



OPEN ACCESS

EDITED BY

Andrew Scott,
Royal Brompton Hospital, United Kingdom

REVIEWED BY

Redha Boubertakh,
National Heart Centre Singapore, Singapore
Johnathan Le,
The University of Utah, United States

*CORRESPONDENCE

Changyu Sun
✉ csyfc@missouri.edu

RECEIVED 21 September 2024

ACCEPTED 10 January 2025

PUBLISHED 24 January 2025

CITATION

Sun C, Goyal N, Wang Y, Tharp DL, Kumar S and Altes TA (2025) Conditional diffusion-generated super-resolution for myocardial perfusion MRI.
Front. Cardiovasc. Med. 12:1499593.
doi: 10.3389/fcvm.2025.1499593

COPYRIGHT

© 2025 Sun, Goyal, Wang, Tharp, Kumar and Altes. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Conditional diffusion-generated super-resolution for myocardial perfusion MRI

Changyu Sun^{1,2*}, Neha Goyal³, Yu Wang¹, Darla L. Tharp⁴, Senthil Kumar³ and Talissa A. Altes²

¹Department of Chemical and Biomedical Engineering, University of Missouri, Columbia, MO, United States, ²Department of Radiology, University of Missouri, Columbia, MO, United States, ³Department of Medicine, University of Missouri, Columbia, MO, United States, ⁴Department of Biomedical Sciences, University of Missouri, Columbia, MO, United States

Introduction: Myocardial perfusion MRI is important for diagnosing coronary artery disease, but current clinical methods face challenges in balancing spatial resolution, temporal resolution, and slice coverage. Achieving broader slice coverage and higher temporal resolution is essential for accurately detecting abnormalities across different slice locations but remains difficult due to constraints in acquisition speed and heart rate variability. While techniques like parallel imaging and compressed sensing have significantly advanced perfusion imaging, they still suffer from noise amplification, residual artifacts, and potential temporal blurring due to the rapid transit of dynamic contrast vs. the temporal constraints of the reconstruction.

Methods: This study introduces a conditional diffusion-based generative model for myocardial perfusion MRI super resolution, addressing the trade-offs between spatiotemporal resolution and slice coverage. We adapted Denoising Diffusion Probabilistic Models (DDPM) to enhance low-resolution perfusion images into high-resolution outputs without requiring temporal regularization. The forward diffusion process introduces Gaussian noise incrementally, while the reverse process employs a U-Net architecture to progressively denoise the images, conditioned on the low-resolution input image.

Results: We trained and validated the model on a retrospective dataset of dynamic contrast-enhanced (DCE) perfusion MRI, consisting of both stress and rest images from 47 patients with heart disease. Our results showed significant image quality improvements, with a 5.1% reduction in nRMSE, a 1.1% increase in PSNR, and a 2.2% boost in SSIM compared to GAN-based super-resolution method ($P < 0.05$ for all metrics using paired t -test) in retrospective study. For the 9 prospective subjects, we achieved a total nominal acceleration of 8.5-fold across 5–6 slices through a combination of low-resolution acquisition and GRAPPA. PerfGen outperformed GAN-based approach in sharpness (4.36 ± 0.38 vs. 4.89 ± 0.22) and overall image quality (4.14 ± 0.28 vs. 4.89 ± 0.22), as assessed by two experts in a blinded evaluation ($P < 0.05$) in prospective study.

Discussion: This work demonstrates the capability of diffusion-based generative models in generating high-resolution myocardial perfusion MRI from conditional low-resolution images. This approach has shown the potentials to accelerate myocardial perfusion MRI while enhancing slice coverage and temporal resolution, offering a promising alternative to existing methods.

KEYWORDS

super-resolution, myocardial perfusion MRI, deep learning, diffusion probabilistic models (DDPM), conditional generative model, dynamic contrast-enhanced MRI (DCE MRI)

1 Introduction

Improving myocardial perfusion MRI is critical for assessing perfusion defects, requiring a balance between high spatial resolution, temporal fidelity, and slice coverage (1–3). Clinically, sufficient spatial resolution is necessary to detect subtle perfusion abnormalities but achieving enough spatial resolution (<3.0 mm) (4) and extensive slice coverage is particularly challenging under high heart rate conditions. The need to capture more slices (≥ 3 slices) within a short acquisition window further complicates the ability to fully resolve both motion and perfusion dynamics (5).

Recent techniques such as parallel imaging and compressed sensing (6–8), using both Cartesian (2) and non-Cartesian sampling (7, 9), have made progress in accelerating acquisition and increasing resolution in myocardial perfusion MRI. However, there remains open questions regarding the trade-offs between spatial and temporal fidelity, motion correction (10), as well as the potential for residual artifacts (11). Moreover, these methods typically require the complete acquisition of the entire temporal series to apply temporal regularization (1, 2) and motion correction, which can hinder the ability to display real-time images during contrast inflow and washout.

Given the ongoing challenges, there remains a need for alternative strategies to increase imaging speed for high temporal resolution and expanded slice coverage while simultaneously maintaining spatiotemporal fidelity. Low-resolution (LR) acquisitions inherently allow for faster imaging and higher signal-to-noise ratio (SNR), which can be crucial for capturing rapid contrast changes and minimizing the effects of cardiac and respiratory motion. By leveraging super-resolution (SR) methods (12, 13), these images can be enhanced to achieve higher spatial fidelity, offering a balance between imaging speed and diagnostic quality.

However, low-resolution perfusion may suffer from reduced spatial details and fidelity, as well as more severe dark rim artifacts and partial volume effects (14). These artifacts can interfere with accurate perfusion analysis and affect diagnostic outcomes. To address this, strategies must be developed to compensate for the loss of spatial resolution and mitigate artifacts. Recent advances in deep learning, particularly generative models, provide a promising way for enhancing the quality of LR images in cardiac MR imaging (15–18). However, Generative Adversarial Networks (GAN) are prone to experience unstable training and mode collapse issues (19). In contrast, diffusion models have proven to produce high-quality images with robust training stability and superior image quality (12, 20–21). Additionally, diffusion generative models offer a robust mechanism for improving spatial resolution of myocardial perfusion MRI without relying on temporal regularization. Previous studies have investigated GAN-based generative models on cardiac MRI (15, 18), but applying generative models to myocardial perfusion MRI has not been explored. By conditioning the diffusion model on low-resolution perfusion images, it is possible to enhance image detail while retaining the benefits of rapid acquisition, high temporal resolution and expanded slice coverage.

We propose to develop a conditional diffusion-based generative model for myocardial perfusion MRI super resolution, termed PerfGen, that leverages existing clinical imaging protocols and data to generate myocardial perfusion images conditioned by low-resolution images. This study explores the proof-of-the-concept that diffusion generative models can be integrated with myocardial perfusion MRI to synthesize high-resolution (HR) perfusion images and demonstrated its feasibility to accelerate the acquisition. This model provides an alternative solution that balances spatial resolution, temporal fidelity, and slice coverage, offering a new way for efficient and high-quality myocardial perfusion MRI.

2 Materials and methods

2.1 Data acquisition and preprocessing

All patients provided informed consent, and all studies were performed in accordance with protocols approved by our institutional review board.

2.1.1 Retrospective myocardial perfusion data

Dynamic contrast-enhanced (DCE) perfusion data were collected from 47 heart disease patients using standard clinical MRI protocols at the University of Missouri-Columbia Hospital. The dataset was divided into an 80:20 split, with 38 patients for training and 9 patients for testing. Each subject had 3 short-axis slices (base, mid, and apex), with all temporal frames used, resulting in a total of 8,040 images for training and 1,830 images for testing. A mixed dataset with both rest and stress perfusion data were collected using gadolinium contrast for perfusion and Regadenoson for stress with free breathing acquisition. Prospective electrocardiogram triggering was used for all patients. Within the training group, 15 patients underwent rest perfusion only, and 23 underwent stress perfusion only. For the testing group, 3 subjects were assessed under rest conditions and 6 under stress. All testing data and most training data were acquired on a 1.5 T MAGNETOM Aera (Siemens Healthineers), except for 4 subjects from the training group were imaged using a 3 T MAGNETOM Vida (Siemens Healthineers).

Imaging parameters for the gradient echo perfusion sequence included a repetition time of 2.2–2.3 ms, echo time of 1.08 ms, flip angles between 12° and 15°, resolution of 2.3–2.4 mm \times 2.3–2.4 mm, 60–80 temporal measurements, and a GRAPPA acceleration rate of 2. We use chest and spine phased-array receiver coils (20–34 channels) with an acquisition matrix of 160 \times 120–160, a temporal resolution of 138–184 ms per slice, a saturation pulse delay of 100–120 ms, and acquire 3 slices per R-wave peak to R-wave peak (RR) interval.

2.1.2 Prospective myocardial perfusion data

Nine DCE rest perfusion patient data were collected at University of Missouri-Columbia Hospital using a 3 T MAGNETOM Vida (Siemens Healthineers), with 8 acquired using GRAPPA-3 and 1 using GRAPPA-2. Prospective electrocardiogram triggering was

used for all patients. Five to six short-axis myocardial perfusion slices were acquired per RR interval during free breathing. Imaging parameters of the gradient echo perfusion sequence included a repetition time of 2.2–2.3 ms, echo time of 1.08 ms, flip angles between 12° and 15°, resolution of 2.3–2.4 mm × 2.3–2.4 mm, 60 temporal measurements, and a GRAPPA acceleration rate of 2–3. Late gadolinium enhancement (LGE) imaging was used as a reference for validating the super-resolved perfusion defects, particularly in the presence of late enhanced regions. The acquisition matrix size is 160 × 48–62 following a 35%–36% low-resolution acceleration, with 16–22 actual phase encoding lines acquired using GRAPPA-3. The temporal resolution was 36.8–50.6 ms per slice, with an inversion time of 100 ms for the saturation pulse, resulting in a total acquisition time of 118.4–125.3 ms per slice. For the GRAPPA-2 data, the acquisition matrix size is 176 × 62 following a 35% low-resolution acceleration, with 32 actual phase encoding lines acquired using GRAPPA-2. The temporal resolution was 73.6 ms per slice, resulting in a total acquisition time of 136.8 ms.

