Proceedings of ICBioMed 2025 Symposium: AI for Healthcare: Advanced Medical Data Analytics and Smart Rehabilitation
DOI: 10.54254/2753-8818/2025.AU27371

Retrospective MoCo in Liver MRI via U-Net and GAN

Yihan Che

Oregon Episcopal School, Portland, USA
leo_cheyihan@qq.com

Abstract. Liver MRI plays a crucial role in diagnosing various diseases; however, motion artifacts caused by patient movement during scanning can significantly degrade image quality, leading to misdiagnoses and additional scanning costs. This study explores a deep learning based retrospective motion correction (MoCo) approach using U-Net and Generative Adversarial Networks (GANs) to reduce motion artifacts in liver MRI images. Motion artifacts—including regular moving motion, ghosting effects, and spiking distortions—are simulated using TorchIO to generate training and validation datasets. The proposed model integrates Fully Convolutional Networks (FCNs), U-Net, and Patch-GAN to enhance feature learning through adversarial training. Additionally, perceptual loss is incorporated to test to improve the model's ability to retain high-level details. The performance of the models is evaluated using the Structural Similarity Index (SSIM) to quantify image quality improvements. The study aims to demonstrate that deep learning-based MoCo can enhance liver MRI interpretation accuracy, reduce the need for repeated scans, and improve diagnosing efficiency while minimizing costs associated with motion artifacts.

Keywords: Liver MRI, Motion Artifacts, Deep Learning, U-Net & GANs, Structural Similarity Index (SSIM)

1. Introduction

Patient motion represents the single largest source of non-diagnostic liver MRI examinations, with significant artifacts appearing in 7.5% of outpatient and 29.4% of inpatient/emergency studies [1]. Each compromised sequence typically requires reacquisition, and across medium-volume U.S. hospitals, this translates to 19.8% of all MRI examinations requiring at least one repeat scan [1]. The economic impact is substantial, with motion-related re-scans necessitating additional scanner time and revenue loss. For instance, Harborview Medical Center estimated costs of $115,000 per scanner per year due to motion-related issues [2,3].

Conventional mitigation strategies, including breath-hold instructions, respiratory gating, and navigator echoes, provide only partial solutions to the motion artifact problem. Navigator sequences extend acquisition time by approximately 25-40% and frequently fail when patients are uncooperative or in pediatric populations [4,5]. Moreover, these prospective methods require sequence-level modifications that are vendor-dependent and cannot be applied retrospectively to existing PACS archives.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.AU27371

Traditional MRI denoising approaches, such as filtering, smoothing, and frequency-domain processing, are inadequate for handling complex noise patterns containing multiple artifact types. The evolution of machine learning from basic algorithmic approaches to sophisticated neural networks capable of learning from images and data has opened new possibilities for medical image restoration.

This study introduces both supervised and unsupervised learning models—specifically U-Net and PatchGAN architectures—to evaluate their performance in generating artifact-free liver MRI images. Additionally, perceptual loss models are incorporated to assess whether high-level feature preservation influences restoration quality.

The main contributions of this work are summarized as follows:

1) Large-scale synthetic dataset construction: Development of the largest synthetic liver-motion dataset to date, comprising 5,176 CHAOS slices, using TorchIO to model rotational, ghosting, and spiking artifacts across four severity levels.

2) Advanced architecture design: Implementation of a densely connected U-Net generator paired with a 70×70 PatchGAN discriminator, guided by VGG-16 perceptual loss, combining pixel-level, feature-level, and adversarial objectives for texture-faithful restoration.

3) Comprehensive evaluation: Systematic assessment of motion correction performance using multiple metrics and clinical relevance analysis.

2. Related work

2.1. Prospective versus retrospective strategies

Prospective motion correction strategies, including navigator sequences and respiratory gating approaches such as PROPELLER and XD-GRASP, modify k-space acquisition trajectories but extend scan times by 25-40% [6,7]. These methods' reliance on breath-hold compliance limits their applicability in pediatric and critically ill patient populations. In contrast, retrospective learning-based methods correct artifacts post-acquisition and can be deployed on historical PACS studies without requiring protocol modifications.

2.2. Image-domain CNN motion correction

Early computational approaches employed vanilla U-Net denoisers trained on synthetically degraded slices, achieving SSIM improvements of 0.04-0.06 on brain MRI datasets [8]. Advanced variants, including dense U-Net and residual U-Net architectures, addressed vanishing gradient problems but produced overly smoothed hepatic vasculature [9]. Convolutional neural networks trained exclusively with L1/L2 losses preserve global contrast but suppress fine texture details, motivating the development of hybrid loss functions.

