)$ learns to approximate the gradient of the log data density $\nabla_{\mathbf{x}_t} \log p(\mathbf{x}_t)$, allowing sampling from $p_\theta(\mathbf{x}_0)$ via iterative denoising.

However, hybrid approaches have further extended the diffusion approaches in MRI reconstruction by combining several architectures (such as GANs, Transformers or Mamba

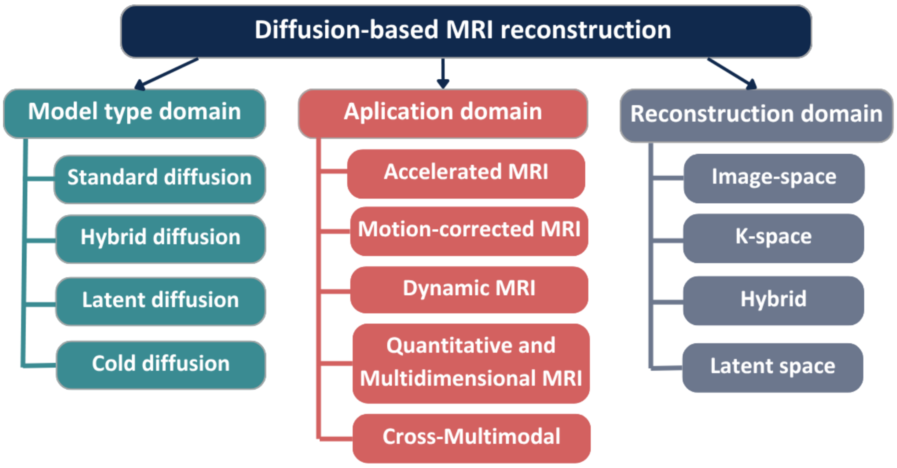

Figure 3. Illustration of classification scheme of diffusion-based MRI reconstruction studies. All reviewed studies are systematically categorized along three dimensions: (1) model type (standard, hybrid, latent and cold diffusion), (2) application domain (accelerated MRI, motion-corrected, dynamic, quantitative/multidimensional and cross-modal MRI tasks) and (3) reconstruction domain (image-space, k-space, hybrid and latent-space approaches).

networks) to provide improved perceptual realism, attention-driven feature fusion and temporal consistency. Physics-informed and model-driven diffusion strategies embed explicit MRI forward operators $\mathcal{A}$ and data consistency terms $|\mathcal{A}(\mathbf{x}) - \mathbf{y}|^2$ within each diffusion iteration that ensures fidelity to observed data, which is correlated with reconstruction stability. Based on these advances, latent diffusion models (LDMs) adapt the diffusion process into a compact latent space $\mathcal{Z}$ computed from an autoencoder $E(\cdot)$, where the encapsulated representation $z_0 = E(\mathbf{x}_0)$ contains significant structural and semantic information. This decrease in dimensionality allows efficient learning of global dependencies with fine anatomical details. Furthermore, the new cold diffusion paradigm generalizes traditional diffusion models by replacing random Gaussian corruption with deterministic or domain-specific degradation, defined as: $D_t(\mathbf{x}_0)$ which can be k-space undersampling and image blurring. The reverse process then learns the underlying inverse mapping D_t^{-1}, obtaining physical coherence with the underlying imaging model. Overall, diffusion-based methodology unifies stochastic sampling, energy-based optimization and inverse problem regularization from the perspectives of a probabilistic generative approaches and classical variational reconstruction theory. An illustration of methods based on model type distribution is shown in Figure 5.

3.1.1. Standard diffusion models

Within the class of standard Denoising Diffusion Probabilistic Models (DDPMs) approaches, several works advance the field (Table 2). Oh et al. [13] train an annealed score-based model solely on motion-free MR images and couple Come-Closer-Diffuse-Faster (CCDF) with iterative forward-reverse diffusion and annealed data-consistency weighting, achieving faster inference without sacrificing high-frequency anatomical detail. Moving from reconstruction to detection, Damudi et al. [15] streamline sampling to a single step for unsupervised brain lesion localization. By replacing Gaussian corruption with partial diffusion and simplex noise, they retain the benefits of diffusion priors while cutting inference time by orders of magnitude.

Extending standard diffusion to multi-modal settings, Xie et al. [16] propose MC-Diffusion, a Bayesian joint reconstruction framework in which a shared score function couples PET and MRI so that both modalities are denoised coherently via forward corruption and reverse restoration on their joint distribution. In a similar direction, Sarkar et

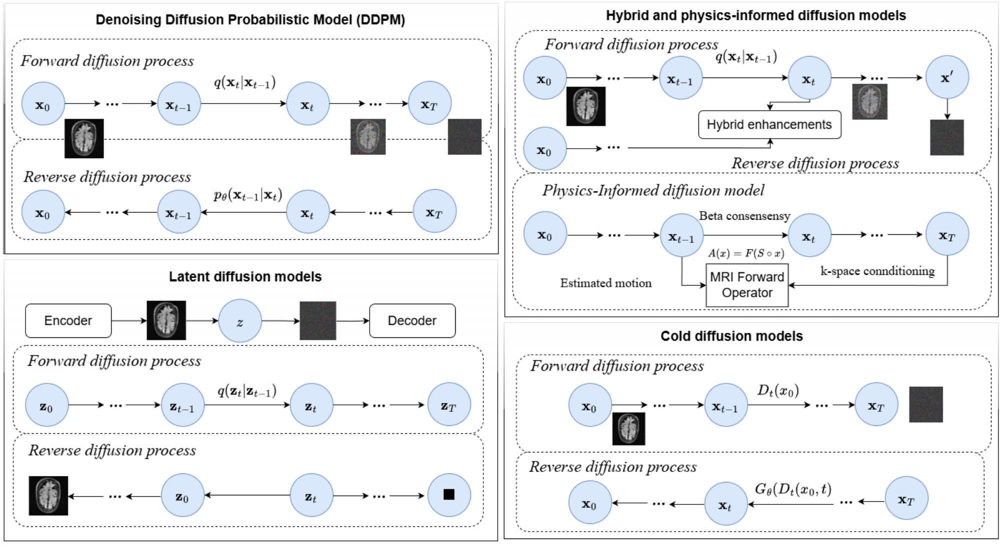

Figure 4. Overview of diffusion-based generative models for MRI reconstruction. The image illustrates four major categories of diffusion approaches: (1) DDPMs, which rely on stochastic forward noising and learned reverse denoising processes, (2) hybrid and physics-informed diffusion models, which incorporate adversarial, attention-based or implicit neural components as well as MRI acquisition physics through data consistency and forward operators, (3) latent diffusion models (LDMs), where diffusion is performed in a compressed latent space obtained via an encoder–decoder architecture, and (4) cold diffusion models, which replace stochastic noise with deterministic domain-specific degradations and learn their inversion.

al. [17] focus on data-consistent sampling through AutoDPS, which performs diffusion posterior sampling (DPS) conditioned directly on corrupted measurements. Standard diffusion is also leveraged for quantitative reconstruction. He et al. [18] introduce accelerated quantitative susceptibility mapping (ACE-QSM), a DDPM tailored to quantitative susceptibility mapping that restores high-frequency susceptibility information lost with short-TE acquisitions. Here, a wavelet-based stabilization further improves temporal consistency across echoes. Further, Zhang et al. [19] present a domain-conditioned and temporal-guided diffusion (dDiMo) that augments score-based denoising with spatiotemporal priors and a nonlinear conjugate-gradient step inside each reverse iteration.

Beyond single-model designs, Geng et al. [20] introduced Detail-Preserving Multiple Diffusion Model (DP-MDM), that is an ensemble of multiple diffusion models in parallel across k-space scales. Here, one prior targets global structure while others specialize in high-frequency recovery via virtual masks. A cascaded fusion then yields reconstructions preserve details. Li et al. [21] tackles efficiency concerns by reformulating DPS to avoid back-propagating through the score network when enforcing measurement consistency. Zhong et al. [22] introduces approach that learns a joint spatiotemporal score so that denoising acts coherently across space and time. Finally, Huang et al. [11] adapt DDPMs to diffusion MRI by reconstructing fiber-orientation distributions (FODs) using a volume order-aware encoder and frequency balanced cross-attention across spherical-harmonic orders. This maintains geometric and frequency consistency in 4D signal domain.

3.1.2. Hybrid and physics-informed diffusion models

Recent research focuses on integration of stochastic diffusion processes with explicit physical modeling to enhance data fidelity and interpretability in MRI reconstruction (Table 3). Levac et al. [10] introduced a physics-informed hybrid diffusion framework for retrospective motion correction. Their Bayesian posterior formulation jointly estimates motion parameters and clean images. This extends DPS with motion-aware forward

Table 2. Classification of standard diffusion-based MRI reconstruction studies.

Author (Year)	Model Type	Core contribution
Oh et al. (2024) [13]	Score-based DDPM	Annealed diffusion (CCDF) with forward–reverse sampling for motion artifact correction; unsupervised, physics-aware training.
Damudi et al. (2024) [15]	DDPM (single-step)	Single-step simplex-noise diffusion for unsupervised lesion detection; lightweight one-step design.
Xie et al. (2024) [16]	Score-based joint diffusion	MC-Diffusion for PET–MRI joint reconstruction using shared score-based priors.
Sarkar et al. (2025) [17]	Diffusion Posterior Sampling	AutoDPS: posterior diffusion sampling for unsupervised MRI restoration; blind corruption estimation.
He et al. (2025) [18]	DDPM (QSM)	ACE-QSM: score-based diffusion for quantitative susceptibility mapping with wavelet stabilization.
Zhang et al. (2025) [19]	Domain-conditioned DDPM	dDiMo: spatiotemporal priors with CG optimization in reverse diffusion for dynamic MRI.
Geng et al. (2024) [20]	Multi-diffusion model	DP-MDM: multiple score networks for detail-preserving MRI reconstruction at multiple scales.
Li et al. (2025) [21]	Posterior-sampling DDPM	Efficient diffusion posterior sampling (DPS) via gradient-descent likelihood updates.
Zhong et al. (2023) [22]	Spatiotemporal DDPM	Dynamic diffusion (dDiMo) jointly modeling spatial–temporal correlations.
Huang et al. (2024) [11]	DDPM (dMRI FOD)	Volume-order-aware encoding and cross-attention for fiber orientation distribution restoration.

operators. Here, the motion is treated as an additional latent variable, which blends the generative power of diffusion priors with deterministic MRI physics. Following this, hybrid approaches, such as one proposed by Zhao et al. [23], introduce DiffGAN, which fuses diffusion models with GANs. The incorporation of a Local Vision Transformer (LVT) enhances fine-grained feature extraction, while adversarial optimization improves perceptual realism and sampling efficiency. This highlights how adversarial learning can complement diffusion processes for structure-aware MRI synthesis. In another direction, Guan et al. [24] propose the Global-to-Local Diffusion Model (GLDM), a hybrid score-based approach that combines global structural priors with low-rank and data-consistency regularization. By progressively refining coarse global reconstructions into high-detail outputs, GLDM unites probabilistic generative inference with physics-based regularization. This enables zero-shot generalization without fully sampled data. Chu et al. [25] further expand the hybrid paradigm with DiffINR, which embeds Implicit Neural Representations (INRs) into diffusion posterior sampling. The INR acts as a continuous-space physical model that enforces k-space consistency and stabilizes reconstruction even under high acceleration factors. Similarly, Shin et al. [26] address inference stability through ELF-Diff, a physics-aligned framework that integrates adaptive low-frequency mixing and ensemble refinement. Their method achieves faster and more reliable reconstructions across diverse undersampling patterns. Temporal and dynamic modeling are also advancing within this class. Wang et al. [27] propose a dDiMo that fuses spatial, temporal and frequency priors via 3D CNN-based noise estimation and conjugate gradient optimization. By conditioning on k-space data and temporal dynamics, dDiMo bridges conventional diffusion sampling and model-driven reconstruction.

