Deep Learning for Artifacts Reduction: Motion Correction in Medical Imaging: A Comprehensive Literature Review

Ms. Sukshma R. D[1], Dr. Thimmaraju S. N[2]
[1]*Research Scholar VTU RRC MYSURU*
[2]*Professor Department of CS&E VTU RRC MYSURU*

Abstract—**Magnetic Resonance Imaging (MRI) provides unmatched soft-tissue contrast for the diagnosis of neurological, musculoskeletal, cardiovascular, and oncological conditions, but its comparatively long acquisition time makes it highly sensitive to voluntary and involuntary patient motion. Motion-induced artifacts blurring, ghosting, ringing, and geometric distortion degrade diagnostic confidence and frequently necessitate costly re-scans. Over the past decade, artificial intelligence (AI) and deep learning (DL) have transformed retrospective MRI motion correction and image reconstruction, giving rise to convolutional, adversarial, transformer-based, diffusion-based, and physics-guided k-space learning paradigms. This paper presents a structured, thematically organized review of 40 recent IEEE-indexed studies on AI-driven MRI motion correction and reconstruction. A comparative analysis contrasts representative architectures across datasets, reported quantitative performance, advantages, and limitations. Persistent challenges concerning cross-scanner generalization, hallucinated anatomy, computational cost, and the scarcity of paired real-motion clinical data are identified, and future directions combining physics-informed constraints with generative priors are proposed to guide the development of clinically trustworthy, artifact-robust MRI reconstruction systems.**

Index Terms—**Magnetic Resonance Imaging; Motion Artifact Correction; Deep Learning; Generative Adversarial Networks; Vision Transformers; Diffusion Models; Image Reconstruction; k-Space Learning; Compressed Sensing.**

I. INTRODUCTION

Magnetic Resonance Imaging (MRI) has become one of the most valuable non-invasive imaging modalities in modern healthcare because of its exceptional soft-tissue contrast and absence of ionizing radiation. Unlike computed tomography or radiography, MRI reveals subtle anatomical and pathological detail that is essential for accurate diagnosis and treatment planning in neurological, cardiovascular, musculoskeletal, and oncological conditions. However, the relatively long and sequential nature of k-space data acquisition makes MRI intrinsically vulnerable to patient motion. Even minor voluntary or involuntary movement head motion, respiration, swallowing, cardiac pulsation, or physiological tremor can corrupt spatial encoding and manifest as blurring, ghosting, ringing, or geometric distortion in the reconstructed image [1]– [3].

Motion artifacts are of particular concern in neuroimaging, in pediatric and geriatric populations, and in emergency settings where patients cannot be reliably still for the entire acquisition window. These artifacts may degrade the perceived image quality and obscure subtle pathology increasing the chance of diagnostic error and expensive repeat examinations. In general, motion mitigation techniques can be categorized into prospective strategies that track and compensate for motion during acquisition using external sensors, navigator echoes, or specialized sequences, and retrospective strategies that correct the corrupted data or image after acquisition [2], [3]. Prospective methods generally require additional hardware and scanner modifications, making them difficult to implement in routine clinical practice, so retrospective correction is preferred due to its flexibility and compatibility with current imaging workflows [34].

The rapid maturation of deep learning has fundamentally reshaped retrospective MRI motion correction and reconstruction. Convolutional neural networks (CNNs), generative adversarial networks (GANs), vision transformers (ViTs), diffusion probabilistic models, and physics-guided data-consistency frameworks have each demonstrated the capacity to suppress motion-induced degradation while attempting to preserve anatomical fidelity. Learning directly in k-space, together with the incorporation of data-consistency and equivariance constraints, has further allowed models to retain measured information while limiting the hallucination of anatomically implausible structures. Nevertheless, robust generalization across heterogeneous scanners, acquisition protocols, and patient populations, together with a reliable balance between reconstruction accuracy and computational efficiency, remains an open research problem [26], [33], [34].

This paper synthesizes 40 recent IEEE-indexed studies to provide a PhD-level, thematically organized account of AI-driven MRI motion artifact correction and high-quality image reconstruction. The remainder of the paper is organized as follows. Section II describes the review methodology and search strategy. Section III introduces background concepts relevant to MRI acquisition and motion artifacts. Section IV presents the thematic literature review across classical, convolutional, adversarial, transformer, diffusion, physics-guided, and generalization-oriented approaches, together with existing review articles. Section V provides a comparative analysis of representative methods. Section VI identifies open research gaps and challenges, Section VII outlines future research directions, and Section VIII concludes the paper.

II. REVIEW METHODOLOGY

This review followed a structured, PRISMA-inspired identification–screening–eligibility–inclusion process to compile a representative and technically diverse literature base focused on AI-based MRI motion correction and reconstruction.

Identification: A systematic search was conducted primarily on IEEE Xplore, supplemented by cross-referencing of citation lists, using combinations of the keywords "MRI motion correction," "MRI reconstruction," "deep learning," "generative adversarial network," "vision transformer," "diffusion model," "k-space," and "compressed sensing," covering publications from 1999 (to capture the foundational registration-based literature) through 2026.

