

Supplementary Document S1

1. data pre-processing: creating a robust training set for prostate MRI enhancement

A crucial step in this study involved the creation of a robust training dataset comprised of T2W images realistically mimicking various MRI artifacts. This approach, utilizing the TorchIO library for medical image augmentation, aimed to enhance the robustness and generalizability of the trained model by exposing it to a wide range of image degradations commonly encountered in clinical practice.¹

1.1. Motion artifacts: simulating the impact of patient movement

Motion artifacts leads to blurring or ghosting that can obscure anatomical details.² To make our deep learning model more robust to these artifacts, we simulated realistic motion during the MRI scan. This simulation involves randomly generating a series of small movements, representing the kind of positional shifts a patient might make.

Each simulated movement is mathematically represented as a 3D transformation. These transformations include both rotation (turning) and translation (shifting) components, and their parameters are randomly varied to create a diverse range of plausible motions. To ensure that these simulated movements don't unrealistically displace the prostate from its average position within the image, each movement is adjusted using a "de-meaning" process. This process ensures that the simulated motion primarily degrades image quality through blurring or ghosting, rather than causing a significant shift in the prostate's overall location.

The simulated movements are then applied to the original image data, which is re-sampled to maintain smooth image features despite the introduced motion. Finally, the motion-corrupted image is synthesized by processing the image data in k-space, the frequency domain representation of the MRI signal. This process incorporates the temporal dynamics of the simulated movements, resulting in a realistic depiction of a motion-affected T2-weighted prostate MRI.

1.2. Gaussian noise: simulating inherent acquisition noise

Inherent noise is an unavoidable part of MRI acquisition, creating random fluctuations in signal intensity that can obscure subtle details within prostate images. To

make our synthetically degraded prostate images more realistic, we incorporated random Gaussian noise, simulating this inherent noise.

Instead of using fixed noise levels, we varied the amount of noise added to each image. This mimics the range of noise levels encountered in real-world prostate MRI scans. For each image, the parameters controlling the noise distribution were randomly selected, ensuring our model is exposed to a variety of noise profiles during training.

This random noise is added to each individual voxel within the image. The amount of noise added is determined by scaling a random value with a randomly chosen standard deviation and shifting it by a randomly selected mean. This process creates realistic noise patterns, reflecting the noise characteristics inherent to real-world prostate MRI.

1.3. Gaussian blur: simulating loss of sharpness in prostate MRI

Blurring is often caused by factors like imperfect scanner focus or slight patient movements. This loss of sharpness can make it difficult to see fine anatomical details, potentially affecting diagnosis. To prepare our model for this real-world challenge, we introduced Gaussian blur into our synthetic image degradation process.

We didn't apply the same amount of blurring to each image. Instead, we randomly varied the degree of blurring, mimicking the range of sharpness variations seen in real prostate MRIs. This exposes our model to a wider range of blurring artifacts during training, making it more robust to blurry images in real-world scenarios.

The blurring is implemented by convolving each image with a Gaussian filter. The size of this filter, which controls how much blurring is applied, is randomly chosen and scaled based on the resolution of each image. This ensures the blurring effect is appropriately applied relative to the size of the details in the prostate image.

1.4. Ghosting artifacts: simulating periodic motion effects

Ghosting artifacts are often caused by rhythmic motions like blood flow or bowel movement. These artifacts appear as faint copies or "ghosts" of anatomical structures, shifted along a specific direction in the image. To simulate ghosting artifacts, we manipulate the image data in its frequency domain representation, known as k-space.

In k-space, periodic motions like those causing ghosting affect specific frequency bands. We simulate this by selectively suppressing the strength of certain frequencies in k-space. The amount of suppression controls the intensity of the ghosting effect, while the spacing between the suppressed frequencies determines how often the ghosting pattern repeats. By adjusting these parameters, we can create a wide range of ghosting artifacts with varying appearances.

Once the k-space data has been modified to include the simulated ghosting, we transform the data back to its original spatial representation, resulting in a prostate MRI image containing realistic ghosting artifacts.

1.5. Spike artifacts: simulating radio-frequency interference

Spike artifacts, often called herringbone or corduroy artifacts, can create unwanted stripes in MRI images. These artifacts are caused by radio-frequency interference during the scan, which introduces spikes in the k-space representation of the image.

To simulate these artifacts, we directly add spikes to the k-space data of the prostate MRI. Each spike's location in k-space determines the direction and frequency of the stripe pattern that will appear in the final image.