2.1.3 Low-resolution data preparation

To simulate LR perfusion data from HR perfusion images for model training, we used the following steps to generate LR and HR pairs (Figure 1). Fast Fourier Transform was applied to the original HR magnitude image to convert to k-space domain for retrospective experiments. The center 30%–50% of the phase-encoding lines were used, with the dynamic low-resolution ratio aiming for data augmentation. The outer k-space lines were zero-padded with the data in the readout maintained and converted back to the image domain using an inverse fast Fourier transform followed by taking the absolute value. Both the synthesized LR images and paired HR images were cropped to the same central 96 × 96 matrix size followed by the image normalization. The cropping was performed manually to position

the heart within the cropped region. On-the-fly data augmentation included random vertical flip and horizontal flip, each applied with a probability of 50%. HR perfusion images served as the reference, and the LR synthetic images were enhanced using zero-padding and PerfGen. In the prospective study, the multi-coil complex-valued k-space data was truncated by setting the phase resolution to 35% in the sequence.

2.2 Conditional generation with denoising diffusion probabilistic models

Given a dataset of LR-HR perfusion MRI pairs, $D = (x_i, y_i)_{i=1}^N$, which are samples from an unknown conditional distribution of high-resolution myocardial perfusion MRI domain, a parametric approximation of $p(y|x)$ was learned through a stochastic iterative refinement process that maps the source LR image x to target HR image \hat{y}_0 . We adapted the Denoising Diffusion Probabilistic Models (DDPM) and Image Super-Resolution via Iterative Refinement (SR3) (12) to generate HR MR perfusion images from LR image through diffusion process.

Figure 2A provides an illustration of a conditional diffusion-based model to map Gaussian noise y_T to a HR image \hat{y}_0 , conditioned on the source LR image x . The forward diffusion process q follows the Markov process to gradually add Gaussian noise to the HR perfusion image y_0 step by step until the image converges to a pure Gaussian distribution y_T . The reverse process p utilizes a U-Net model (22), trained to conditionally denoise the image to reconstruct a HR perfusion image \hat{y}_0 using the LR perfusion image x as the guidance.

2.2.1 Forward diffusion process

The forward diffusion process gradually added Gaussian noise to the HR perfusion image y_0 over T iterations until

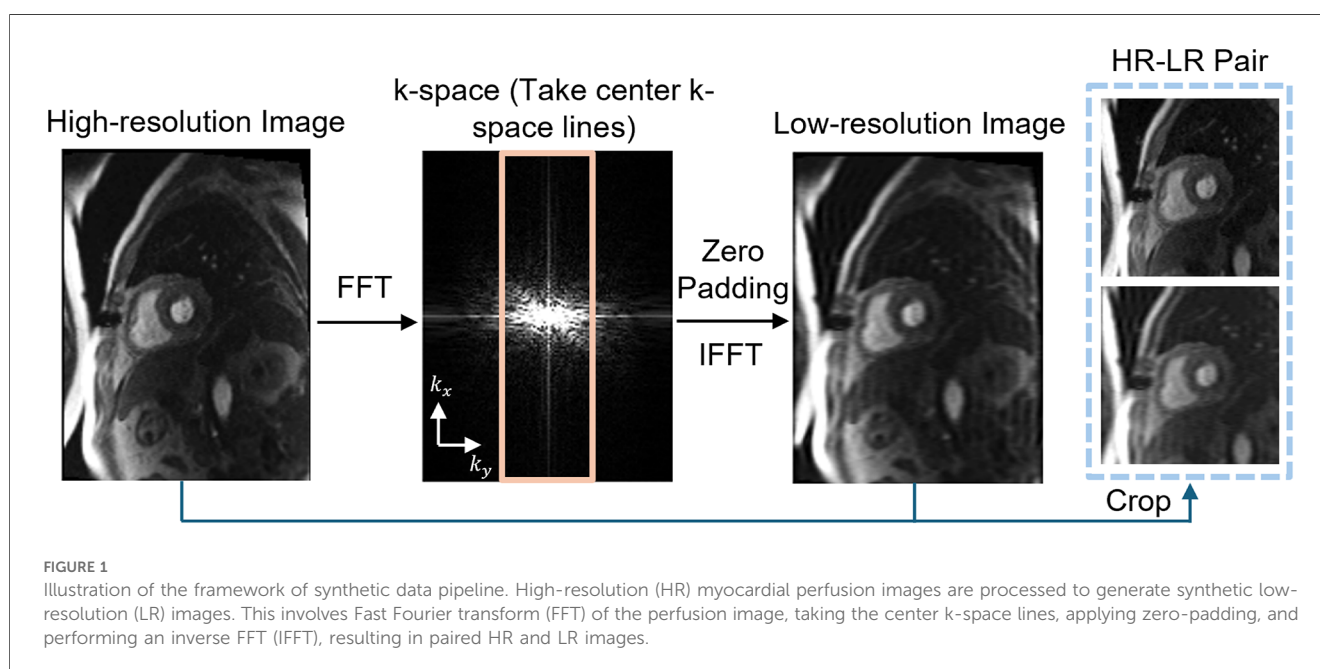


FIGURE 1

Illustration of the framework of synthetic data pipeline. High-resolution (HR) myocardial perfusion images are processed to generate synthetic low-resolution (LR) images. This involves Fast Fourier transform (FFT) of the perfusion image, taking the center k-space lines, applying zero-padding, and performing an inverse FFT (IFFT), resulting in paired HR and LR images.

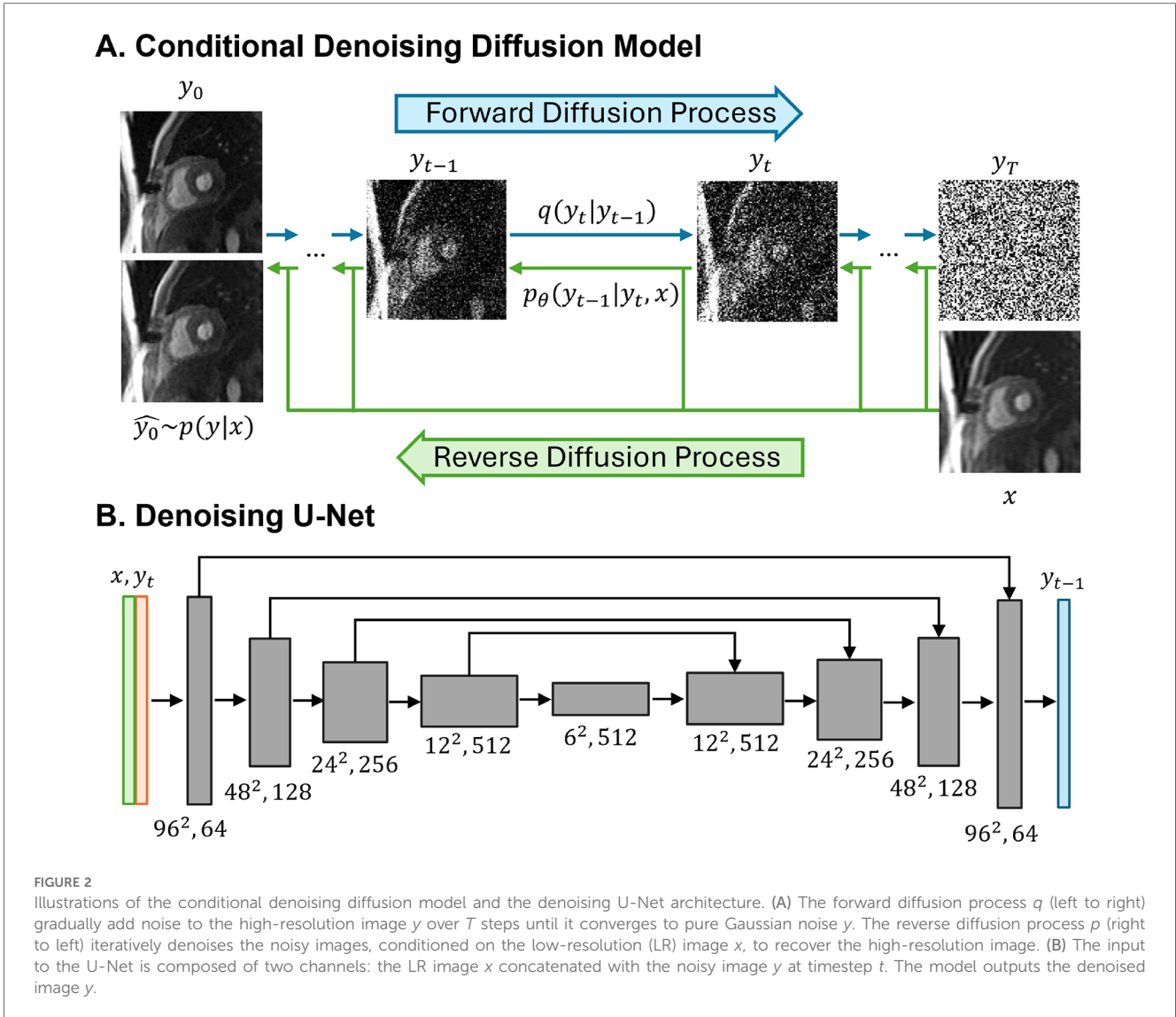


FIGURE 2

Illustrations of the conditional denoising diffusion model and the denoising U-Net architecture. (A) The forward diffusion process q (left to right) gradually add noise to the high-resolution image y over T steps until it converges to pure Gaussian noise y . The reverse diffusion process p (right to left) iteratively denoises the noisy images, conditioned on the low-resolution (LR) image x , to recover the high-resolution image. (B) The input to the U-Net is composed of two channels: the LR image x concatenated with the noisy image y at timestep t . The model outputs the denoised image y .

