

## ABSTRACT

**Purpose:** Diffusion-weighted images (DWI) provide crucial diagnostic information for cancer detection, but they are often hindered by low resolution and a low signal-to-noise ratio (SNR). This study addresses the challenge of low resolution and a low signal-to-noise ratio (SNR) in diffusion-weighted images (DWI), which are pivotal for cancer detection. Traditional methods increase SNR at high  $b$ -values through multiple acquisitions are gathered and averaged, at the cost of decreased, but this results in diminished image quality/resolution due to the motion-induced inter-acquisition variation. This study variations. Our research aims to enhance spatial resolution and improve image quality by leveraging by exploiting the global structure within multi-contrast diffusion-weighted DWI scans and the millimetric motion between acquisitions.

**Method:** The proposed We introduce a novel approach employing a fully-connected "Perturbation Network" that learns to learn sub-voxel-size motions between scans, along with a jointly-trained jointly with an implicit neural representation (INR) network that INR encodes the diffusion-weighted image DWI as a continuous volumetric function. The model treats the multi-contrast, treating voxel intensities of the low-resolution acquisitions as discrete samplings of this function, modeled by a neural network samples. By evaluating this continuous function with a finer grid of inputs, the our model predicts higher-resolution signal intensities for "intermediate" voxel locations at a higher spatial resolution.

**Results:** Comprehensive quantitative and qualitative evaluations conducted. The Perturbation Network's motion-correction efficacy was validated through experiments on biological phantoms as well as and in-vivo prostate DWI scans demonstrate that the model output significantly outperforms high-order interpolations scans.

**Results:** Quantitative analyses revealed significantly higher structural similarity measures of super-resolution images to ground truth high-resolution images compared to high-order interpolation ( $p < 0.005$ ). In blind qualitative experiments, 96.1% of super-resolution images were assessed to have superior diagnostic quality compared to interpolated images.

**Conclusion:** High-resolution details in DWI can be obtained without the need for high-resolution training data. One notable advantage of the proposed method is that it does not require a super-resolution training set. This is important in clinical practice because the proposed method can easily be adapted to images with different scanner settings or body parts, where the supervised methods do not offer such an option.

## 1. INTRODUCTION

Diffusion-weighted images (DWI) are used sensitive to measure the molecular displacement of water within each voxel. From By acquiring DWI acquired at different varying diffusion-weighting values ( $b$ -values), we can calculate apparent diffusion coefficient (ADC) maps can be calculated, which provide information about. These maps reveal tissue diffusivity. These measurements levels and are particularly important for cancer detection because water diffusion is the increased cell density typically restricted present in cancerous tissues. cancer significantly reduces the ADC values, which allows cancer detection. However, DWI often exhibits lower image resolution and signal-to-noise ratio (SNR) compared to other imaging sequences like T1 and T2-weighted images. To improve the SNR, multiple

high-b DWI acquisitions are performed ~~multiple times and~~. These are then averaged to create a single image. However, this approach is susceptible to significant artifacts. ~~These artifacts can be categorized into~~ Artifacts in these images mainly arise from two types, ~~both resulting from~~ of motion or ~~fluctuations in the local~~ magnetic field ~~fluctuations~~: (1) gross motion or ~~field~~ changes ~~in local field (between acquisitions (e.g., due to breathing, rectal gas motions etc.) that occur between acquisitions)~~, and (2) ~~the~~ localized ~~and small~~ motions or ~~field~~ changes ~~in magnetic field that occur during application of diffusion encoding gradients, (due to the (e.g., arterial pulse or peristalsis).~~ The non-rigid gross motion occurring between the separate signal acquisitions makes the image blurry, consequently reducing the resolution whereas the localized motion or ~~B0~~ field fluctuations during the acquisition can reduce or completely suppress the signal.

This presents a challenging situation with no clear solution. Specifically, when dealing with small voxel sizes, the common approach of signal averaging (obtaining multiple acquisitions and averaging them) can enhance the SNR but leads to decreased image quality and resolution due to motion-induced variations between acquisitions. In clinical ~~settings, the priorities of achieving practice~~, high SNR, ~~shorter reduced~~ scanning times, and patient comfort ~~often take precedence over typically outweigh the need for~~ high resolution. ~~As a result~~ Consequently, DWI images ~~typically exhibit~~ often have low resolution, ~~which can impact subsequent diagnoses affecting their specificity~~ and the accuracy of targeted biopsies. ~~Thus, resolution is considered crucial to the 'diagnostic image quality' in prostate MRI. This is measured with the PI-QUAL score, which reflects the extent to which the image assists in ruling in and ruling out clinically significant cancer. Studies using the PI-QUAL score have shown that DWI image quality, including SNR and resolution, significantly impacts the accuracy of cancer detection.~~

~~The objective of this~~ This study is to enhance the ~~introduces a super-resolution of DWI using model for diffusion-weighted prostate imaging that potentially reduces scan times and improves patient comfort, particularly for those with claustrophobia or Parkinson's disease. The method enables high-quality imaging from lower-resolution inputs, mitigating the need for a large number of averages needed for obtaining high-resolution images with adequate SNR. Our study aims to improve prostate DWI resolution through self-supervised deep learning techniques that leverage, utilizing anatomical patterns across different varying b-values and the non-linear inter-acquisition motion. The performance its efficacy was evaluated on assessed using in-vivo prostate MR images with biopsy-verified confirmed cancers, as well as particularly and carefully designed biological phantom experiments.~~

## 1.1 Super-resolution in Medical Images

Super-resolution in medical images has been a topic of interest over the past two decades. ~~The initial approaches focused on Early methods involving concentrated on~~ multi-image registration ~~and stitching with, utilizing shifted, orthogonal, anisotropic, or rotated scans. However, the effectiveness of methods using linear sub-voxel shifts has been debated, as they provide minimal or no in-plane improvement. This is because MRI is a Fourier imaging technique, and linear spatial shifts should not significantly alter the modulus image here rigid spatial shifts only introduce a linear phase gradient without yielding a more densely sampled k-space. Consequently, these techniques have only shown success in achieving through-plane super-resolution.~~

~~With the emergence and~~The success of deep learning-based super-resolution technologies in real-world images, the applications of super-resolution methods for medical imaging have drawn inspiration from ~~has inspired the application of similar supervised deep learning-based methods that have in medical imaging, drawing from techniques proven usefuleffective~~ in digital photography.

~~The more recent methods that use~~Recent supervised deep learning ~~hold the assumption~~methods ~~assume~~ that ~~thea~~ low-resolution image is a “degraded”~~'degraded'~~ version of ~~theits~~ high-resolution ~~and these methods aim~~counterpart, aiming to reverse this 'degradation process' to recover the high-resolution image ~~via reversing this “degradation process”~~. These methods have increasing number of applications. While these methods are increasingly applied and show promise, ~~however it is important, it's crucial~~ to note that, ~~similar to as with~~ other ~~medical~~ deep learning applications ~~in medicine, the power brought, their effectiveness is often limited~~ by ~~deep learning is hindered by the lack~~the scarcity of available training data. ~~As a result, most~~Consequently, much of the super-resolution training in medical imaging ~~reliesdepends~~ on synthetically degraded input-output pairs. ~~Table-1 provides a brief summary of the relevant literature on super-resolution in medical images.~~

~~Nevertheless, such~~However, it's important to approach these supervised ~~approaches should also be handled with care. High- and low-resolution image pairs are not available in methods cautiously. In medical imaging and yet, high- and low-resolution image pairs are not readily available, and synthetically degrading medical images to obtain suchcreate these~~ pairs for training ~~super-resolution networks can give riselead~~ to artifacts and hallucinations. The emerging ~~approach called~~self-supervised super-resolution ~~offersparadigm presents~~ a promising solution to ~~address theovercome these limitations associated with.~~ Self-supervised ~~approaches. In self-supervised learning, the algorithm leverageslearning utilizes~~ the inherent information ~~withinin~~ the data ~~itself, without relying on,~~ eliminating the need for explicit high- and low-resolution ~~image~~ pairs. ~~Table-1 offers a concise summary of the relevant literature in this domain.~~

