Toward Improving Knee MRI Image Quality for Single Image Super-resolution

Hafssa Mediani[1,*], Salma Tayeb[2], Soufiana Mekouar[3], Mohammed Majid Himmi[4]

[1,2,4]*LIMIARF, Department of Physics, Faculty of Sciences, Mohammed V University in Rabat, Rabat, Morocco*

[3]*Department of Geomorphology and Geomatics, Scientific Institute, Mohammed V University in Rabat, Rabat, Morocco*

(Received: March 20, 2025; Revised: June 12, 2025; Accepted: September 15, 2025; Available online: October 04, 2025)

Abstract

This study aims to improve image super-resolution techniques by balancing distortion reduction with perceptual quality improvement. It introduces a new framework called Toward Improving Super-Resolution, which focuses on producing high-quality knee magnetic resonance images. The framework uses a lightweight, resolution-independent, feedforward convolutional network with 266,000 parameters, which includes smoothing and denoising preprocessing and Leaky Rectified Linear Unit activations for stable training. The model builds on a baseline deep learning architecture to improve training stability and visual quality while maintaining computational efficiency. The fast Magnetic Resonance Imaging knee dataset was compared to established super-resolution methods like Super-Resolution Convolutional Neural Network, Very Deep Super-Resolution, and Enhanced Deep Super-Resolution. Toward Improving Super-Resolution achieved a peak signal-to-noise ratio of 38.405 ± 0.129 decibels and a structural similarity index of 0.9815 ± 0.0021, surpassing Super-Resolution Convolutional Neural Network, Very Deep Super-Resolution, and Enhanced Deep Super-Resolution. It maintained high performance at scales three and four, demonstrating accuracy and statistical robustness. The study shows that the proposed framework can enhance diagnostic imaging, reduce the need for repeated scans, and speed up clinical decision-making.

Keywords: Magnetic Resonance Imaging (MRI), Super-Resolution (SR), Deep Neural Network, Low Resolution (LR), High Resolution (HR)

1. Introduction

Image SR enhances image resolution, transforming blurred and Low-Resolution (LR) images into clear, HR versions. It improves visual perception and detail and can be achieved using either single or multiple images. HR images are crucial for various applications, including object detection [1], face recognition [2], medical diagnostics [3], remote sensing [4], astronomy [5], and forensics [6]. They are vital for accurate disease diagnosis, enabling the detection of small structures [2] and pathologies, like microvasculature changes around tumors [7] and subtle soft exudates in retinal diseases [8].

Even with advancements in imaging technology, many medical images still suffer from poor spatial resolution because of equipment and imaging limitations. Identifying small anatomical features and early disease detection is challenging. Furthermore, methods like X-ray and Computed Tomography (CT) scans expose patients to radiation, which carries risks [9]. Reducing radiation doses can increase noise and further diminish image quality. MRI is preferred over CT scans because it uses non-ionizing radiation, making it safer for repeated diagnostic imaging [10]. However, MRI scans often have lower resolution and longer acquisition times, which can limit their clinical usefulness, especially in imaging complex regions like the knee.

Various interpolation-based methods have been used to enhance medical image resolution, but they often compromise edge sharpness and contrast. SR techniques improve image quality and assist in disease diagnosis, yet they remain challenging due to their ill-posed nature and complexity. Traditional performance metrics might not fully correspond with human perception.

***Corresponding author: Hafssa Mediani (hafssamedianni@gmail.com)**
DOI: https://doi.org/10.47738/jads.v6i4.926

The Super-Resolution Convolutional Neural Network (SRCNN) [11] model successfully tackles these challenges and demonstrates strong performance. As a groundbreaking deep learning method for single-image super-resolution, it improves LR images by learning end-to-end mapping to high resolution versions.

The main advantages of SRCNN are its simplicity and effectiveness. With only three convolutional layers, SRCNN is easy to understand and implement. Despite its simplicity, it significantly outperforms traditional methods in image SR tasks.

SRCNN laid the groundwork for deep learning-based SR but has limitations due to its shallow depth. This restricts its ability to capture complex patterns and textures. Subsequent models like Very Deep Super Resolution (VDSR) [12] and Enhanced Deep Super-Resolution (EDSR) [13] have addressed these issues with deeper architectures, residual connections, and adversarial training.

However, both VDSR and EDSR have limitations when applied to knee MRI data. VDSR's deep 20-layer architecture complicates training and requires significant computational resources, making it difficult to deploy in real-time or resource-limited clinical environments. EDSR, while removing batch normalization layers to improve performance, remains large and computationally intensive. Additionally, both models may not effectively address the low signal-to-noise ratio and specific texture characteristics of knee MRI images.

To address these issues, we build on the SRCNN architecture, aiming to maintain its simplicity while improving accuracy. Inspired by SRCNN's effectiveness, we keep the model non-complex to simplify training. To balance performance and efficiency, we designed a lightweight, fully convolutional feedforward network comprising four layers without skip connections or pooling. We optimize the model through careful analysis and modifications, using the appropriate loss function. Experimental results show that the modified model produces better outcomes.

Our method uses bicubic interpolation to upscale LR images before inputting them into our model. We also prepare MRI data by applying filters and blurs during preprocessing to improve data quality and ensure it is suitable for the model. We tested our model on the new fastMRI dataset from NYU Langone, focusing on knee single-coil MRI images. The SR network demonstrated state-of-the-art performance in enhancing the quality of these diagnostic images, as measured by Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM).