the image converges to a Gaussian distribution via the diffusion kernel (Equation 1). Equation 2 provides the complete generation process.

$$q(y_t | y_{t-1}) = N(y_t | \sqrt{\alpha_t}y_{t-1}, (1 - \alpha_t)I), \tag{1}$$

$$q(y_{1:T} | y_0) = \prod_{t=1}^T q(y_t | y_{t-1}), \tag{2}$$

where α_t , $1 \leq t \leq T$ are variance schedule subject to $0 < \alpha_t < 1$, I is the identity matrix. T is set to 2,000, and the added Gaussian noise to the HR image generated a sequence of noisy images with increasing noise level $y \in [y_1, y_2, \dots, y_T]$.

Specifically, $q(y_t)$ can be obtained directly from y_0 at any time step without iterations where $\gamma_t = \prod_{i=1}^t \alpha_i$ (Equation 3).

$$q(y_t | y_0) = N(y_t | \sqrt{\gamma_t}y_0, (1 - \gamma_t)I), \tag{3}$$

2.2.2 Reverse Diffusion Process

The reverse diffusion process is defined as a reverse Markov process, starting from Gaussian noise y_T and progressively denoised to reconstruct the HR perfusion images \hat{y}_0 (Equations 4–6):

$$p_\theta(y_{0:T} | x) = p(y_T) \prod_{t=1}^T p_\theta(y_{t-1} | y_t, x), \tag{4}$$

$$p(y_T) = N(y_T | 0, I), \tag{5}$$

$$p_\theta(y_{t-1} | y_t, x) = N(y_{t-1} | \mu_\theta(x, y_t, \gamma_t), \sigma_t^2 I) \tag{6}$$

$p_\theta(y_{t-1} | y_t, x)$ is the posterior distribution to be learned, distribution variance σ_t^2 is fixed to be $1 - \alpha_t$, distribution mean $\mu_\theta(x, y_t, \gamma_t)$ is reparametrized as (Equation 7):

$$\mu_\theta(x, y_t, \alpha_t) = \frac{1}{\sqrt{\alpha_t}} \left[y_t - \frac{1 - \alpha_t}{\sqrt{1 - \gamma_t}} f_\theta(x, y_t, \gamma_t) \right] \tag{7}$$

where $f_{\theta}(x, y_t, \gamma_t)$ is the denoising model which takes the source LR perfusion image x and a noisy image y_t to predict the noise ϵ .

After the parametrization, each denoising step in the reverse process will be (Equation 8):

$$y_{t-1} \leftarrow \frac{1}{\sqrt{\alpha_t}} \left[y_t - \frac{1 - \alpha_t}{\sqrt{1 - \gamma_t}} f_{\theta}(x, y_t, \gamma_t) \right] + \sqrt{1 - \alpha_t} \epsilon_t \quad (8)$$

where $f_{\theta}(x, y_t, \gamma_t)$ is the denoising model, ϵ_t is the predicted noise at step t with $\epsilon_t \sim N(0, I)$.

2.2.3 Model implementation

We adapted the SR3-DDPM model to super-resolve a 2D low-resolution MR perfusion image into a HR image. The denoiser is achieved using a U-Net model and the optimization that employs KL-divergence to maximize the likelihood of the generated HR images \hat{y}_0 and the ground truth HR image y_0 . L1-loss between the noise predicted by the network and the amount of noise added was used, and the objective function for training f_{θ} was defined as (Equation 9):

$$\operatorname{argmin}_{\theta} L = \|f_{\theta}(x, \sqrt{\gamma}y_0 + \sqrt{1 - \gamma}\epsilon, \gamma) - \epsilon\|_1 \quad (9)$$

where f_{θ} represents the denoising U-Net model, x is the LR image, y is the corresponding HR image, ϵ is the added noise with $\epsilon \sim N(0, I)$, γ is a scalar parameter related to the variance scheduler with $\gamma_t = \prod_{i=1}^t \alpha_i$.

The model starts with pure Gaussian noise and a LR perfusion image, using the corresponding HR perfusion image as the ground truth. The model will iteratively refine the noisy output through a U-Net model trained to denoise at various noise levels and generate images with the desired HR perfusion data distribution. By using the LR perfusion MR image to condition the generation process, the SR image is specifically determined to maintain anatomical consistency similar to the original LR perfusion images.

In the U-Net architecture (Figure 2B), the input comprised two channels representing the LR image and the noisy image, and one output channel, representing the generated less noisy HR images. The LR and HR pairs in our synthesis pipeline maintained the same matrix size, and the conditioning LR image was used at the shallowest level of the U-Net by channel-wise concatenating with the noisy image. Both the LR image and noisy image at time step t were encoded through a convolutional layer followed by two linear layers for further encoding. The U-Net structure was composed of convolution, group normalization, Swish activation, residual connections and pooling layers. The U-Net structure consisted of five levels, with the number of channels in each level being [64,128,256,512,512]. Each level contained two ResNet blocks (23) with a dropout rate of 0.2. At the bottleneck, an additional self-attention was applied after the convolution layers. The self-attention module employs convolutional layers to compute the query, key, and value

representations for spatial attention. It is followed by another convolutional layer to refine the output and is interleaved with the original ResNet block at the bottleneck for enhanced feature representation. Detailed model architecture was depicted in Supplementary Figure S1.

PerfGen was implemented using Python and PyTorch on two 48GB NVIDIA A6000 GPUs. PerfGen had 92M trainable parameters. The model was trained for 50,000 iterations with AdamW optimizer with a learning rate of $3e-5$ and a batch size of 128. During inference, we used DDPM sampling with full inference steps ($T = 2,000$).

2.3 Model evaluation

For synthetic data, to compare PerfGen with GAN-based super-resolution model trained on cardiac MRI (15), normalized Root-Mean-Square-Error (nRMSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) were calculated, using original HR images as reference. The metrics were evaluated within the 96×96 field of view, focusing specifically on the heart region. The nRMSE (24) was calculated as (Equation 10):

$$nRMSE = \frac{1}{\max(I) - \min(I)} \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} (\tilde{I}_i - I_i)^2} \quad (10)$$

where \tilde{I} is SR images super-resolved by PerfGen or GAN, I is the reference HR image, N is the total number of pixels, \tilde{I}_i and I_i are the pixel intensities at position i in the SR and HR images, respectively. PSNR (25) was calculated as (Equation 11):

$$PSNR = 10 \times \log_{10} \left(\frac{255^2}{\frac{1}{N} \sum_{i=0}^{N-1} (\tilde{I}_i - I_i)^2} \right) \quad (11)$$

SSIM (26) was calculated as (Equation 12):

$$SSIM = \frac{(2\mu_{\tilde{I}} \mu_I + c_1)(2\sigma_{\tilde{I}I}^{\sim} + c_2)}{(\mu_{\tilde{I}}^2 + \mu_I^2 + c_1)(\sigma_{\tilde{I}}^2 + \sigma_I^2 + c_2)} \quad (12)$$

where $\mu_{\tilde{I}}$ and μ_I are the average and variance of \tilde{I} and I , $\sigma_{\tilde{I}I}^{\sim}$ is the covariance of \tilde{I} and I , and c_1 and c_2 are constants to prevent division by a near-zero denominator.

Differences between GAN super-resolved images and PerfGen super-resolved images were statistically tested using a paired t -test ($P < 0.05$). For prospective data, images super-resolved by two methods were qualitatively compared with LGE images at matched slice position to identify perfusion defects. One cardiologist and one radiologist scored prospectively acquired datasets on a 1–5 scale (1 being the worst and 5 being the best), assessing perfusion image sharpness and overall quality relative to clinical perfusion image standards. Differences between methods were assessed with the Wilcoxon signed-rank test.