Screening and Eligibility: Records were retained if they (i) addressed MRI-specific motion artifact correction and/or image reconstruction, (ii) employed a machine-learning or deep-learning methodology or provided a systematic review of such methods, and (iii) were published in an IEEE-affiliated journal, transaction, or conference proceeding. Studies addressing unrelated imaging modalities, non-technical clinical commentary, or purely hardware-based prospective motion tracking without a learning component were excluded.

Inclusion: The final corpus comprises 40 studies spanning IEEE Transactions on Medical Imaging, IEEE Transactions on Circuits and Systems for Video Technology, IEEE Transactions on Computational Imaging, IEEE Access, and major IEEE conference venues including ISBI, ICIP, ICASSP, EMBC, BioCAS, CAMSAP, and WACV. Table I summarizes the review pipeline.

Table I. Review Pipeline Summary

Stage	Description	Count
Identificat ion.	Records located via IEEE Xplore keyword search and citation tracing	~120
Screening	Duplicate and off-topic (non-MRI, non-DL) records removed	~70
Eligibility	Full-text assessed for methodological relevance and IEEE indexing	46
Included	Studies retained for thematic synthesis and comparative analysis	40

III. BACKGROUND

MRI raw data are acquired in the spatial-frequency domain, termed k-space, and the final image is obtained through an inverse Fourier (or more general non-linear) transform. Because k-space is filled progressively over an extended acquisition window, any patient motion during this interval introduces phase and magnitude inconsistencies across the sampled trajectory. Depending on when and how motion occurs relative to the sampling pattern, artifacts manifest as rigid-body ghosting and blurring (in-plane translation/rotation), non-rigid deformation (respiratory or cardiac motion), or signal dropout and geometric warping (through-plane motion) [1], [2], [34].

Motion mitigation techniques are conventionally divided into prospective and retrospective categories. Prospective methods compensate for motion during acquisition using external optical or navigator-based tracking and real-time sequence adjustment; while effective, they require specialized hardware and increase scan complexity. Retrospective methods instead operate on the acquired, motion-corrupted k-space or image data after acquisition is complete, and are the primary focus of the AI literature reviewed here because of their compatibility with standard clinical protocols and their amenability to data-driven learning [3], [34]. Quantitative evaluation across this literature is dominated by full-reference image-quality metrics Structural Similarity Index Measure (SSIM), Peak Signal-to-Noise Ratio (PSNR), Normalized Mean Square Error (NMSE), and Root Mean Square Error (RMSE) computed against a motion-free reference or simulated ground truth.

IV. LITERATURE REVIEW

Recent progress in AI and DL has been organized below into eight thematic categories reflecting the dominant architectural and methodological paradigms adopted for MRI motion correction and reconstruction.

A. Classical and Model-Based Motion Correction

Prior to the widespread adoption of deep learning, motion and deformation correction in MRI relied on explicit registration and optimization-based models. The free-form deformation (FFD) approach of Rueckert et al. combined a global affine transform with a B-spline local deformation field, optimized against a mutual-information similarity measure, to register contrast-enhanced breast MRI and remains a foundational reference for non-rigid registration methodology [1]. Rizzuti et al. proposed a joint retrospective motion correction and reconstruction scheme that exploits an uncorrupted reference contrast (e.g., an unaffected T1 acquisition) together with weighted total-variation regularization to correct a motion-affected companion contrast under a shared-anatomy assumption [2]. More recently, Ortiz-Gonzalez et al. combined a 3D TV-L1 optical-flow algorithm with a learned correction network to propagate expert-annotated regions of interest across the cardiac cycle in cine MRI, bridging classical variational motion estimation with learned refinement [3]. These model-based approaches established the mathematical vocabulary data consistency, regularization, and deformation fields that underpins much of the subsequent deep-learning literature.

B. CNN-Based Motion Artifact Correction

Convolutional neural networks were the first deep architectures widely applied to retrospective MRI motion correction owing to their strong local feature-extraction capability. Tharun and Agateware proposed a CNN incorporating a fuzzy pooling layer to enhance cardiac cine MR images prior to segmentation, improving robustness to motion-induced blurring [4]. Dabrowski et al. introduced SISMIK, a deep-learning motion estimator operating directly in k-space combined with a model-based correction step, reporting rotation-angle RMSE around 0.55° and reconstruction PSNR of approximately 37.8 dB with median SSIM of 0.98 on brain MRI, while explicitly avoiding hallucinated structures through its model-based formulation [5]. Renders et al. proposed DELTA-MRI, which directly estimates deformation fields from longitudinally acquired, under sampled k-space data rather than treating registration as a post-processing step, targeting accelerated longitudinal imaging [6]. Su et al. addressed abdominal diffusion-weighted MRI by optimizing respiratory-phase binning using dynamic programming, reducing missing slices without extending acquisition time [7]. Collectively, this body of work demonstrates that CNN- and k-space-based estimators can achieve high-fidelity motion parameter estimation, but their performance depends strongly on the realism of the training motion model.