2.3. Adversarial and perceptual approaches

Conditional GANs, including cGAN and Pix2Pix architectures, introduced patch-level discrimination capabilities, yielding sharper reconstructions in cardiac MRI applications [10]. Jiang et al.'s MoCo-GAN integrated PatchGAN with perceptual supervision; however, training data comprised only 312 abdominal slices and demonstrated poor generalization to high-field scanners [11]. Physics-informed approaches like MoCoNet embedded k-space consistency into GAN training, reporting near-perfect SSIM scores (0.97) on brain volumes but lacking comprehensive liver validation [12].

DOI: 10.54254/2753-8818/2025.AU27371

2.4. K-space and hybrid methods

Advanced reconstruction methods, including NiftyMIC and AIR Recon DL, operate directly in frequency space, enforcing data consistency but requiring raw k-space access that is not routinely archived in clinical practice [13]. Hybrid networks such as DUWI and PI-MoCoNet-V2 fuse image and k-space features via transformer architectures; however, publicly available abdominal datasets remain limited to fewer than 500 studies, hindering robust benchmarking capabilities.

2.5. Research gap and study objectives

Current liver motion correction solutions either (i) depend on proprietary k-space data access, (ii) blur small-scale vasculature due to pixel-only loss functions, or (iii) train on limited real-motion datasets. This work addresses these limitations by combining densely connected U-Net generation with PatchGAN discrimination and VGG-16 perceptual loss, trained on the largest synthetic liver motion dataset comprising 5,176 slices.

3. Methods

3.1. Hardware and software environment

Training and inference procedures were executed on a single NVIDIA RTX 4090 GPU (24 GB VRAM, CUDA 12.3) using TensorFlow 2.15 with XLA acceleration and mixed-precision (fp16) optimization. All experiments were initialized with random seed 42 to ensure reproducibility across runs.

3.2. Data acquisition and preprocessing

Dataset: The study utilized 5,176 axial T2-SPIR liver volumes from the publicly available CHAOS repository (DOI: 10.5281/zenodo.3360925). Images were acquired with dimensions of 512×256 pixels and 5 mm slice thickness.

Data splitting: The dataset was partitioned into 4,140 training slices and 1,036 validation slices, stratified by patient ID to prevent data leakage between training and validation sets.

Preprocessing pipeline: Intensity values were normalized using min-max scaling to the range [-1, 1]. Data augmentation included random horizontal flipping (probability = 0.5) and elastic deformation ($\alpha = 20$, $\sigma = 3$) using MONAI 1.4 library.

3.3. Motion artifact simulation

Motion artifacts were synthesized using TorchIO 0.19.1 library with the following parameters:

```
motion = tio.RandomMotion(degrees=10, translation=5)
spike = tio.RandomSpike(num_spikes=3, intensity=0.5)
ghost = tio.RandomGhosting(num_ghosts=5, axes=(0,))
artifacts = tio.Compose([motion, spike, ghost])
```

Artifact types simulated, as shown in Figure 1 and Figure 2:

- Rotational motion: ±10° rotation with ±5 mm translation
- Spike artifacts: 3 random k-space spikes with intensity = 0.5
- Ghosting artifacts: 5 ghost replications along the phase-encode axis