2. Related Works

The concept of SR techniques is converting LR images into HR images to obtain clear images with high visual quality. It is crucial for applications of medical imaging, as accurate diagnosis and treatment planning require HR images that are easier to analyze and better interpret.

Learning-based methods use knowledge from an image database to achieve SR. They can learn complex mappings from LR to HR images. They generally outperform traditional methods in producing superior results due to their ability to leverage machine learning approaches. The three major learning-based SR techniques are neighbor embedding (NE), sparse coding, and deep learning approaches. NE assumes small patches of LR and HR make low-dimensional non-linear manifolds with identical local geometry. Uses locally linear embedding (LLE) [14] to create HR patches [15]. Includes techniques like non-negative NE and least square approximation and Dual Geometric Neighbor Embedding (DGNE) [16].

Sparse Coding HR properties are retrieved by considering that LR and HR characteristics share similar sparse components of reconstruction [17], [18]. Techniques like orthogonal matching pursuit (OMP) and Principal Component Analysis (PCA) are used to reduce the dimensionality of LR characteristics [19], [20]. Methods include combining manifold regularization with sparse support regression [21], and Graph Embedding Super-Resolution (GESR) [22]. Deep Learning Approaches utilize neural networks for representation learning to directly map the relationship between input and output from data. These methods are data-driven and recover essential features for desired SR [23], [24].

Linear Connections are a type of neural network architecture consisting of a basic configuration characterized by a straightforward design that includes only one signal flow path without crossing multiple links or branches. The input flows directly from the first layer to the next subsequent layers sequentially, in which the multiple convolutional layers

are stacked on top of each other. Certain linear networks are learned to replicate residual images, which is the difference between LR images and HR images, and are classified as linear networks because of their structure [25], [26], [27].

We mention below the most prominent methods of linear connections. Firstly, SRCNN uses sequential convolutional layers and relies on pre-upsampled input, typically obtained using traditional interpolation methods [28], [29]. Besides, VDSR is a deep Convolutional Neural Networks (CNN) structure with 20 weighted layers, that learns residual mapping to provide a distinction between LR and HR images [12]. In addition, Denoising Convolutional Neural Network (DnCNN) Learns to directly forecast high-frequency residuals using convolutional, batch normalization, and Rectified Linear Unit (ReLU) layers [26]. Furthermore, Image Restoration Convolutional Neural Network (IRCNN) utilizes a series of CNN-based denoisers for image deblurring, denoising, and SR tasks [27]. Also, Fast Super Resolution Convolutional Neural Network (FSRCNN) enhances speed and performance over SRCNN with a basic design of four convolutional layers and one deconvolutional layer [30]. Another method, Enhanced Burst Super Resolution (EBSR) separates the multi-frame SR problem into alignment, fusion, and reconstruction parts using modules like FEPCD and CNLF [31]. Meanwhile, Wavelet Multiscale Convolutional Neural Network (WMCNN) trains CNNs separately to approximate wavelet multiscale representations for HR aerial image reconstruction [32]. Similarly, Efficient Subpixel Convolutional Neural Network (ESPCN) extracts feature information directly from LR images and generates HR images using a sub-pixel convolution layer [33]. Finally, Efficient Video Super-Resolution Network (EVSRNet) uses neural architecture search for real-time video super-resolution, balancing quality and efficiency [34].

Despite advances in these methods, applying them directly to knee MRI images presents challenges due to the unique noise, texture, and anatomical features of this imaging modality. Specifically, deep models like VDSR and EDSR, while achieving high performance, are computationally expensive and often require large datasets for effective training, which limits their clinical usability. Our proposed model, Toward Improving Super-Resolution (TISR), addresses these limitations by adopting a lightweight architecture inspired by SRCNN, carefully optimized to balance simplicity and accuracy, making it a promising solution suitable for enhancing knee MRI images in real-world healthcare settings.

3. Methodology

In this section, we outline the architecture of the proposed model. We show the important components of our architecture and the crucial modifications that exhibit improved computational efficiency. The principal steps of our model start with the preprocessing phase, which applies blurs and filters on LR and HR MRI diagnostic images respectively, followed by selecting the appropriate activation function and finalizing the complete version of our model.

3.1. Filters and Blurs in Data Preprocessing

Filtering and blurring are essential preprocessing steps to enhance model performance and robustness in MRI SR tasks figure 1. They facilitate the creation of cleaner and more consistent input data, resulting in improved learning stability and reconstruction accuracy during both training and inference. We use bilateral filtering to reduce noise in HR MRI images while preserving edges and fine anatomical structures. This edge-aware denoising is crucial in SR because it prevents the model from learning blurred or inaccurate representations of important regions. Bilateral filtering combines spatial and intensity information to maintain structural integrity:

$$I_{filtred}(x,y) = \frac{1}{W_p} \sum_{(i,j)\in\Omega} I(i,j) \cdot e^{-\frac{(i-x)^2+(j-y)^2}{2\sigma_d^2}} \cdot e^{-\frac{|I(i,j)-I(x,y)|^2}{2\sigma_r^2}} \quad (1)$$

$$I_{filtred}(x,y) = \frac{1}{W_p} \sum_{(i,j)\in\Omega} I(i,j) \cdot e^{-\frac{(i-x)^2+(j-y)^2}{2\sigma_d^2}} \cdot e^{-\frac{|I(i,j)-I(x,y)|^2}{2\sigma_r^2}} \quad (2)$$

Here, σd and σr control the spatial and intensity sensitivities, ensuring that edge features are preserved—vital for maintaining anatomical accuracy.