3 Results

3.1 Model validation with synthetic data

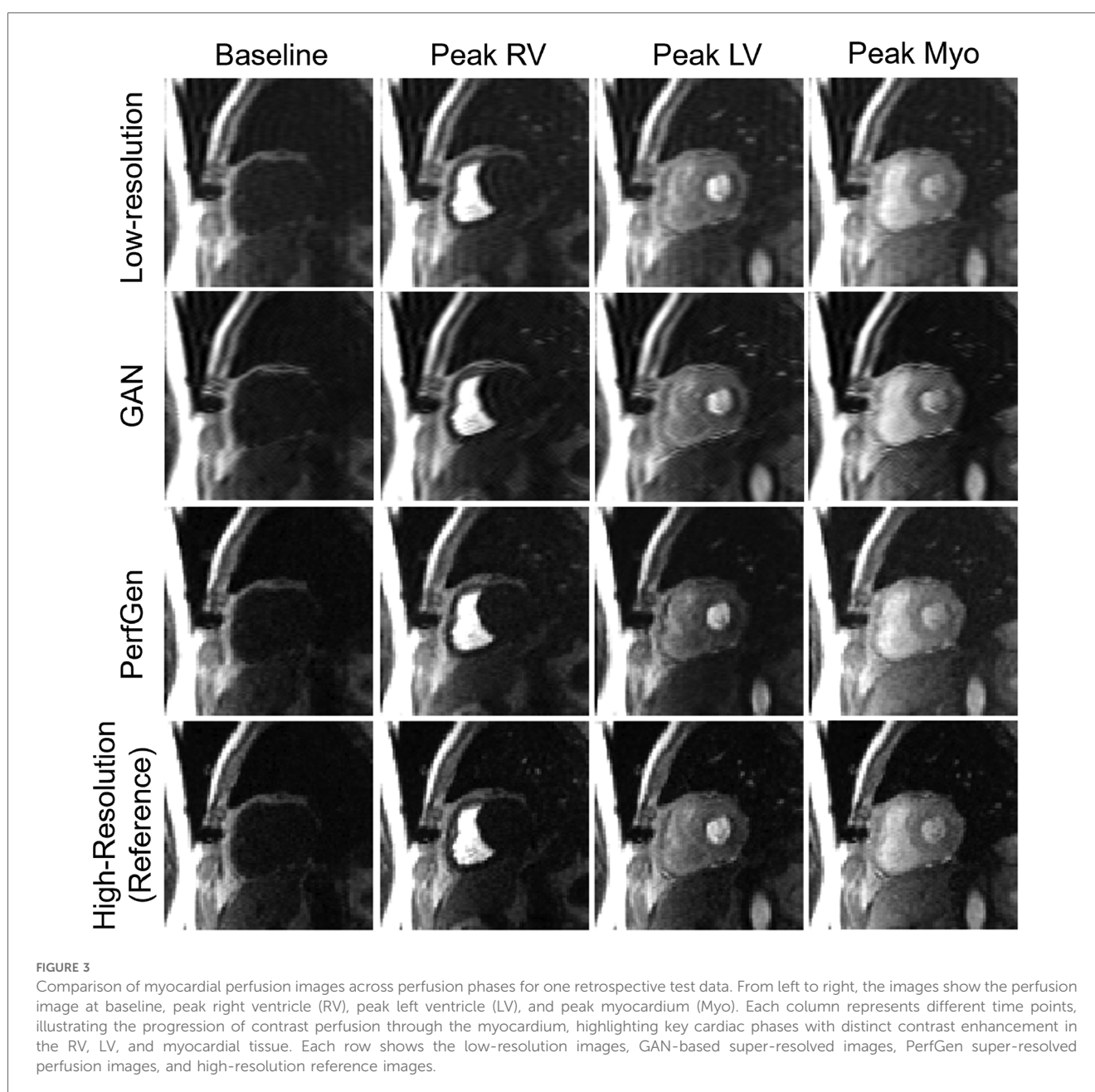
3.1.1 Qualitative comparison

Figure 3 compares the myocardial perfusion images across different phases of contrast enhancement during first-pass perfusion using synthetic test data. The PerfGen super-resolved images are compared to LR, GAN super-resolved and HR reference images at baseline, peak right ventricle (RV), peak left ventricle (LV), and peak myocardium (Myo). The results show an improvement in the image resolution and contrast for the PerfGen super-resolved images, allowing for enhanced visualization of contrast perfusion through the myocardium. The

enhanced detail provided by PerfGen aligned better with the HR reference images than LR and GAN methods.

Figure 4A further demonstrates the evaluation of the PerfGen by comparing myocardial perfusion images at the basal, midventricular, and apical slice locations. The PerfGen super-resolved images show enhanced resolution and structural details across all slice locations compared to the HR reference. Figure 4B shows the signal-t plot illustrating the changes in the LV myocardial region, LV blood pool, and RV blood pool. The spatially super-resolved images demonstrate better alignment with the reference high-resolution spatial images compared to the acquired LR spatial images and GAN-based super-resolved images.

Figure 5 presents example cross-sectional time-intensity profiles from the subject shown in Figure 4, comparing LR images, GAN



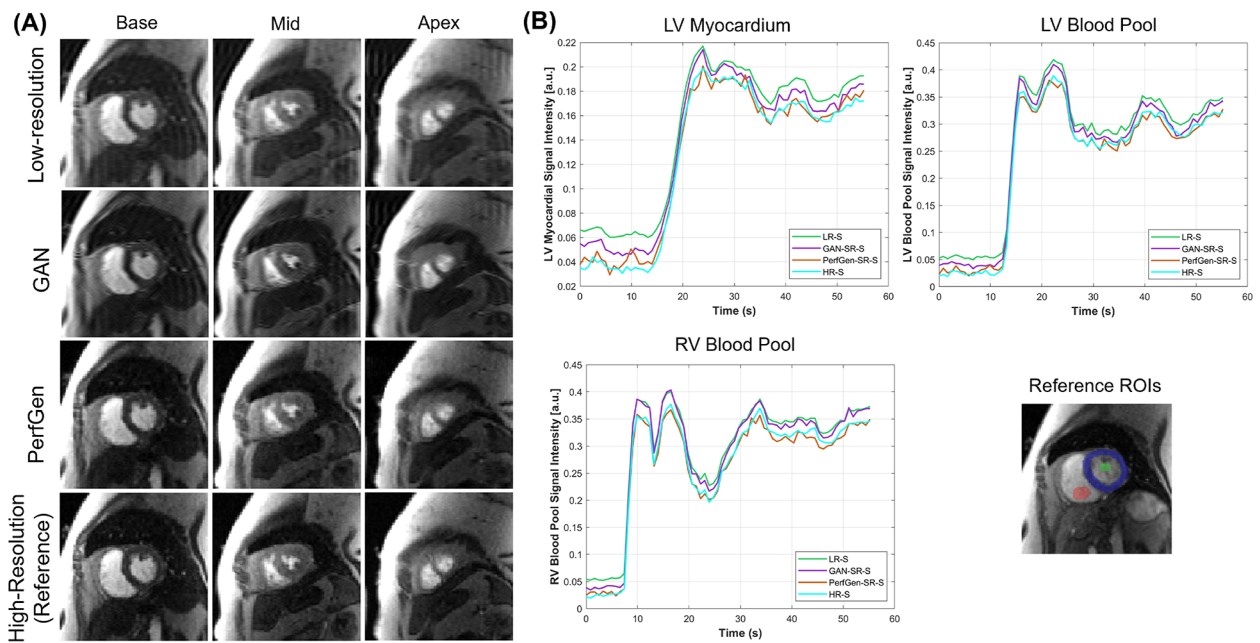


FIGURE 4
 (A) Comparison of GAN-based super-resolved images, PerfGen super-resolved images and high-resolution (HR) perfusion images across different slice locations for one synthetic test data. From left to right, the myocardial perfusion images are shown at the base, midventricular, and apical slice locations. The top row showed low-resolution (LR) images, the following rows showed GAN-based and PerfGen super-resolved images, and the bottom row showed the corresponding HR reference images. This comparison highlights the effectiveness of PerfGen in enhancing image resolution and better alignment with reference images than GAN-based approach across various slice locations of the heart. (B) The signal-t plots illustrate the signal intensity of the basal slice in (A) in terms of the left-ventricular (LV) myocardial region, LV blood pool and right-ventricular (RV) blood pool changes over time. The spatially super-resolved images by PerfGen (PerfGen-SR-S) aligns better with the reference HR spatial images (HR-S) than the acquired LR spatial images (LR-S) and the GAN-based super-resolved images (GAN-SR-S).

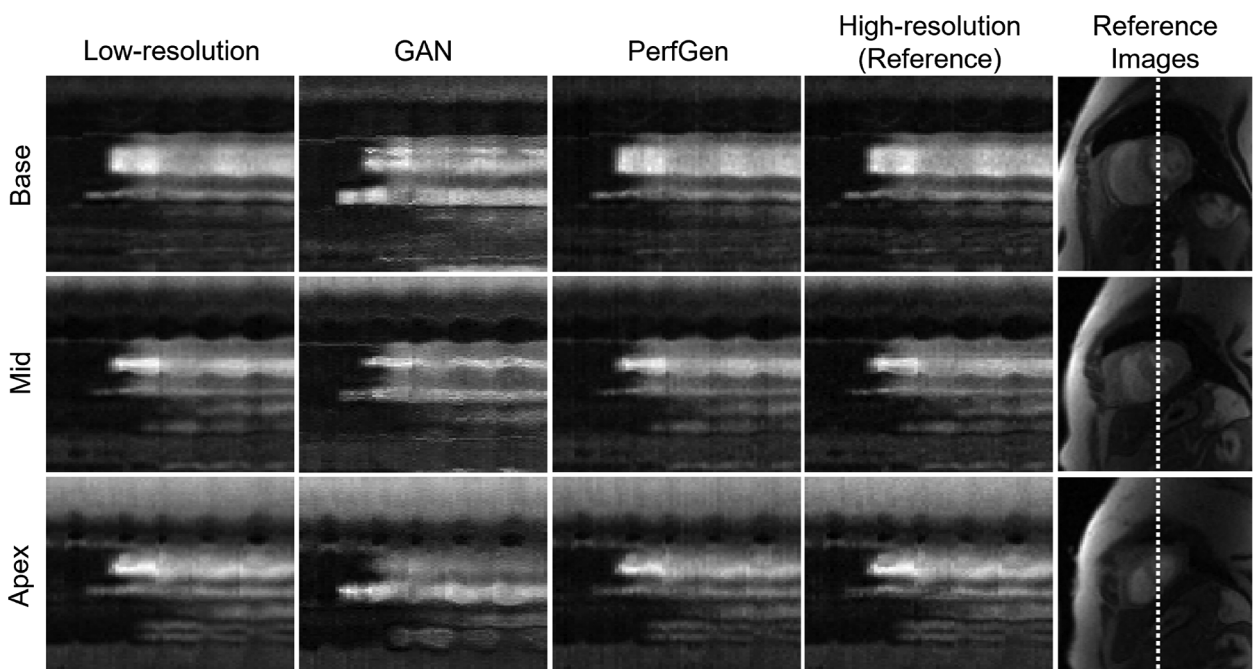


FIGURE 5
 Comparison of cross-section profile along time plots corresponding to low-resolution (LR), GAN, PerfGen and reference high-resolution (HR) for the patient in Figure 4 were shown. PerfGen presents better temporal fidelity than zero-padded LR images and GAN-based super-resolved images compared to the reference HR x-t plots.

super-resolved images, PerfGen super-resolved images, and reference HR images. The PerfGen super-resolved method shows better temporal fidelity than the LR images and GAN super-resolved images compared with the reference HR plots. **Figure 6** shows the stress

perfusion images of a patient with inducible myocardial ischemia. PerfGen demonstrates better alignment with HR reference images than LR images and GAN-based super-resolved images in terms of the overall image quality and the accurate perfusion defect detection.