Furthermore, hybrid approaches are extended by frameworks that explicitly merge neural physics models with diffusion priors. Ahmed et al. [28] introduce PINN-DADif, which integrates Physics-Informed Neural Networks (PINNs) with adaptive diffusion. Qiao et al. [29] complement this concept with SGP-MDN, which reuses pretrained score networks as auxiliary generative priors inside a physics-constrained model-driven network. This improves artifact suppression and sample efficiency. At the patient-specific level, Uh et al. [30] design an adaptive diffusion model that dynamically adjusts network parameters to integrate individual MRI priors during inference. This improves reconstruction fidelity for subjects with unique anatomical or pathological variations. Moreover, adversarial and transformer-based hybrids continue to gain on research interest. Zhang et al. [31] propose a diffusion-adversarial framework augmented by local transformer attention. This enhances high-frequency reconstruction and contextual awareness and produces sharper and more accurate MRI images. Zhao et al. [32] introduce the Mamba Diffusion Model (MDM), which unites diffusion with structured state-space modeling. By embedding Mamba modules that capture long-range dependencies, MDM offers perceptually aware, globally coherent MRI reconstructions. In parallel, Guan et al. [33] propose DMSE, which fuses subset-k-space priors and global data-driven constraints within a probabilistic diffusion formulation. Noise robustness and denoising-driven hybrids are also explored. Aali et al. [34] demonstrate that combining DPMs with model-based deep learning (MoDL) and self-supervised GSURE denoising enhances generalization under noise and multi-coil settings. In cardiac applications, Qiu et al. [35] and Shah et al. [14] independently develop quadratic conditional diffusion models (DBSR) that integrate explicit physical priors such as blur kernel estimation and attention-guided denoising in order to achieve high fidelity and super resolution in cardiac MRI. Similarly, Hou et al. [36] present Fast and Reliable Score-based Generative Model (FRSGM), a diffusion ADMM hybrid, where multiple ensemble denoisers act as stochastic score estimators within each optimization loop. This combination of probabilistic inference and constrained optimization achieves rapid convergence and robust data consistency.

Furthermore, Zhang et al. [37] introduce Diffusion-QSM, where they embed the QSM forward model within diffusion sampling to enforce physical authenticity and enhance magnetic susceptibility estimation. Similarly, Chen et al. [38] develop JSMoCo, a joint score-motion-coil framework that estimates motion parameters and coil sensitivities alongside reconstruction, which encodes MRI physics within the diffusion process. Earlier works such as Hu et al. [39] and Daducci et al. [40] laid conceptual foundations for such integration by embedding MRI forward operators and Bayesian physics formulations within diffusion models. They demonstrated that data consistency and generative flexibility can coexist within a unified reconstruction framework. Finally, Li et al. [41] proposed SNAFusion-MM that is a teacher-student hybrid model that aggregates 2D diffusion priors into coherent 3D volumes. Their method maintains high generative fidelity at a fraction of the computational cost.

Table 3. Classification of hybrid and physics informed diffusion-based MRI reconstruction studies.

Author (Year)	Model Type	Core contribution
Levac et al. (2024) [10]	Hybrid (Bayesian DPS)	Motion-aware hybrid diffusion with joint posterior sampling of image and motion parameters.
Zhao et al. (2024) [23]	DiffGAN (Hybrid)	Adversarial–diffusion with local vision transformer (LVT) for perceptual realism.
Guan et al. (2024) [24]	GLDM (Hybrid)	Global-to-local diffusion with low-rank and DC regularization for zero-shot MRI.
Chu et al. (2025) [25]	DiffINR	Diffusion integrated with implicit neural representations (INR) for stable posterior sampling.
Shin et al. (2025) [26]	ELF-Diff	Standard score-based diffusion model enhanced with ensemble and adaptive frequency mixing strategies
Wang et al. (2025) [27]	dDiMo	A domain-conditioned and temporal-guided diffusion model that integrates spatial, temporal, and frequency-domain information through 3D CNN-based noise estimation.
Ahmed et al. (2025) [28]	PINN-DADif	Physics-informed adaptive diffusion integrating PINNs with stochastic denoising priors.
Qiao et al. (2025) [29]	SGP-MDN	Model-driven network guided by pretrained score priors with physics constraints.
Uh et al. (2025) [30]	Adaptive Diffusion	Patient-specific adaptive diffusion incorporating individualized priors during inference.
Zhang et al. (2024) [31]	Transformer + GAN Hybrid	Local transformer attention within adversarial diffusion for MRI SR.
Zhao et al. (2025) [32]	Mamba-Diffusion	Hybrid diffusion–state-space architecture (VSSM + SIF-Mamba) for perception-aware modeling.
Guan et al. (2025) [33]	DMSE (Hybrid)	Distribution-matching subset-k-space embedding combining stochastic priors and deterministic physics.
Aali et al. (2025) [34]	DPM + MoDL	GSURE-based denoising to enhance hybrid diffusion and MoDL performance on noisy data.
Qiu et al. (2024) [35]	DBSR (Hybrid)	Quadratic conditional diffusion with blur-kernel estimation for cardiac MRI SR.
Hou et al. (2025) [36]	FRSGM (ADMM Hybrid)	ADMM-integrated diffusion with ensemble denoisers ensuring convergence and stability.
Shah et al. (2024) [14]	Quadratic Conditional DDPM	Conditional diffusion guided by blur kernels for controllable cardiac MRI SR.
Zhang et al. (2025) [37]	Diffusion-QSM	Physics-informed diffusion for susceptibility mapping with resampling refinement.
Chen et al. (2025) [38]	JSMoCo	Joint estimation of motion and coil maps via physics-informed diffusion with Gibbs sampling.
Hu et al. (2016) [39]	Physics-Integrated DDPM	Embedded MRI encoding operator in score-based diffusion for data-consistent reconstruction.
Daducci et al. (2016) [40]	Bayesian Hybrid Diffusion	Probabilistic–physical model fusion for uncertainty-aware reconstruction.
Li et al. (2025) [41]	SNAFusion-MIX	Multi-step teacher–student hybrid diffusion for efficient 3D reconstruction.

3.1.3. Latent diffusion models

Latent diffusion models (LDMs) represent a key evolution in generative MRI and neural decoding research (Table 4). They enabled complex reconstruction tasks within compact and semantically meaningful latent spaces. These models reduce computational demands while preserving fine-grained image fidelity and multimodal consistency by decoupling the generative process from the high-dimensional pixel space.

For example, Liu et al. [42] introduced Mind-Bridge, an LDM framework designed for reconstructing visual images from fMRI signals. The approach unites a Depth Structure Variational Autoencoder (DSVAE) to capture spatial and structural information, a Very Deep VAE (VDVAE) for semantic encoding and a Versatile Diffusion backbone guided by multimodal contrastive language-image pre-training (CLIP) embeddings. Edge estimation via the Canny operator further enhances structural fidelity, resulting in reconstructions that preserve both semantic and geometric details. Building on similar principles of cross-modal decoding, Wang et al. [43] developed an LDM architecture that combines a multimodal Masked Autoencoder (MAE) with diffusion-based reconstruction. Their model learns shared latent representations through joint training on paired fMRI-image data and subsequently fine-tunes a conditional diffusion model to recover visually coherent natural scenes. This multimodal approach allows efficient integration of visual and neural information and outperforms approaches based on only one modality in decoding perceived images.

Extending the latent diffusion paradigm to the frequency domain, Lu et al. [44] proposed the Latent-k-space Refinement Diffusion Model (LRDM), which performs diffusion directly in a low-dimensional latent representation of the k-space domain. Their two-stage method first generates coarse priors in latent k-space and then refines high-frequency details through a secondary diffusion stage. Since it is a hierarchical approach, it significantly reduces diffusion steps (requiring as few as four iterations) while maintaining high reconstruction fidelity and computational efficiency. Moreover, Zhao [45] introduced an LDM framework with disentangled representation learning for multidimensional MRI reconstruction. Through a dual encoder-decoder structure, geometry and contrast information are separated into distinct latent spaces, each governed by its own diffusion process. This separation allows precise control over individual image attributes, which improves the interpretability and flexibility of the reconstruction process.

Furthermore, Kalantari et al. [46] proposed a generative framework that combines visual and semantic representations of human brain activity. Their method integrates Bootstrapping Language-Image Pretraining (BLIP) for caption-based semantic decoding and a conditional LDM for image synthesis. By unifying deep generative networks for perceptual decoding with semantic conditioning via BLIP, the model achieves reconstructions that align both conceptually and visually with the original images, offering new insight into neural representation modeling. Finally, Li et al. [47] presented NeuralDiffuser, which redefines diffusion-based reconstruction through a neuroscience-inspired latent architecture built upon a pretrained stable diffusion backbone. The model introduces primary visual feature guidance, combining top-down semantic cues and bottom-up perceptual gradients extracted from CLIP encoders to enrich the diffusion trajectory. Their method enhances both semantic coherence and structural precision while producing more accurate reconstructions of the perceptual characteristics encoded in brain signals.

Table 4. Classification of latent diffusion-based MRI reconstruction studies.

Author (Year)	Model Type	Core contribution
Li et al. (2024) [42]	Mind-Bridge LDM	CLIP-guided LDM with structural and semantic alignment for brain decoding.
Wang et al. (2024) [43]	MAE-LDM (fMRI)	Multimodal masked autoencoder with conditional LDM for neural decoding.
Lu et al. (2025) [44]	LRDM	Latent-k-space refinement diffusion with hierarchical two-stage reconstruction.
Zhao et al. (2024) [45]	Disentangled LDM	Dual-encoder LDM separating geometry and contrast for multidimensional MRI.
Kalantari et al. (2025) [46]	BLIP-LDM	BLIP-conditioned latent diffusion for visual/semantic decoding from fMRI.
Li et al. (2024) [47]	NeuralDiffuser	Latent-guided diffusion integrating top-down semantics and low-level features.

3.1.4. Cold diffusion models

Cold diffusion models represent a deterministic rethinking of the diffusion paradigm (Table 5). Cold diffusion replaces stochastic noise-based corruption with physically grounded, structured degradation processes. These models retain the iterative nature of diffusion frameworks but eliminate random sampling in favor of domain-specific transformations, particularly in the k-space domain. They provide a more interpretable, stable, and physics-consistent approach to MRI reconstruction under high acceleration or limited data conditions.