We randomize both the number and location of these spikes to simulate the unpredictable nature of real-world spike artifacts. The number of spikes controls how severe the artifact is, while their random locations create stripes in various directions and positions within the image. The intensity of each spike is also randomized, resulting in stripes with varying prominence. After adding the spikes to the k-space data, we transform the data back to its normal spatial representation, creating a prostate MRI image with realistic spike artifacts.

1.6. Bias field inhomogeneities: simulating magnetic field imperfections

Bias field inhomogeneity is a common artifact in MRI, caused by imperfections in the scanner's magnetic field.³ This artifact creates gradual changes in image brightness across the prostate, making some areas appear brighter or darker than others, even if the tissues are the same. This can make it harder to distinguish between different tissues and interpret the image accurately.

We simulated this artifact using a mathematical model based on polynomials. This

model can create smooth, gradual changes in image brightness similar to those seen in real bias field artifacts. The model uses a set of coefficients to control the intensity variations, and by randomly generating these coefficients, we create a variety of bias field patterns.

The complexity of these patterns can be adjusted by changing the order of the polynomial used in the model. A higher-order polynomial allows for more intricate and spatially varying intensity changes. The generated bias field pattern is then applied to the original prostate MRI image, simulating the effect of magnetic field imperfections.

1.7. Combining noise types: enhancing training realism

To create a more challenging and realistic training scenario, each image in the training and validation sets was augmented with every type of synthetic degradation individually (motion, Gaussian noise, blur, ghosting, spike artifacts, bias field inhomogeneity). This ensured the model learned to handle each artifact in isolation.

In addition to individual augmentations, combined noise augmentations were also applied to enhance the model's robustness and generalizability. For 20% of the images in the training and validation sets, we randomly selected a combination of two or more of the aforementioned noise types and applied them together to the same image. This approach aimed to simulate the more complex

and diverse noise profiles that are representative of real-world clinical scenarios where multiple artifacts can co-occur. The specific combinations of noise types were randomly selected for each image, ensuring a wide variety of combined degradations within the training data.

For the test set, to evaluate the model's ability to generalize to unseen combinations of artifacts, these noise types were combined randomly with varying ratios. This rigorous testing procedure helped assess the model's performance under conditions that more closely reflect real-world clinical prostate MRI.

2. The Pix2Pix model: image-to-image translation for prostate MRI enhancement

To address the challenge of improving the quality of degraded T2-weighted prostate MRI images, we employed the Pix2Pix model, a conditional Generative Adversarial Network (cGAN) renowned for its efficacy in image-to-image translation tasks.⁴ Unlike Cycle-GAN, which relies on unpaired data from two different domains, Pix2Pix leverages paired images for training. This makes it particularly suitable for our study, where we have access to ground truth data in the form of original, high-quality images corresponding to the synthetically degraded images.

The Pix2Pix architecture comprises two key components: a generator and a discriminator, which are simultaneously trained in an adversarial manner. The generator, structured

as a U-Net,⁵ takes the degraded T2-weighted image as input and endeavors to generate a high-quality image that closely resembles the original, artifact-free image. The U-Net architecture, with its encoder-decoder structure and skip connections, is particularly advantageous in medical image processing. The encoder progressively downsamples the input image, extracting features at multiple scales, while the decoder upsamples the encoded representation, reconstructing the output image. The skip connections, linking corresponding encoder and decoder layers, facilitate the direct flow of low-level information across the network, preserving crucial anatomical details and preventing excessive blurring often associated with traditional encoder-decoder networks.

The discriminator, on the other hand, employs a PatchGAN classifier,⁴ which evaluates individual $N \times N$ image patches rather than the entire image at once. This patch-based approach encourages the generator to focus on generating realistic high-frequency details crucial for maintaining the clarity and realism of medical images. During training, the generator and discriminator are engaged in a continuous adversarial loop. The generator strives to produce increasingly realistic images to deceive the discriminator, while the discriminator becomes more adept at discerning real images from the synthetic images generated by the generator. The proposed Pix2Pix architecture is given in Figure 1.