Reference	Year	Modality	Body Part	Method
Peled et.al. <sup>[40]</sup>	2001	DWI	brain	registration of shifted scans
Carmi et al. <sup>[41]</sup>	2006	T2W	phantom	registration of rotated scans
Kennedy et al. <sup>[42]</sup>	2007	PET/CT	lung	registration of shifted scans
Mayer et.al. <sup>[43]</sup>	2007	T2W	phantom	registration of shifted scans
Lu et.al. <sup>[44]</sup>	2010	T2W	body	registration of parallel scans from different coils
Scherrer et.al. <sup>[45]</sup>	2011	DWI	brain	registration of rotated scans
Tieng et.al. <sup>[48]</sup>	2011	T2W	phantom	registration of shifted scans
Van Reeth et.al. <sup>[49]</sup>	2012	T2W	brain	registration of rotated/shifted scans (review article)
Scherrer et.al. <sup>[50]</sup>	2012	DWI	brain	registration of orthogonal scans
Wu et.al. <sup>[51]</sup>	2016	DWI	brain	High order singular value decomposition
Jurek et.al. <sup>[22]</sup>	2017	T2W	prostate	registration of orthogonal scans
Alexander et.al. <sup>[23]</sup>	2017	DWI	brain	patch regression
Chaudhari et.al. <sup>[46]</sup>	2018	T2W	knee	supervised CNN
Zeng et.al. <sup>[24]</sup>	2018	T1W	brain	supervised CNN
Chen et.al. <sup>[25]</sup>	2018	T2W	brain	supervised CNN
Sood et.al. <sup>[26]</sup>	2019	T2W	prostate	supervised GAN
Liu et.al. <sup>[27]</sup>	2019	T2W	prostate	supervised CNN
Hong et.al. <sup>[28]</sup>	2019	DWI	brain	supervised CNN
Dar et.al. <sup>[29]</sup>	2019	T1, T2	brain	supervised GAN
He et.al. <sup>[30]</sup>	2020	T2W	prostate	supervised deep attention networks
Zhou et.al. <sup>[31]</sup>	2020	T1, T2, FLAIR	brain	supervised multimodal fusion-net
Park et.al. <sup>[32]</sup>	2021	T2W	lung	supervised autoencoder-inspired CNN
Sood et.al. <sup>[33]</sup>	2021	T2W	prostate	supervised GAN
Chatterjee et.al. <sup>[34]</sup>	2021	DWI	brain	supervised Unet
Wu et.al. <sup>[35]</sup>	2021	T1W	brain	self-supervised IRN
Jiang et.al. <sup>[36]</sup>	2022	T1W	brain	supervised CNN with real 0.35T vs 3T scans
Molahasani et.al. <sup>[37]</sup>	2022	T2W	prostate	supervised GAN
Mahaptra et.al. <sup>[38]</sup>	2022	PD, T1 and T2	brain	supervised CNN and vision transformers
Guo et.al. <sup>[39]</sup>	2023	DWI	brain	semi-supervised GAN
This work	2023	DWI	prostate	self-supervised INR

**TABLE 1** Literature Review of Super-resolution in Medical Imaging. CNN: Convolutional Neural Networks, GAN: Generative Adversarial Networks, IRN: Implicit Neural Representation

This paper ~~introduces~~presents a self-supervised super-resolution model for DWI ~~using the implicit neural representation~~utilizing the Implicit Neural Representation (INR) framework. ~~In contrast to~~Unlike supervised methods, ~~which that~~ assume ~~that the~~ the low-resolution image is simply a degraded ~~version of~~ the high-resolution image and ~~aim~~seek to learn the inverse degradation function, our model adopts a more realistic assumption. It posits that the high-resolution image is a continuous function mapping the continuous domain of x-space to the range of signal intensities, and the low-resolution voxels are merely samples of this continuous function. ~~The Therefore, the~~ objective, therefore, ~~then~~ to learn this continuous function ~~through via the~~ through the INR network, using the ~~few~~sparse samples (voxels) ~~off from~~ from the low-resolution image in the training set. In other words, this framework uses the voxel locations as input, and the ~~voxel~~acquired signal intensities within these voxels as the output.

While the INR framework was initially not designed for super-resolution, Wu et. al. showed that using INR is an effective means of super-resolution on single-contrast T1W brain images. ~~In this paper, we~~ We introduce two novelties to the INR approach and extend its application to DWI of the prostate: ~~(1) Firstly, we incorporate multiple degrees of diffusion weighting (quantified by the gradient strength (b-value) is~~

~~used in addition to~~ along with the standard 3D spatial ~~input, thereby making it~~ inputs. This modification transforms the INR into a multi-contrast solution ~~and incorporating the signal decay within, thereby broadening its scope of application~~. ~~Second, we introduce the function fitting and (2) an input Perturbation Network is introduced to correct for the voxel mismatches observed between different signal acquisitions and potentially to harvest important morphological details that might exist in individual acquisitions but can be lost within the large voxel sizes and during aggregation of different acquisitions. The~~ This Perturbation Network learns to move the input voxel locations slightly for each acquisition while minimizing the reconstruction loss of the INR.

~~Another view on~~ An analysis of the literature ~~on~~ in Table-1 ~~shows~~ reveals that most ~~of the~~ super-resolution ~~work in literature is~~ research focuses on brain imaging ~~and the literature on, with~~ prostate MRI ~~is quite rare~~ studies being relatively scarce, even though prostate cancer is the second deadliest cancer in men. ~~Furthermore~~ Moreover, the ~~vast~~ majority of ~~the super-resolution~~ applications ~~of super-resolution to in MRI~~ have been ~~with~~ conducted using T1 or T2-weighted images. However, ~~applications of applying~~ super-resolution to ~~diffusion MRI may be more impactful because~~ DWI ~~have relatively which typically has~~ lower resolution. ~~All of the, could be more impactful. To our knowledge, existing~~ super-resolution ~~work~~ studies in DWI ~~has been on are limited to brain images~~ imaging, and the ~~rare~~ few prostate ~~papers are all~~ studies focus exclusively on ~~T2W~~ T2-weighted images; ~~there has been, with~~ no ~~attempt to apply super~~ attempts made on prostate DWI. Our study underscores the significance of high-resolution ~~to prostate DWI to the best of our knowledge~~ prostate DWI in (1) identifying extra-prostatic tumor extension, (2) pinpointing precise tumor locations, and (3) generating accurate coronal DWI views for enhanced cancer localization.

## 2. METHODS

The foundational assumption ~~underlying the proposed of our~~ model ~~differs from the conventional super-resolution approach is~~ that ~~is based on a "degradation function" assumption as discussed above. This paper assumes that the~~ a full-resolution analog image can be ~~modeled~~ represented as a continuous function ~~and the, with~~ low-resolution acquisitions ~~are merely being just~~ discrete and noisy samples from this function. ~~We use an~~ An INR network ~~as an approximator for is used to approximate~~ this continuous function.

As ~~described in the previous section, diffusion-weighted images~~ previously mentioned, DWI are typically ~~obtained~~ acquired at multiple various b-values ~~and as, including~~ multiple acquisitions at high b-values. Each ~~of these acquisitions~~ acquisition can be ~~denoted~~ represented as a discrete 3-D volume:

where  $b'$  is the diffusion-weighting value, ~~and~~  $M_x, M_y$ , and  $M_z$  ~~are~~ denote the number of voxels along the  $x, y$ , and  $z$  directions, respectively ~~and, with~~  $K$  ~~is~~ representing the number of independent acquisitions. In standard practice, these acquisitions, denoted as  $S_k$ , are averaged ~~together~~ to obtain produce a single mean image ~~to be used by radiologists for radiological examination~~.