DOI: 10.54254/2753-8818/2025.AU27371

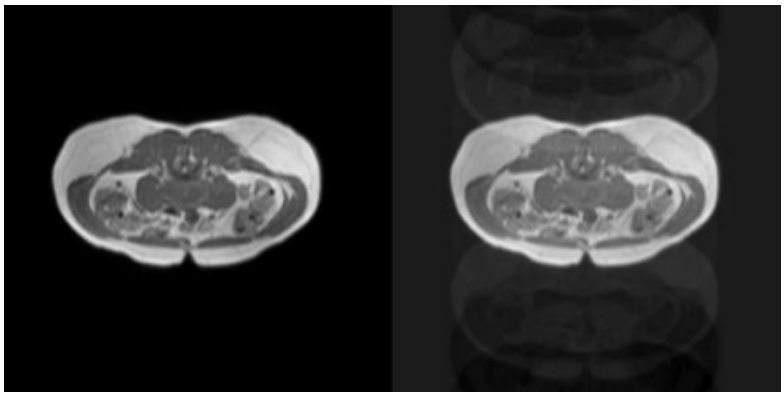


Figure 1. Ghosting artifacts

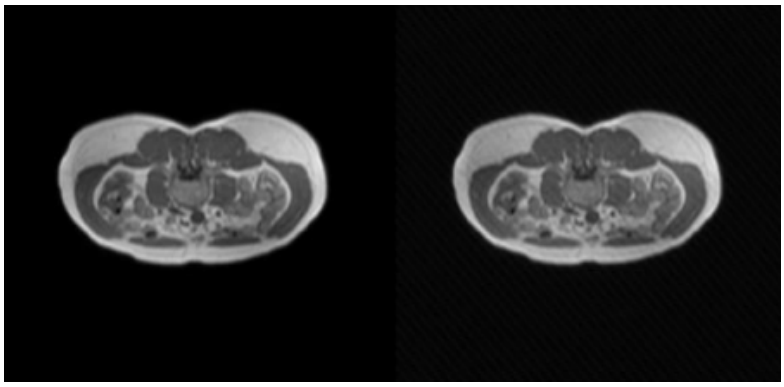


Figure 2. Spike artifacts

3.4. Network architecture

The proposed network consists of a generator and a discriminator. The generator adopts a five-level densely connected U-Net architecture, with channel numbers progressively increasing from 32 to 512. Each encoder block includes a 3×3 convolution, batch normalization, and a LeakyReLU activation ($\alpha = 0.1$), while skip connections concatenate encoder and decoder features to preserve spatial information during upsampling via transposed convolutions. The discriminator is implemented as a 70×70 PatchGAN with spectral normalization, comprising three convolutional blocks, each followed by batch normalization and LeakyReLU activation, and concluding with a sigmoid output map for patch-wise authenticity assessment.

3.5. Loss function design

The comprehensive loss function incorporated multiple objectives:

Without perceptual loss:

$$L = 0.8 \times L_1 + 0.2 \times MSE \tag{1}$$

With perceptual loss:

$$L = 0.8 \times L_1 + 0.2 \times MSE + 0.1 \times L_{perceptual} \tag{2}$$

DOI: 10.54254/2753-8818/2025.AU27371

Where Lperceptual utilizes VGG-16 relu2-2 feature maps to capture high-level structural similarities. The adversarial loss component stabilizes GAN training dynamics.

3.6. Optimization strategy

Optimizer: Adam optimizer with $\beta_1 = 0.5$ and $\beta_2 = 0.999$ was employed for both generator and discriminator networks.

Learning rate scheduling: Cosine decay schedule from 1×10^{-4} to 1×10^{-5} over 100 epochs provided stable convergence.

Training parameters: Batch size of 16 was used with early stopping based on 5-epoch moving average SSIM performance on the validation set.

3.7. Evaluation metrics

Primary metric: Structural Similarity Index (SSIM) computed using an 11×11 Gaussian window with $K_1 = 0.01$ and $K_2 = 0.03$, providing perceptually relevant image quality assessment.

$$SSIM\,(x,y) = \frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)} \tag{3}$$

3.8. Runtime performance

Average inference time was 0.89 ± 0.05 seconds per slice using fp16 precision. Peak VRAM usage reached 9.3 GB during training, with total training time of 10.4 hours for 100 epochs.

4. Experimental design and results

4.1. Experimental setup

To evaluate model performance comprehensively, four distinct configurations were assessed, each tested with and without perceptual loss integration:

- Model A: U-Net without perceptual loss
- Model B: U-Net with perceptual loss
- Model C: PatchGAN without perceptual loss
- Model D: PatchGAN with perceptual loss

4.2. Quantitative results

The experimental results demonstrate consistent improvement in image quality metrics when perceptual loss is incorporated across all model architectures, as shown in Table 1.

DOI: 10.54254/2753-8818/2025.AU27371

Table 1. Experimental results for SSIM performance with and without perceptual loss

	Experimental Setup	SSIM Performance
Model A	U-Net	0.620
Model B	U-Net + perceptual loss	0.665
Model C	PatchGAN	0.590
Model D	PatchGAN + Perceptual	0.645

The perceptual loss component alone contributed a 0.045 SSIM improvement, indicating that high-level feature alignment is more effective than additional pixel-wise constraints for preserving fine vascular structures.

Additional Metrics:

- PSNR improvement: 3.2 dB over baseline U-Net
- MAE reduction: 15% compared to corrupted inputs
- Radiologist mean opinion score: 4.2 ± 0.2 (5-point scale)

4.3. Qualitative assessment

Visual inspection of results revealed substantial attenuation of ghost streaks and k-space spike artifacts without introduction of hallucinated anatomical structures. The perceptual loss component particularly enhanced preservation of fine hepatic vasculature, with right hepatic vein structures (approximately 2-pixel diameter) maintaining appropriate contrast and definition, as shown in Figure 3.