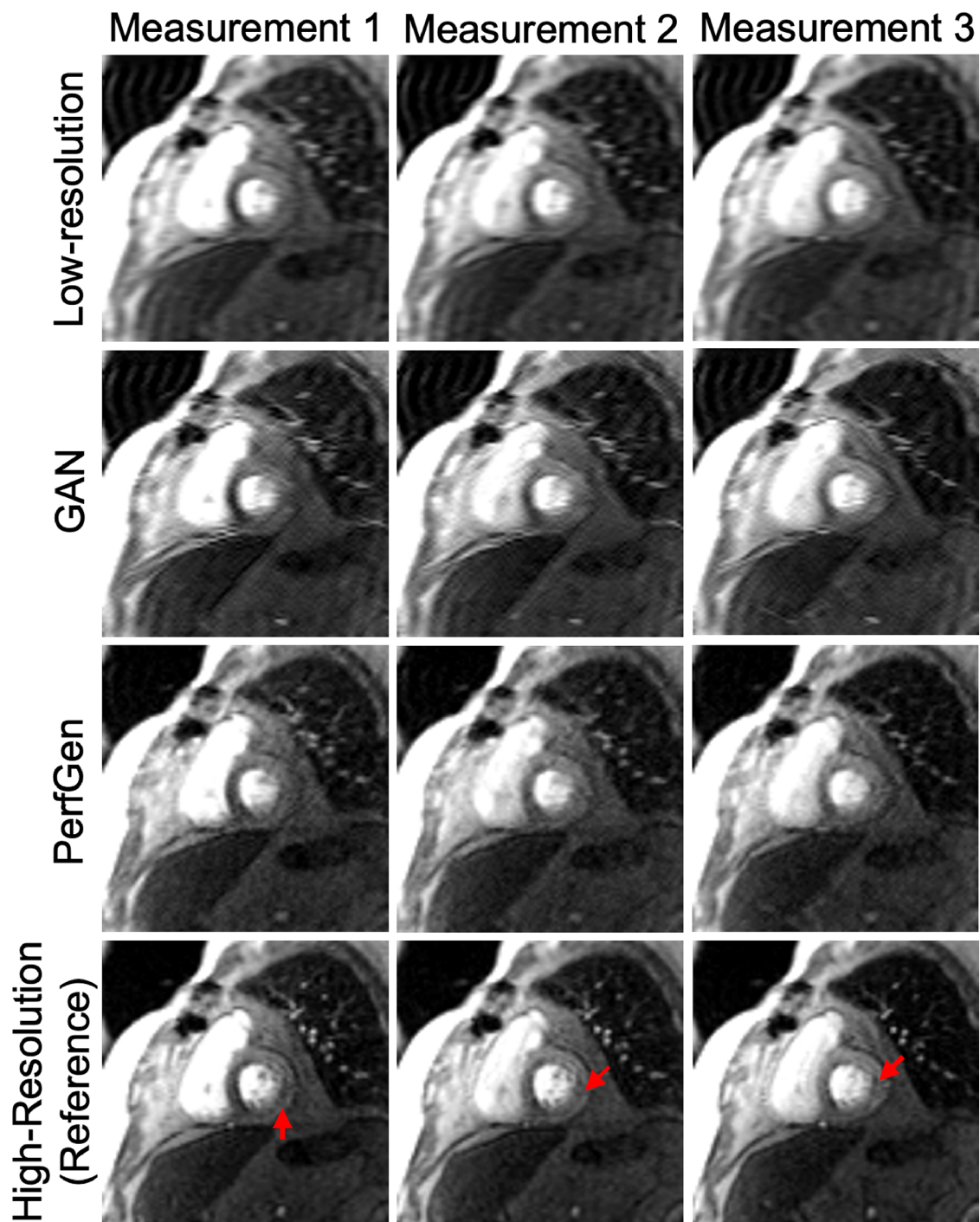


FIGURE 6

Example images from a retrospective patient undergoing regadenoson stress perfusion imaging with inducible myocardial ischemia with a moderate-sized defect in the basal septum and a small defect in the inferolateral wall. Low-resolution (LR) images, GAN-based super-resolution, PerfGen, and high-resolution (HR) images were compared as reference for image quality and their ability to illustrate perfusion defects. LR images, used as the baseline, appear blurred but indicates perfusion defects in the septum. The GAN-based super-resolution method improves image sharpness but shows slight mismatches with the HR reference in terms of image details in myocardial regions, with potentially reduced fidelity in depicting the small inferolateral wall perfusion defects (indicated by red arrows). PerfGen demonstrates greater visual similarity to the HR reference, more accurately highlighting perfusion defects in similar regions with higher fidelity.

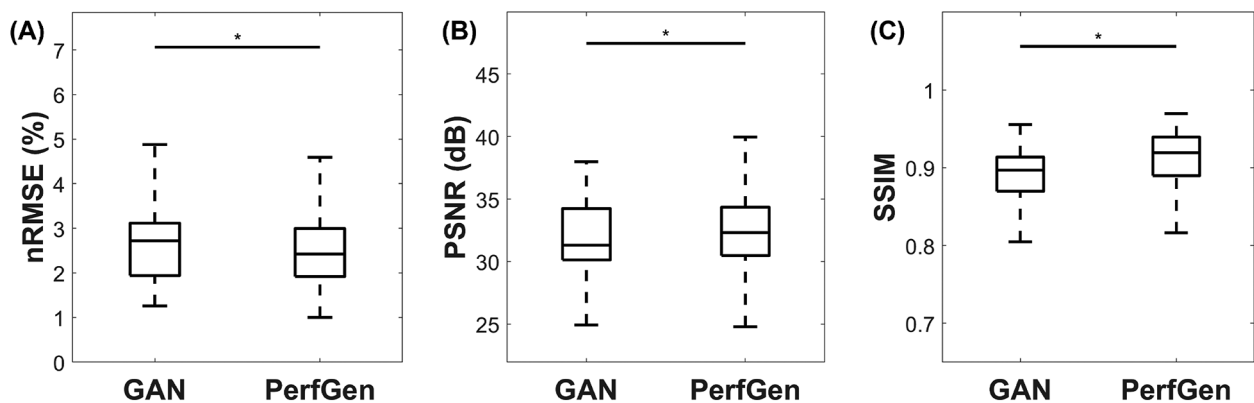


FIGURE 7

Quantitative comparison of GAN and PerfGen super-resolved myocardial perfusion images for nine tested datasets. Boxplots show (A) normalized root mean square error (nRMSE), (B) peak signal-to-noise ratio (PSNR), and (C) structural similarity index (SSIM) for GAN and PerfGen images. PerfGen demonstrates a significant improvement in all metrics, with lower nRMSE, higher PSNR, and higher SSIM compared to super-resolved images by GAN, with a reduction in nRMSE by approximately 5.1%, an increase in PSNR by 1.1%, and an improvement in SSIM by 2.2%. Statistical significance is indicated by the asterisks.

3.1.2 Quantitative comparison

Figure 7 presents a quantitative comparison between GAN and PerfGen super-resolved myocardial perfusion images of nine testing datasets. The PerfGen method significantly outperformed the GAN-based approach across all evaluated metrics. Specifically, PerfGen achieved a 5.1% reduction in nRMSE (mean nRMSE: $2.68 \pm 0.85\%$ for GAN vs. $2.55 \pm 0.84\%$ for PerfGen, respectively), a 1.1% increase in PSNR (mean PSNR: 31.89 ± 2.82 dB for GAN vs. 32.24 ± 2.77 dB for PerfGen, respectively), and a 2.2% improvement in SSIM (mean SSIM: 0.87 ± 0.15 vs. 0.89 ± 0.16 for PerfGen, respectively). These improvements are statistically significant, as indicated by the asterisks in Figure 7, demonstrating the superior performance of PerfGen in enhancing image quality.

3.2 Model validation with prospectively acquired data

Figure 8 compares the super-resolved perfusion images by PerfGen and GAN to both LR perfusion images and LGE images. The super-resolved images show perfusion defects that closely match the defects observed in the LGE images at corresponding slice locations, providing proof-of-concept that PerfGen can potentially identify the super-resolve perfusion defects from LR images. PerfGen demonstrates better alignment with LGE compared to GAN-based super-resolved perfusion images. Furthermore, the combination of a LR acquisition with 35% phase lines and GRAPPA-2 allowed the acquisition of five slices, with a 2.86-fold improvement in temporal resolution compared to clinical routine settings, demonstrating the improved image quality with improved slice coverage.