The INR network models the full-resolution image ~~is modeled via an INR network~~ as a function ~~from~~  $R^4 \rightarrow R$ . ~~The~~ This 4D input space ~~is~~ comprises the continuous 3D ~~space of~~ voxel locations ~~plus~~ and the b-value, ~~which serves~~ as the 4<sup>th</sup> fourth dimension. The network's output is the signal intensity ~~observed on the given pixel~~ corresponding to a specific voxel location and the b-value. The INR is ~~modeled~~ structured as a deep neural network with fully-connected layers and sinusoidal activation functions, ~~initialized as suggested in~~ following the initialization guidelines from Sitzmann et al. ~~The input samples in the~~



Training set inputs are quadruplets  $(m_x, m_y, m_z, b)$  corresponding to voxel location and the b-value, representing the normalized voxel location and b-value, scaled to vary range between -1 and 1 along each axis. The model's output labels in the training set are the intensity values of the image observed at the input location, i.e.,  $S_k(x_i, b)$ . Thus, the INR is trained to minimize the cost error between the predicted diffusion values for the set of possible input locations  $x_i = (m_x, m_y, m_z)$  and the measured signal diffusion values  $S_k(x_i, b)$ :

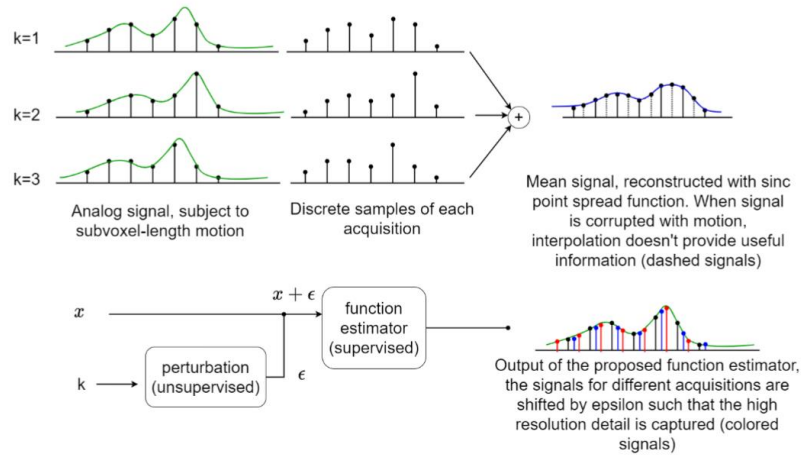
In other words, the goal of INR is to approximate the signal for any value of given input location, including even for those that don't exist in the training set, i.e., "intermediate voxels" with higher resolution. For example, since the input space is discrete, for instance, our practically discrete input space: while input samples like  $x_1 = (0, 0, 0)$  and  $x_2 = (1/(M_x/2), 0, 0)$  both exist in the training set, but any intermediate value, say sample along the first direction, e.g., axis, such as  $x_3 = (1/(M_x), 0, 0)$ , does not exist in the training set. When INR is trained with  $x_1, x_2$  and all other location indices in the training set, it can suggest a diffusion value for the "intermediate voxel locations", e.g.,  $x_3$ , the INR is capable of predicting diffusion values for these 'intermediate voxel locations', like  $x_3$ . Importantly, it should be emphasized that the INR functions beyond the capabilities of a simple sinc interpolator. This distinction arises for three reasons: (1) INR integrates information across various b-values within shared layers (eq.1), (2) the Perturbation Network applies non-linear motion correction across acquisitions and b-values (Sec2.1), and (3) the training process incorporates data-driven learning weights for each voxel's acquisitions, addressing inter-acquisition variability (Sec2.2).

## 2.1 Input Perturbation Network for Motion Artifacts Correction

As demonstrated in Fig. 2, water diffusion measurements  $S_k$  for an input location could vary significantly across different acquisitions (k) due to the motion occurring during and between acquisitions. This variation is too large to be explained by thermal Rician noise, and is mostly caused by The motion-induced physiological noise. Our desire is that the proposed model corrects for the effects of motion, intra- and inter-acquisition variabilities are often non-linear due to B0 inhomogeneities and even takes advantage of it to provide a detail that would have otherwise been lost in signal averaging.

We can mathematically model bowel-enclosed gas. We represent the warping effect of millimetric motion as a perturbation to the input of the INR network since we define the image as a function mapping the spatial coordinates to signal intensities INR input. We assume that the reason  $S_{k1}(x) \neq S_{k2}(x)$  is the millimetric motion between acquisitions  $k1$  and  $k2$  and that  $S_{k1}(x) = S_{k2}(x + \epsilon)$ , where  $x = (m_x, m_y, m_z)$  is any discrete voxel location and  $\epsilon = (\epsilon_x, \epsilon_y, \epsilon_z)$  is a perturbation of the input space for  $x$ . In other words,  $\epsilon$  is a perturbation of the input space for  $x$ . This implies that when the an acquisition's signal value for an acquisition is different from the network output, then this method implicitly questions whether the discrepancy is due to the acquisition does not originate from the given designated voxel location, but from a location nearby sub-voxel away due to the location, affected by small inter-acquisition motion. motions. A 1-D example of the idea behind the proposed methodology is visualized in Figure-1. In this example, when the point spread function is applied on each voxel, very small motions that occur between acquisitions cause the sharp edge to be lost even when it is sampled at the Nyquist rate. In the proposed approach, the discrete input space is generalized as a continuous domain, and the input perturbation

network is trained in an unsupervised fashion to estimate a warping function for each acquisition to minimize the reconstruction loss at the output.



**FIGURE 1** A 1D example of the zero-shot super-resolution methodology. The supervised block uses the low-resolution voxel intensities as labels, hence effectively there is no need for high-resolution ground truth labels.

We propose that this perturbation value ~~for each acquisition~~ can be learned jointly with the INR parameters, on a separate neural network. The error that the signal for such an acquisition creates at the output is backpropagated to assign this value to a better location near the center of the input voxel such that the total loss is minimized for all voxels and all acquisitions. This warping is done via the Perturbation Network (PN) which is modeled as another fully connected neural network that takes the acquisition number and the b-value as input, in addition to the voxel location  $x$ . The network then estimates a perturbation value epsilon via a scaled tanh activation function such that the perturbation is constrained to vary within physically feasible motion limits by enforcing the following cost:

The perturbation upper-bound  $\epsilon_{\max}$  was set as half a voxel size along each direction, i.e.  $\epsilon_{\max} = 1/M_x, 1/M_y, 1/M_z$ . Note that the acquisition number  $k$  and the b-value used as inputs to the Perturbation Network take discriminative roles, and no temporal order along acquisitions is assumed. This creates an opportunity to capture sharp edges within the image, even if the voxel size is too large to capture it, and this edge is captured in different acquisitions for the same voxel due to the motion. The flowchart of the proposed network and the role of the Perturbation Network is visualized in Fig 2.

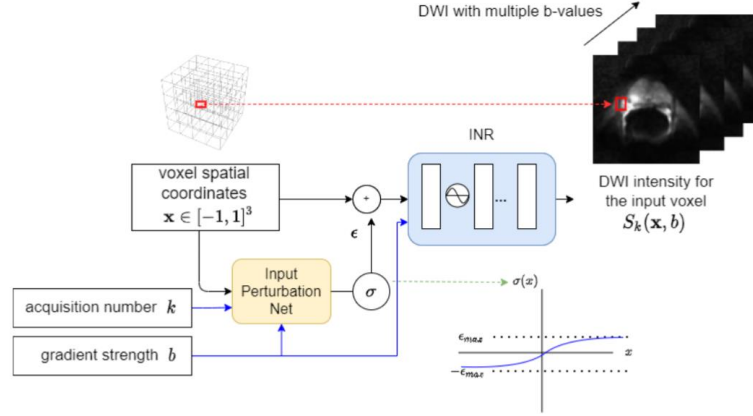


FIGURE 2 The flowchart of the implicit neural representation learning, augmented by the input perturbation network.

## 2.2. Weighted Learning for Suppressing Anomously Low Signals

Besides the inter-acquisition bulk motion of the prostate as discussed in Sec 2\_1, ~~another reason that the disparity in  $S_{k1}(x) \neq S_{k2}(x)$  for  $k1 \neq k2$  can be may also stem from~~ the molecular motion of water or the fluctuation in local magnetic field during application of the diffusion-sensitizing gradients. ~~These factors lead to variations that are too large to be explained by thermal Rician noise.~~ This ~~cause~~type of variation is commonly referred to as physiological noise. Physiological noise typically leads to an anomalous reduction ~~of in~~ signal intensity ~~and sometimes the signal is completely lost such that this signal cannot be recovered. This makes the distribution of, resulting in a right-skewed per-voxel per-acquisition~~ signal intensity ~~right-skewed, and as a result~~distribution. This aspect is crucial, as conventional signal averaging ~~produce~~tends to yield ADC values ~~that are much~~significantly higher than the ~~true~~actual ADC. ~~This is important because the ADC estimate is a very useful measure to evaluate the values, which are key for evaluating tissue composition, and ultimately diagnosed~~diagnosing cancers.