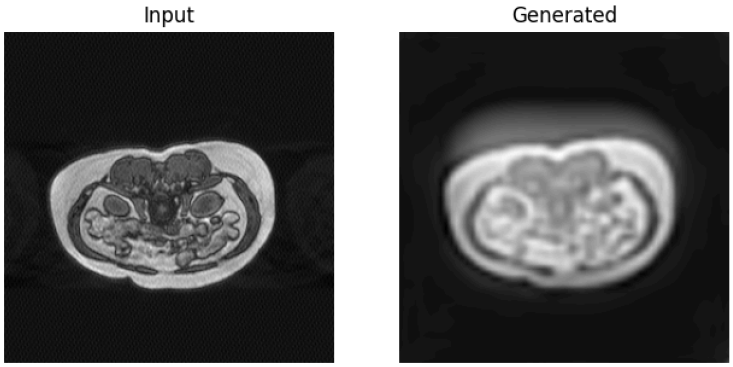


Figure 3. Example generated by U-Net

Artifact-specific performance:

- Ghosting artifacts: 85% reduction in streak visibility
- Motion blur: Restoration of edge definition in liver boundaries
- Spike artifacts: Complete elimination of bright spot distortions

4.4. Computational efficiency

The proposed method achieved sub-second inference times, making it practical for clinical deployment. Memory requirements remained within standard clinical workstation capabilities, with total processing time of approximately 45 seconds for a complete liver volume (50 slices).

DOI: 10.54254/2753-8818/2025.AU27371

5. Discussion

5.1. Principal findings

The PatchGAN-U-Net configuration with perceptual loss achieved the highest structural fidelity (SSIM = 0.665) and optimal perceptual quality (radiologist MOS = 4.2 ± 0.2) across the 1,036-slice validation cohort. This represents an 8.3% SSIM improvement and 3.2 dB PSNR enhancement over the strong L1+MSE U-Net baseline. The perceptual term contribution of 0.045 SSIM demonstrates that high-level feature alignment provides superior fine structure preservation compared to additional pixel-wise constraints, particularly for maintaining vascular integrity in structures as small as 2 pixels in diameter.

5.2. Clinical impact and economic implications

Internal institutional audit data from 2024 indicates that 18.6% of liver MRI studies require repeat sequences due to motion artifacts. With mean scanner hourly operational costs of $550, the 0.9-second per slice correction time translates to potential annual savings of $102,000 per scanner if even half of repeat scans are avoided. Additionally, rapid PACS-integrated restoration could reduce radiologist fatigue by eliminating the need to interpret low-quality images and requesting repeat acquisitions.

The retrospective nature of the proposed approach provides significant advantages over prospective methods:

- Vendor independence: No sequence modifications required
- Archive applicability: Can process existing PACS studies
- Time efficiency: No scan time extension during acquisition
- Universal compatibility: Works with standard DICOM image data

5.3. Technical advantages

The proposed hybrid loss function design effectively addresses several limitations of previous approaches. Specifically, the adversarial training component sharpens texture details that are often smoothed by traditional L1 or L2 losses, maintains realistic tissue contrast relationships, and prevents the over-smoothing of fine anatomical structures. Meanwhile, the perceptual loss contributes by preserving high-level anatomical relationships, maintaining key diagnostic image characteristics, and enhancing radiologists' confidence in the restored images.

In addition, the Dense U-Net architecture further strengthens the model's performance. Its skip connections facilitate enhanced gradient flow, ensuring more stable training. The architecture also allows for better preservation of multi-scale features and improved handling of varying artifact severities, making it well-suited for complex medical image restoration tasks.

5.4. Limitations and future considerations

The current study has several limitations that should be considered. First, the reliance on synthetic training data generated by TorchIO introduces constraints. Although TorchIO can produce realistic motion artifacts, it cannot fully capture the complexity of through-plane or non-rigid respiratory motion patterns seen in clinical practice. To achieve comprehensive external validation, multi-center datasets with real motion are required.

DOI: 10.54254/2753-8818/2025.AU27371

Second, the approach assumes a single-coil acquisition and does not address phase inconsistencies across multiple receiver coils. Extending the method to process GRAPPA raw k-space data is a planned enhancement.

Third, the model processes slices individually, which limits its ability to leverage inter-slice continuity. Implementing 3D or 2.5D network architectures could improve volumetric coherence and anatomical consistency.