Figure 9A provides evaluation of PerfGen in enhancing myocardial perfusion images across five different perfusion slices.

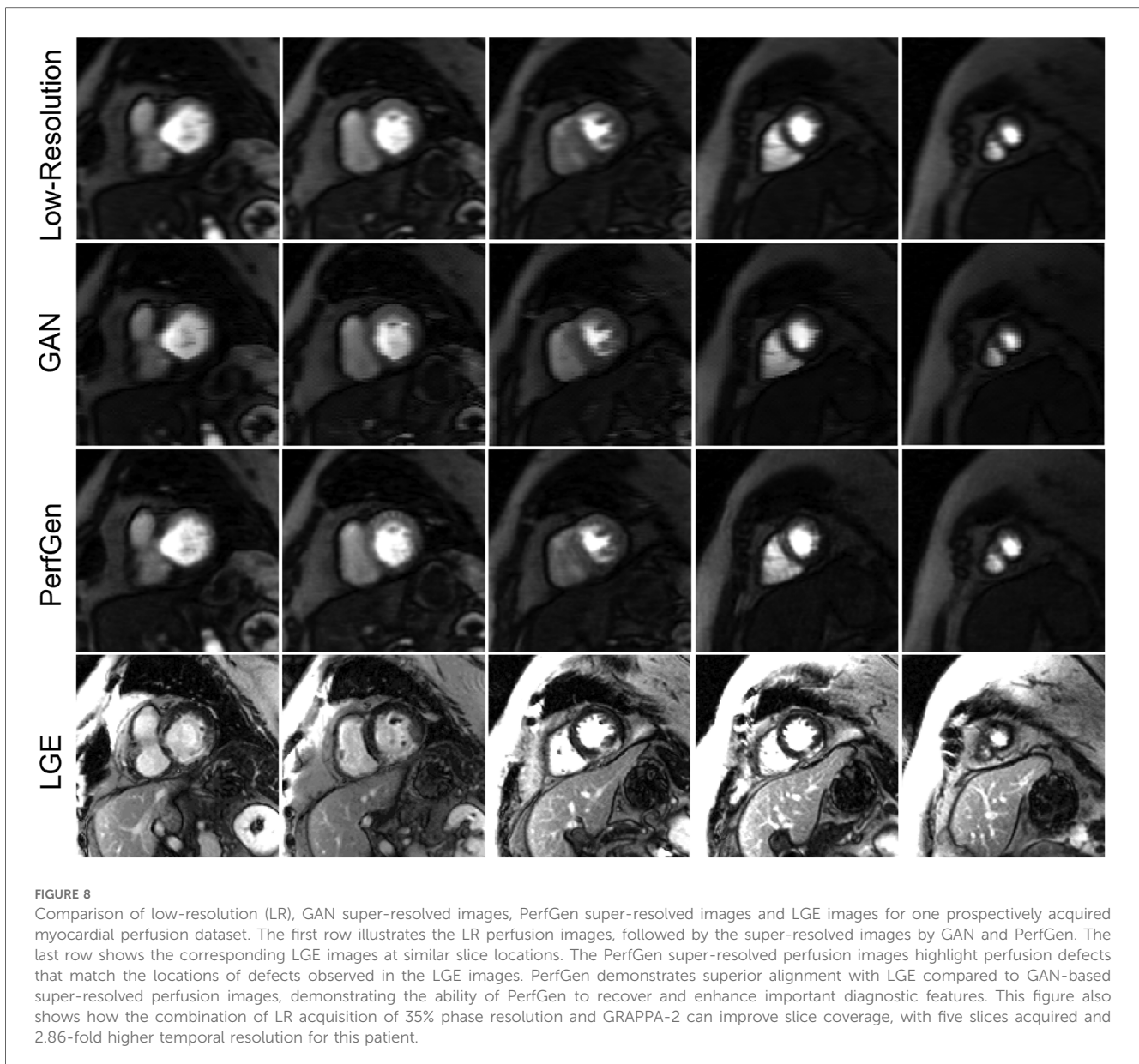
Super-resolved images are compared with LR images using zero-padding for five prospectively acquired slice, showing improved contrast and image detail across various perfusion phases by PerfGen. The super-resolution method demonstrates improved visualization of image details that are not as obvious in the LR images or GAN super-resolved images acquired using GRAPPA-3 and 35% phase resolution. Figure 9B shows the signal-time plots of the intensity changes over time in the LV myocardial region, LV blood pool and RV blood pool. The plots demonstrate that achieving a 4.3-fold increase in temporal resolution and five slices covered in myocardial perfusion MRI, compared to the clinically used GRAPPA-2 and three slices, improves the ability to capture sharp transitions in contrast during myocardial perfusion as compared against a synthetic low temporal resolution curve using rate-2 acceleration. Supplementary Video 1 showed the movies of six slices from one prospective patient acquired with GRAPPA-3 and 35% phase encoding lines, comparing LR, GAN-based super-resolved and PerfGen super-resolved images.

For the nine prospectively acquired subjects, all slices were evaluated by two experts in Figure 10. For sharpness, the scores were 4.36 ± 0.38 for GAN and 4.89 ± 0.22 for PerfGen ($P < 0.05$); for overall image quality, the scores were 4.14 ± 0.28 for GAN and 4.89 ± 0.22 for PerfGen ($P < 0.05$).

PerfGen not only improves spatial resolution but also enhances critical image features, such as perfusion defects, that align well with LGE reference images, providing proof of concept for demonstrating the potential of super-resolution techniques in diagnostic accuracy in myocardial perfusion imaging.

4 Discussion

With the growing interest in myocardial perfusion MRI in identifying myocardial ischemia, there is an increased need for high



spatiotemporal resolution and expanded slice coverage to accurately monitor dynamic changes in blood flow and myocardial perfusion. This poses a challenge in achieving acquisition speed to capture rapid change and fine details without sacrificing the quality and accuracy necessary for effective diagnosis. Low-resolution acquisition is an alternative approach that inherently allows for acceleration and higher SNR. However, the reduction in spatial and temporal details may degrade the image quality, influence the diagnosis accuracy and potentially impact the subsequent quantitative analysis.

In this study, we demonstrated that existing clinical perfusion MRI images can be effectively used to train a conditional diffusion generative model for super-resolution. We proposed a super-resolution pipeline that utilizes low-resolution myocardial perfusion MRI as the guidance after initial reconstruction by GRAPPA (27), which is also potentially applicable to compressed

sensing (28) or unrolled network (29) outputs, offering a complementary approach to the existing workflows. When combined with GRAPPA (factor 2–3) in prospective acquisitions, this method offers a nominal 5.7–8.5 folds acceleration, allowing for better slice coverage and improved temporal resolution. This approach not only accelerates acquisition but also mitigates the loss of contrast and details typically associated with low-resolution imaging. We validated our model on an infarction patient using reference LGE images, with the perfusion defects showing consistent with the scar regions in LGE.

While diffusion generative models have demonstrated the training stability and high-quality image generation across various vision tasks (12, 20, 30), their application in generating myocardial perfusion MRI and integration with cardiac MRI have yet to be explored. The study demonstrates that the diffusion generative model produces myocardial perfusion images