 ISSN 2723-6471

Conversely, we apply Gaussian blurring to LR MRI images to eliminate high-frequency noise and reduce irrelevant details. This enables the SR model to focus on reconstructing meaningful structures rather than amplifying noise. The 2D Gaussian kernel is defined as :

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{3}$$

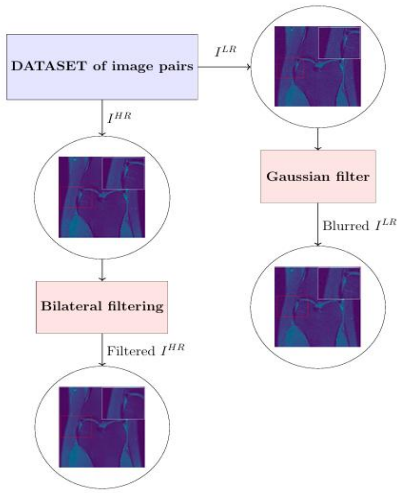

Figure 1. The Data Preprocessing Phase: Apply Bilateral Filtering for the HR MRI Images, and Gaussian Filter for LR MRI Images

The standard deviation σ determines the level of smoothing applied. Together, these filters enhance SR performance by ensuring that both HR targets and LR inputs highlight anatomically relevant content. In early experiments, neglecting these filters delayed training convergence and slightly degraded SSIM results, suggesting their beneficial role. A formal study of ablation is planned for future research, although it is not part of the current paper.

3.2.Activation Function

In the standard SRCNN architecture, ReLU activation is applied after the first and second convolutional layers to introduce non-linearity. This non-linearity allows the network to learn complex mappings from LR to HR images. In our proposed architecture figure 2, we replace ReLU with Leaky Rectified Linear Unit (LeakyReLU) as the activation function. This choice is driven by the nature of MRI data, which often contains low-contrast, subtle tissue gradients. In such regions, ReLU's zero activation for negative inputs can cause neurons to become inactive and result in the loss of important features. LeakyReLU addresses this by allowing a small, non-zero gradient (αx) when $x < 0$, ensuring neurons remain responsive even in low-activation zones. This helps preserve weak but clinically relevant features during reconstruction. The LeakyReLU function is defined as:

$$LeakyReLU\ (x) = \begin{cases} x & if\ x \geq 0 \\ \alpha x & if\ x < 0 \end{cases} \tag{4}$$

α is small constant (e.g., 0.01) that controls the slope for negative inputs.

ISSN 2723-6471

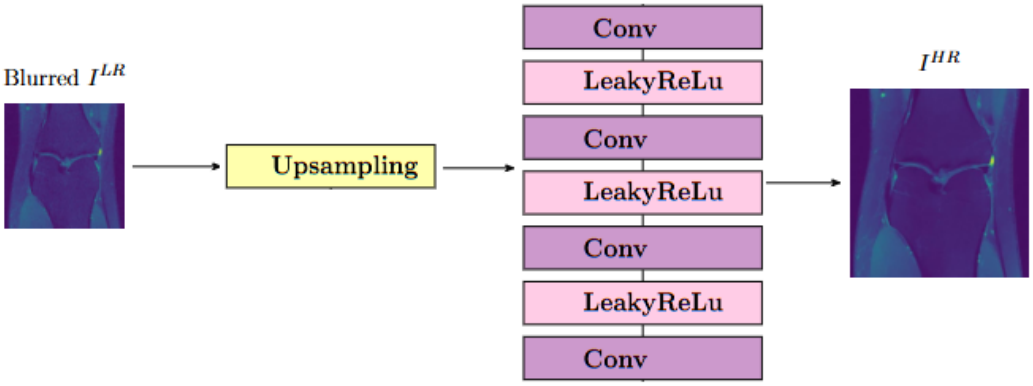

Figure 2. The Architecture of the Proposed SR Network (TISR)

Although ReLU is computationally efficient, it can suffer from the "dying ReLU" problem, where neurons become permanently inactive. In preliminary experiments, replacing ReLU with LeakyReLU in our network improved convergence stability and resulted in slightly higher SSIM and PSNR values on validation MRI slices. This indicates that maintaining gradient flow with LeakyReLU is especially helpful for enhancing structural sensitivity in super-resolving medical images. While formal comparative analysis is planned for future work, these early results support LeakyReLU as a suitable activation function in this context.

3.3.Proposed Model Architecture

Our proposed CNN-based method (TISR) architecture figure 2 consists of a four-layer convolutional neural network, in which LeakyReLU activation is used after the first three convolutional layers to introduce non-linearity into the model as a way to allow the network to capture and learn more complex mappings from LR to HR MRI images. The process begins with applying blurs and filters to our knee MRI dataset of LR and HR image pairs. To smooth and blur LR images, the Gaussian kernel is moved across them through convolution, and then each pixel value is replaced by a weighted average of its neighbors. In addition, reducing noise from the HR images without damaging edges or losing significant diagnostic information makes it useful for preprocessing images before further analysis.