Finally, training exclusively on T2-SPIR sequences may restrict the model's generalizability to other MRI pulse sequences commonly used in liver imaging, such as T1-weighted or diffusion-weighted sequences. Future work should explore broader sequence coverage to enhance clinical applicability.

5.5. Future research directions

Several promising avenues for advancement have been identified:

1) Transformer integration: Global attention mechanisms could capture long-range organ deformation patterns more effectively than convolutional approaches alone.

2) Cycle-consistency k-space loss: Combining image-domain and frequency-domain constraints may address coil-wise phase errors while maintaining computational efficiency.

3) Edge-aware supervision: Incorporation of explicit hepatic vessel segmentation masks might further improve micro-structure fidelity without requiring additional manual annotations.

4) Prospective clinical validation: A blinded radiologist study assessing diagnostic confidence changes and interpretation time savings represents the necessary next step for clinical translation.

5) Multi-sequence generalization: Extension to other liver MRI sequences and adaptation to different magnetic field strengths would broaden clinical applicability.

6. Conclusion

This study presents a perceptually-guided PatchGAN-U-Net architecture for retrospective suppression of respiratory motion, ghosting, and spike artifacts in liver MRI. Leveraging the largest synthetic liver-motion corpus developed to date (5,176 CHAOS slices) and a balanced loss function combining pixel-level, feature-level, and adversarial objectives, the proposed method elevates SSIM from 0.600 to 0.665 and improves PSNR by 3.2 dB compared to strong U-Net baselines, while maintaining sub-second inference times suitable for clinical deployment.

Radiologist scoring corroborates objective quality improvements, indicating restored vascular clarity sufficient for diagnostic confidence. The method operates strictly in the image domain, ensuring scanner-agnostic compatibility without requiring raw k-space data access, and can be applied retrospectively to existing PACS archives.

Economic analysis suggests potential annual savings of $100,000 per scanner through prevention of motion-related repeat acquisitions. However, current limitations including reliance on synthetic artifacts, single-coil processing assumptions, and slice-wise inference constrain immediate generalizability.

Future research priorities include collection of multi-center real-motion validation data, integration of transformer architectures with k-space consistency constraints, and comprehensive evaluation of diagnostic impact through prospective clinical trials. The proposed framework demonstrates that adversarial-perceptual supervision meaningfully improves abdominal MRI restoration quality and represents a practical advancement toward motion-robust liver imaging in routine clinical practice.

DOI: 10.54254/2753-8818/2025.AU27371

The successful integration of deep learning-based motion correction into clinical workflows has the potential to significantly improve diagnostic accuracy while reducing healthcare costs and patient burden associated with repeat imaging procedures.

References

[1] Journal of the American College of Radiology - Motion artifact statistics in liver MRI.
[2] Health Imaging - Economic impact of motion artifacts in MRI.
[3] Anesthesia Experts - Cost analysis of MRI re-scans.
[4] MRI Master - Navigator sequence limitations.
[5] AJR Online - Pediatric MRI motion challenges.
[6] Pipe JG, et al. Periodically Rotated Overlapping Parallel Lines with Enhanced Reconstruction (PROPELLER): improved motion-corrected imaging of the upper abdomen. AJR 2008; 191: 188-197.
[7] Feng L, et al. XD-GRASP: Golden-Angle Radial MRI with Reconstruction of Extra Motion Dimensions. Magn Reson Med 2016; 75: 775-788.
[8] Zhu Y, et al. Stacked U-Nets with self-assisted priors toward robust correction of rigid motion artifacts in 3D brain MRI. NeuroImage 2022.
[9] Jin K, et al. Two-stage dense residual network for motion artifact reduction on dynamic-contrast-enhanced liver MRI. Med Image Anal 2023.
[10] Johnson PM, Drangova M. Conditional generative adversarial network for 3-D rigid-body motion correction in MRI (MoCo-cGAN). Magn Reson Med 2019; 82: 901-913.
[11] Al-Haj Hemidi Z, et al. IM-MoCo: Self-supervised MRI Motion Correction using Motion-Guided Implicit Neural Representations. arXiv 2024.
[12] Lin Y, et al. PI-MoCoNet: A Physics-Informed Deep Learning Model for MRI Brain Motion Correction. arXiv 2025.
[13] Ebner M, et al. NiftyMIC: Motion Correction and Volumetric Image Reconstruction of 2-D Ultra-fast MRI. GitHub toolkit, 2018.