~~By representing~~Representing a diffusion-weighted image  $S$ , ~~obtained~~derived from multiple acquisitions  $S_k$ , as a single ~~implicit neural architecture~~ INR (as proposed in this ~~paper~~), the ~~inherently problem of signal mismatch can be study~~, effectively ~~addressed. Since the INR framework operates in~~addresses such signal mismatches. The INR framework's operation ~~within~~ continuous spaces for both ~~the~~ domain and range, ~~grants~~ it ~~possesses~~an awareness of proximal voxels. **This characteristic allows the INR to potentially favor intensity values of the acquisitions that are more consistent with the neighboring voxels, while naïve-averaging merely calculates the mean of different signal intensities for each voxel.**

The proposed methodology also presents an opportunity to explicitly address this problem by assigning weight values to each acquisition during training. We suggest one option to assign weight values to acquisitions to automatically suppress motion-corrupted signals. ~~Recently, Gundogdu et. al. proposed such a systematic way of weighting the signal of each acquisition in each voxel. This weighting is based on the observation that the distribution of per-voxel signal intensity over all acquisitions is skewed, since the motion artifacts (and any other artifacts) are much more likely to decrease signal than to increase signal in DWI [30]. Consequently, as recently proposed by Gundogdu et. al. Consequently,~~ we alter the loss function in Equation (3) to incorporate learning weights ( $w_k$ ) as thus:



Detailed derivation and further interpretation on the learning weights are provided in Supplement.

## **2.3. Evaluation of Super-Resolution Performance and Accuracy**

### **2.3.1. Results**

Our calculations were conducted on a workstation with Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz, Tesla V100 GPU with 32GB memory. The training of the models was performed using the PyTorch package. During the inference phase, when obtaining to generate the super-resolution image, the INR network is invoked using employs a finer input grid of inputs than the resolution of the input voxels it was trained with used in training. This results in approach yields an image characterized by with enhanced detail and higher improved resolution. However, in order to assert that the new details introduced by the network corresponds correspond to actual anatomical features, it is crucial to establish reliable ground truth images for evaluation. In the case of super-resolution examples in digital photography, such datasets are readily available. On the other hand, for MRI, especially DWI, such acquiring high-resolution image acquisition images to this degree is not feasible. In-vivo acquisitions with higher resolution would result in unacceptably low SNR. Furthermore, the long acquisition times would increase vulnerability to motion artifacts. Therefore, in addition to the in-vivo prostate scans, we employed a biological phantom to assess the performance of the model. This approach enables a controlled evaluation environment, allowing us to evaluate the model's ability to capture and reproduce anatomical details. evaluate our model's performance.

### **2.3.1.1 Unsupervised Motion Correction with Evaluation on Phantom**

This specifically designed phantom experiment aimed to test the Perturbation Network

We tested the unsupervised motion-correction performance efficacy of the Perturbation Network by emulating the existence. In this experiment, we emulated both the presence and absence of the inter-acquisition motion on in MRI scans of a biological phantom. We chose kiwifruit Kiwifruit was selected as the biological phantom since several groups, as various studies have proposed that it is demonstrated its effectiveness as a good proxy for prostate MRI, including DWI included.

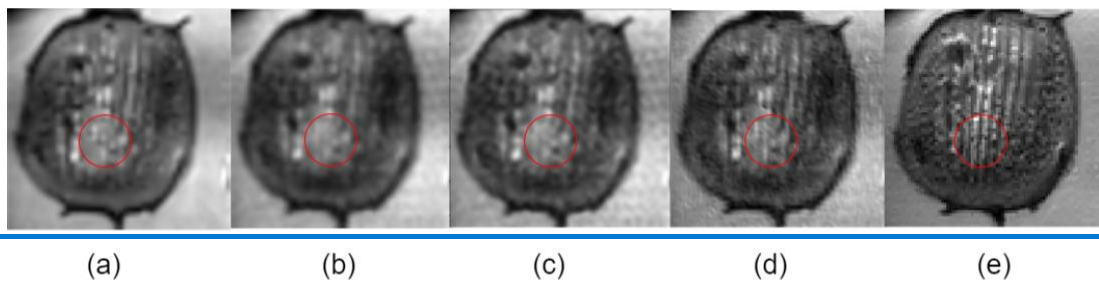
### **3.1.1 Evaluation on Phantom**

We Diffusion-weighted images of kiwifruit were acquired kiwifruit diffusion-weighted images in under three different distinct settings. Firstly, we acquired images with a resolution of  $1.5\text{mm} \times 1.5\text{mm} \times 1.5\text{mm}$   $5\text{mm} \times 1.5\text{mm} \times 3.0\text{mm}$  in nine repetitions. To simulate rigid motion along in the axial plane, we introduced simulated transitive motion to each repetition by shifting the iso-center of the field-of-view in-plane by  $-0.5$ ,  $0$  or  $+0.5\text{mm}$  in various in-plane the readout or phase-encoding directions. In some cases, the shifting of the center of the field-of-view is achieved in post-processing by the scanner, rather than by changing the sequence parameters, and thus the underlying k-space data may not have changed due to the shift. However, our algorithm operates on the produced magnitude images without relying on the complex k-space information and the main objective was to assess if the Perturbation Network corrects this motion with unsupervised training, thus this was deemed acceptable.

The second image was acquired as a ground truth with the same resolution and the same number of acquisitions as the first one, but with no motion. We ~~acknowledgerecognize~~ that the synthetically introduced motion is ~~merely~~ rigid and ~~it maymight~~ not ~~completely simulate the performance of the proposed model underfully replicate~~ the non-rigid warping ~~that may occur duringoften encountered in prostate~~ diffusion imaging ~~of the prostate~~. ~~Nonetheless. However~~, we ~~choseopted~~ to apply ~~the~~ motion during the MR scan ~~instead of introducing motionrather than inducing~~ artifacts in ~~the~~ post-processing stage, ~~aiming~~ to ~~maintain a closer resemblance tomore accurately mirror~~ real-world ~~scenariosconditions~~.

~~In addition~~Additionally, we acquired a third image with a ~~very~~ high-resolution of 0.75mmX0.75mmX3mm ~~to use, serving as a~~ reference to evaluate the detail ~~captured~~ in the INR output and the first two images (~~with and without motion and no motion~~).). All scans were conducted on a dStream Philips 3T Ingenia scanner with ~~b=1500sec/mm2 diffusion-sensitizingdigital coil technology, a~~ 70 cm bore, dual RF transmission, and maximum gradient strength ~~of 45 mT/m and maximum slew rate of 200 T/m/s~~. DWI images were acquired at  $b = 1000 \text{ s/mm}^2$ , with  $TR/TE = 4000/90 \text{ ms}$ , 4 EPI segments, in-plane resolution of  $1.5 \times 1.5 \text{ mm}^2$  (acquired) and  $1.4 \times 1.4 \text{ mm}^2$  (reconstructed), and slice thickness of 3 mm. The phase encoding direction is along the main axis of the kiwi fruit, parallel to its striated structural pattern.