Shen et al. [48] pioneered the use of a k-space cold diffusion model. They introduced a deterministic forward process that removes the need for Gaussian noise, which is typically employed in conventional diffusion frameworks. Instead, the model defines degradation through progressive k-space undersampling masks, where each step systematically omits portions of the frequency domain. The reverse network, implemented via a U-Net denoiser, is trained to reconstruct fully sampled k-space data from these structured degradations. This closely aligns the diffusion process with real MRI acquisition physics, leading to improved generalization and robustness while maintaining the interpretability of the generative process. Building on this, Cui et al. [12] proposed a physics-informed cold diffusion framework that deepens the integration between diffusion modeling and physical k-space transformations. Their model replaces random Gaussian perturbations with a reverse heat diffusion process that deterministically interpolates high-frequency (HF) components from low-frequency (LF) k-space inputs. This directly models the physics of frequency recovery. By incorporating a data-fidelity constraint within each iteration, the approach ensures consistency with acquired measurements while preserving interpretability and reconstruction stability. They demonstrate that cold diffusion can unify physics-based reasoning and generative inference within a single deterministic framework. Extending these principles further, Dong et al. [49] developed AMK-CDiffNet, an adaptive multi-scale cold diffusion model. Like the previous two methods, AMK-CDiffNet replaces stochastic noise with structured, domain-specific degradations. The noise is implemented as progressive undersampling in k-space. However, it advances the paradigm through a multi-scale fusion mechanism and enables hierarchical feature aggregation across different frequency bands. Because of this, their method can capture both global anatomical context and fine spatial details, which improves reconstruction fidelity across varying resolutions. Moreover, by aligning the diffusion process with multi-scale k-space physics, the model achieves higher sampling efficiency and greater numerical stability than traditional stochastic approaches.

Table 5. Classification of cold diffusion-based MRI reconstruction studies.

Author (Year)	Model Type	Core contribution
Shen et al. (2024) [48]	CDiff	Deterministic k-space degradation and data-consistent reverse network for MRI.
Cui et al. (2024) [12]	Physics-informed CDiff	Reverse heat diffusion modeling of HF–LF interpolation in k-space.
Dong et al. (2025) [49]	AMK-CDiffNet	Adaptive multi-scale cold diffusion with progressive undersampling and fusion.

3.2. Application Domain Distribution

The diversity of diffusion-based magnetic resonance imaging applications stems from the way generative priors are adapted to different forms of inverse problems. MRI reconstruction can generally be modeled as:

$$\mathbf{y} = \mathcal{A}(\mathbf{x}) + \mathbf{n},$$

where $\mathbf{y}$ is the undersampled k-space data, $\mathcal{A}$ is the encoding operator (including Fourier transforms and coil sensitivities) and $\mathbf{n}$ is noise. Diffusion models serve as data priors $p_\theta(\mathbf{x})$ that regularize the reconstruction of $\mathbf{x}$ when $\mathbf{y}$ is given.

In accelerated MRI, diffusion priors replace traditional sparsity-based or compressed sensing regularizers. This allows learning of powerful nonlinear manifolds that enable high-fidelity reconstruction from undersampled k-space data. These models reduce acquisition times while preserving diagnostic accuracy and structural integrity, which provides a

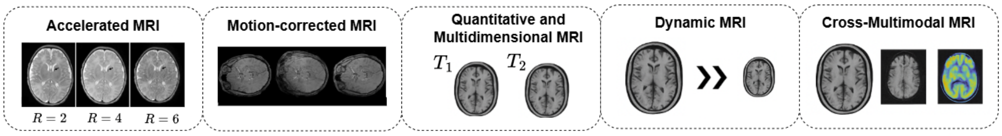

Figure 5. Illustration of major MRI application domains. The figure highlights five key areas where modern reconstruction and generative modeling techniques are applied: accelerated MRI (varying undersampling factors $R = 2, 4, 6$), motion-corrected MRI (recovering images degraded by patient motion), quantitative and multidimensional MRI (e.g., T_1, T_2), dynamic MRI (capturing temporal anatomical changes), and cross-multimodal MRI (integrating complementary MRI contrasts or external modalities such as PET).

data-driven alternative to classical optimization-based reconstruction. In motion-corrected MRI, generative diffusion frameworks are extended to account for both rigid and non-rigid motion by integrating learned priors with explicit physical motion models. Through joint estimation of motion parameters and clean anatomical representations, these methods effectively mitigate artifacts caused by patient movement. This enables robust and artifact-free reconstructions even in long or free-breathing acquisitions. Dynamic MRI further benefits from diffusion models that incorporate temporal priors and spatiotemporal score functions $s_\theta(\mathbf{x}_t, t)$ to capture the evolution of moving anatomy. Such formulations promote temporal coherence and spatial consistency in cardiac, pulmonary and functional MRI sequences. For quantitative and multidimensional MRI, diffusion models extend beyond structural imaging to recover tissue-specific parameters such as T_1, T_2, or magnetic susceptibility maps. By embedding biophysical constraints within the generative process, these models enhance both reconstruction accuracy and acquisition efficiency, producing interpretable parameter maps that align with clinical diagnostic needs. Finally, in cross-modal and multimodal imaging, conditional diffusion models $p_\theta(\mathbf{x}|\mathbf{c})$ learn joint distributions across complementary modalities, such as PET–MRI, CT–MRI or fMRI image pairs. By conditioning on latent neural or structural features $\mathbf{c}$, they enable synthesis or reconstruction of one modality from another, effectively bridging perceptual, functional and anatomical representations. An illustration of methods is shown in Figure 5.

3.2.1. Accelerated MRI

Diffusion-based approaches for accelerated MRI reconstruction offer powerful generative priors that restore high-fidelity images from undersampled k-space data while reducing scan time (Table 6). These approaches address critical challenges such as aliasing artifacts, motion distortion and loss of high-frequency details which are common issues in fast MRI acquisitions.

Early efforts, such as those by Zhao et al. [23], focus on restoring high-quality MRI images directly from undersampled measurements. This resulted in reduced aliasing artifacts and acquisition time with preserved diagnostic fidelity. Building upon this, Shen et al. [48] introduced a cold diffusion framework that replaces Gaussian noise with deterministic k-space undersampling masks, producing artifact-free reconstructions. Similarly, Ahmed et al. [28] proposed PINN-DADif, a physics-informed hybrid diffusion model that jointly enforces k-space data fidelity and image-space regularization. In parallel, Cui et al. [12] advanced deterministic modeling through reverse heat diffusion in k-space, effectively interpolating high-frequency details from low-frequency measurements. Damudi et al. [15] introduced a simplex diffusion framework for rapid MRI anomaly localization and artifact reduction, focusing on single-step denoising and reconstruction from undersampled acquisitions. Their method improves computational efficiency and stability under limited sampling. Qiao et al. [29] proposed the SGP-MDN, which integrates pretrained score priors with MR physics to improve artifact suppression and consistency. Their method demonstrates how hybrid diffusion-physics coupling can enhance image quality and reconstruction fidelity from highly undersampled k-space data.

Furthermore, hybrid and dynamic reconstruction paradigms are also proposed. For example, Guan et al. [24] developed the Global-to-Local Diffusion Model (GLDM) for cardiac cine MRI, which enables robust dynamic sequence reconstruction. Thus is obtained through zero-shot learning across varied acquisition trajectories. Chu et al. [25] followed with DiffINR, which integrates diffusion priors with implicit neural representations (INRs). Shin et al. [26] proposed ELF-Diff, which stabilizes reconstructions under different sampling masks and mitigates hallucinations through adaptive low-frequency mixing. Hou et al. [36] further combined stochastic diffusion with ADMM optimization in FRSGM, yielding faster convergence and higher fidelity under high acceleration factors.

Further advancements have targeted noise robustness, 3D reconstruction and super-resolution. Aali et al. [34] showed that self-supervised denoising significantly improves diffusion based reconstruction robustness in noisy multi-coil scenarios. Li et al. [41] introduced SNAFusion-MM, a multi-view hybrid diffusion framework for efficient 3D MRI reconstruction. Their method distills 2D priors into volumetric consistency. In the context of super-resolution, Qiu et al. [35] and Zhao et al. [32] developed conditional and perception-aware diffusion models to restore high-resolution cardiac and brain MR images affected by blur and noise. Dong et al. [49], proposed AMK-CDiffNet, which achieves high edge preservation. Meanwhile, Lu et al. [44] explored latent k-space diffusion and achieved high-quality reconstructions with drastically reduced computational costs.

Finally, classical and quantitative MRI studies continue to validate diffusion's potential in broader contexts. Geng et al. [20] introduced a multi-scale ensemble of diffusion models for detail-preserving acceleration. Li et al. [21] proposed an efficient posterior-sampling method to integrate measurement guided gradients directly in diffusion reconstruction loops. Hu et al. [39] and Daducci et al. [40] bridged early physics-based diffusion modeling with modern generative inference, showing diffusion's versatility from structural MRI to quantitative susceptibility mapping.

Table 6. Classification of reviewed studies by application domain in accelerated MRI.

Author (Year)	Application domain	Core contribution
Zhao et al. (2024) [23]	Accelerated MRI / DiffGAN	Adversarial–diffusion model that restores high-quality MRI from undersampled data. It reduces aliasing artifacts and preserves diagnostic fidelity.
Shen et al. (2024) [48]	Accelerated MRI / Cold Diffusion	Deterministic k-space undersampling replacing Gaussian noise. It achieves artifact-free and stable reconstructions.
Ahmed et al. (2025) [28]	Accelerated MRI / PINN-DADif	Physics-informed diffusion combining k-space data fidelity with image-space regularization to enhance interpretability and stability.
Cui et al. (2024) [12]	Accelerated MRI / Physics-informed cold diffusion	Reverse heat diffusion in k-space interpolating high-frequency details from low-frequency data. Enables deterministic and physically consistent reconstruction.
Guan et al. (2024) [24]	Accelerated MRI / GLDM	Global-to-Local Diffusion Model enabling zero-shot generalization across acquisition trajectories.
Chu et al. (2025) [25]	Accelerated MRI / DiffINR	Integration of diffusion priors with implicit neural representations for robust reconstruction under high acceleration factors.
Shin et al. (2025) [26]	Accelerated MRI / ELF-Diff	Adaptive low-frequency mixing improves reconstruction stability and reduces hallucination artifacts under different sampling masks.
Hou et al. (2025) [36]	Accelerated MRI / FRSGM	Combines diffusion and ADMM optimization for faster convergence and high fidelity at high acceleration factors.
Aali et al. (2025) [34]	Accelerated MRI / DPM + MoDL	Self-supervised denoising improves diffusion-based reconstruction robustness in noisy multi-coil MRI acquisitions.
Li et al. (2025) [41]	3D Accelerated MRI / SNAFusion-MM	Multi-view hybrid diffusion framework fusing 2D priors into consistent 3D volumes for efficient volumetric reconstruction.
Qiu et al. (2024) [35]	Accelerated MRI / DBSR	Quadratic conditional diffusion with blur-kernel estimation achieving high-resolution cardiac MRI reconstruction.
Zhao et al. (2025) [32]	Accelerated MRI / Mamba Diffusion	Perception-aware diffusion model with global context modeling for high-resolution brain MRI restoration.
Dong et al. (2025) [49]	Accelerated MRI / AMK-CDiffNet	Adaptive multi-scale cold diffusion achieving strong edge preservation and improved numerical stability.
Lu et al. (2025) [44]	Accelerated MRI / LRDM	Latent k-space refinement diffusion with hierarchical two-stage reconstruction. It achieves high fidelity with low computational cost.
Geng et al. (2024) [20]	Accelerated MRI / DP-MDM	Multi-scale ensemble of diffusion models preserving fine structural details.
Li et al. (2025) [21]	Accelerated MRI / DPS	Efficient diffusion posterior sampling that incorporates measurement-guided gradients to enhance reconstruction accuracy.
Hu et al. (2016) [39]	Accelerated MRI / Physics-Integrated DDPM	Early embedding of MRI encoding operators into diffusion sampling for data-consistent reconstruction.
Daducci et al. (2016) [40]	Accelerated MRI / Bayesian Hybrid Diffusion	Bayesian fusion of probabilistic diffusion and physical modeling for uncertainty-aware MRI reconstruction.
Damudi et al. (2024) [15]	Accelerated MRI / Simplex diffusion	Proposes a single-step diffusion method for anomaly localization and artifact suppression in undersampled MRI, emphasizing computational efficiency and stability for rapid reconstruction.
Qiao et al. (2024) [29]	Accelerated MRI / SGP-MDN	Integrates pretrained score priors with MRI physics constraints to enhance artifact removal and consistency in highly undersampled k-space reconstruction.