The training phase starts with a blurry LR MRI image as input, which is then resized to the HR volume using an interpolation step as illustrated in stage 1 of figure 3. This initial step ensures that our network operates on an image with the required output dimensions, so the HR network is initialized based on the interpolated estimate of the HR image. It creates a HR mesh that our model can refine and improve. The idea is to increase image size while reducing noise and produce smoother, more attractive results by using a larger area of pixels (4x4 grid) and cubic polynomials to estimate new pixel values. Interpolation techniques like bicubic are known to preserve more image detail compared to simpler resampling approaches.

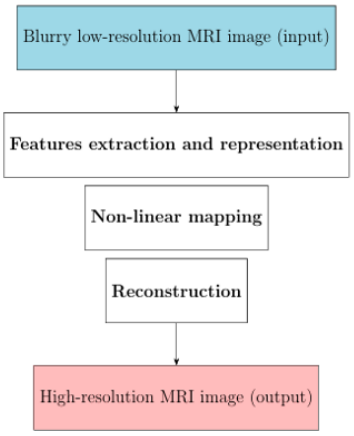

Figure 3. The Principle Stages of Our Proposed Model TISR

As shown in the second stage of figure 3, the TISR model processes the incoming picture using three primary subcomponents: (1) feature extraction and representation to capture spatial and structural elements from the input; (2) nonlinear mapping layers that facilitate the transformation between LR and HR domains; and (3) a reconstruction module that synthesizes the enhanced HR image. This stage exemplifies the fundamental aspects of our architecture and highlights the innovative contributions of our approach. As a result, learning is simplified because TISR has a better starting point for learning the mapping from LR images to HR images. This makes the training process more stable and effective compared to starting from a random or zero-filled HR image. This is followed by passing the interpolated input through the TISR model to produce the final ultra-resolution image.

The training phase is based on using a dataset of LR blurred image pairs and HR filters to train a neural network that optimizes the entire process in one go, learning complex mappings directly from the data. The network is capable of learning through feature representations, cross-layer transformations, and nonlinear activation. The network then directly outputs the HR image based on the learned weights and features. As illustrated in the third stage of Figure3, the output is a HR MRI image characterized by enhanced sharpness, improved anatomical detail, and reduced noise, making it highly suitable for diagnostic applications. Resorting to the use of neural networks has helped us improve our super resolution model and its computational efficiency to enhance the quality of diagnostic images by leveraging their powerful feature extraction and learning capabilities.

3.4.Image reconstruction

Deep learning techniques, especially CNNs, are effective at reducing natural noise and removing noise patterns in MRI super-resolution. Unlike traditional reconstruction methods, our approach uses a specialized architecture that combines multi-scale patch extraction with adaptive, learned feature representations designed to preserve subtle anatomical details. This improves image recovery by focusing on clinically important structures rather than just removing noise.

In our pipeline, the patch extraction and representation stage involves extracting overlapping patches from the LR image, which are then encoded into high-dimensional feature vectors using learnable convolutional filters tailored for MRI data. This adaptive encoding differs from standard fixed patch processing by dynamically emphasizing edges and textures critical for diagnosis. After feature extraction, a non-linear mapping process uses cascaded convolutional layers with residual connections to transform LR patch representations into HR feature vectors. This step goes beyond typical nonlinear mapping by explicitly adding domain-specific constraints and attention mechanisms to improve detail reconstruction in ambiguous areas.

The reconstruction stage combines the HR patches through a weighted averaging method that considers patch overlap and preserves spatial consistency across the entire image. Even without an extra refinement layer, the design of the nonlinear mapping and reconstruction steps is carefully planned to minimize boundary artifacts and maintain overall structural coherence. Overall, our image reconstruction process introduces a novel integration of adaptive patch-based feature extraction, domain-informed nonlinear mapping, and overlap-aware reconstruction, which together enhance the fidelity and diagnostic quality of super-resolved MRI images.

4. Results and Discussion

4.1. Dataset

The Center for Advanced Imaging Innovation and Research (CAI2R) at NYU School of Medicine and NYU Langone Health develops and implements innovative imaging techniques to enhance human health, with a focus on rapid image acquisition and advanced image reconstruction for improved disease understanding and patient care.

In this study, we used the fastMRI [35], [36] Knee Single-Coil dataset created by NYU Langone, which contains raw k-space data and corresponding image reconstructions. Specifically, we extracted 30,000 training images, 7,500 validation images, and 3,000 test images, all in 2D format, from fully sampled 3D knee MRI volumes. Each file represents one knee scan and includes two reconstructed image types: reconstruction_esc (from single-coil simulation) and reconstruction_rss (from multi-coil RSS reconstruction), both generated from the same k-space data.

To create LR and HR training pairs, we used reconstruction_esc as the input because of its Lower Signal-To-Noise Ratio (SNR) and more artifacts, and reconstruction_rss as the HR target due to its improved SNR through Root Sum-of-squares coil combination. To improve the contrast between the input and target, a Gaussian blur was applied to the reconstruction_esc images, and a denoising filter was applied to the reconstruction_rss images.

The dataset contains grayscale (single-channel) 2D MRI slices. No color normalization or data augmentation was used, as anatomical consistency is maintained within each scan and the images are inherently grayscale. All data were de-identified and curated by the fastMRI team for research purposes only and are not intended for clinical diagnosis.