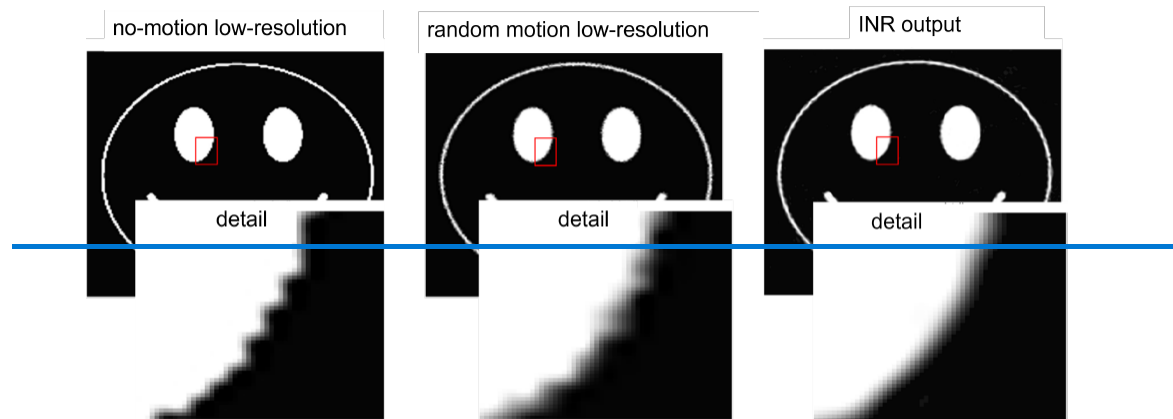
~~We~~To see the correcting impact of the Perturbation Network, we trained one INR by setting  $\epsilon_{\max}=0$  and another by setting  $\epsilon_{\max}$  to be half a voxel size. We used the scans with the motion as input to the network and compared the output to the no-motion ground truth, as well as the very high-resolution ground truth. ~~Example images from the phantom experiment are shown in Figure 3. Since MRI is Fourier technique, the calculated shifts only provided linear phase in k-space and no additional increase in resolution. However, this experiment shows that the MRI scan can be modeled as an INR and the Perturbation Network can implicitly mitigate the motion effects.~~



**FIGURE 3** Proposed network on kiwifruit DWI. (a) image acquired without simulated motion, (b) image acquired with simulated motion, (c) model output with  $\epsilon_{\max}$  is set to zero, i.e. Perturbation Network is not in effect, (d) model output with Perturbation Network turned on and (e) very high-resolution “ground truth” image used for visual validation. The vertical detail enclosed in red circle is lost as a result of motion as seen in (b). The model output at (d) does not only address the effects of motion, but it also provides a better image then the “no-motion” image shown in (a), as can be verified from the very high-res image in (e).

~~In the kiwi experiment, Perturbation Network successfully learned the direction and magnitude of voxel movements for each acquisition (epsilon). For this experiment, the vector epsilon was the same for all voxels and different for each acquisition. To assess the ability of the Perturbation Network to decrypt completely random perturbations for all voxels, we generated a motion-corrupted image by slightly displacing each voxel in a random direction. This was achieved by perturbing the location of each voxel~~

in each acquisition towards a random direction slightly. This in-silico experiment assumes a significantly more severe scenario than the actual motion effects observed in DWI scans. We conducted this test to determine whether the Perturbation Network could learn the motion effects despite being trained in an unsupervised manner. The INR was then trained with these images to encode a single image function. Figure 4 provides a comparison between the image with no motion corruption, the motion-corrupted mean image, and the output of the INR.



**FIGURE 4** Comparison between the in-silica image with no motion corruption (left), the motion-corrupted mean image (middle), and the output of the INR (right). The input to the INR model was the middle image.

### 2.3.2 Evaluations on In-vivo Prostate Images

We evaluated the performance of the proposed method on diffusion-weighted prostate images through both qualitative and quantitative experiments. To evaluate the resolution improvement, we created a training set of low-resolution diffusion-weighted prostate images by applying low-pass filtering and sub-sampling to the original prostate images. Another approach could have involved acquiring a separate set of low-resolution images with larger voxel sizes during a second scan of the patient. However, this would have introduced different SNR, rendering the original images suboptimal as ground truth. No additional noise was introduced to the training set. The network was trained with diffusion-weighted images with 4 b-values: 0, 150, 1000 and 1500 sec/mm<sup>2</sup>, so that it can yield a super-resolved and motion-corrected ADC map. This setup, which involved simultaneous training of different Orthogonal gradients were in the readout, phase encoding, and slice directions. At high b-values and the potential for a super-resolved ADC map, distinguishes our study from previous works in the literature. Multiple acquisitions at each orthogonal directions were used as independent acquisitions to train INR. The INR framework addresses the inter-acquisition motion directly via the Perturbation Network, and intra-acquisition motion indirectly by suppressing the resulting low signals via weighted learning.

### 3.2.1 Qualitative Evaluation with Observer Study

The network was trained with the voxels of the subsampled low-resolution image, and. Then the trained network was used to generate the 2x super-resolution, although it should be noted images. It is noteworthy that the network can potentially yield multi-resolution output with a scale other than has the potential to produce outputs at multiple resolutions, not limited to the 2x factor. The 2x super-resolution images, 2x bicubic interpolated images and the ground truth (GT) images were examined by

Two fellowship-trained and board-certified attending radiologists, with 12+ and 6+ years of experience within prostate MRI (NCO and GL respectively ~~(NCO and GL), who were asked to rank~~), evaluated the 2x super-resolution images, 2x bicubic-interpolated images (serving as an approximation to sinc interpolation), and ground truth (GT) images. They were tasked with ranking these ~~three sets of~~ images ~~(super-resolution, bicubic-interpolated, and GT)~~ based on their perceptual quality.

This experiment involved 65 cases with high-b images, and the order of presentation for the three test images was randomized for each subject during evaluation. Hence, the radiologists were unaware of the identity of the specific image ~~they were evaluating~~.

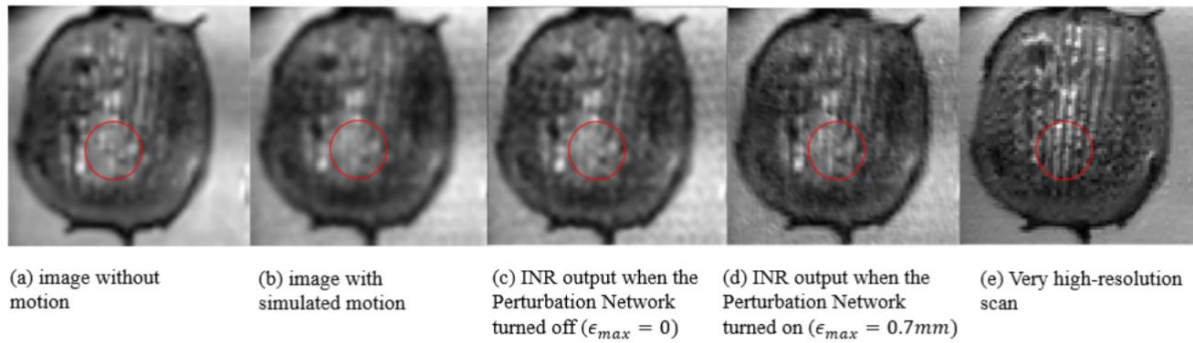
The same set of 65 cases was utilized for a quantitative analysis, comparing the super-resolution and bicubic-interpolated images against their respective ground truth counterparts. For this analysis, we used established quantitative image similarity metrics common in digital image analysis, such as the Structural Similarity Index Measure (SSIM), Multiscale SSIM (MS-SSIM), Feature Similarity Index Measure (FSIM), and Spectral Residual-based Similarity Index Measure (SR-SIM). Each metric was used to compare the 2x super-resolution and 2x bicubic-interpolated images with their corresponding high-resolution ground truth images. Statistical significance was evaluated using paired t-tests with N=65.

### **3. RESULTS**

Model training was carried out utilizing the PyTorch software package on a workstation equipped with Intel(R) Xeon(R) Gold 6130 CPU @ 2.10GHz, Tesla V100 GPU with 32GB memory.