C. Generative Adversarial Network-Based Approaches

Adversarial learning has been extensively explored to synthesize visually realistic, artifact-free MRI reconstructions, particularly when paired motion-free/motion-corrupted training data are unavailable. Armanious et al. proposed an unsupervised adversarial framework for correcting severe rigid brain motion artifacts using a cycle-consistency formulation with an MS-SSIM-based loss, removing the requirement for co-registered paired data [8]. Saju et al. extended this idea with an Ensemble CycleGAN combining three motion-type-specific CycleGAN models; on their test set, ensembling raised SSIM from a motion-corrupted baseline of 0.834 to 0.951 and correspondingly reduced NMSE relative to a single CycleGAN model [9]. Chen et al. proposed DCGAN-MS, a disentangled CycleGAN guided by multi-mask k-space subsampling that reduces motion-artifact complexity prior to image-domain translation, reporting PSNR improvements from approximately 26.1 dB to 31.1 dB and SSIM values consistently above 0.8 across evaluated cases [10]. Tsai et al. combined a GAN framework with a channel-attention transformer generator, hierarchically processing motion-corrupted images and reporting a gain of 0.435 dB in PSNR and 0.071 in SSIM over baseline correction [11]. Liu and Yaghoobi instead targeted the accelerated-acquisition reconstruction problem directly, proposing an attentive-selection GAN that integrates large-field contextual features to recover fine anatomical detail from under sampled compressed-sensing k-space data [12]. Across this category, adversarial training consistently improves perceptual realism and quantitative fidelity relative to naive CNN baselines, but the risk of GAN-specific hallucination of plausible-looking yet anatomically incorrect structures remains a recurring concern raised throughout the literature [34].

D. Vision Transformer and Self-Supervised Learning Approaches

Several studies have adapted self-attention mechanisms to MRI motion correction and reconstruction, with the success of transformer architectures in general computer vision tasks. Zhang et al. assessed a ViT-based motion-correction model with masked-image-modeling self-supervised pretraining and achieved statistically significant improvements in PSNR and SSIM ($p < 0.005$) compared with a tailored U-Net baseline, with pretraining speeding up convergence and enhancing the final reconstruction quality [13]. To circumvent this limitation of the swin Transformer, Du and Han proposed a Multi-Modal Vision Transformer with Variable and Shifted Windows (MVSWT) for accelerated MRI reconstruction which shows PSNR gain of up to 14.36 dB and SSIM gain of up to 0.297 over zero-filled reconstruction with significant reduction in NMSE [14]. Singh developed a motion-robust pipeline for brain-tumor diagnosis, comparing optimized Swin and hybrid transformer backbones with interpretable explainable-AI layers, with an emphasis on diagnostic robustness in motion-corrupted acquisitions, not merely on pixel-level reconstruction fidelity [15]. These studies show that the long-range dependency modeling of transformers, particularly when combined with self-supervised pretraining, can lead to a significant improvement in reconstruction quality, but at the cost of increased computation and data than convolutional counterparts.

E. Diffusion Model-Based Motion Correction and Reconstruction

Denoising diffusion probabilistic models have rapidly become one of the most active research directions in MRI reconstruction owing to their strong generative prior and iterative refinement capability. Angella et al. critically assessed diffusion models for 2D brain-MRI motion correction on a public benchmark, highlighting that although diffusion models produce high-quality reconstructions, they remain susceptible to hallucination, which carries direct diagnostic risk [16]. The same group subsequently proposed DIMA, an unsupervised diffusion-based correction framework for brain MRI that reported SSIM performance comparable to, and in several anatomical views statistically indistinguishable from, a supervised ground-truth benchmark (e.g., SSIM 0.788 ± 0.036 vs. 0.820 ± 0.030), while substantially outperforming CycleGAN and Pix2Pix baselines [17]. Liu et al. applied a Denoising Diffusion Probabilistic Model to cardiac cine MRI motion correction using a U-Net-based iterative refinement process, achieving an SSIM of 0.887 and PSNR of 28.81 dB, significantly outperforming comparison methods ($p<0.0001$) [18]. Zhang et al. proposed TC-DiffRecon, combining a diffusion model with a modified MF-UNet and a texture-coordination module to counteract the image-fragmentation tendency of unconditional diffusion sampling, reporting a PSNR drop from 29.73 dB to 27.56 dB and an SSIM drop from 0.739 to 0.618 when the texture-coordination component is ablated demonstrating its contribution to reconstruction coherence [19]. Bangun et al. introduced R3DM, a regularized 3D latent-diffusion approach for volumetric MRI reconstruction that outperforms 2D diffusion baselines on fastMRI knee and BRATS brain data by exploiting inter-slice correlation [20]. Liang et al. proposed a sequential diffusion-guided deep image prior that mitigates the noise-overfitting weakness of unsupervised deep-image-prior reconstruction by periodically re-anchoring the optimization to a pretrained diffusion prior [21]. Mirza and Çukur, and later Mirza, Aydın, and Çukur, proposed super-resolution and frequency-based soft-diffusion models, respectively, that inject spatial-frequency-aware noise schedules into the diffusion process to improve high-frequency anatomical detail recovery for accelerated MRI [22], [23]. Finally, Goud and Spandith proposed PATOMIR, which embeds a pretrained 3D latent-diffusion prior as a plug-and-play regularizer within a physics-informed unrolled optimization network, targeting uncertainty-aware volumetric medical image reconstruction [24]. Across this category, diffusion-based methods consistently deliver strong perceptual and quantitative performance but at the cost of iterative sampling latency and a persistent hallucination risk that several of the reviewed studies explicitly measure and attempt to mitigate.