4.2. Training details

The algorithm was trained on approximately 30,000 images from the FastMRI dataset, which provides high-quality diagnostic images of various scenes and tissues suitable for our task. We're working with about 700 H5 files, each containing between 30 and 38 slides or frames. Each slide is a 320 x 320-pixel image. As shown in table 1, The blurred LR images are fed to the model, which are then resized by a factor of 4 prior to processing. The network employs a fully convolutional architecture at scales 2, 3, or 4 with four convolutional layers. The first layer features a 3×3 kernel, 128 filters, and padding of 1. Two middle layers use 9×9 kernels with padding of 4, while the last layer has a 3×3 kernel and padding of 1. All layers have a stride of 1 and symmetric padding to maintain consistent spatial resolution. To promote stable training and minimize vanishing gradient issues, LeakyReLU activations ($\alpha = 0.01$, `inplace=False`) are applied after each convolution. The architecture is purely feedforward, with no pooling layers, normalization, or skip connections, making it a lightweight model with approximately 266,000 trainable parameters. The model was trained using the Adam optimizer with a learning rate of 1e-4, a batch size of 16, over 30 epochs, employing 8 parallel data-loading workers. Reproducibility was ensured by setting a fixed random seed of 123, and the architecture remains resolution-independent while maintaining accuracy for 320×320 input images. Its design is flexible, supporting different input and output channel sizes (`num_channels`) for various imaging modalities.

Table 1. Describing the Components and Details of TISR Algorithm

Algorithm	EDSR	VDSR	SRCNN	TISR (our)
Architecture	SRResNet	VGG-net	End-to-end mapping	End-to-end mapping
Datasets	CDIV2K BSD100 Urban100	MatConvNet	91 images ImageNet	FastMRI
Input	LR	bicubic	bicubic	Filters and Blurs + bicubic
Activation function	-	ReLU	ReLU	LeakyReLU
Filters	256	64	64	128
Filter size	5 x 5	3 × 3	9 × 9, 5 × 5	3 x 3, 9 x 9
Number of layers	65	20	3	4
Epochs	5	10, 80	20	30
Performance matrices	PSNR, SSIM	PSNR, SSIM	PSNR, SSIM	PSNR, SSIM
Scale Factors	2,3,4	2,3,4	2,3,4	2 ,3, 4

The original datasets for each approach are also displayed in table 1. HR natural image datasets (DIV2K, BSD100, and Urban100) that offer a range of textural and structural characteristics were used to train EDSR. The MatConvNet benchmark, which includes classic LR datasets like Set5 and Set14, served as the foundation for VDSR. Due to its small variety of natural scenes, the 91-image ImageNet compact dataset was frequently utilized in early SR research. SRCNN employed this dataset.

We used metrics of performance such as PSNR and SSIM to evaluate the quality of the super resolution images. In addition, we used the cross-validation method to evaluate TISR performance. The database comprised about 800 images and was divided into 10 sections (10-fold). This method enabled us to verify the robustness and effectiveness of our model, as this method provided results that were very close to the results achieved without using it figure 4.

ISSN 2723-6471

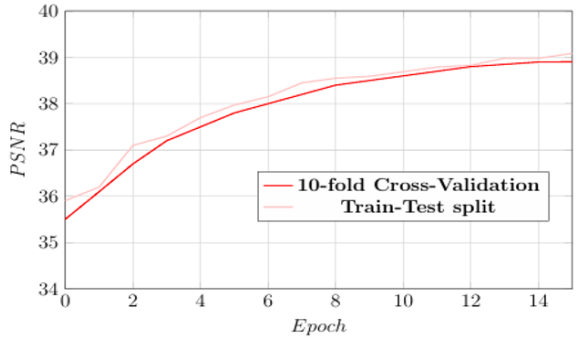

Figure 4. Comparative Evaluation of Model Performance: Analyzing Results of PSNR from Train-Test Split and Cross-Validation Approaches

4.3. Utilized Software

We used HP Proliant DL380 G7 to apply our experiments. It boasts 96.0Gio of memory and an Intel® Xeon (R) CPU E5620 @ 2.40GH × 16 CPU, ensuring powerful and efficient processing. It includes llvmpipe (LLVM 15.0.7,128bits) graphics card for superior graphical performance. The storage capacity includes a 2.7 To, permitting for quick access and ample space. The operating system is Ubuntu 22.04.3 LTS with the GNOME 42.9 interface for a contemporary and consistent user experience.

4.4. Experiments

We take both quantitative and qualitative aspects when evaluating the performance of the proposed TISR model, against existing state-of-the-art SR methods applied to FastMRI images, whose implementations are publicly available codes provided by their authors. As shown in table 2, the proposed TISR achieves the highest scores in most evaluation matrices. It achieves a PSNR of 38.405 decibels (dB) at scale 2 and an SSIM of 0.9815, superior to other algorithms, of which SRCNN has a PSNR of 30.575 and an SSIM of 0.7362 dB, and VDSR has a PSNR of 32.092 and an SSIM of 0.8073 dB. Also, EDSR has a PSNR of 32.407 and an SSIM of 0.8314 dB. Additionally, on a scale 3 it outperforms other architectures with a PSNR of 35.935 dB and an SSIM of 0.9356. Finally, it achieved the highest PSNR (scale 4) of 34.175 dB and an SSIM of 0.9037.