3.2.2. Motion-corrected MRI

Diffusion-based approaches are effective for motion-corrected MRI reconstruction, a domain where reducing motion-induced artifacts is crucial for preserving diagnostic quality, particularly in dynamic or long-duration scans (Table 7). These methods integrate generative priors with motion modeling and physics-informed constraints to correct both rigid and non-rigid displacements while maintaining anatomical fidelity and data consistency.

For example, Oh et al. [13] proposed an annealed score-based diffusion framework tailored for motion artifact reduction in brain and liver MRI. Their method addresses both rigid and non-rigid motion by combining iterative forward-reverse diffusion with adaptive data-consistency weighting. This resulted in enhanced structural sharpness and reduced motion blur. Building on this, Levac et al. [10] developed a physics-informed hybrid diffusion framework that jointly reconstructs motion-free images and estimates in-plane motion trajectories from motion-corrupted and subsampled k-space data. Their Bayesian joint posterior formulation incorporates motion as a latent variable. This enables the model to simultaneously correct motion and accelerate acquisition. Similarly, Sarkar et al. [17] introduced AutoDPS, an unsupervised diffusion posterior sampling model designed to restore diagnostic-quality brain MR images affected by complex, real-world motion and undersampling artifacts. By learning to infer degradation parameters on-the-fly and embedding them within a data-consistent diffusion process, the approach effectively corrects motion distortions without requiring paired supervision. Uh et al. [30] presented an adaptive diffusion modeling approach incorporating subject-specific priors to enhance robustness against patient motion and acquisition variability. Their method primarily addresses motion-corrected MRI, leveraging adaptive guidance during reverse diffusion to maintain temporal and spatial consistency across dynamic scans. Finally, Chen et al. [38] introduced JSMoCo, a physics-informed diffusion framework that simultaneously estimates motion parameters and coil sensitivity maps within a multi-coil acquisition setup. By combining score-based diffusion priors with explicit modeling of the MRI forward process, their method performs self-calibrated motion correction and image reconstruction in a unified loop.

Table 7. Classification of reviewed studies by application domain in motion-corrected MRI.

Author (Year)	Application domain	Core contribution
Oh et al. (2024) [13]	Artifact suppression	Annealed forward–reverse diffusion for reduced motion blur.
Levac et al. (2024) [10]	Motion-corrected MRI	Joint motion trajectory estimation and reconstruction via hybrid diffusion.
Sarkar et al. (2025) [17]	Motion artifact correction	AutoDPS: unsupervised posterior sampling with motion parameter estimation.
Chen et al. (2025) [38]	Motion-corrected accelerated MRI	JSMoCo: joint estimation of motion and coil sensitivity maps via diffusion priors.
Uh et al. (2025) [30]	Motion-corrected accelerated MRI	Adaptive Diffusion Model

3.2.3. Dynamic MRI

This class of methods extends diffusion priors beyond static imaging and embeds temporal guidance, motion modeling and cross-contrast learning to address the challenges inherent in dynamic or multi-parametric MRI acquisitions (Table 8).

For example, Guan et al. [33] introduced a Distribution Matching with Subset-k-Space Embedding (DMSE) framework for multi-contrast MRI reconstruction. Although primarily focused on multi-contrast T1-, T2-, and PD-weighted imaging, their approach shares key principles with dynamic MRI. By exploiting inter-contrast redundancy to reduce scan time and suppress motion-related artifacts. This model effectively leverages complementary information among contrasts and enhances both reconstruction fidelity and generalization across diverse sampling conditions and anatomical datasets. Zhang et al. [19] proposed the domain-conditioned and temporal-guided diffusion model (dDiMo), explicitly tailored for dynamic MRI. Their framework targets accelerated cardiac cine and free-breathing lung imaging across both Cartesian and golden-angle radial trajectories. Proposed approach incorporates temporal priors and conjugate-gradient optimization with every reverse diffusion step. This allows robust inter-frame coherence and spatial consistency and performance gains. Similarly, Wang et al. [27] proposed temporally guided approach for motion-resolved MRI reconstruction. Their method integrates spatiotemporal priors into the diffusion process, which enables effective suppression of artifacts while simultaneously correcting motion. The model demonstrates superior generalization across cardiac and pulmonary datasets and diverse undersampling rates. Shah et al. [14] developed a DBSR for cardiac MRI super-resolution, embedding temporal blur kernels to improve fidelity in time-resolved data. In their method, the diffusion processes capture inter-frame coherence and high-frequency motion details for cardiac cine enhancement. Zhong et al. [22] presented a spatiotemporal score-based diffusion model for cardiac cine MRI, focusing on temporal smoothness and motion continuity. Their approach jointly models spatial and temporal dependencies through a unified probabilistic framework. This integration enables consistent reconstruction of dynamic sequences with reduced temporal flicker and improved motion fidelity. Finally. Zhang et al. [50] presents a super-resolution (SR) approach to enhance the inherently low spatial resolution of multishell diffusion MRI using a GAN with deep residual channel attention.

Table 8. Classification of reviewed studies by application domain in dynamic MRI.

Author (Year)	Application domain	Core contribution
Guan et al. (2024) [33]	Dynamic cardiac MRI	Global-to-local diffusion with low-rank regularization.
Zhang et al. (2025) [19]	Dynamic MRI	Domain-conditioned diffusion (dDiMo) for spatiotemporal coherence.
Wang et al. (2025) [27]	Dynamic MRI	Temporal-guided diffusion for motion-resolved reconstruction.
Zhong et al. (2023) [22]	Dynamic MRI	Spatiotemporal DDPM with temporal smoothness constraints.
Shah et al. (2024) [14]	Dynamic MRI	Temporally conditioned diffusion guided by blur kernels, capturing inter-frame coherence and motion-related detail.
Zhang et al. (2025) [19]	Dynamic MRI	Jointly reconstructs accelerated Cartesian and non-Cartesian dynamic data using x-t and k-t priors

3.2.4. Quantitative and multidimensional MRI

In quantitative and multidimensional MRI the goal is to recover meaningful tissue parameters, contrast maps or directional diffusion profiles (Table 9). For example, He et al. [18] proposed ACE-QSM, which is an accelerated quantitative MRI approach adapted for QSM in neuroimaging. ACE-QSM leverages a score-based denoising diffusion probabilistic model to restore high-frequency details lost in short-echo acquisitions. The method enables accurate quantification of iron deposition, demyelination and lesion receptivity and shortens acquisition times. Similarly, Zhang et al. [31] proposed approach that combines transformers with attention and adversarial regularization. This model tackles high-fidelity reconstruction of static anatomical MRI across multiple datasets and anatomical regions.

Regarding multidimensional MRI reconstruction, Zhao et al. [45] proposed a LDM that improves T_1/T_2 mapping from undersampled acquisitions. They introduced dual-encoder architecture that separates geometry and contrast into independent latent subspaces. This allows diffusion model to learn distinct generative priors which improves the reconstruction of quantitative maps and ensuring spatial coherence and contrast fidelity. Huang et al. [11] targeted diffusion MRI (dMRI) reconstruction with a focus on recovering FODs in regions severely affected by susceptibility-induced distortion. Their approach introduces volume-order-aware encoding and frequency-balanced cross-attention to handle complex four-dimensional data.

3.2.5. Cross-multimodal MRI

In the cross-modal domain the goal is to learn shared representations across imaging modalities or even across different data types such as medical images, neural signals

Table 9. Classification of reviewed studies by application domain in quantitative and multidimensional MRI.

Author (Year)	Application domain	Core contribution
He et al. (2025) [18]	Quantitative MRI	Short-TE diffusion model for QSM with wavelet stabilization.
Zhang et al. (2025) [31]	Quantitative MRI	Diffusion-QSM with physics-informed refinement.
Zhao et al. (2024) [45]	Quantitative MRI	LDM with disentangled T_1/T_2 feature spaces.
Huang et al. (2024) [11]	Diffusion MRI	FOD-diffusion for fiber orientation restoration with cross-attention.

and/or language (Table 10). For example, Xie et al. [16] proposed approach toward true multi-modal medical image fusion through joint PET-MRI reconstruction. Their model learns a shared latent distribution that encodes both anatomical structure from MRI and functional information from PET. This addresses modality-specific limitations such as low signal-to-noise ratio in PET and extended acquisition times in MRI.

Sveral works have adapted latent and diffusion-based approaches. Liu et al. [42] introduced Mind-Bridge, a latent diffusion framework that translates fMRI signals into natural scene images with high structural and semantic fidelity. Similarly, Wang et al. [43] combines fMRI signals with paired visual data to enable semantically meaningful scene reconstruction. Their multimodal masked autoencoder (MAE) and conditional latent diffusion framework capture correspondences between neural activity and visual features. Moreover, Li et al. [47] proposed NeuralDiffuser, which unifies image, text and neural representations. Since it operates in a latent space aligned with pre-trained stable diffusion and CLIP encoders, this approach integrates voxel-level fMRI activations with multimodal embeddings. The architecture combines top-down semantic guidance with bottom-up perceptual cues. This improves interpretability and accuracy over previous fMRI-to-image approaches. Finally, Kalantari et al. [46] advanced the field by tackling not only seen but also imagined and conceptually generated images from brain signals. Their model integrates fMRI data with language vision models such as BLIP and CLIP. This enables a richer cross-modal alignment that spans perception and semantics.

Table 10. Classification of reviewed studies by application domain in cross-multimodal MRI.

Author (Year)	Application domain	Core contribution
Xie et al. (2024) [16]	PET–MRI	Joint score-based diffusion for multi-modal PET–MRI enhancement.
Liu et al. (2024) [42]	fMRI decoding	Mind-Bridge: CLIP-guided latent diffusion for fMRI→image translation.
Wang et al. (2024) [43]	fMRI decoding	Multimodal MAE–LDM for brain visual scene reconstruction.
Li et al. (2025) [47]	fMRI decoding	NeuralDiffuser integrating top-down and bottom-up latent features.
Kalantari et al. (2025) [46]	fMRI decoding	BLIP-conditioned latent diffusion for semantic–visual decoding.