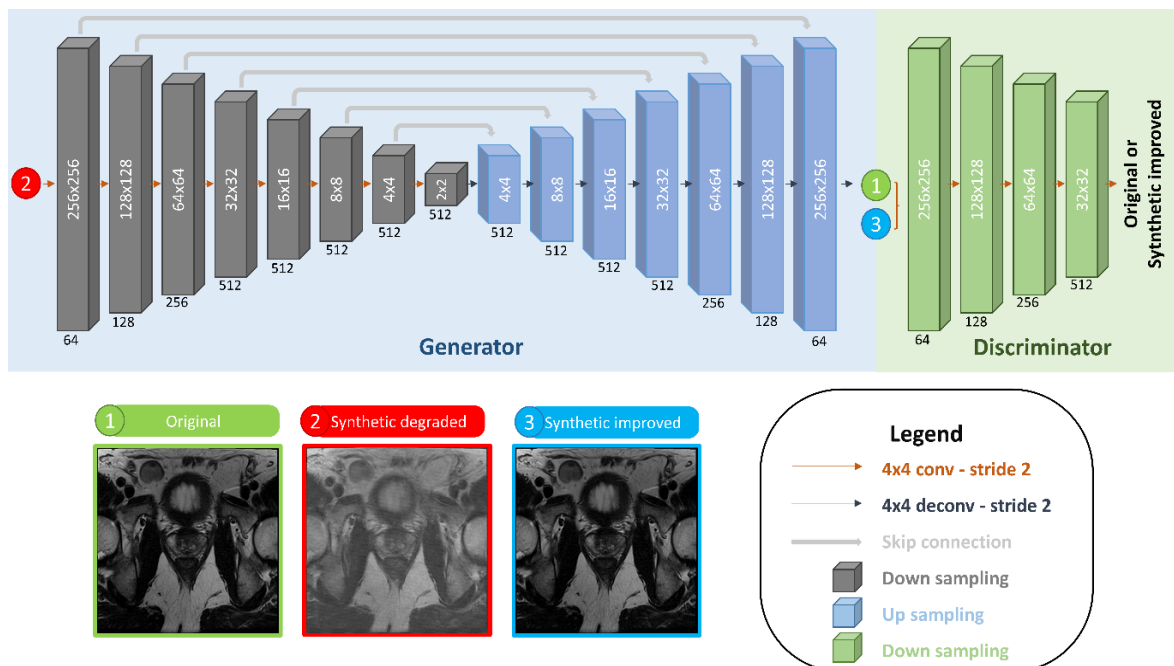


Figure 1. The proposed Pix2Pix architecture in current work.

A U-Net architecture represented by gray and blue blocks generate a synthetically improved image in generator. PatchGAN architecture shown in green blocks evaluate image patches to differentiate between original and synthetic images in discriminator and promotes high-frequency detail generation in the synthetic images. The image examples at the bottom of Figure 1. showcase (1) the original image, (2) the synthetically degraded image, and (3) the synthetically improved image respectively. Moreover, the numbers on the blocks denote the number of feature maps at each layer, and the image dimensions are displayed above each block. Each convolution operations, skip connections, down and up sampling operations are attended to corresponding arrows and blocks with specified colors under legend section in Figure 1.

This adversarial process is guided by a combined loss function encompassing both an adversarial loss and an L1 loss. The adversarial loss, determined by the discriminator's ability to classify image patches as real or

fake, ensures that the generated images are visually indistinguishable from the real images in the training dataset. The L1 loss, calculated as the mean absolute error between the generated image and the ground truth target image, promotes structural similarity, preventing excessive blurring and preserving the anatomical integrity crucial for accurate diagnosis.

We optimized the Pix2Pix model using the Adam optimizer with a learning rate of 0.0002 for both the generator and the discriminator. The model was trained for 200 epochs, with a heavier emphasis placed on the L1 loss during the early stages of training. This prioritizes the generation of structurally accurate images over purely visually realistic ones, particularly crucial in the context of medical imaging, where diagnostic accuracy is paramount.

References

1. Pérez-García F, Sparks R, Ourselin S. TorchIO: a python library for efficient loading, preprocessing, augmentation and patch-

based sampling of medical images in deep learning. *Comput Methods Programs Biomed.* 2021;208:106236. [\[CrossRef\]](#)

2. Shaw R, Sudre C, Ourselin S, Cardoso MJ. MRI k-Space motion artefact augmentation: model robustness and task-specific uncertainty, in: *Proceedings of The 2nd International Conference on Medical Imaging with Deep Learning*. PMLR. 2019:427-436. Accessed July 1, 2024. [\[CrossRef\]](#)
3. Van Leemput K, Maes F, Vandermeulen D, Suetens P. Automated model-based bias field correction of MR images of the brain. *IEEE Trans Med Imaging.* 1999;18(10):885-896. [\[CrossRef\]](#)
4. Isola P, Zhu JY, Zhou T, Efros AA. Image-to-image translation with conditional adversarial networks. 2018. [\[CrossRef\]](#)
5. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation, in: N. Navab, J. Hornegger, W.M. Wells, A.F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Springer International Publishing, Cham, 2015:234-241. [\[CrossRef\]](#)