Table 2. Comparison Between Our Proposed Model with the State-Of-The-Art

Scale	EDSR	VDSR	SRCNN	**TISR** (our)
2	32.407 / 0.8314	32.092 / 0.8073	30.575 / 0.7362	38.405 / 0.9815
3	31.604 / 0.8024	30.432 / 0.7563	29.108 / 0.7315	35.935 / 0.9356
4	29.543 / 0.7813	28.071 / 0.7326	26.051 0.7084	34.175 / 0.9037

To provide greater statistical rigor, we report the standard deviation of PSNR and SSIM metrics across five independent runs with different random seeds. For the scale 2, the mean PSNR and SSIM of TISR are 38.405 ± 0.129 dB and 0.9815 ± 0.0021, respectively. This low variance indicates stable and consistent performance. Similar patterns are observed for scales 3 and 4, with standard deviations not exceeding 0.15 dB for PSNR and 0.0025 for SSIM, confirming the statistical reliability of the results.

From figure 5 and figure 6, it can be observed that good results are obtained by our proposed model among all previous algorithms, in which the PSNR starts to rise from the first epoch and continues to advance, unlike other methods where the PSNR values begin to stabilize upon reaching the fifth epoch of the PSNR index and to the sixth epoch of the SSIM indicator.

 ISSN 2723-6471

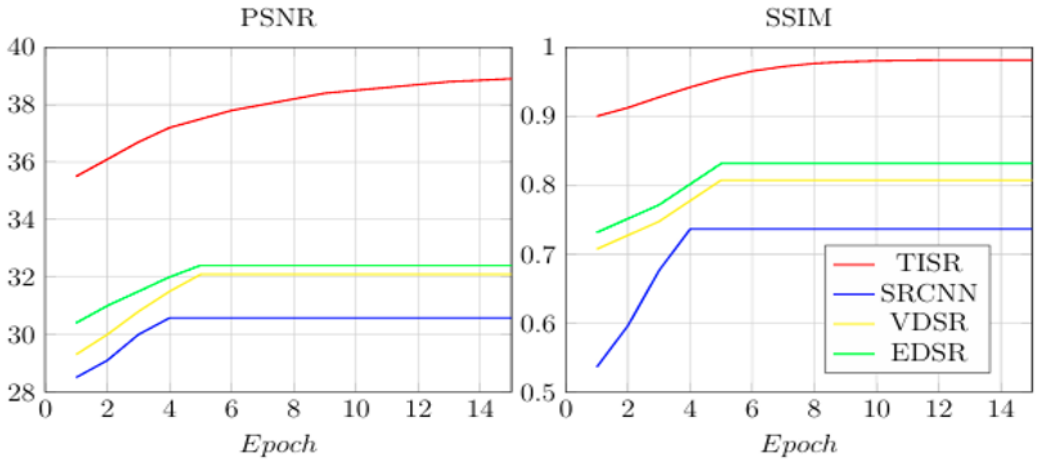

Figure 5. Comparisons between the different models used in our study using FastMRI as a dataset for PSNR and SSIM

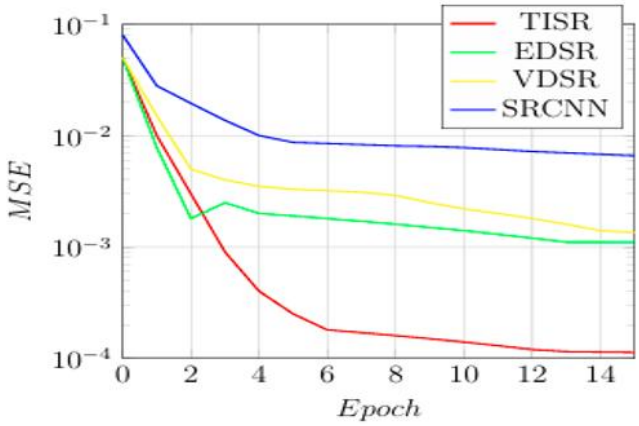

Figure 6. Model Comparisons using the FastMRI Dataset for MSE

Moreover, the low Mean Squared Error (MSE) values obtained by TISR indicate minimal error between the original images and the reconstructed images. These results, which are corroborated by mean performance indicators and the variances that go along with them, strengthen the validity of the enhancements made by our approach and lessen the possibility that the gains were the result of random chance. As illustrated in figure 7, the comparative analysis demonstrates that our model consistently outperforms other super-resolution approaches at a scale factor of 4.

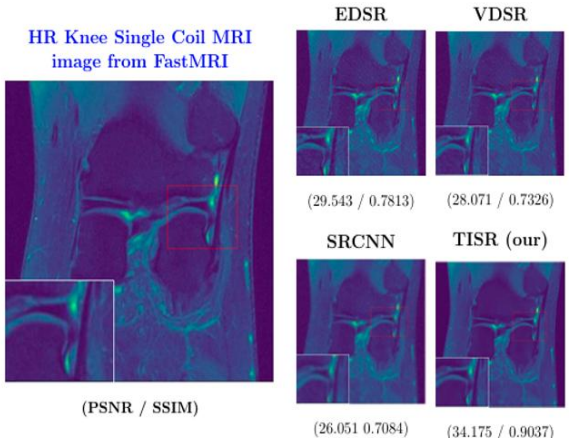

Figure 7. Comparative Analysis of Our Model with other Approaches to SR (scale 4)