3.3. Domain of Reconstruction

The domain in which the diffusion process is applied determines interaction between physical data fidelity and learned priors. In image-space diffusion, the forward operator acts as pixel-wise Gaussian corruption, where the network learns the conditional probability $p_\theta(\mathbf{x}_0|\mathbf{x}_t)$. This is straightforward and leverages visual features directly but may not strictly enforce k-space consistency. In k-space diffusion, the process is redefined over Fourier coefficients:

$$q(\mathbf{k}_t|\mathbf{k}_{t-1}) = \mathcal{N}(\sqrt{1-\beta_t}\mathbf{k}_{t-1}, \beta_t\mathbf{I}),$$

where $\mathbf{k}_t$ denotes the noisy k-space data. The reverse process learns to reconstruct the true distribution $p(\mathbf{k}_0)$, effectively predicting missing frequency components consistent with MRI acquisition physics. Hybrid diffusion alternates between k-space and image-space representations within each denoising step, ensuring that low-frequency global structures

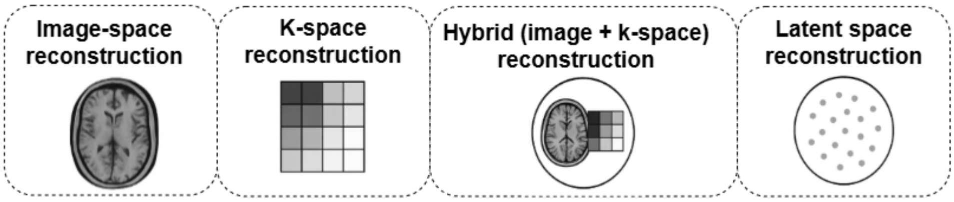

Figure 6. Overview of reconstruction domains in MRI methods. The figure illustrates four primary reconstruction domains: image-space reconstruction (operates directly on spatial MRI images), k-space reconstruction (processes data in the frequency domain), hybrid reconstruction, (jointly leverages both image and k-space representations) and latent-space reconstruction (data are mapped into a compressed latent manifold for more efficient or robust reconstruction).

and high-frequency textural information are simultaneously preserved. These models often use Fourier transform $\mathcal{F}$ and its inverse $\mathcal{F}^{-1}$ within the diffusion iterations:

$$\mathbf{x}_{t+1} = \mathcal{F}^{-1}(D_\theta(\mathcal{F}(\mathbf{x}_t))),$$

where D_θ denotes the learned diffusion denoiser. Finally, latent-space diffusion introduces an encoder–decoder pair (E, D) that maps the image domain into a compact latent space $z = E(\mathbf{x})$. Diffusion is then performed over z_t rather than $\mathbf{x}_t$, which reduces dimensionality and allows the model to capture global anatomical structures with fewer parameters and iterations. An illustration of domain reconstruction method is shown in Figure 6.

3.3.1. Image-space reconstruction

Studies based on image-space reconstruction perform the diffusion process directly in the image domain (Table 11). This enables efficient denoising, artifact correction, super resolution and selective incorporation of k-space consistency to maintain physical realism. For example, Oh et al. [13] introduced a hybrid image-space reconstruction approach in which the diffusion process operates on spatial representations. Their method selectively enforces k-space consistency. In each iteration, low-frequency k-space information is preserved and high-frequency details are restored. This produces motion-corrected images that combine generative flexibility with physics-informed accuracy. Similarly, Shin et al. [26] adopt an image-space diffusion strategy guided by data consistency optimization to mitigate hallucinations. Their approach ensures anatomical precision through harmonized priors and k-space fidelity and achieves stable and realistic reconstructions across diverse sampling patterns. Furthermore, Damudi et al. [15] apply diffusion and denoising directly to pixel-space brain MRI data. Here, the anomalies are inferred as deviations between the corrupted and reconstructed images. Similarly, He et al. [18] employ a DDPM for QSM. This model restores high-frequency susceptibility details and embends time-frequency regularization to maintain data consistency. Both studies illustrate the effectiveness of image-domain diffusion in preserving anatomical fidelity without relying on heavy k-space modeling. Extending this principle, Zhao et al. [23] and Zhang et al. [31] incorporate adversarial and attention-based mechanisms to further enhance spatial realism. Zhao et al. [23] introduces a hybrid adversarial-diffusion approach in which the forward diffusion adds structured noise and a GAN-based reverse process reconstructs artifact-free, high-resolution images. In parallel, Zhang et al. integrate transformer-based attention with adversarial regularization to improve local coherence and perceptual quality, achieving superior restoration across multiple MRI datasets.

Several image-space diffusion models explicitly bridge generative reconstruction with physical measurement constraints. Qiao et al. [29] combine score-based priors with data-consistency losses to produce reconstructions that balance structural realism and quantitative fidelity. Aali et al. [34] employ self-supervised denoising to enhance the robustness of diffusion-based reconstructions. Similarly, Zhao et al. [32] and Qiu et al. [35] address the challenge of blind degradation and super-resolution by conditioning the diffusion process on learned blur kernels. This allows perceptual sharpness and faithful restoration in cardiac and neuroimaging contexts. Shah et al. [14] proposed DBSR designed for cardiac MRI super-resolution within the image domain. Their approach formulates the diffusion process as a conditional transformation guided by learned blur kernels and spatial priors, enabling progressive enhancement of image sharpness and temporal consistency across cardiac frames. The model reconstructs high-resolution images directly from low-resolution or motion-corrupted inputs. Operating entirely in image space, DBSR highlights the potential of conditional diffusion modeling for improving both spatial resolution and temporal coherence in dynamic cardiac imaging.

Finally, a growing class of 3D and quantitative image-space frameworks demonstrates how diffusion priors can generalize beyond 2D reconstruction. Li et al. [41] propose SNAFusion-MM, which employs multi-view 2D diffusion priors fused across axial, sagittal and coronal planes to reconstruct coherent 3D volumes. Their approach maintains consistency with physical measurements. Zhang et al. [37] extend this concept to physics-informed quantitative MRI through Diffusion-QSM, which synthesizes susceptibility maps directly from image data while incorporating MRI forward operators to ensure fidelity and interpretability.

Table 11. Classification of reviewed studies according to their reconstruction domain, focusing on image-space-based approaches.

Author (Year)	Reconstruction Domain	Core Reconstruction Strategy
Oh et al. (2024) [13]	Image-space	Diffusion denoising in image space with embedded k-space consistency weighting to suppress motion artifacts.
Damudi et al. (2024) [15]	Image-space	One-step simplex diffusion applied directly to image intensities for lesion detection.
Zhao et al. (2024) [23]	Image-space	DiffGAN combines diffusion-based denoising and GAN refinement on image-space representations.
Xie et al. (2024) [16]	Image-space	PET–MRI joint reconstruction using shared diffusion priors across modalities.
He et al. (2025) [18]	Image-space	DDPM restores QSM susceptibility maps using wavelet-stabilized time–frequency priors.
Shin et al. (2025) [26]	Image-space	ELF-Diff denoises undersampled MRI via image-domain diffusion with iterative data consistency optimization.
Qiao et al. (2025) [29]	Image-space	Score-based priors refine undersampled reconstructions via image-domain denoising.
Zhang et al. (2024) [31]	Image-space	Adversarial diffusion with local transformer attention enhances perceptual realism.
Zhao et al. (2025) [23]	Image-space	Mamba-enhanced diffusion model reconstructs images using blur-aware latent conditioning.
Aali et al. (2025) [34]	Image-space	Image-domain denoising and diffusion reconstruction trained on self-supervised denoised data.
Qiu et al. (2024) [35]	Image-space	Conditional blur-kernel diffusion reconstructs cardiac MRI with motion-compensated priors.
Shah et al. (2024) [14]	Image-space	Conditional quadratic diffusion (DBSR) for cardiac MRI super-resolution.
Zhang et al. (2025) [37]	Image-space	Diffusion-QSM reconstructs susceptibility maps via image-domain diffusion and physical priors.
Li et al. (2025) [41]	Image-space	Multi-view 2D priors fused into consistent 3D reconstructions.

3.3.2. K-space reconstruction

Methods emphasizing diffusion modeling directly in the frequency, i.e., k-space, domain are justified as reconstruction directly reflects the nature of MRI signal capture physics (Table 12). Their goal is to recover missing or corrupted frequency components prior to image creation, resulting in a higher level of data accuracy and better high-frequency detail preservation than in purely image-domain diffusion models. For example, Cui et al. [12] introduce a deterministic cold diffusion framework that performs all of its functions in the k-space domain. For this process, degradation and restoration are defined in frequency space, not pixel space. The forward process uses progressive undersampling to mimic a physical loss of acquisition, while the reverse process reconstructs lost HF elements using a heat-diffusion-inspired approach. The method, treating LF data as a deterministic prior and imposing physics-based constraints to ensure data fidelity, is effective at mitigating hallucination artifacts and improving texture and edge preservation, with robust generalization across out-of-distribution datasets. As follow this type of paradigm, Guan et al. [33] extend diffusion-based modeling to multi-coil k-space reconstruction. Their method learns priors directly from subset k-space distributions instead of image intensities, integrating diffusion sampling with low-rank regularization and iterative data consistency enforcement. By alternating k-space diffusion and inverse Fourier transformation, the model reconstructs high-quality images with strong structural coherence. This framework elegantly bridges probabilistic diffusion priors with deterministic physics-based reconstruction principles, highlighting the advantages of operating within the measurement domain itself. Xie et al. [16] introduced MC-Diffusion, a Bayesian PET-MRI joint reconstruction framework designed to overcome PET's low signal-to-noise ratio and MRI's long acquisition times, which often require undersampled k-space data. By reformulating joint reconstruction as a joint regularization problem with separate PET and MRI fidelity terms, the method leverages a joint score-based diffusion model to learn the shared probability distribution between the two modalities.

Furthermore, Geng et al. [20] propose the DP-MDM, which explicitly learns frequency-space priors under physical acquisition constraints. The model incorporates predictor–corrector diffusion steps followed by data consistency modules, ensuring fidelity to raw MRI measurements. The introduction of circular and radial binary masks enables multi-scale modeling of spatial-frequency relationships, enhancing both global and local feature recovery. By directly reconstructing from k-space signals, DP-MDM achieves superior interpretability, robustness, and control over structural detail preservation compared with image-based or hybrid diffusion frameworks. Finally, Dong et al. [49] propose AMK-CDiffNet that works exclusively in the k-space domain. Different from stochastic noise-based approaches, AMK-CDiffNet uses a deterministic forward process simulating progressive undersampling, paired with an inverse diffusion mechanism that hierarchically restores lost information. An adaptive multi-scale architecture of the model is available, which is conducive for cross-frequency learning that is successful with both fine details retrieval and global anatomical context at the same time. This results in reconstructions that achieve exceptional artifact suppression and edge definition while remaining computationally efficient.

Table 12. Classification of reviewed studies according to their reconstruction domain, focusing on k-space-based approaches.

Author (Year)	Reconstruction Domain	Core Reconstruction Strategy
Cui et al. (2024) [12]	K-space	Cold diffusion via deterministic heat-based interpolation restoring missing high-frequency k-space data.
Guan et al. (2025) [24]	K-space	DMSE operates on subset-k-space embeddings with low-rank regularization for multi-contrast MRI.
Geng et al. (2024) [20]	K-space	Multi-diffusion ensemble learning in frequency domain for detail-preserving MRI.
Dong et al. (2025) [49]	K-space	AMK-CDiffNet deterministic cold diffusion directly on k-space with multi-scale fusion.