F. Physics-Guided, Compressed-Sensing, and k-Space Learning Methods

Conversely, a parallel and complementary line of research embeds the physical forward model of the MRI directly into the network architecture to constrain the reconstructions to be consistent with the acquired k-space measurements. In [25], Pezzotti et al. proposed an adaptive-intelligence unrolled reconstruction network for undersampled knee MRI, and evaluated it in terms of NMSE, PSNR and SSIM, considering the fastMRI challenge ranking criteria. One of the first attempts to generalize motion correction across motions, modalities, imaging planes and scanner centers with a single model was proposed by Wang et al. in the Equivariant Imaging Prior (EIP) framework [26]. They formulate multi-scenario motion correction as a mask-varying, self-supervised compressed-sensing problem. In [27], Yaman et al. performed a systematic comparison of unrolled physics-driven network architectures for MRI reconstruction, studying the effect of the choice of regularize network within an iterative data-consistency loop on reconstruction accuracy and computational cost. Zabihi et al. proposed SepUnet, which employs depth-wise separable convolutions in the backbone of U-Net for fastMRI reconstruction to reduce the parameters significantly without any proportional loss in accuracy [28]. Chen et al. proposed the Pyramid Convolutional RNN (PC-RNN), based on three convolutional-RNN modules working at different scales to iteratively reconstruct under sampled MRI in a coarse-to-fine manner without explicit coil sensitivity maps and getting SSIM values > 0.93 with 12 compressed coils at 4× acceleration [29]. Ilıcak et al. [30] proposed DUNDD, a dual-domain physics-guided 3D network with attention-based fusion that targets the low signal-to-noise regime of portable, low-field MRI systems by jointly exploiting k-space and image-domain information. Lin proposed a deep learning-based multi-slice reconstruction algorithm which can exploit inter-slice redundancy to improve the consistency of the reconstruction [31]. Hussain et al. proposed an AI-driven feature-extraction approach for compressed-sensing k-space reconstruction that reduces scan time and preserves diagnostic image quality [32]. Together, these physics-guided methods show that directly embedding the MRI acquisition model into network design significantly improves data fidelity and interpretability compared to pure image-domain learning.

G. Generalization, Robustness, and Disentanglement Strategies

Because clinical MRI motion patterns vary substantially across patients, anatomical regions, and scanners, a growing subset of the literature explicitly targets generalization. Wang et al. proposed a disentangled-embedding framework for retrospective motion correction that separates motion-related and anatomy-related latent representations, enabling a single model to generalize across multiple motion types and body regions rather than requiring application-specific retraining [33]. This disentanglement strategy, together with the equivariant imaging prior of Wang et al. [26] and the model-based k-space estimation of SISMIK [5], represents the most direct algorithmic response to the generalization gap repeatedly identified as a limiting factor for clinical translation.

H. Review Articles and Broader Surveys

Several studies have adapted self-attention mechanisms to MRI motion correction and reconstruction, with the success of transformer architectures in general computer vision tasks. Zhang et al. assessed a ViT-based motion-correction model with masked-image-modeling self-supervised pretraining and achieved statistically significant improvements in PSNR and SSIM ($p < 0.005$) compared with a tailored U-Net baseline, with pretraining speeding up convergence and enhancing the final reconstruction quality [13]. To circumvent this limitation of the swin Transformer, Du and Han proposed a Multi-Modal Vision Transformer with Variable and Shifted Windows (MVSWT) for accelerated MRI reconstruction which shows PSNR gain of up to 14.36 dB and SSIM gain of up to 0.297 over zero-filled reconstruction with significant reduction in NMSE [14]. Singh developed a motion-robust pipeline for brain-tumor diagnosis, comparing optimized Swin and hybrid transformer backbones with interpretable explainable-AI layers, with an emphasis on diagnostic robustness in motion-corrupted acquisitions, not merely on pixel-level reconstruction fidelity [15]. These studies show that the long-range dependency modeling of transformers, particularly when combined with self-supervised pretraining, can lead to a significant improvement in reconstruction quality, but at the cost of increased computation and data than convolutional counterparts.

V. COMPARATIVE ANALYSIS

Table II compares a representative subset of the reviewed methods across dataset scope, core architecture, reported quantitative outcome, principal advantage, and principal limitation. The comparison illustrates a clear architectural progression from CNN encoder–decoder designs toward adversarial, transformer, and diffusion-based generative models, accompanied by an increasing emphasis on physics-based data-consistency constraints to control hallucination risk.