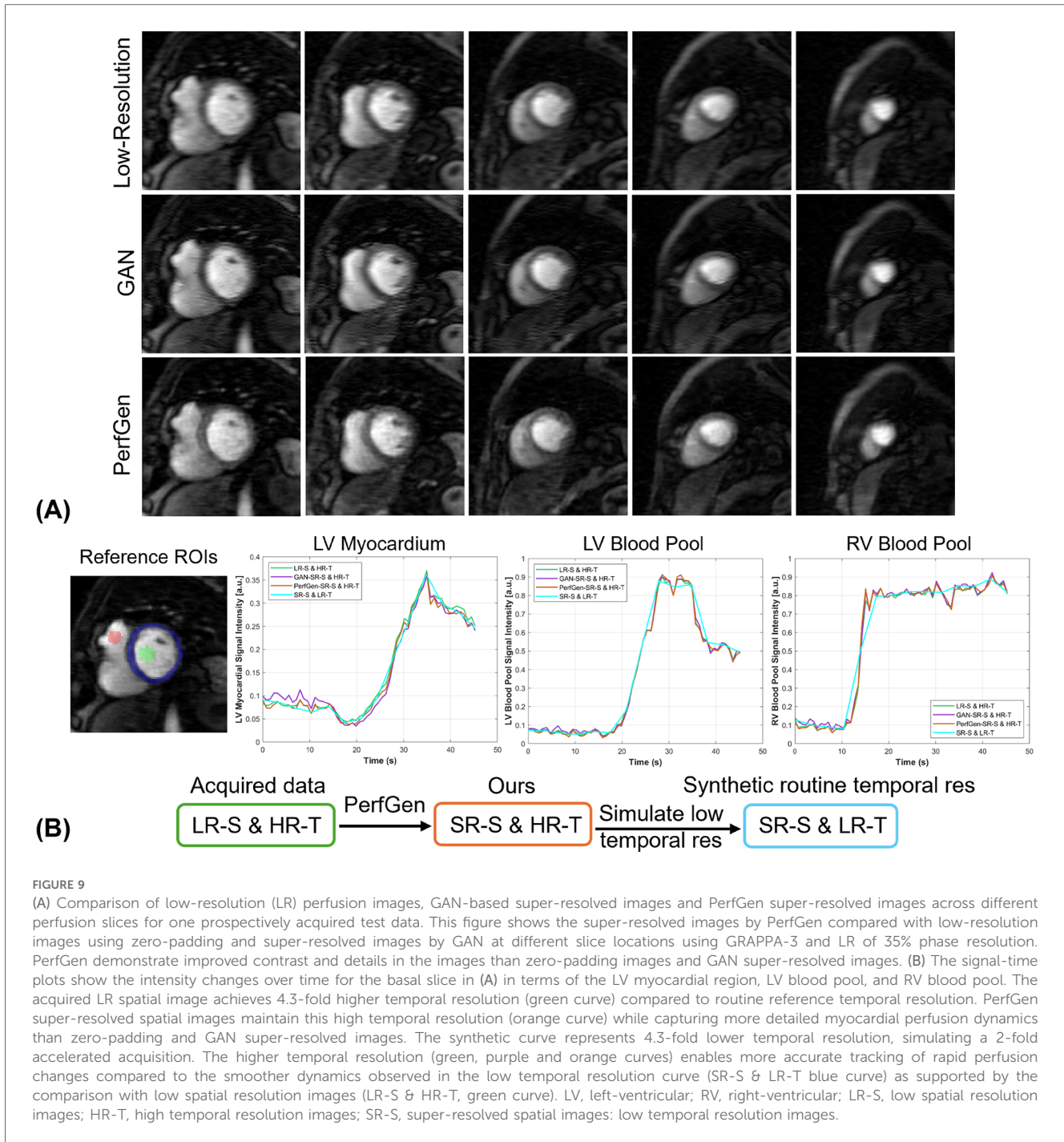


FIGURE 9
 (A) Comparison of low-resolution (LR) perfusion images, GAN-based super-resolved images and PerfGen super-resolved images across different perfusion slices for one prospectively acquired test data. This figure shows the super-resolved images by PerfGen compared with low-resolution images using zero-padding and super-resolved images by GAN at different slice locations using GRAPPA-3 and LR of 35% phase resolution. PerfGen demonstrate improved contrast and details in the images than zero-padding images and GAN super-resolved images. (B) The signal-time plots show the intensity changes over time for the basal slice in (A) in terms of the LV myocardial region, LV blood pool, and RV blood pool. The acquired LR spatial image achieves 4.3-fold higher temporal resolution (green curve) compared to routine reference temporal resolution. PerfGen super-resolved spatial images maintain this high temporal resolution (orange curve) while capturing more detailed myocardial perfusion dynamics than zero-padding and GAN super-resolved images. The synthetic curve represents 4.3-fold lower temporal resolution, simulating a 2-fold accelerated acquisition. The higher temporal resolution (green, purple and orange curves) enables more accurate tracking of rapid perfusion changes compared to the smoother dynamics observed in the low temporal resolution curve (SR-S & LR-T blue curve) as supported by the comparison with low spatial resolution images (LR-S & HR-T, green curve). LV, left-ventricular; RV, right-ventricular; LR-S, low spatial resolution images; HR-T, high temporal resolution images; SR-S, super-resolved spatial images; low temporal resolution images.

comparable to routine GRAPPA-2 perfusion images and outperform the compared GAN-based method, highlighting its potential to enhance temporal resolution and slice coverage for clinical use. This super-resolution approach provides several key advantages: (1) it effectively generates fine image details, outperforming one existing GAN-based super-resolution method, (2) the combination of low-resolution acquisition and GRAPPA reduces the risk of residual artifacts from highly accelerated undersampling (8.5-fold), (3) the GRAPPA-reconstruction by the vendor allows for real-time visualization for perfusion imaging, enabling real-time monitoring of contrast dynamics, and (4) the

super-resolution process operates independently of the full temporal series, allowing for efficient image-by-image processing and minimize the potential loss of temporal fidelity.

Our results showed higher temporal resolution than the clinically used GRAPPA-2, where the higher temporal resolution enables better capture of fast perfusion dynamics. This enhancement can reduce temporal blurring, provide more precise time-intensity curves for quantitative analysis, and allow for more accurate assessment of myocardial ischemia. Additionally, higher temporal resolution can mitigate motion artifacts caused by cardiac and respiratory motion, resulting in images

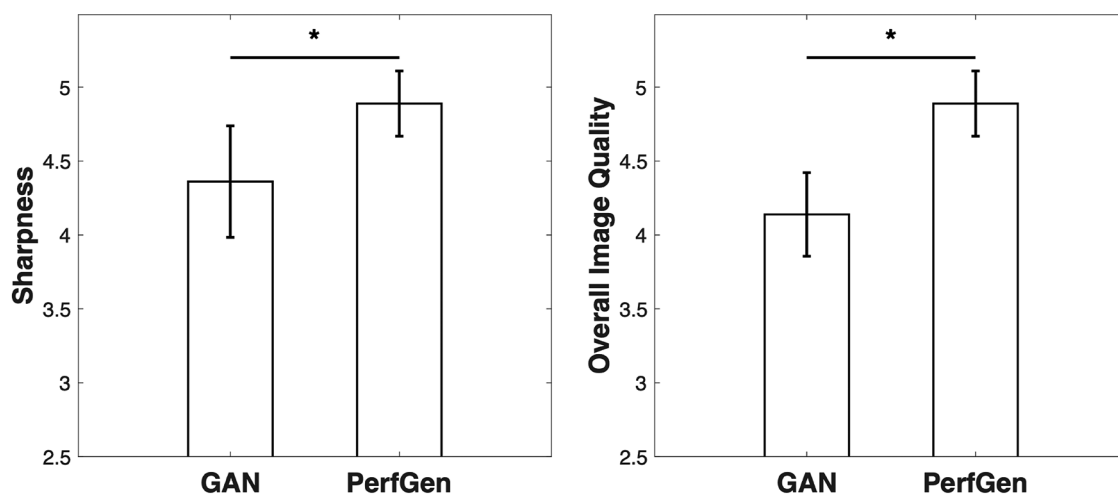


FIGURE 10

Quantitative comparisons of GAN and PerfGen methods for super-resolution of prospective perfusion images, evaluated on 9 prospective subjects and scored by two experts. PerfGen demonstrated significantly better performance than GAN in sharpness and overall image quality. Sharpness scores were 4.36 ± 0.38 for GAN and 4.89 ± 0.22 for PerfGen ($P < 0.05$), while overall image quality scores were 4.14 ± 0.28 for GAN and 4.89 ± 0.22 for PerfGen ($P < 0.05$).

with better image quality and potentially more reliable diagnostic outcomes.

One of the potential limitations of this approach is that the super-resolved output cannot theoretically exceed the spatial resolution of the reference images used for training. A broader dataset would be beneficial for a more thorough model training. Although GRAPPA-reconstruction allows for real-time visualization, PerfGen requires additional time to enhance image quality and does not currently support real-time processing. While the diffusion generative model aims to learn the distribution of high-resolution perfusion images and generate high-quality outputs, and proof-of-concept studies highlight the promise of the PerfGen model, accurately detecting small perfusion defects remains challenging due to partial volume effects and dark rim artifacts. Addressing these potential limitations will require further training on larger datasets, optimized network architectures, and robust training strategies. Additionally, further studies are necessary to validate these findings in diverse clinical scenarios, particularly for assessing ischemic perfusion defects in stress perfusion MRI. It is important to note that the comparison between our method and GAN-based method involves a single diffusion generative model and one published GAN-based approach trained for cardiac MRI super-resolution which has been applied to various cardiac MRI applications (15, 31). This comparison is not intended as a comprehensive theoretical comparison between GANs vs. diffusion models. The relative performance of these models may also depend on factors such as the datasets used, the specific applications, and other implementation details. Currently, our findings are preliminary and serve as a proof-of-concept; additional clinical validation is necessary to assess the reliability of this approach. Future efforts could also explore integrating the super-resolution method with physics-guided self-supervised learning reconstruction networks for high-resolution perfusion

MRI (32), and extend this approach with quantitative analysis (33–35). Further studies will be needed to establish the clinical utility and validate the diagnostic values of this method.

Overall, this work shows the capability of the conditional diffusion generative model in high-resolution myocardial perfusion MRI generation and demonstrates its feasibility to accelerate myocardial perfusion MRI acquisition, increase temporal resolution and slice coverage, and improve image quality without introducing significant artifacts or blurring.

Data availability statement

The raw data presented in this article are not readily available due to patient privacy. Requests to access the dataset should be directed to csyfc@missouri.edu.