3.3.3. Hybrid (image + k-space) reconstruction

Hybrid diffusion methods integrate the benefits of image-space and k-space modeling (Table 13). They combine generative diffusion priors with explicit physical constraints derived from MRI acquisition models. This allows reconstructions to retain both visual realism and physical fidelity. Hybrid reconstruction approaches are capable of achieving better artifact suppression, high structural integrity and diagnostic accuracy by alternating between denoising in the spatial domain and ensuring measurement consistency in the frequency domain. For example, Levac et al. [10] pioneer this direction with a physics-informed hybrid diffusion framework for motion-corrected MRI. Their model alternates between image-space inference and k-space consistency enforcement, where diffusion priors iteratively refine motion-free reconstructions while forward operators parameterized by motion variables ensure alignment with physical measurements. Sarkar et al.[17] take DPS even further into a hybrid domain that incorporates k-space undersampling patterns into image-space inference through the modeling of k-space undersampling behavior. The accuracy of their model in sampling is ensured by this method, which keeps the flexibility of diffusion priors and has enabled accurate and artifact-free 3D MRI reconstruction under severe undersampling. Similarly, Guan et al. [24] proposed GLDM, which integrates stochastic differential equations on k-space data and image-domain optimization via low-rank and data-consistency regularization. It preserves physical realism while improving structural detail preservation and also enforces the dual-domain synergy regarding probabilistic and deterministic elements. Moreover, Chu et al. [25] propose DiffINR that performs generative denoising in image space while enforcing physics-consistent updates in k-space, achieving robust performance under extreme acceleration factors. Likewise, Zhang et al. [19] and Wang et al. [27] extend hybrid diffusion modeling to dynamic MRI, embedding spatiotemporal (x–t) and frequency-temporal (k–t) priors into the diffusion process. Their dDiMo framework anchors diffusion in k-space but includes image-domain guidance for temporal alignment, resulting in motion-resolved and temporally coherent reconstructions.

A number of works further improve hybrid diffusion using physics-based priors and optimization principles. Shen et al.[48] introduce a k-space-to-image cold diffusion model that learns degradation and restoration in frequency space while computing losses in image space. This guarantees interpretability and robustness across sampling schemes. Ahmed et al.[28], based on a hybrid PINN-regularized method, combine image-domain diffusion priors with k-space physics constraints. Similarly, Uh et al.[30] propose a patient-specific adaptive diffusion model using k-space conditioning and image-space refinement as alternates. This allowed accurate and anatomy consistent reconstructions tailored to individual subjects. As for optimization-driven hybrids, Hou et al.[36] use iterative ADMM updates alternating between image denoising and k-space projection. Their FRSGM integrates diffusion-based priors with physics-consistent solvers that build semantically detailed reconstructions. Lu et al. [44] do a similar latent-k-space approach, in which diffusion takes place in compressed k-space latents and finetunes high-frequency spatial details in image space; this approach achieves high fidelity at lower computational costs.

Huang et al.[11] apply diffusion modeling to the FOD domain which effectively connects volumetric spatial reconstruction and spherical harmonic signal modeling for high-dimensional diffusion MRI. Similarly, Li et al.[21] fuse image-space generative modeling with measurement-guided gradient descent in k-space to achieve reconstructions that retain perceptual and physical coherence. Chen et al.[38] proposed a self-calibrating Gibbs sampling framework, referred to as JSMoCo, which jointly estimates motion, coil sensitivities and image reconstructions across k-space and image domains. Hu et al.[39] find that by alternating diffusion-based denoising between the two domains, stability and convergence are boosted, and Daducci et al.[40] generalize this principle to the q-space and spatial domains of microstructural MRI to enhance generalization over acquisition protocols. Finally, Zhong et al.[22] embed spatial and temporal diffusion in a joint x–t representation, capturing both dynamic motion and anatomical consistency in the reconstruction of cardiac cine MRI.

3.3.4. Latent-space reconstruction

Latent-space reconstruction optimizes computational speed, semantic coherence and multimodal integration (Table 14). These models, working on compressed latent representations that encode high-level structure and meaning, strike a more optimal balance between computational ease, interpretability and fidelity of reconstruction. For example, Liu et al. [42] also proposed Mind-Bridge to encode brain voxel data into structural and semantic latent embeddings. These embeddings are decoded using a Versatile Diffusion backbone that is conditioned on CLIP text-image prompts. The resulting reconstructed outputs retain spatial detail and semantic meaning.

Table 13. Classification of reviewed studies according to their reconstruction domain, focusing on hybrid-space-based approaches.

Author (Year)	Reconstruction Domain	Core Reconstruction Strategy
Levac et al. (2024) [10]	Hybrid (image + k-space)	Alternates image-domain inference with k-space motion modeling for joint motion and reconstruction.
Sarkar et al. (2025) [17]	Hybrid (image + k-space)	AutoDPS posterior sampling links diffusion priors to undersampled measurements for artifact-free MRI.
Chu et al. (2024) [25]	Hybrid (image + k-space)	DiffINR couples image diffusion priors with INR-based physical modeling in k-space.
Zhang et al. (2025) [19]	Hybrid (k-space + x–t)	Dynamic diffusion reconstruction integrating k–t and x–t spatiotemporal priors.
Wang et al. (2025) [27]	Hybrid (k-space + x–t)	Spatiotemporal dDiMo diffusion with conjugate-gradient refinement for dynamic MRI.
Shen et al. (2024) [48]	Hybrid (k-space + image)	Cold diffusion using deterministic k-space degradation with image-domain supervision.
Ahmed et al. (2025) [28]	Hybrid (image + k-space)	PINN-DADif combines image-space denoising and physics-constrained k-space regularization.
Uh et al. (2025) [30]	Hybrid (image + k-space)	Patient-specific adaptive diffusion integrating k-space conditioning and image-space priors.
Hou et al. (2025) [36]	Hybrid (image + k-space)	FRSGM alternates diffusion-based image denoising with ADMM k-space updates.
Lu et al. (2025) [44]	Hybrid (k-space + image)	Latent-k-space diffusion with image refinement to recover high-frequency detail.
Li et al. (2025) [21]	Hybrid (image + k-space)	Diffusion posterior sampling with measurement-guided gradient descent.
Chen et al. (2025) [38]	Hybrid (image + k-space)	JSMoCo jointly models motion and coil sensitivity in diffusion-based reconstruction.
Hu et al. (2016) [39]	Hybrid (image + k-space)	Alternating iterative denoising enforcing k-space fidelity and image-domain consistency.
Daducci et al. (2016) [40]	Hybrid (q-space + image)	Bayesian reconstruction merging q-space diffusion and spatial-domain regularization.
Zhong et al. (2023) [22]	Hybrid (spatiotemporal x–t)	dDiMo learns spatiotemporal diffusion distribution across spatial–temporal axes.
Guan et al. (2024) [24]	Hybrid-space	HyDiffNet alternates between k-space consistency and image-space denoising.
Huang et al. (2024)[11]	Hybrid-space	CD^2M integrates k-space and image-space through a shared latent diffusion process with domain-consistency constraints.

Table 14. Classification of reviewed studies according to their reconstruction domain, focusing on latent-space-based approaches.

Author (Year)	Reconstruction Domain	Core Reconstruction Strategy / Domain Interaction
Liu et al. (2024) [42]	Latent-space	Mind-Bridge reconstructs visual images from fMRI via latent multimodal diffusion.
Wang et al. (2024) [43]	Latent-space	MAE–LDM performs multimodal latent reconstruction from fMRI embeddings.
Li et al. (2024) [47]	Latent-space	Stable Diffusion latent decoding guided by fMRI features and CLIP embeddings.
Zhao et al. (2024) [45]	Latent-space	Disentangled latent T1/T2 mapping reconstruction with physical consistency.
Kalantari et al. (2025) [46]	Latent-space	BLIP-conditioned latent diffusion reconstructs semantic visual content from fMRI.

Building upon this, Wang et al. [43] also investigate latent-space reconstruction with a MAE that jointly encodes both fMRI and image data into a common latent feature space. Encoded fMRI embeddings are then conditioned on the diffusion process, resulting in highly faithful, semantically stable visual reconstructions. Since operations within the compact latent domain entail significantly fewer diffusion steps when compared to image-level models, this method improves both the reconstruction fidelity and computational cost. Further, Li et al. [47] introduced NeuralDiffuser based on the latent space of a pre-trained stable diffusion framework. Combining together fMRI features with VQ-VAE and CLIP-derived latent representations of the features helps the model encapsulate both global structure and high-level semantic information. A new guiding mechanism fine-tunes the reverse diffusion path to maintain anatomical accuracy and perceptual realism.

Most latent-space frameworks emphasize brain decoding, where most existing frameworks are limited to brain decoding, Zhao et al. [45] widen the scope to multidimensional MRI. Their method learns diffusion priors over disentangled latent representations corresponding to geometry and contrast, rather than direct image intensities. Through the trade-off between physics-informed k-space consistency enforcement and latent-space regularization, the framework is able to provide effective reconstruction on multi-contrast MRI data. The combination of generative priors and physical fidelity demonstrates the adaptability of latent diffusion strategies that extend beyond cognitive imaging fields. Finally, Kalantari et al. [46] proposed a LDM that works on combining visual and semantic decoding between images and fMRI data to reconstruct both seen and imagined images. In their model, visual characteristics decoded from neural signals are projected into the latent space, denoised using a conditional LDM, and further refined by semantic features extracted with BLIP captioning. Such a design ensures that reconstructed images remain visually realistic, while still conceptually mirroring the underlying brain representations.

4. Discussion and Conclusion

This review presents a systematic study and critical chategorization on the diffusion-based methods for MRI reconstruction, including the standard, hybrid, latent and cold diffusion paradigms. We have systematically categorized 40 recent works by model type, application and reconstruction domain. We identify that diffusion models evolve up from only image-domain denoising frameworks to physics-informed, multi-domain generative architectures in an almost instant. They have shown superior results in terms of preserving anatomical fidelity, reducing artifacts and accelerating image reconstruction when compared to classical compressed sensing and deep learning methods (Table 15). Standard diffusion models demonstrated strong generative abilities for artifact removal, denoising and image restoration across a variety of MRI modalities. Their iterative sampling process allows flexible modeling of complex noise and artifact pattern and enables good adaptability to heterogeneous datasets and sampling strategies, including 3D imaging. However, these models remain limited by their high computational cost and sensitivity to sampling hyperparameters. While recent innovations such as one-step sampling and posterior-guided diffusion have improved efficiency, further architectural optimization and integration with parallel imaging or compressed sensing methods are still needed for practical clinical use.

Hybrid and physics-informed diffusion models combine stochastic generative learning with the deterministic physics of MRI reconstruction. By embedding explicit data-consistency operators or incorporating constraints through approaches such as PINNs, ADMM, or INR-guided priors, these frameworks enhance both interpretability and adherence to MRI acquisition principles. Such designs effectively reduce hallucination artifacts and improve generalization across different sampling patterns and acceleration factors. However, integrating physical operators also increases optimization complexity and computational cost. Future hybrid architectures may adopt adaptive coupling strategies that dynamically balance the influence of learned and physical priors during inference.

LDMs extend diffusion modeling into compact, semantically meaningful latent spaces, achieving low training and sampling costs while maintaining strong perceptual and structural fidelity. Their impressive results in fMRI-to-image reconstruction and cross-modal decoding highlight the potential of diffusion models to bridge the gap between neuroscience and clinical imaging. However, current LDMs rely heavily on pretrained backbone models such as Stable Diffusion or CLIP-based encoders, which may limit their ability to generalize to medical domains with limited data diversity. Developing domain-specific latent representations and fine-tuned diffusion backbones trained directly on medical datasets could greatly improve both the accuracy and interpretability of these models.