Table II. Comparison Of Representative Ai-Based Mri Motion Correction and Reconstruction Methods

Ref.	Architecture / Method	Dataset / Application	Reported Outcome	Key Advantage	Key Limitation
[25]	Unrolled adaptive-intelligence CNN	fastMRI knee (undersampled k-space)	Evaluated via NMSE, PSNR, SSIM under fastMRI ranking	Strong data-consistency integration	Limited to knee anatomy protocols
[8]	Unsupervised adversarial CycleGAN	Brain MRI, simulated rigid motion	MS-SSIM-based unpaired correction	No paired data required	Risk of adversarial hallucination
[9]	Ensemble CycleGAN (3 motion types)	Brain MRI, simulated rigid motion	SSIM 0.834→0.951; NMSE reduced	Motion-type specialization boosts accuracy	Higher training/inference cost (3 models)
[10]	Disentangled CycleGAN + multi-mask k-space	Brain MRI	PSNR 26.1→31.1 dB; SSIM > 0.8	Reduces artifact complexity pre-translation	Complex multi-stage pipeline
[11]	GAN with channel-attention transformer	Brain MRI, simulated motion	+0.435 dB PSNR; +0.071 SSIM	Hierarchical channel attention	Evaluated on limited motion severity range
[13]	ViT with masked-image-modeling SSL	Brain MRI, k-space-resampled motion	Significant PSNR/SSIM gains ($p<0.005$) vs U-Net	Pretraining improves convergence & accuracy	Large data/compute requirement
[14]	Multi-modal shifted-window ViT (MVSWT)	Accelerated MRI reconstruction	+14.36 dB PSNR; +0.297 SSIM vs zero-filled	Flexible local/long-range attention	Sensitive to acceleration factor increase
[17]	Unsupervised diffusion (DIMA)	Brain MRI motion correction	SSIM ≈0.79–0.86, near supervised benchmark	Outperforms CycleGAN/Pix2Pix baselines	Iterative sampling latency
[18]	DDPM with U-Net denoiser	Cardiac cine MRI	SSIM 0.887; PSNR 28.81 dB ($p<0.0001$)	Strong perceptual + quantitative fidelity	High inference-time cost
[19]	Diffusion + modified MF-UNet (TC-DiffRecon)	Accelerated MRI reconstruction	PSNR/SSIM degrade under texture-ablation (29.7→27.6 dB)	Texture-coordination reduces fragmentation	Performance sensitive to acceleration factor
[20]	Regularized 3D latent diffusion (R3DM)	fastMRI knee, BRATS brain, plant-root data	Outperforms 2D diffusion baselines (SSIM/PSNR tables)	Exploits inter-slice 3D correlation	Higher memory/computation for 3D volumes

[26]	Equivariant imaging prior (EIP), self-supervised CS	Multi-scenario motion correction	Single model generalizes across motions/scanners	Strong cross-scenario generalization	Self-supervised training complexity
[5]	k-space DNN motion estimation (SISMIK)	Brain MRI, in-plane rigid motion	RMSE ≈0.55°; PSNR 37.8 dB; SSIM 0.98	Model-based step avoids hallucination	Restricted to in-plane rigid motion
[28]	Depthwise-separable U-Net (SepUnet)	fastMRI reconstruction	Comparable accuracy with far fewer parameters	Lightweight, efficient inference	Trade-off vs. full-capacity CNNs at high AF
[29]	Pyramid convolutional RNN (PC-RNN)	Multi-coil undersampled MRI	SSIM > 0.93 at 4× acceleration (12 coils)	Coarse-to-fine multi-scale learning	No explicit coil-sensitivity modeling
[33]	Disentangled embeddings (motion/anatomy)	Multi-region, multi-motion-type MRI	Generalizes across motion types/regions	Single model, broad applicability	Disentanglement quality data-dependent

As summarized in Table II, adversarial and diffusion-based methods generally report the strongest perceptual and quantitative gains but incur the greatest computational and hallucination-risk costs, whereas physics-guided and k-space-based methods (e.g., SISMIK [5], EIP [26], SepUnet [28]) trade a portion of peak reconstruction quality for improved interpretability, generalization, and computational efficiency. This trade-off directly motivates the hybrid, physics-informed generative directions discussed in Section VII.

VI. RESEARCH GAPS AND CHALLENGES

Despite substantial progress, this review identifies several recurring and unresolved research gaps across the surveyed literature.

1) Fragmented pipelines: Almost all the reviewed studies consider motion detection, artifact removal, reconstruction, or evaluation as separate issues instead of components of a unified, jointly-optimized pipeline, which limits the end-to-end performance and consistency [2], [34].

2) Limited cross-scanner generalization: Most models are trained and validated on a single dataset/scanner/simulated motion model; performance often deteriorates when applied to unseen protocols, field strengths or patient populations, despite emerging generalization-oriented frameworks [26], [33].

3) Hallucination and clinical trustworthiness: Generative approaches, particularly GAN- and diffusion-based models, can synthesize visually plausible but anatomically incorrect structures, a risk explicitly documented in several of the reviewed diffusion studies [16], [17], [19].

4) Scarcity of paired real-motion clinical data: Because acquiring co-registered motion-corrupted and motion-free scans from the same patient is clinically impractical, nearly all reviewed studies rely on simulated rigid or non-rigid motion, raising questions about generalization to authentic, complex physiological motion [7], [9].

5) Computational efficiency versus accuracy: Iterative diffusion sampling and large transformer backbones achieve strong fidelity but incur inference latency that constrains real-time or point-of-care deployment, particularly for portable and low-field MRI systems [24], [30].