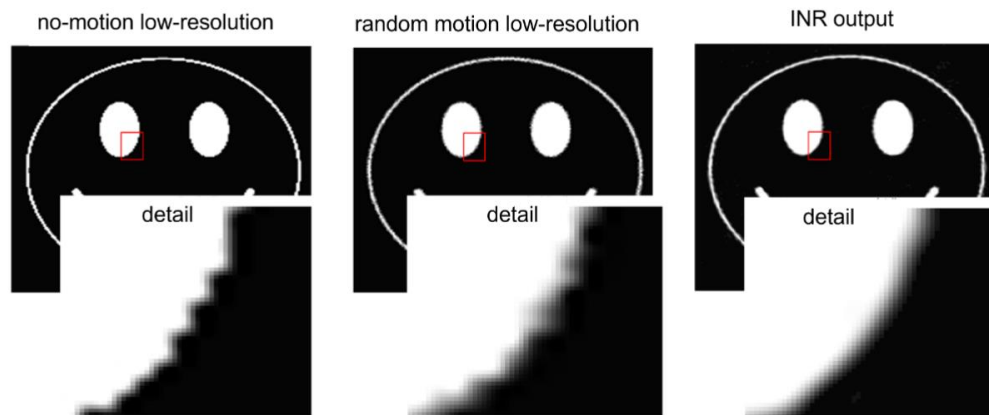
#### **3.1. Unsupervised Motion Correction with the Perturbation Network**

Figure-3 displays representative images from the phantom experiment. Since MRI is a Fourier technique, the calculated shifts resulted in a linear phase change in k-space without contributing to increased resolution. The purpose of this experiment was to demonstrate that MRI scans can be modeled using an INR, and the Perturbation Network can effectively mitigate motion effects. The Perturbation Network successfully inferred the direction and magnitude of voxel movements (epsilon) for each acquisition and applied the necessary corrections.



**FIGURE 3** Proposed network on kiwifruit DWI. (a) image acquired without simulated motion, (b) image acquired with simulated motion, (c) model output with  $\epsilon_{max}$  is set to zero, i.e. Perturbation Network is not in effect, (d) model output with Perturbation Network turned on and (e) very high-resolution “ground truth” image used for visual validation. The vertical detail enclosed in red circle is lost as a result of motion as seen in (b). The model output at (d) does not only address the effects of motion, but it also provides a better image than the “no-motion” image shown in (a), as can be verified from the very high-res image in (e).

In the kiwifruit experiment, the perturbation vector epsilon remained consistent across all voxels but varied for each acquisition. To assess the Perturbation Network’s ability to decipher completely random perturbations for all voxels, we created a motion-corrupted image by slightly displacing each voxel in random directions. This was achieved by slightly perturbing the location of each voxel in each acquisition towards a random direction. This in-silico experiment simulated a scenario significantly more severe than the actual motion effects typically observed in DWI scans. We conducted this test to determine whether the Perturbation Network could learn such nonlinear motion effects despite being shown to them. On trained in an unsupervised manner. Figure 4 provides a comparison between the image without motion corruption, the motion-corrupted average image, and the INR output.



**FIGURE 4** Comparison between the in-silica image with no motion corruption (left), the motion-corrupted mean image (middle), and the output of the INR (right). The input to the INR model was the middle image.



### 3.2 Evaluations on In-Vivo Prostate Images

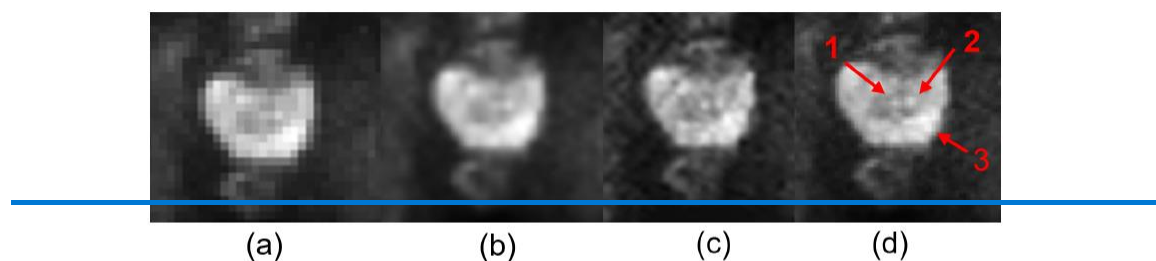
In the observer study, 96.1% of the super-resolution images were rated ~~to have better~~as having superior perceptual quality ~~than compared~~ the interpolated ~~image (93.8% by NCO and 98.4% by GL) images~~. Moreover, ~~on average~~ 40.7% of the super-resolution images were voted to have better quality ~~(55.3% by NCO and 26.1% by GL)~~ than the ~~GT image, which had twice~~ground truth images, despite the ground truth images having double the resolution and ~~four times more voxels than~~quadruple the voxel count of the model's input. This ~~second~~latter finding is particularly ~~intriguing~~noteworthy as it indicates that the Perturbation Network significantly contributes to the improvement in image quality in these cases. The ~~most likely~~ reason for this ~~outcome~~ is that the model ~~input sees~~is trained by K=12 individual acquisitions ~~despite seeing~~ only a quarter of the number of voxels as input (compared to ground truth), ~~yet each input is trained by K=24 acquisitions~~. It is worth noting that the interpolated and ground truth images were also generated using the mean of the same number of acquisitions, implying no expected SNR gain with the algorithm. When the Perturbation Network is not in effect, the main INR learns to average these acquisitions for the observed locations in the training set and attempts to predict the intermediate voxels based on the continuity of the INR function. However, considering that the model can produce outputs that visually outperform the high-resolution ground truth, we can argue that the Perturbation Network learns to reposition certain features away from the voxel center, thereby providing higher resolution information and accounting for sub-voxel motion. This effect is most prominent when dealing with sharp features that exhibit relatively high SNR.

#### 3.2.2 Quantitative Evaluation with Image Similarity

~~The same set of 65 cases was utilized for a quantitative comparison between the super-resolution and interpolated images based on their corresponding ground truth counterparts. We employed well-established quantitative image similarity measures widely used in the field of digital image analysis, including the structural similarity index measure (SSIM), multiscale SSIM (MS-SSIM), feature similarity index measure (FSIM), and spectral residual-based similarity index measure (SR-SIM). For each metric, we compared the 2x super-resolution and 2x bicubic-interpolated images to their respective high-resolution ground truth images. Statistical significance was evaluated using paired t-tests with N=65. The results indicated that the super-resolution images~~In quantitative analyses, super-resolution INR outputs exhibited significantly higher similarity to the ground truth images compared to the bicubic-interpolated images ( $p < 0.005$ ) across all quantitative similarity metrics (SSIM, MS-SSIM, FSIM, and SR-SIM).

All Bands					High Frequency Bands			
Metric	bicubic vs GT	SR vs GT	Gain (%)	p-value	bicubic vs GT	SR vs GT	Gain (%)	p-value
SSIM	88.5	89.3	0.9	<0.001	52.5	56.3	7.5	<0.001
MS-SSIM	88.6	90.1	1.8	<0.001	72.0	75.8	5.4	<0.001
SR-SIM	87.8	89.0	1.5	<0.005	87.6	89.1	1.8	<0.005
FSIM	90.2	91.9	1.9	<0.001	88.5	90.5	2.3	<0.001

TABLE 2 Image similarity measures of super-resolution compared to bicubic interpolation. SR: super-resolution, GT: ground truth