An alternative approach, known as cold diffusion, replaces stochastic Gaussian noise with deterministic degradation operators that mimic real MRI acquisition processes, such as progressive k-space undersampling. This design enhances reconstruction stability, eliminates randomness in sampling, and provides clear physical interpretability. Early studies have shown that cold diffusion can achieve faster convergence and higher fidelity than traditional stochastic methods, although its application to more complex settings—such as multi-coil, dynamic or 3D MRI remains underexplored. Combining deterministic forward processes with learned latent priors could offer a promising balance between physical realism and generative flexibility. Across all model categories, a clear trend emerges toward domain-integrated generative reconstruction, where probabilistic modeling, physics-informed constraints and latent embedding work together to optimize realism, accuracy and efficiency.

In spite of these promising developments, diffusion-based MRI reconstruction is still facing a number of practical and methodological challenges (Figure 7). First, the

Table 15. Advantages and disadvantages of different diffusion-based model types for MRI reconstruction.

Model Type	Advantages	Disadvantages
Standard Diffusion Models	Strong generative capacity for artifact removal, denoising, and image restoration across MRI modalities. Capable of modeling complex, non-Gaussian noise distributions through iterative sampling. Adapt well to heterogeneous datasets and 3D imaging tasks.	High computational cost due to many sampling steps. Sensitive to sampling schedules and hyperparameters. Limited clinical practicality without integration of acceleration or parallel imaging methods.
Hybrid and Physics-Informed Diffusion Models	Combine stochastic generative learning with deterministic MRI physics for improved realism and data fidelity. Reduce hallucination artifacts and improve generalization across sampling patterns and acceleration factors. Enhanced interpretability through explicit data-consistency terms or physics-based constraints.	Added optimization and computational complexity from physical operator integration. Balancing learned and physical priors remains challenging. Risk of longer training or inference times compared to purely data-driven models.
Latent Diffusion Models	Operate in compact latent spaces, reducing training and sampling cost. Maintain structural and perceptual fidelity while improving computational efficiency. Effective in multimodal and fMRI-to-image reconstruction tasks.	Dependence on pretrained backbones that may not generalize to medical data. Limited by small and domain-specific medical datasets. Require fine-tuning or domain adaptation to achieve optimal performance.
Cold Diffusion Models	Replace stochastic noise with deterministic degradation, improving stability and interpretability. Offer faster convergence and higher fidelity reconstruction. Physically aligned with real MRI acquisition processes.	Applicability to complex scenarios not yet fully validated. Reduced generative flexibility compared to stochastic methods. May need hybridization with latent or physics-based approaches for broader usability.

majority of existing models are trained on supervised or semi-supervised data obtained mostly from publicly accessible datasets (e.g., fastMRI or IXI). This often do not present diversity in scanner hardware, pulse sequences, or anatomical regions. As a result, these models struggle to generalize to real clinical environments and perform reliably on rare or pathological cases. Adapting models across different domains and datasets remains essential for ensuring that diffusion-based methods can reliably translate into real clinical practice. Second, diffusion models often require hundreds of reverse sampling steps which requires a heavy computational load on real-time MRI reconstruction. Although recent studies have explored faster strategies such as one step sampling, implicit diffusion solvers and hybrid latent-physical priors, these methods still fall short of meeting clinical speed requirements. Future research should focus on developing adaptive step-size controllers, neural ODE/SDE reformulations and quantized inference pipelines to enable efficient real-time and resource limited deployment. Third, interpretability and uncertainty estimation are still underdeveloped. Most current methods do not provide reliable confidence intervals or physical error bounds for reconstructed images, making their diagnostic reliability difficult to assess. Incorporating Bayesian uncertainty estimation, Monte Carlo diffusion sampling, or calibrated predictive intervals could greatly enhance the trustworthiness and clinical acceptance of these models. Finally, as diffusion frameworks evolve toward multi-domain, hybrid, and latent architectures, new challenges arise in maintaining consistency across imaging modalities and acquisition geometries. Future research should explore multi-contrast co-learning, self-supervised physics-informed pretraining, and energy-efficient generative algorithms optimized for embedded MRI reconstruction hardware. Addressing these open challenges will be essential for realizing the full potential of diffusion-based reconstruction as a clinically viable alternative to traditional deep learning and model-based MRI methods.

Author Contributions: M.H.,I.G, J.P, K.M: Conceptualization, methodology, development, writing original draft, editing; M.H., I.G.: Conceptualization, methodology, development, writing original draft, supervision; K.M.: Editing, validation, visualization; J.P, I.G.: Editing, validation, visualization. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: This work was supported by the Croatian Science Foundation under the project number IP-2024-05-9492.

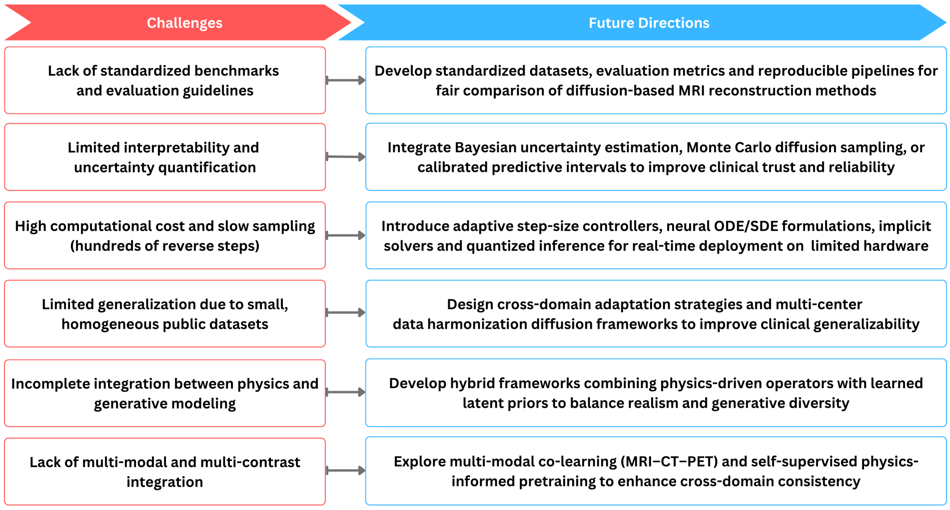

Figure 7. Summary of current challenges and corresponding future research directions in diffusion-based MRI reconstruction. The key methodological, computational and clinical obstacles are identified in recent literature, along with proposed strategies aimed at improving generalization, interpretability, efficiency and real-world clinical applicability of diffusion-driven reconstruction frameworks.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

BLIP	Boot strapping language image pretraining
CCDF	Come closer diffuse faster
CLIP	Contrastive language image pretraining
CS	Compressed sensing
DBSR	Quadratic conditional diffusion model
DDPMs	Denoising diffusion probabilistic models
dDiMo	Domain conditioned and temporal guided diffusion
dMRI	Diffusion MRI
DMSE	Distribution matching with subset k-space embedding
DPMs	Diffusion probabilistic models
DPS	Diffusion posterior sampling
DSVAE	Depth structure variational autoencoder
FODs	Fiber oriented distributions
GANs	Generative adversarial networks
GLDM	Global to local diffusion model
INRs	Implicit neural representation
LDMs	Latent diffusion models
LRDM	Latent k-space refinement diffusion model
LVT	Local vision transformer
MAE	Multi modal masked autoencoder
MDM	Mamba diffusion model
MoDL	Model based deep learning
MRI	Magnetic resonance imaging
PI	Parallel imaging
PINNs	Physics informed neural network
QSM	Quantitative susceptibility mapping
SDEs	Stohastic differential equations
VDVAE	Very deep variational autoencoder