Ethics statement

The studies involving humans were approved by University of Missouri-Columbia Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

CS: Conceptualization, Data curation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. NG: Validation, Writing – review & editing. YW: Methodology, Writing – review & editing. DT: Validation, Writing – review &

editing, SK: Writing – review & editing, Validation. TA: Resources, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Sun C, Robinson A, Wang Y, Bilchick KC, Kramer CM, Weller D, et al. A slice-low-rank plus sparse (slice-L+S) reconstruction method for k-t undersampled multiband first-pass myocardial perfusion MRI. *Magn Reson Med.* (2022) 88(3):1140–55. doi: 10.1002/mrm.292281
- Otazo R, Kim D, Axel L, Sodickson DK. Combination of compressed sensing and parallel imaging for highly accelerated first-pass cardiac perfusion MRI. *Magn Reson Med.* (2010) 64(3):767–76. doi: 10.1002/mrm.22463
- Maredia N, Radjenovic A, Kozerke S, Larghat A, Greenwood JP, Plein S. Effect of improving spatial or temporal resolution on image quality and quantitative perfusion assessment with k-t SENSE acceleration in first-pass CMR myocardial perfusion imaging. *Magn Reson Med.* (2010) 64(6):1616–24. doi: 10.1002/mrm.22493
- Kramer CM, Barkhausen J, Bucciarelli-Ducci C, Flamm SD, Kim RJ, Nagel E. Standardized cardiovascular magnetic resonance imaging (CMR) protocols: 2020 update. *J Cardiovasc Magn R.* (2020) 22(1):17. doi: 10.1186/s12968-020-00607-1
- Hamirani YS, Kramer CM. Cardiac MRI assessment of myocardial perfusion. *Future Cardiol.* (2014) 10(3):349–58. doi: 10.2217/fca.14.18
- Adluru G, Awate SP, Tasdizen T, Whitaker RT, Dibella EV. Temporally constrained reconstruction of dynamic cardiac perfusion MRI. *Magn Reson Med.* (2007) 57(6):1027–36. doi: 10.1002/mrm.21248
- Naresh NK, Haji-Valizadeh H, Aouad PJ, Barrett MJ, Chow K, Ragin AB, et al. Accelerated, first-pass cardiac perfusion pulse sequence with radial k-space sampling, compressed sensing, and k-space weighted image contrast reconstruction tailored for visual analysis and quantification of myocardial blood flow. *Magn Reson Med.* (2019) 81(4):2632–43. doi: 10.1002/mrm.27573
- Sakuma H, Ishida M. Advances in myocardial perfusion MR imaging: physiological implications, the importance of quantitative analysis, and impact on patient care in coronary artery disease. *Magn Reson Med Sci.* (2022) 21(1):195–211. doi: 10.2463/mrms.rev.2021-0033
- Salerno M, Sica CT, Kramer CM, Meyer CH. Optimization of spiral-based pulse sequences for first-pass myocardial perfusion imaging. *Magn Reson Med.* (2011) 65(6):1602–10. doi: 10.1002/mrm.22746
- Pontre B, Cowan BR, DiBella E, Kulaseharan S, Likhite D, Noorman N, et al. An open benchmark challenge for motion correction of myocardial perfusion MRI. *IEEE J Biomed Health Inform.* (2017) 21(5):1315–26. doi: 10.1109/JBHI.2016.2597145
- Wissmann L, Gotschy A, Santelli C, Tezcan KC, Hamada S, Manka R, et al. Analysis of spatiotemporal fidelity in quantitative 3D first-pass perfusion cardiovascular magnetic resonance. *J Cardiovasc Magn Reson.* (2017) 19(1):11. doi: 10.1186/s12968-017-0324-z
- Saharia C, Ho J, Chan W, Salimans T, Fleet DJ, Norouzi M. Image super-resolution via iterative refinement. *IEEE Trans Pattern Anal Mach Intell.* (2023) 45(4):4713–26. doi: 10.1109/TPAMI.2022.3204461
- Wang XT, Yu K, Wu SX, Gu JJ, Liu YH, Dong C, et al. ESRGAN: enhanced super-resolution generative adversarial networks. *Computer Vision - Ecvv 2018 Workshops.* (2019) 11133(Pt V):63–79. doi: 10.1007/978-3-030-11021-5_5
- Di Bella EV, Parker DL, Sinusas AJ. On the dark rim artifact in dynamic contrast-enhanced MRI myocardial perfusion studies. *Magn Reson Med.* (2005) 54(5):1295–9. doi: 10.1002/mrm.20666
- Yoon S, Nakamori S, Amyar A, Assana S, Cirillo J, Morales MA, et al. Accelerated cardiac MRI cine with use of resolution enhancement generative adversarial inline neural network. *Radiology.* (2023) 307(5):e222878. doi: 10.1148/radiol.222878
- Huang C, Petibon Y, Ouyang J, Reese TG, Ahlman MA, Bluemke DA, et al. Accelerated acquisition of tagged MRI for cardiac motion correction in simultaneous PET-MR: phantom and patient studies. *Med Phys.* (2015) 42(2):1087–97. doi: 10.1118/1.4906247
- Küstner T, Munoz C, Psenicny A, Bustin A, Fuin N, Qi H, et al. Deep-learning based super-resolution for 3D isotropic coronary MR angiography in less than a minute. *Magn Reson Med.* (2021) 86(5):2837–52. doi: 10.1002/mrm.28911
- Xia Y, Ravikumar N, Greenwood JP, Neubauer S, Petersen SE, Frangi AF. Super-resolution of cardiac MR cine imaging using conditional GANs and unsupervised transfer learning. *Med Image Anal.* (2021) 71:102037. doi: 10.1016/j.media.2021.102037
- Gui J, Sun Z, Wen Y, Tao D, Ye J. A review on generative adversarial networks: algorithms, theory, and applications. *IEEE Trans Knowl Data Eng.* (2023) 35(4):3313–32. doi: 10.1109/TKDE.2021.3130191
- Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models. *Proceedings of the 34th International Conference on Neural Information Processing Systems; Vancouver, BC, Canada: Curran Associates Inc.* (2020). p. 574.
- Chung H, Ye JC. Score-based diffusion models for accelerated MRI. *Med Image Anal.* (2022) 80:102479. doi: 10.1016/j.media.2022.102479
- Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *Med Image Comput Assist Interv.* (2015) 9351:234–41. doi: 10.1007/978-3-319-24574-4_28

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcvm.2025.1499593/full#supplementary-material>

SUPPLEMENTARY TABLE 1

Details of contrast agent and stress agent protocol

SUPPLEMENTARY FIGURE 1

Detailed depiction of the denoising U-Net architecture. The network takes in a two-channel input, where one channel represents the low-resolution condition image and the other contains the noisy image at time step t . The output is the denoised image at the subsequent time step $t-1$. The U-Net architecture consists of convolutional layers, group normalization, Swish activation, residual connections, and pooling layers. A convolution layer and two linear layers were used as the image encoder for the inputs. At the bottleneck, an additional self-attention layer was applied.

SUPPLEMENTARY VIDEO 1

Comparison of zero-padded low-resolution, GAN-based super-resolved and PerfGen super-resolved myocardial perfusion images across six slices and perfusion measurements acquired with 35% phase encoding lines and GRAPPA-3. The first row shows the zero-padded low-resolution images, while the following rows present the corresponding GAN-based super-resolved images and PerfGen super-resolved images. The super-resolved images by PerfGen demonstrate improved spatial resolution and contrast, recovering key perfusion features that may be less apparent in the low-resolution images and GAN super-resolved images. This movie highlights the potential of PerfGen to enhance image quality and diagnostic accuracy for myocardial perfusion imaging.

23. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016).
24. Lustig M, Pauly JM. SPIRiT: iterative self-consistent parallel imaging reconstruction from arbitrary k-space. *Magn Reson Med.* (2010) 64(2):457–71. doi: 10.1002/mrm.22428
25. Hore A, Ziou D. Image quality metrics: PSNR vs. SSIM. *2010 20th International Conference on Pattern Recognition* (2010) p. 2366–9.
26. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Process.* (2004) 13(4):600–12. doi: 10.1109/TIP.2003.819861
27. Griswold MA, Jakob PM, Heidemann RM, Nittka M, Jellus V, Wang J, et al. Generalized autocalibrating partially parallel acquisitions (GRAPPA). *Magn Reson Med.* (2002) 47(6):1202–10. doi: 10.1002/mrm.10171
28. Lustig M, Donoho D, Pauly JM. Sparse MRI: the application of compressed sensing for rapid MR imaging. *Magn Reson Med.* (2007) 58(6):1182–95. doi: 10.1002/mrm.21391
29. Demirel OB, Yaman B, Shenoy C, Moeller S, Weingärtner S, Akçakaya M. Signal intensity informed multi-coil encoding operator for physics-guided deep learning reconstruction of highly accelerated myocardial perfusion CMR. *Magn Reson Med.* (2023) 89(1):308–21. doi: 10.1002/mrm.29453
30. Ho J, Saharia C, Chan W, Fleet DJ, Norouzi M, Salimans T. Cascaded diffusion models for high fidelity image generation. *J Mach Learn Res.* (2022) 23(47):1–33.
31. Demirel OB, Ghanbari F, Hoeger CW, Tsao CW, Carty A, Ngo LH, et al. Late gadolinium enhancement CMR with generative AI. *J Cardiovasc Magn Reson.* (2024) 27(1):101127. doi: 10.1016/j.jocmr.2024.101127
32. Sun C, Kumar S, Altes T. CMR 2-55 - Free-breathing high-resolution perfusion MRI with over 11-fold acceleration using self-supervised learning. *J Cardiovasc Magn R.* (2024) 26:100137. doi: 10.1016/j.jocmr.2024.100137
33. Sun C, Kumar S, Altes T. Kiosk 6R-FA-01 - Spatiotemporal FBactal analysis and radiomics of free-breathing stress myocardial perfusion. *J Cardiovasc Magn R.* (2024) 26:100725. doi: 10.1016/j.jocmr.2024.100725
34. Li R, Edalati M, Muccigrosso D, Lau JMC, Laforest R, Woodard PK, et al. A simplified method to correct saturation of arterial input function for cardiac magnetic resonance first-pass perfusion imaging: validation with simultaneously acquired PET. *J Cardiovasc Magn Reson.* (2023) 25(1):35. doi: 10.1186/s12968-023-00945-w
35. Xue H, Brown LAE, NIELLES-Vallespin S, Plein S, Kellman P. Automatic in-line quantitative myocardial perfusion mapping: processing algorithm and implementation. *Magn Reson Med.* (2020) 83(2):712–30. doi: 10.1002/mrm.27954