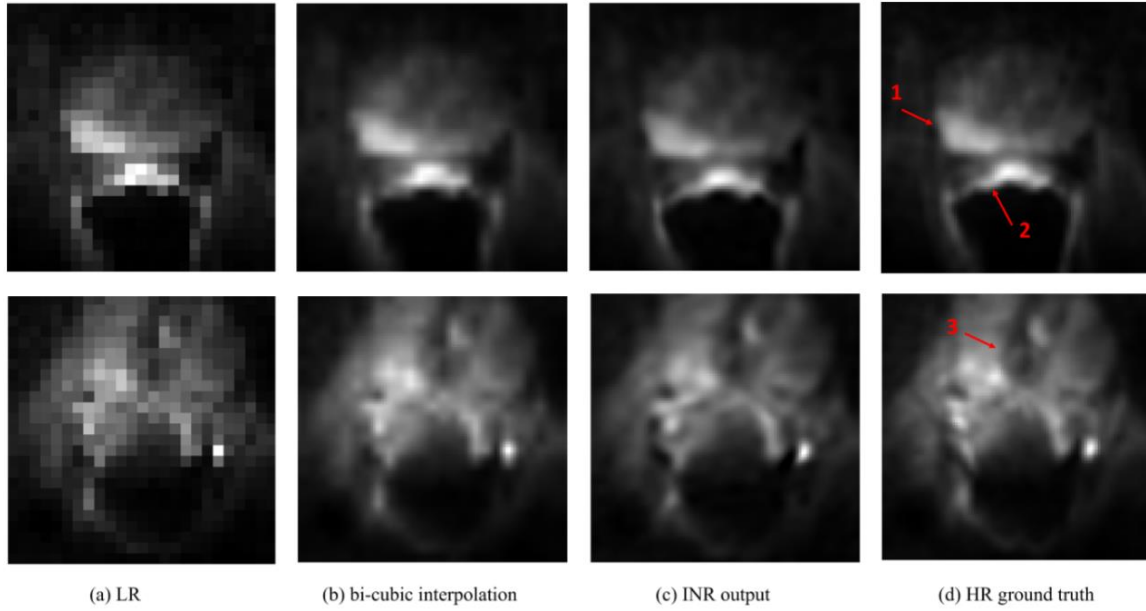


**FIGURE 5** Performance of the proposed method on a sample prostate image. (a) LR image, (b) 2x bi-cubic interpolated image, (c) the output of the model with 2x super-resolution and (d) the HR ground truth image. The red arrows marked with 1 and 2 depict the ejaculatory ducts and the 3 is a peripheral zone cancer. It can be seen that the SR output delineates some anatomical detail like ejaculatory ducts better than the interpolation. Furthermore, cancer contrast is increased with the contribution of the weighted training.

Naturally, ~~most image~~ areas ~~in many images~~ are smooth regions and consist of low-resolution components that are similar in super-resolution, bicubic-interpolated and ground truth images. Consequently, these measurements yield very high similarity values ~~in the aforementioned measurements~~. However, our ~~main interest lies in assessing primary focus is on determining~~ whether the high-resolution details introduced by the super-resolution model ~~also exist are present~~ in the ground truth images. To investigate this, we applied a high-pass filter to all super-resolution, ground truth, and bicubic images, effectively removing the low spatial frequency components, and focused solely on the high spatial frequency details. This analysis revealed that most of the gain obtained by super resolution is in high spatial resolution components, e.g., edges etc. ~~The Table 2 presents the~~ similarity scores of the super-resolution ~~image~~ and ~~the~~ bicubic-interpolated mean ~~image images~~ (baseline) ~~with respect relative~~ to the high-resolution ground truth ~~are given in Table 1. Figure 4. Figure 5~~ showcases a typical example from this evaluation set. ~~More Additional~~ details ~~about on~~ the similarity scores and ~~more further image~~ examples ~~of images are provided can be found~~ in the Supplement.

Metric	All Bands				High Frequency Bands			
	bicubic vs GT	SR vs GT	Gain (%)	p-value	bicubic vs GT	SR vs GT	Gain (%)	p-value
SSIM	88.5	89.3	0.9	<0.001	52.5	56.3	7.5	<0.001
MS-SSIM	88.6	90.1	1.8	<0.001	72.0	75.8	5.4	<0.001
SR-SIM	87.8	89.0	1.5	<0.005	87.6	89.1	1.8	<0.005
FSIM	90.2	91.9	1.9	<0.001	88.5	90.5	2.3	<0.001

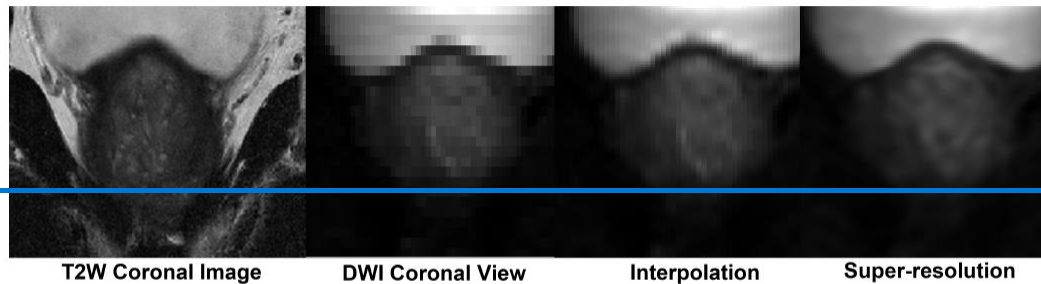
**TABLE 2** Image similarity measures of super-resolution compared to bicubic interpolation. SR: super-resolution, GT: ground truth



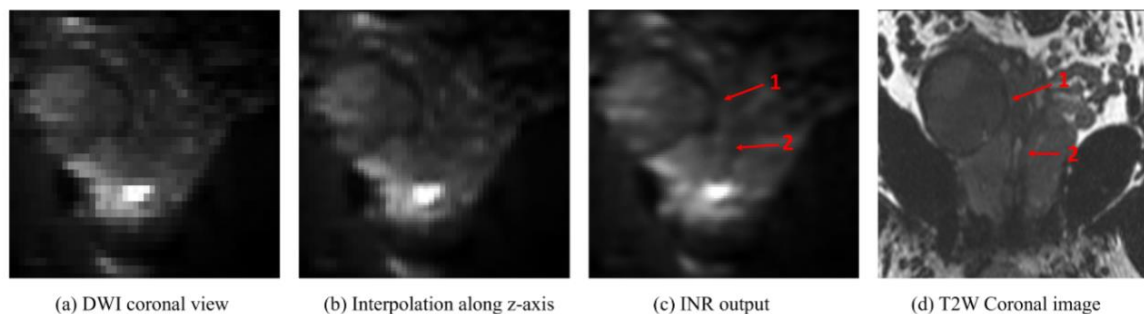
**FIGURE 5** Performance of the proposed method on two cases: (a) LR image, (b) 2x bi-cubic interpolated image, (c) the output of the model with 2x super-resolution, and (d) the HR ground truth image. The red arrows show anatomic details that INR enhances. Arrow 1 shows the extraprostatic extension of the cancer, arrow 2 shows the rectal wall, and arrow 3 shows that the boundaries of the urethra and an adjacent suspicious lesion can be seen better in SR image.

### 3.3 Super-resolution Performance Along Slice-select Direction

The proposed methodology treats the DWI as 3D volumetric objects, aiming to enhance super-resolution not only in the axial view but also along the slice select direction. Typically, the z-axis slices are taken thicker than the axial resolution to improve SNR gain. ~~As a result~~ Consequently, the coronal or sagittal views of DWI, despite their potential significance, are often overlooked in clinical practice, ~~despite their potential importance, particularly for MRI-guided biopsy.~~ By reconstructing ~~Reconstructing~~ INR with finer details along the slice-select direction, ~~it becomes possible to visualize~~ enables the visualization of coronal or sagittal DWI views ~~of DWI that are obtained with,~~ derived from axial slices. ~~In~~ Figure-5, ~~we present~~ demonstrates the super-resolution reconstructed coronal view of a sample diffusion-weighted image alongside the corresponding slice of the T2W coronal image ~~captured from the same slice.~~



**FIGURE 6** Super-resolution reconstructed coronal view of one sample image compared to the interpolated image. T2W coronal image that corresponds to the same slice is given on the left for comparison.



**FIGURE 6** Super-resolution reconstructed coronal view of one sample image compared to the interpolated image. T2W coronal image that corresponds to the same slice is given on the right for reference. Arrow 1 shows the boundaries and extent of the cancer whereas Arrow 2 shows the prostatic urethra which was not visible in interpolation but visible in super-resolution image.

### 3.4 Contrast Enhancement via Weighted Training

To assess the contrast enhancement in ADC values due to weighted training, we trained the super-resolution network using all the voxels from with the high-resolution ground truth GT image. We then compared and applied the learning weights as detailed in the model's output with the mean image. To evaluate the performance of Supplement. ADC maps calculated using the model, we examined output are compared to the ADC maps of GT images. The analyses were done on a cohort of 12 patients with biopsy-verified cancers and measured by measuring both the SNR on cancer and the cancer-to-healthy-prostate contrast ratio (CR). Statistical analysis was conducted using paired t-tests.

The CR significantly improved on average from 1.9 to 2.4 ( $p < 0.01$ ) when comparing the model output with the mean input image. Similarly, the SNR showed a significant average increase from 1.87 to 1.92 ( $p < 0.01$ ), despite no noise suppression being applied.