6) Inconsistent, narrow evaluation protocols: Reported metrics vary in scope (SSIM/PSNR alone versus broader anatomical-fidelity or downstream diagnostic-task evaluation), making direct cross-study comparison difficult and underscoring the need for standardized, clinically meaningful benchmarks [34], [35].

VII. FUTURE DIRECTIONS

Based on the identified gaps, we propose the following directions to guide further research towards clinically deployable AI-based MRI motion correction and reconstruction systems.

1) Physics-informed generative hybrids: following PATOMIR [24] and EIP [26], which combine diffusion or adversarial generative priors with explicit data-consistency and equivariance constraints to preserve perceptual quality while directly bounding hallucination risk.

2) Unified end-to-end frameworks: Jointly optimizing motion estimation, artifact correction, reconstruction, and quality assessment within a single differentiable pipeline rather than as decoupled stages.

3) Foundation-model-style self-supervised pretraining: Extending masked-image-modeling and self-supervised strategies shown effective for ViT-based correction [13] to large, multi-institutional, multi-contrast pretraining corpora to improve cross-scanner generalization.
4) Uncertainty-aware and explainable reconstruction: Incorporating calibrated uncertainty estimation (as in PATOMIR [24]) and explainability layers (as in [15]) so that radiologists can identify regions of lower reconstruction confidence.
5) Lightweight and real-time architectures: Advancing efficient designs such as SepUnet [28] and dual-domain portable-MRI networks [30] to enable deployment on low-field and point-of-care scanners with constrained computational budgets.
6) Standardized, multi-institutional validation: Establishing shared benchmarks with real (not only simulated) motion-corrupted acquisitions and consistent reporting of SSIM, PSNR, NMSE, RMSE, inference time, and downstream diagnostic accuracy to enable rigorous cross-study comparison, in line with recommendations from existing review literature [34]– [37].

VIII. CONCLUSION

This review has synthesized 40 IEEE-indexed studies on artificial-intelligence-based MRI motion artifact correction and high-quality image reconstruction, organized thematically across classical model-based methods, convolutional networks, generative adversarial networks, vision transformers, diffusion models, physics-guided k-space learning, and generalization-oriented strategies. The comparative analysis shows a clear trajectory from purely image-domain correction toward physics-constrained, generative, and self-supervised architectures that jointly pursue artifact suppression and anatomical fidelity. However, fragmented processing pipelines, limited generalization across scanners and protocols, the persistent risk of hallucinated anatomy, scarcity of real paired motion data, and inconsistent evaluation protocols remain significant barriers to routine clinical translation. Future research that integrates physics-informed constraints with generative priors, pursues unified end-to-end and uncertainty-aware architectures, and adopts standardized multi-institutional validation is expected to bridge the gap between rapid algorithmic progress and clinically trustworthy, deployable MRI motion correction and reconstruction systems.

REFERENCES

[1] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. G. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid registration using free-form deformations: Application to breast MR images," *IEEE Transactions on Medical Imaging*, vol. 18, no. 8, pp. 712–721, Aug. 1999.

[2] G. Rizzuti, A. Sbrizzi, and T. van Leeuwen, "Joint retrospective motion correction and reconstruction for brain MRI with a reference contrast," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 490–504, 2022, doi: 10.1109/TCI.2022.3183383.

[3] Ortiz-Gonzalez, E. Kobler, S. Simon, L. Bischoff, S. Nowak, A. Isaak, W. Block, A. M. Sprinkart, U. Attenberger, J. A. Luetkens, E. Bayro-Corrochano, and A. Effland, "Optical flow-guided cine MRI segmentation with learned corrections," *IEEE Transactions on Medical Imaging*, vol. 43, no. 3, pp. 940–953, Mar. 2024, doi: 10.1109/TMI.2023.3325766.

[4] T. S. B. and J. S., "Image enhancement of cardiac MR motion image for high-quality segmentation using combined fuzzy pooling layer in convolutional neural networks," in *Proc. 2023 12th International Conference on Advanced Computing (ICoAC)*, 2023, doi: 10.1109/ICoAC59537.2023.10249287.

[5] O. Dabrowski, J.-L. Falcone, A. Klauser, J. Songeon, M. Kocher, B. Chopard, F. Lazeyras, and S. Courvoisier, "SISMIK for brain MRI: Deep-learning-based motion estimation and model-based motion correction in k-space," *IEEE Transactions on Medical Imaging*, vol. 44, no. 1, pp. 396–408, Jan. 2025, doi: 10.1109/TMI.2024.3446450.

[6] J. Renders, B. Shafieizargar, M. Verhoye, J. De Beenhouwer, A. J. den Dekker, and J. Sijbers, "Delta-MRI: Direct deformation estimation from longitudinally acquired k-space data," in *Proc. 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, 2023, doi: 10.1109/ISBI53787.2023.10230697.

[7] M. Su, C. Ariyurek, S. Vasylechko, O. Afacan, and S. Kurugol, "Optimizing respiratory phase binning for improved motion correction in abdominal diffusion-weighted MRI," in *Proc. 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*, 2025, doi: 10.1109/ISBI60581.2025.10981053.