5. Discussion

These good numerical measurements indicate that our proposed algorithm produces sharper images with less distortion, low noise, and high resolution, which are close to the original images compared to the images produced by state-of-the-art HR methods. This indicates that the model effectively maintains structural details, contrast, and textures essential in clinical imaging. This is evidenced by the qualitative comparison in figure 7, where the reconstructions from SRCNN, VDSR, and EDSR exhibit considerable blurring, reduced contrast, and a loss of intricate anatomical details, as indicated by their lower PSNR and SSIM values. Conversely, the proposed TISR method attains the highest PSNR and SSIM, resulting in sharper edges and a more accurate texture representation. This quantitative advantage highlights the high fidelity and strong resemblance between the original image and the one reconstructed by our model.

TISR tends to reduce noise and artifacts while preserving anatomical boundaries, leading to cleaner reconstructions. However, although visual quality is important, diagnostic accuracy remains the main focus in medical imaging. Therefore, our priority is on structural fidelity and clinical interpretability rather than aesthetics alone. Despite strong performance, some limitations persist. In certain cases, especially with very low-quality or aliased inputs, minor artifacts or over-smoothing may occur, possibly affecting fine anatomical details. Although such issues were rare in our evaluation, they underscore the need for careful clinical use and further validation. A thorough error analysis and visual inspection showed that hallucinations were minimal but more likely in highly degraded or ambiguous regions.

6. Conclusions

We have designed a new deep-learning approach to enhance a knee MRI image for single-image super-resolution. Our method features an uncomplicated model to simplify training while improving accuracy. Our proposed CNN-based method (TISR) introduces non-linearity to allow the network to capture and learn more complex mappings from LR to HR MRI images. We also apply filters and blurs during preprocessing to improve data quality and ensure fit to the model.

Based on quantitative and qualitative analyses, TISR has outperformed current state-of-the-art methods. However, we acknowledge that the present study does not include clinical assessment or radiologist-based evaluation. Therefore, our results, though promising, should be viewed as preliminary from a clinical perspective. We plan to incorporate radiologist-led evaluations in future work to more accurately assess diagnostic value.

We believe in enhancing our models by exploring additional filters and using different training strategies. Due to its simplicity, robustness, and ability to generate smooth, flawless images, our method is specifically useful for applications that prioritize visual aesthetics, whether related to medical or other images. Nevertheless, in medical imaging applications, we stress the importance of thorough clinical validation before deployment.

7. Declarations

7.1. Author Contributions

Conceptualization: H.M., S.T.; Methodology: H.M., S.M.; Software: H.M.; Validation: S.T., M.M.H.; Formal Analysis: H.M.; Investigation: H.M.; Resources: S.T., M.M.H.; Data Curation: H.M.; Writing – Original Draft Preparation: H.M.; Writing – Review and Editing: S.T., S.M., M.M.H.; Visualization: H.M.; All authors have read and agreed to the published version of the manuscript.

7.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

7.3. Funding

The authors carried out all aspects of this research independently and have no acknowledgments to declare. This study did not involve any procedures that required ethical approval. Furthermore, this research did not receive any specific funding. The authors declare no competing interests.

7.4. Institutional Review Board Statement

Not applicable.

7.5. Informed Consent Statement

Not applicable.

7.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-based convolutional networks for accurate object detection and segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 1, pp. 142–158, 2016, doi: 10.1109/TPAMI.2015.2437384.

[2] Y. Bai, Y. Zhang, M. Ding, and B. Ghanem, "SOD-MTGAN: Small object detection via multi-task generative adversarial network," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, vol. 2018, no. 1, pp. 1-9, 2018.

[3] H. Greenspan, "Super-resolution in medical imaging," *Comput. J.*, vol. 52, no. 1, pp. 43–63, 2008, doi: 10.1093/comjnl/bxm075.

[4] T. M. Lillesand, R. W. Kiefer, and J. W. Chipman, Remote Sensing and Image Interpretation, 7th ed. Hoboken, NJ, USA: Wiley, 2015.

[5] A. P. Lobanov, "Resolution limits in astronomical images," *arXiv preprint arXiv:astro-ph/0503225*, vol. 2005, no. 1, pp. 1-18, 2005.

[6] A. Swaminathan, M. Wu, and K. R. Liu, "Digital image forensics via intrinsic fingerprints," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 101–117, 2008, doi: 10.1109/TIFS.2007.916010.

[7] F. Lin, J. D. Rojas, and P. A. Dayton, "Super resolution contrast ultrasound imaging: Analysis of imaging resolution and application to imaging tumor angiogenesis," in Proc. *IEEE Int. Ultrasonics Symp. (IUS)*, vol. 2016, no. 1, pp. 1–4, 2016, doi: 10.1109/ULTSYM.2016.7728735.

[8] D. Mahapatra, B. Bozorgtabar, and R. Garnavi, "Image super-resolution using progressive generative adversarial networks for medical image analysis," *Comput. Med. Imaging Graph.*, vol. 71, no. 1, pp. 30–39, 2019, doi: 10.1016/j.compmedimag.2018.10.005.

[9] D. J. Brenner and E. J. Hall, "Computed tomography — an increasing source of radiation exposure," *N. Engl. J. Med.*, vol. 357, no. 22, pp. 2277–2284, 2007, doi: 10.1056/NEJMra072149.