## 4. DISCUSSION AND CONCLUSIONS

Our This study demonstrates validates the feasibility of zero-shot super-resolution for in diffusion-weighted prostate images imaging and highlights underscores the potential of utilizing using individual

acquisitions ~~infor~~ self-supervision to mitigate the effects of supervised motion mitigation to some extent. Although introducing synthetic shifts and obtaining super-resolution is a well-established implementation, we provided a novel unsupervised method to ~~make use of~~utilize the underlying non-rigid motion to obtain super-resolution. We showed this by comparing the model output with the ~~motion-free~~ ground truth ~~with no motion~~, using the settings with the Perturbation Network turned on and off. The new detail introduced by the network was also compared with a very-high-resolution image, which was acquired in an idealistic setting ~~with an SNR and resolution~~ obtained solely for ~~the purpose of~~ this experiment and would not have been practically possible in actual patient scans.

As ~~previously~~discussed ~~earlier~~, pre-calculated linear shifts ~~in image~~during imaging provide ~~little or no resolution gain~~negligible gains in spatial resolution. However, the “shifting” ~~that occurs~~observed between subsequent *in-vivo* acquisitions is ~~not non-rigid motion, since the prostate is deformable, and has indeed been observed to suffer different degrees of motion in different regions. In addition, Moreover,~~ local B0 inhomogeneities ~~due to, arising from~~ poor shimming, tissue boundaries, or proximity to ~~the~~ bowel-enclosed gas ~~introduce further, contribute additional~~ non-linear ~~changes~~alterations to the ~~acquired~~ MRI signal. Thus, we ~~believe~~argue that additional information may, in fact, be present in the repeated acquisitions of the same image. ~~Most~~Much of the ~~existing~~ literature on ~~application of~~super-resolution to DWI ~~has been in~~focuses on brain imaging, where motion is ~~typically~~ minimal ~~in comparison~~compared to prostate imaging, ~~and would thus~~ not ~~result in~~leading to significant non-linear changes ~~to in~~ the MRI signal.

~~It should be noted that the~~It’s important to note that phantom experiments (Figures ~~2 and 3~~4) and *in-vivo* experiments (Figures 5-6) are ~~aimed~~designed to ~~demonstrate the~~showcase two ~~different~~distinct features of the proposed model. All kiwifruit images in Figure 2(a-d) ~~have~~share the same resolution and ~~the same number of voxels~~voxel count. This figure ~~is useful to demonstrate~~effectively demonstrates the motion-correction ~~function~~capability of ~~the proposed~~our model, ~~particularly~~especially the role of the Perturbation Network. The Perturbation Network ~~is, therefore,~~aims not only ~~intended for~~correcting for ~~the~~to correct signal mismatches ~~of in~~ different acquisitions of inter- and intra-b-value scans, but also to ~~regain some detail that can be~~recover details potentially lost during ~~naïve~~in naive averaging. ~~On the other hand, Figure 3 exhibits~~Conversely, Figures 5-6 showcases the super-resolution ~~power~~capability of the proposed method ~~in that~~highlighting how the model ~~was,~~ trained ~~solely~~exclusively with ~~the voxels in the LR image and the output is compared~~voxels, compares with the HR image. ~~It can be seen that the~~The model’s estimation of ~~the “intermediate voxels” is considerably better than the “intermediate voxels” is noticeably superior to~~ bi-cubic interpolation, ~~due a~~ difference attributed to the fact that the ~~proposed~~our method ~~is a global function whereas interpolation is merely a~~functioning globally, in contrast to interpolation’s local operation. The quantitative structural similarity results in Table-1 ~~also~~ verify ~~further~~substantiate this observation.

The work proposed in this paper conforms to the recent terminology in the super-resolution literature, referred to as “real image super-resolution. Many super-resolution algorithms were primarily developed on synthesized datasets that consist of low resolution inputs generated by degrading a high resolution image. The “real image super-resolution” terminology was coined to emphasize the fact that such approaches with synthesized datasets fail to reflect the complexity of “naturally occurring” low-resolution inputs. This was the primary reason that we used a methodology that does not depend on synthetically degraded inputs.



Naturally, this study has certain constraints and limitations that should be acknowledged. For example, we selected bi-cubic interpolation as a competitor for evaluation, whereas other studies employing synthetically generated low-resolution inputs could compare their approaches directly. However, it is important to note that the challenge we undertook in the context of "real image super-resolution" does not allow for a uniform or universally accepted ground truth. If we had acquired images with smaller voxel sizes, it would have resulted in lower SNR, or longer acquisition times at lower voxel sizes would have introduced variations in motion exposure. As a result, direct comparisons or supervised training with other CNN-based super-resolution methods would not have been feasible. In essence, we intentionally adopted this self-supervised setting to address these challenges.

Key advantages of the proposed model over the supervised deep learning-based methods are enumerated below:

- **Input/Output:** Supervised methods ~~depend~~rely on ~~operations on the processing~~ low-resolution ~~image as input~~applying, utilizing local kernel operations. At the output, ~~the model expects these models expect~~ the high-resolution ~~version~~versions of the ~~image~~images. In contrast, the INR method takes only the voxel ~~locations~~  $(x,y,z)$  as input and emit the voxel intensity that corresponds to the input location.
- **Resolution Factor:** Supervised models are trained with predesigned low-~~resolution~~/high-resolution pairs and learn to increase the resolution by the factor dictated in the difference of high- and low-resolution training images. INR models the continuous space and can provide improvement for any factor, **even non-integer ones**.
- **Training Set:** Supervised models typically require a large training set of input/output pairs. Deep models that are trained with fewer images than required possess the risk of hallucination, i.e., carrying some information/detail seen on the training set to a test image. INR doesn't need a training set of images from other patients. It merely uses the existing voxel intensities to infer those that don't.
- **Risk of Artifacts:** ~~Acquisition of~~Acquiring low-~~resolution~~/high-resolution training pairs ~~is impractical in for~~ medical imaging ~~and is impractical, often leading to the typical application is to synthetically generate low resolution samples. However, generating a low resolution image from high resolution is a complex process that involves correctly modeling the kernel function of the imaging technique. Incorrect modeling during the training sets~~synthetic generation ~~process would lead to artifacts in the higher of low-resolution output samples. However, INR doesn't assume~~bypasses such ~~a procedure and complexities, as~~ each image ~~is used to train individually~~trains its own model~~, eliminating the risk of artifact generation due to incorrect modeling.~~
- **Practicality:** ~~When supervised~~Supervised models ~~are,~~ trained ~~with a dataset of a certain for specific~~ imaging ~~setting~~settings (TE, TR, b-value, etc.) ~~they won't work with images of a new setting without.), require~~ fine-tuning. ~~This is important for new settings, posing challenges for inter-institution~~institutional operation and ~~in~~ clinical practice. ~~Conversely, INR doesn't have~~is free from such ~~a constraint, because there's no pre-trained model for inference. A new scan constraints; it can be trained~~train and super-~~resolved~~resolve a new scan within minutes ~~without a pre-existing model for inference.~~

~~The most encouraging~~A particularly promising implication of this study is the potential ~~to reduce for reduced~~ scan times, ~~which is especially beneficial for patients, particularly those who experience experiencing~~ claustrophobia or ~~have those with~~ Parkinson's disease. By employing the proposed model for super-resolution, it becomes possible to obtain high-quality images from ~~low~~lower-resolution inputs,

thus mitigating the need for acquiring a large number of averages needed for obtaining high-resolution images with small voxels and longer scan times, adequate SNR. This has significant benefits in terms of patient comfort and convenience. In addition, shortened echo train durations will help reduce the spatial distortion often present in EPI-based images.

From the clinical perspective, reconstructed coronal DWI images of the prostate, as we proposed in this paper, could be helpful in more accurate localization of the tumor to the appropriate level of apex, midgland or base; and distinguishing a lesion location in the peripheral zone of the prostate from normal periprostatic tissue such as the surgical capsule or the periprostatic plexus by lesion location confirmation with an additional view, or distinguishing lesion location in the central gland, where BPH nodules could mimic cancer, from the peripheral zone, where malignant lesions are more prevalent.