[8] K. Armanious, A. Tanwar, S. Abdulatif, T. Küstner, S. Gatidis, and B. Yang, "Unsupervised adversarial correction of rigid MR motion artifacts," in *Proc. 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, Iowa City, IA, USA, 2020, pp. 1494–1498.

[9] G. A. Saju, M. Akhi, and Y. Chang, "Ensemble CycleGAN for retrospective rigid motion correction in MRI," in *Proc. 2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2024, doi: 10.1109/EMBC53108.2024.10782023.

[10] G. Chen, H. Xie, X. Rao, X. Liu, M. Otikovs, L. Frydman, P. Sun, Z. Zhang, F. Pan, L. Yang, X. Zhou, M. Liu, Q. Bao, and C. Liu, "MRI motion correction through disentangled CycleGAN based on multi-mask k-space subsampling," *IEEE Transactions on Medical Imaging*, vol. 44, no. 4, pp. 1907–1921, Apr. 2025, doi: 10.1109/TMI.2024.3523949.

[11] T.-H. Tsai, Y.-H. Lin, and T.-H. Lin, "Motion artifact correction in MRI using GAN-based channel attention transformer," in *Proc. 2023 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 2023, doi: 10.1109/BioCAS58349.2023.10389083.

[12] J. Liu and M. Yaghoobi, "Fine-grained MRI reconstruction using attentive selection generative adversarial networks," in *Proc. 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, doi: 10.1109/ICASSP39728.2021.9414981.

[13] L. Zhang, X. Wang, E. H. Herskovits, E. R. Melhem, L. Chang, Z. Wang, and T. Ernst, "Reducing motion artifacts in brain MRI using vision transformers and self-supervised learning," in *Proc. 2024 IEEE International Conference on Image Processing (ICIP)*, 2024, doi: 10.1109/ICIP51287.2024.10647626.

[14] H. Du and Q. Han, "Multi-modal vision transformer using variable and shifted windows for MRI reconstruction," in *Proc. 2024 4th International Conference on Electronic Information Engineering and Computer Technology (EIECT)*, 2024, doi: 10.1109/EIECT64462.2024.10866009.

[15] Singh, "A motion-robust framework for MRI-based brain tumor diagnosis: Comparative evaluation of optimized Swin and hybrid transformers with explainable AI," in *Proc. 2025 IEEE MIT Undergraduate Research Technology Conference (URTC)*, 2025, doi: 10.1109/URTC68753.2025.11533011.

[16] P. Angella, V. P. Pastore, and M. Santacesaria, "Assessing the use of diffusion models for motion artifact correction in brain MRI," in *Proc. 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)*, 2025, doi: 10.1109/ISBI60581.2025.10981000.

[17] P. Angella, L. Balbi, F. Ferrando, P. Traverso, R. Varriale, V. P. Pastore, and M. Santacesaria, "DIMA: DIffusing motion artifacts for unsupervised correction in brain MRI images," *IEEE Access*, vol. 13, pp. 198649–198658, 2025, doi: 10.1109/ACCESS.2025.3634749.

[18] Y. Liu, J. Diao, Z. Zhou, H. Qi, and P. Hu, "Cardiac cine MRI motion correction using diffusion models," in *Proc. 2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, 2024, doi: 10.1109/ISBI56570.2024.10635444.

[19] C. Zhang, Y. Chen, Z. Fan, Y. Huang, W. Weng, R. Ge, D. Zeng, and C. Wang, "TC-DiffRecon: Texture coordination MRI reconstruction method based on diffusion model and modified MF-UNet method," in *Proc. 2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, 2024, doi: 10.1109/ISBI56570.2024.10635308.

[20] Bangun, Z. Cao, A. Quercia, H. Scharr, and E. Pfaehler, "MRI reconstruction with regularized 3D diffusion model (R3DM)," in *Proc. 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2025, doi: 10.1109/WACV61041.2025.00078.

[21] S. Liang, I. Alkhouri, Q. Qu, R. Wang, and S. Ravishankar, "Sequential diffusion-guided

deep image prior for medical image reconstruction," in *Proc. ICASSP 2025 – 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2025, doi: 10.1109/ICASSP49660.2025.10890153.

[22] M. U. Mirza and T. Çukur, "Super-resolution diffusion model for accelerated MRI reconstruction," in *Proc. 2023 31st Signal Processing and Communications Applications Conference (SIU)*, 2023, doi: 10.1109/SIU59756.2023.10223786.

[23] M. U. Mirza, M. A. Aydın, and T. Çukur, "Frequency-based soft diffusion model for accelerated MRI reconstruction," in *Proc. 2025 33rd Signal Processing and Communications Applications Conference (SIU)*, 2025, doi: 10.1109/SIU66497.2025.11112367.

[24] P. J. Goud and P. Spandith, "PATOMIR: A 3D latent diffusion prior for uncertainty-aware medical image reconstruction," in *Proc. 2026 International Conference on Smart Futuristic Technology (ICSFT)*, 2026, doi: 10.1109/ICSFT66733.2026.11508004.