[10] R. Smith-Bindman, P. W. Chu, H. A. Firdaus, C. Stewart, M. Malekhedayat, S. Alber, and D. L. Miglioretti, "Projected lifetime cancer risks from current computed tomography imaging in the US," *JAMA Intern. Med.*, vol. 185, no. 6, pp. 710–719, 2025, doi: 10.1001/jamainternmed.2025.0505.

[11] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *arXiv* preprint arXiv:1501.00092, vol. 2015, no. 1, pp. 1-12, 2015.

[12] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in Proc. *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2016, no. 1, pp. 1646–1654, 2016, doi: 10.1109/CVPR.2016.182.

[13] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," *arXiv* preprint arXiv:1707.02921, vol. 2017, no. 1, pp. 1-21, 2017.

[14] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000, doi: 10.1126/science.290.5500.2323.

[15] H. Chang, D.-Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in Proc. *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 1, no. 1, pp. I–I, 2004doi: 10.1109/CVPR.2004.1315043.

[16] S. Yang, Z. Wang, L. Zhang, and M. Wang, "Dual-geometric neighbor embedding for image super resolution with sparse tensor," *IEEE Trans. Image Process.*, vol. 23, no. 7, pp. 2793–2803, 2014, doi: 10.1109/TIP.2014.2319742.

[17] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861–2873, 2010, doi: 10.1109/TIP.2010.2050625.

[18] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in Proc. *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR),* vol. 2008, no. 1, pp. 1–8, 2008, doi: 10.1109/CVPR.2008.4587647.

[19] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse representations," in *Lect. Notes Comput. Sci*., vol. 6920, no. 1, pp. 711–730, 2012, doi: 10.1007/978-3-642-27413-8_47.

[20] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007, doi: 10.1109/TIT.2007.909108.

[21] J. Jiang, R. Hu, Z. Han, K. Huang, and T. Lu, "Efficient single image super-resolution via graph embedding," in Proc. *IEEE Int. Conf. Multimedia Expo (ICME)*, vol. 2012, no. 1, pp. 610–615, 2012, doi: 10.1109/ICME.2012.102.

[22] J. Jiang, R. Hu, Z. Wang, Z. Han, and S. Dong, "Manifold regularized sparse support regression for single image super-resolution," in Proc. *IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 2013, no. 1, pp. 1429–1433, 2013, doi: 10.1109/ICASSP.2013.6637887.

[23] S. Anwar, S. Khan, and N. Barnes, "A deep journey into super-resolution: A survey," *ACM Comput. Surv*., vol. 53, no. 3, pp. 1-12, May 2020, doi: 10.1145/3390462.

[24] Z. Wang, J. Chen, and S. C. H. Hoi, "Deep learning for image super-resolution: A survey," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 43, no. 10, pp. 3365–3387, 2021, doi: 10.1109/TPAMI.2020.2982166.

[25] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," *arXiv* preprint arXiv:1511.04587, vol. 2016, no. 1, pp. 1-12, 2016.

[26] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process*., vol. 26, no. 7, pp. 3142–3155, 2017, doi: 10.1109/TIP.2017.2662206.

[27] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep CNN denoiser prior for image restoration," *arXiv* preprint arXiv:1704.03264, vol. 2017, no. 1, pp. 1-12, 2017.

[28] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a deep convolutional network for image super-resolution," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, vol. 2014, no. 1, pp. 1-8, 2014.

[29] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell*., vol. 38, no. 2, pp. 295–307, 2016, doi: 10.1109/TPAMI.2015.2439281.

[30] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," *arXiv* preprint arXiv:1608.00367, vol. 2016, no. 1, pp. 1-12, 2016.

[31] Z. Luo, L. Yu, X. Mo, Y. Li, L. Jia, H. Fan, J. Sun, and S. Liu, "EBSR: Feature enhanced burst super-resolution with deformable alignment," in Proc. *IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, vol. 2021, no. 1 +, pp. 471–478, 2021, doi: 10.1109/CVPRW53098.2021.00058.

[32] T. Wang, W. Sun, H. Qi, and P. Ren, "Aerial image super resolution via wavelet multiscale convolutional neural networks," *IEEE Geosci. Remote Sens. Lett*., vol. 15, no. 5, pp. 769–773, 2018, doi: 10.1109/LGRS.2018.2810893.

[33] W. Shi, J. Caballero, F. Huszár, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network," *arXiv* preprint arXiv:1609.05158, vol. 2016, no. 1, pp. 1-12, 2016.

[34] S. Liu et al., "EVSRNet: Efficient video super-resolution with neural architecture search," in Proc. *IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, vol. 2021, no. 1, pp. 2480–2485, 2021, doi: 10.1109/CVPRW53098.2021.00281.

[35] J. Zbontar et al., "fastMRI: An open dataset and benchmarks for accelerated MRI," *arXiv* preprint arXiv:1811.08839, vol. 2019, no. 1, pp. 1-19, 2019.

[36] F. Knoll, "fastMRI: A publicly available raw k-space and DICOM dataset of knee images for accelerated MR image reconstruction using machine learning," *Radiol. Artif. Intell*., vol. 2, no. 1, p. 1-17, 2020, doi: 10.1148/ryai.2020190007