References

1. Hassan, F.A.M.; Megally, H.I.; Moubark, M. Role of magnetic resonance imaging in characterization of superficial soft tissue masses. *Egyptian Journal of Radiology and Nuclear Medicine* **2025**.
2. Li, X.; Liu, C.; Zhu, L.; Wang, M.; Liu, Y.; Li, S.; Deng, Q.; Zhou, J. The Role of High-Resolution Magnetic Resonance Imaging in Cerebrovascular Disease: A Narrative Review. *Brain Sciences* **2023**, *13*.
3. Lustig, M.; Donoho, D.; Pauly, J.M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine* **2007**, *58*, 1182–1195. https://doi.org/10.1002/mrm.21391.
4. Pruessmann, K.P.; Weiger, M.; Scheidegger, M.B.; Boesiger, P. SENSE: Sensitivity encoding for fast MRI. *Magnetic Resonance in Medicine* **1999**, *42*, 952–962. https://doi.org/10.1002/(SICI)1522-2594(199911)42:5<952::AID-MRM16>3.0.CO;2-S.
5. Griswold, M.A.; Jakob, P.M.; Heidemann, R.M.; Nittka, M.; Jellus, V.; Wang, J.; Kiefer, B.; Haase, A. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magnetic Resonance in Medicine* **2002**, *47*, 1202–1210. https://doi.org/10.1002/mrm.10171.
6. Hammernik, K.; Klatzer, T.; Kobler, E.; Recht, M.P.; Sodickson, D.K.; Pock, T.; Knoll, F. Learning a variational network for reconstruction of accelerated MRI data. *Magnetic Resonance in Medicine* **2018**, *79*, 3055–3071. https://doi.org/10.1002/mrm.26977.
7. Knoll, F.; Zbontar, J.; Sriram, A.; Muckley, M.; Bruno, M.; Defazio, A.; Parente, M.; Geras, K.J.; Katsnelson, J.; Chandarana, H.; et al. fastMRI: A Publicly Available Raw k-Space and DICOM Dataset of Knee Images for Accelerated MR Image Reconstruction Using Machine Learning. *Radiology. Artificial intelligence* **2020**, *2 1*, e190007.
8. Fan, Y.; Liao, H.; Huang, S.; Luo, Y.; Fu, H.; Qi, H. A Survey of Emerging Applications of Diffusion Probabilistic Models in MRI. *ArXiv* **2023**, *abs/2311.11383*.
9. Abdullah.; Huang, T.; Lee, I.; Ahn, E. Computationally Efficient Diffusion Models in Medical Imaging: A Comprehensive Review. *ArXiv* **2025**, *abs/2505.07866*.
10. Levac, B.; Kumar, S.; Jalal, A.; Tamir, J.I. Accelerated motion correction with deep generative diffusion models. *Magnetic Resonance in Medicine* **2024**, *92*, 853–868. https://doi.org/10.1002/MRM.30082.
11. Huang, S.; Zhong, L.; Shi, Y. Diffusion Model-based FOD Restoration from High Distortion in dMRI. *ArXiv* **2024**.
12. Cui, Z.X.; Liu, C.; Fan, X.; Cao, C.; Cheng, J.; Zhu, Q.; Liu, Y.; Jia, S.; Wang, H.; Zhu, Y.; et al. Physics-Informed DeepMRI: k-Space Interpolation Meets Heat Diffusion. *IEEE Transactions on Medical Imaging* **2024**, *43*, 3503–3520. https://doi.org/10.1109/TMI.2024.3462988.
13. Oh, G.; Jung, S.; Lee, J.E.; Ye, J.C. Annealed Score-Based Diffusion Model for MR Motion Artifact Reduction. *IEEE Transactions on Computational Imaging* **2024**, *10*, 43–53. https://doi.org/10.1109/TCI.2023.3347917.
14. Shah, J.; Che, Y.; Sohankar, J.; Luo, J.; Li, B.; Su, Y.; Wu, T. Enhancing Amyloid PET Quantification: MRI-Guided Super-Resolution Using Latent Diffusion Models. *Life* **2024**, *14*. https://doi.org/10.3390/LIFE14121580.
15. Damudi, M.Z.; Kini, A.S. Single-Step Sampling Approach for Unsupervised Anomaly Detection of Brain MRI Using Denoising Diffusion Models. *International Journal of Biomedical Imaging* **2024**, *2024*. https://doi.org/10.1155/IJBI/2352602.
16. Xie, T.; Cui, Z.X.; Luo, C.; Wang, H.; Liu, C.; Zhang, Y.; Wang, X.; Zhu, Y.; Chen, G.; Liang, D.; et al. Joint diffusion: mutual consistency-driven diffusion model for PET-MRI co-reconstruction. *Physics in Medicine and Biology* **2024**, *69*. https://doi.org/10.1088/1361-6560/AD6117.
17. Sarkar, A.; Das, A.; Ram, K.; Ramanarayanan, S.; Joel, S.E.; Sivaprakasam, M. AutoDPS: An unsupervised diffusion model based method for multiple degradation removal in MRI. *Computer Methods and Programs in Biomedicine* **2025**, *263*. https://doi.org/10.1016/J.CMPB.2025.108684.
18. He, J.; Fu, B.; Wang, B.; Peng, Y.; Nie, L.; Tang, M.; Xu, R.; Wang, R. ACE-QSM: Accelerating quantitative susceptibility mapping acquisition using diffusion models by reducing repetition time. *Biomedical Signal Processing and Control* **2025**, *110*. https://doi.org/10.1016/J.BSPC.2025.108217.
19. Zhang, L.; Zhou, I.Y.; Montesi, S.B.; Feng, L.; Liu, F. Domain-conditioned and Temporal-guided Diffusion Modeling for Accelerated Dynamic MRI Reconstruction. *ArXiv* **2025**.
20. Geng, M.; Zhu, J.; Hong, R.; Liu, Q.; Liang, D.; Liu, Q. DP-MDM: detail-preserving MR reconstruction via multiple diffusion models. *Physics in Medicine & Biology* **2024**, *70*.
21. Li, J.; Wang, C. DIFFUSION POSTERIOR SAMPLING FOR MAGNETIC RESONANCE IMAGING. *Inverse Problems and Imaging* **2025**, *19*, 613–631. https://doi.org/10.3934/IPI.2024047.
22. Zhong, Z.; Ryu, K.; Mao, J.; Sun, K.; Dan, G.; Vasanawala, S.S.; Zhou, X.J. Accelerating High b-Value Diffusion-Weighted MRI Using a Convolutional Recurrent Neural Network (CRNN-DWI). *Bioengineering* **2023**, *10*. https://doi.org/10.3390/BIOENGINEERING10070864.
23. Zhao, X.; Yang, T.; Li, B.; Yang, A.; Yan, Y.; Jiao, C. DiffGAN: An adversarial diffusion model with local transformer for MRI reconstruction. *Magnetic Resonance Imaging* **2024**, *109*, 108–119. https://doi.org/10.1016/J.MRI.2024.03.017.
24. Guan, Y.; Zhang, K.; Qi, Q.; Wang, D.; Ke, Z.; Wang, S.; Liang, D.; Liu, Q. Zero-shot Dynamic MRI Reconstruction with Global-to-local Diffusion Model. *ArXiv* **2024**, *abs/2411.03723*.

25. Chu, J.; Du, C.; Lin, X.; Zhang, X.; Wang, L.; Zhang, Y.; Wei, H. Highly accelerated MRI via implicit neural representation guided posterior sampling of diffusion models. *Medical Image Analysis* **2025**, *100*. https://doi.org/10.1016/J.MEDIA.2024.103398.
26. Shin, Y.; Son, G.; Hwang, D.; Eo, T. Ensemble and low-frequency mixing with diffusion models for accelerated MRI reconstruction. *Medical Image Analysis* **2025**, *101*. https://doi.org/10.1016/J.MEDIA.2025.103477.
27. Wang, J.; Shi, Z.; Gu, X.; Yang, Y.; Sun, J. Diffusion-based arbitrary-scale magnetic resonance image super-resolution via progressive k-space reconstruction and denoising. *Medical image analysis* **2025**, *107 Pt A*, 103814.
28. Ahmed, S.; Feng, J.; Mehmood, A.; Ali, M.U.; Yaqub, M.; Manan, M.A.; Raheem, A. PINN-DADif: Physics-informed deep adaptive diffusion network for robust and efficient MRI reconstruction. *Digital Signal Processing: A Review Journal* **2025**, *160*. https://doi.org/10.1016/J.DSP.2025.105085.
29. Qiao, X.; Li, W.; Xiao, B.; Huang, Y.; Yang, L. Score-based generative priors-guided model-driven Network for MRI reconstruction. *Biomedical Signal Processing and Control* **2025**, *105*. https://doi.org/10.1016/J.BSPC.2025.107564.
30. Uh, J.; Likitlersuang, J.; Dresner, A.; Merchant, T.E.; ho Hua, C. Improving reconstruction of patient-specific abnormalities in AI-driven fast MRI with an individually adapted diffusion model. *Medical Physics* **2025**, *52*. https://doi.org/10.1002/MP.17955.
31. Zhang, C.; Wang, P.; He, J.; Wu, Q.; Xie, S.; Li, B.; Hao, X.; Wang, S.; Zhang, H.; Hao, Z.; et al. Super-resolution reconstruction improves multishell diffusion: using radiomics to predict adult-type diffuse glioma IDH and grade. *Frontiers in Oncology* **2024**, *14*.
32. Zhao, X.; Yang, X.; Song, Z. Mamba-Enhanced Diffusion Model for Perception-Aware Blind Super-Resolution of Magnetic Resonance Imaging. *IEEE Journal of Biomedical and Health Informatics* **2025**. https://doi.org/10.1109/JBHI.2025.3611232.
33. Guan, Y.; Lu, Y.; Cheng, J.; Wei, H.; Wang, S.; Liu, Q. Distribution matching with subset-k-space embedding for multi-contrast MRI reconstruction. *Medical physics* **2025**, *52*, e18056. https://doi.org/10.1002/MP.18056.
34. Aali, A.; Arvinte, M.; Kumar, S.; Arefeen, Y.I.; Tamir, J.I. Robust multi-coil MRI reconstruction via self-supervised denoising. *Magnetic Resonance in Medicine* **2025**. https://doi.org/10.1002/MRM.30591.
35. Qiu, D.; Cheng, Y.; Wong, K.K.; Zhang, W.; Yi, Z.; Wang, X. DBSR: Quadratic Conditional Diffusion Model for Blind Cardiac MRI Super-Resolution. *IEEE Transactions on Multimedia* **2024**, *26*, 11358–11371. https://doi.org/10.1109/TMM.2024.3453059.
36. Hou, R.; Li, F.; Zeng, T. Fast and Reliable Score-Based Generative Model for Parallel MRI. *IEEE Transactions on Neural Networks and Learning Systems* **2025**, *36*, 953–966. https://doi.org/10.1109/TNNLS.2023.3333538.
37. Zhang, M.; Liu, C.; Zhang, Y.; Wei, H. Diffusion-QSM: Diffusion model with timetravel and resampling refinement for quantitative susceptibility mapping. *IEEE Transactions on Biomedical Engineering* **2025**. https://doi.org/10.1109/TBME.2025.3605587.
38. Chen, L.; Tian, X.; Wu, J.; jun Feng, R.; Lao, G.; Zhang, Y.; Liao, H.; Wei, H. Joint coil sensitivity and motion correction in parallel MRI with a self-calibrating score-based diffusion model. *Medical image analysis* **2025**, *102*, 103502.
39. Hu, C.; Hua, X.; Ying, J.; Thompson, P.M.; Fakhri, G.E.; Li, Q. Localizing Sources of Brain Disease Progression with Network Diffusion Model. *IEEE Journal on Selected Topics in Signal Processing* **2016**, *10*, 1214–1225. https://doi.org/10.1109/JSTSP.2016.2601695.
40. Daducci, A.; Palú, A.D.; Descoteaux, M.; Thiran, J.P. Microstructure informed tractography: Pitfalls and open challenges. *Frontiers in Neuroscience* **2016**, *10*. https://doi.org/10.3389/FNINS.2016.00247.
41. Li, X.; Cai, T.; Dan, J.; Ma, S.; Shang, K.; Butala, M.D.; Wang, G. SNAFusion-MM: Distilling sparse sampled measurements by 2D axial diffusion priors with multi-step matching for 3D inverse problem. *Inf. Fusion* **2025**, *124*, 103323.
42. Liu, Q.; Zhu, H.; Chen, N.; Huang, B.; Lu, W.; Wang, Y. Mind-bridge: reconstructing visual images based on diffusion model from human brain activity. *Signal, Image and Video Processing* **2024**, *18*, 953–963. https://doi.org/10.1007/S11760-024-03207-Z.
43. Wang, Y. Efficient Neural Decoding Based on Multimodal Training. *Brain Sciences* **2024**, *14*. https://doi.org/10.3390/BRAINSCI14100988.
44. Lu, Y.; Xie, X.; Wang, S.; Liu, Q. Latent-k-space of Refinement Diffusion Model for Accelerated MRI Reconstruction. *Biomedical physics & engineering express* **2025**.
45. Zhao, R.; Wang, Z.; Lam, F. Learning Disentangled Representation for Multidimensional MR Image Reconstruction. *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* **2024**, pp. 1–4.
46. Kalantari, F.; Faez, K.; Amindavar, H.; Nazari, S. Improved image reconstruction from brain activity through automatic image captioning. *Scientific Reports* **2025**, *15*. https://doi.org/10.1038/S41598-025-89242-3.
47. Li, H.; Wu, H.; Chen, B. NeuralDiffuser: Neuroscience-Inspired Diffusion Guidance for fMRI Visual Reconstruction. *IEEE Transactions on Image Processing* **2024**, *34*, 552–565.
48. Shen, G.; Li, M.; Farris, C.W.; Anderson, S.; Zhang, X. Learning to reconstruct accelerated MRI through K-space cold diffusion without noise. *Scientific Reports* **2024**, *14*. https://doi.org/10.1038/S41598-024-72820-2.
49. Dong, B.; Wu, G.; Feng, X.; Liu, Y.; Han, J. AMK-CDiffNet: Adaptive-multiscale K-space cold diffusion network for fast MRI reconstruction. *Expert Systems with Applications* **2026**, *297*. https://doi.org/10.1016/J.ESWA.2025.129384.
50. Zhang, C.; Wang, P.; He, J.; Wu, Q.; Xie, S.; Li, B.; Hao, X.; Wang, S.; Zhang, H.; Hao, Z.; et al. Super-resolution reconstruction improves multishell diffusion: using radiomics to predict adult-type diffuse glioma IDH and grade. *Frontiers in Oncology* **2024**, *14*. https://doi.org/10.3389/FONC.2024.1435204.