[25] N. Pezzotti, S. Yousefi, M. S. Elmahdy, J. H. F. van Gemert, C. Schuelke, M. Doneva, T. Nielsen, S. Kastryulin, B. P. F. Lelieveldt, M. J. P. van Osch, E. de Weerdt, and M. Staring, "An adaptive intelligence algorithm for undersampled knee MRI reconstruction," *IEEE Access*, vol. 8, pp. 204825–204838, 2020, doi: 10.1109/ACCESS.2020.3034287.

[26] Z. Wang, M. Ran, Z. Yang, H. Yu, J. Jing, T. Wang, J. Lu, and Y. Zhang, "Generalizable MRI motion correction via compressed sensing equivariant imaging prior," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 12, pp. 12550–12567, Dec. 2024, doi: 10.1109/TCSVT.2024.3432751.

[27] B. Yaman, S. A. H. Hosseini, S. Moeller, and M. Akçakaya, "Comparison of neural network architectures for physics-driven deep learning MRI reconstruction," in *Proc. 2019 53rd Asilomar Conference on Signals, Systems, and Computers*, 2019.

[28] S. Zabihi, E. Rahimian, A. Asif, and A. Mohammadi, "SepUnet: Depthwise separable convolution integrated U-Net for MRI reconstruction," in *Proc. 2021 IEEE International Conference on Image Processing (ICIP)*, 2021, doi: 10.1109/ICIP42928.2021.9506285.

[29] E. Z. Chen, P. Wang, X. Chen, T. Chen, and S. Sun, "Pyramid convolutional RNN for MRI image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 41, no. 8, pp. 2033–2046, Aug. 2022, doi: 10.1109/TMI.2022.3153849.

[30] E. Ilıcak, B. Imre, C. Najac, R. van den Broek, B. Lena, A. Webb, and M. Staring, "Physics-guided dual-domain network with attention-based fusion for portable MRI reconstruction," in *Proc. 2026 IEEE 23rd International Symposium on Biomedical Imaging (ISBI)*, 2026, doi: 10.1109/ISBI61048.2026.11515406.

[31] Z. Lin, "Research on a multi-slice MRI reconstruction algorithm based on deep learning," in *Proc. 2022 International Conference on Virtual Reality, Human-Computer Interaction and Artificial Intelligence (VRHCIAI)*, 2022, doi: 10.1109/VRHCIAI57205.2022.00052.

[32] R. Hussain, K. I. Masud, and R. Islam, "Revolutionizing MRI reconstruction from compressed sensing k-space measurements: An AI-driven feature extraction approach," in *Proc. 2025 2nd International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM)*, 2025, doi: 10.1109/NCIM65934.2025.11159952.

[33] Q. Wang, V. Ecker, M. Früh, S. Gatidis, and T. Küstner, "Retrospective motion correction in MRI using disentangled embeddings," in *Proc. 2025 IEEE 10th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, 2025, doi: 10.1109/CAMSAP66162.2025.11423939.

[34] V. Spieker, H. Eichhorn, K. Hammernik, D. Rueckert, C. Preibisch, D. C. Karampinos, and J. A. Schnabel, "Deep learning for retrospective motion correction in MRI: A comprehensive review," *IEEE Transactions on Medical Imaging*, vol. 43, no. 2, pp. 846–873, Feb. 2024, doi: 10.1109/TMI.2023.3323215.

[35] K. Ranjana, "Deep learning approaches for accelerated MRI reconstruction: A review," in *Proc. 2026 International Conference on Smart Multidomain Integrated Learning*

Environments (ICSMILE), 2026, doi: 10.1109/ICSMILE69273.2026.11519289.

[36] A. Kumar and J. Kumar, "Advances in MRI reconstruction: Conventional to deep learning-based k-space acquisition strategies," in *Proc. 2025 International Conference on Electrical, Electronics, and Computer Science with Advance Power Technologies (ICE2CPT)*, 2025, doi: 10.1109/ICE2CPT66440.2025.11340332.

[37] N. A. Faizee and N. S. Mitra, "Reducing motion artifacts in MRI: From conventional techniques to deep learning-based solutions," in *Proc. 2026 IEEE International Conference on Convergence in Computing Technology (I3CTCON)*, Lonavala, India, 2026, doi: 10.1109/I3CTCON68242.2026.11507152.

[38] G. García-Jara, A. Jimenez-Molina, E. Reyes, N. Tapia-Rivas, C. Ramos-Gómez, J. de Grazia, and M. Sepúlveda, "Efficient and motion correction-free myocardial perfusion segmentation in small MRI data using deep transfer learning from cine images: A promising framework for clinical implementation," *IEEE Access*, vol. 11, pp. 103177–103188, 2023, doi: 10.1109/ACCESS.2023.3313980.

[39] M. Kathiravan, R. Shyamala Devi, M. Martina Jose Mary, A. Anandha Baskar, R. Gokula Priyan, and J. Gokulanathan, "Hybrid deep learning approach for medical image denoising with EfficientNetB3 and Pix2Pix," in *Proc. 2025 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 2025, doi: 10.1109/SCEECS64059.2025.10940920.

[40] V. R. and V. K., "Multi-domain medical image enhancement and diagnosis: A unified deep generative framework with transferable learning and attention mechanisms," in *Proc. 2025 IEEE 9th International Conference on Information and Communication Technology (CICT)*, 2025, doi: 10.1109/CICT67193.2025.11399216.