Article

Enhancing the Super-Resolution of Medical Images: Introducing the Deep Residual Feature Distillation Channel Attention Network for Optimized Performance and Efficiency

Sabina Umirzakova *, Sevara Mardieva, Shakhnoza Muksimova [ID], Shabir Ahmad [ID] and Taegkeun Whangbo *

Department of Computer Engineering, Gachon University, Sujeong-gu, Seongnam-si 113-120, Gyonggi-do, Republic of Korea; sevara1998@gachon.ac.kr (S.M.)
* Correspondence: sabinatuit@gachon.ac.kr (S.U.); tkwhangbo@gachon.ac.kr (T.W.)

Abstract: In the advancement of medical image super-resolution (SR), the Deep Residual Feature Distillation Channel Attention Network (DRFDCAN) marks a significant step forward. This work presents DRFDCAN, a model that innovates traditional SR approaches by introducing a channel attention block that is tailored for high-frequency features—crucial for the nuanced details in medical diagnostics—while streamlining the network structure for enhanced computational efficiency. DRFD-CAN's architecture adopts a residual-within-residual design to facilitate faster inference and reduce memory demands without compromising the integrity of the image reconstruction. This design strategy, combined with an innovative feature extraction method that emphasizes the utility of the initial layer features, allows for improved image clarity and is particularly effective in optimizing the peak signal-to-noise ratio (PSNR). The proposed work redefines efficiency in SR models, outperforming established frameworks like RFDN by improving model compactness and accelerating inference. The meticulous crafting of a feature extractor that effectively captures edge and texture information exemplifies the model's capacity to render detailed images, necessary for accurate medical analysis. The implications of this study are two-fold: it presents a viable solution for deploying SR technology in real-time medical applications, and it sets a precedent for future models that address the delicate balance between computational efficiency and high-fidelity image reconstruction. This balance is paramount in medical applications where the clarity of images can significantly influence diagnostic outcomes. The DRFDCAN model thus stands as a transformative contribution to the field of medical image super-resolution.

Keywords: medical image super-resolution; RFDN; memory; channel attention

1. Introduction

Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) scanners offer non-invasive methods for creating detailed images of various body organs and parts. These images are commonly used in medical practice to detect and address issues ranging from tumors to internal bleeding. The process of capturing these images can be hindered by factors such as the capabilities of the equipment, the environment, and costs. Inaccurate or low-quality images can affect the assessments made by healthcare experts and computer-assisted systems. In the imaging of the eye's fundus, certain abnormalities, like microaneurysms or hemorrhages, can occupy very minute areas. Additionally, some elements, such as soft exudates and specific growths, may not be easily visible. Therefore, improving the resolution of these medical images is vital. The quality of these medical images plays a significant role in making accurate clinical decisions. Preliminary processing is essential to optimize the performance of Computer-Aided Diagnosis (CAD) systems. It helps improve the quality, reduce noise, and enhance the image contrast. One challenge faced in medical imaging is the often low resolution (LR) of images due to limitations in equipment and time constraints during capture. To overcome this, researchers have

developed various techniques to enhance the image resolution. The goal of SR is to produce a high-resolution (HR) image from its LR version. These techniques span from interpolation methods to reconstruction and learning-based methods. Recently, deep learning techniques have been at the forefront, delivering promising results in multiple medical imaging areas while preserving essential details. The advancement of deep learning algorithms has led Single Image Super-Resolution (SISR) techniques using neural networks to outperform earlier methods like interpolation-based, reconstruction-based, and learning-based approaches on natural images [**?**]. However, adapting deep neural networks to an enhanced resolution in medical images remains a challenge. In clinical settings, these enhanced images are further used for medical analysis, and the available datasets are often limited in size [**?**]. As a result, enhancing the resolution in medical imagery demands innovative approaches. This includes adjusting training datasets, revising loss functions, tweaking evaluation metrics, and rethinking neural network designs, all to retain critical details and spotlight key structures for medical experts [**?**]. CNNs and generative adversarial networks (GANs) have recently shown promise in enhancing the resolution of medical images [**?**]. However, integrating vision transformers (ViTs), which have demonstrated top-tier performance in natural image refinement and medical image analysis [**?**], remains a challenge. It is important to delve into the strengths, potential, efficiency, and constraints of using ViTs for medical image SR. Additionally, well-established CNN design techniques, like localization processes, residual links, and feature blending, can potentially be incorporated into combined CNN-ViT models to enhance the results. For instance, drawing from CNNs' shared weights and localization functions, a modified ViT known as the Swin Transformer has been introduced for advanced image tasks [**?**]. These innovative Swin layers have been utilized in tasks related to natural image refinement [**?**] and delineation [**?**].

Leveraging prior insights from related medical imaging tasks, like segmentation, can enhance SR endeavors. On one side, hurdles remain in assessing the quality of enhanced images, particularly in the medical realm [**?**]. Typically, the quality assessment of enhanced natural images covers the reconstruction accuracy and human perception. With SISR techniques nearing the confines of signal fidelity metrics, the importance of perceptual quality measures has grown [**?**]. However, medical images present unique challenges, having artifacts primarily stemming from the imaging system hardware and patient movement—issues seldom found in natural images. Metrics like PSNR and structural similarity (SSIM) are commonly used in medical image SR research. Still, relying solely on quality assessment techniques tailored for natural images may not be dependable for medical SR tasks. To complement, researchers often gauge the quality of enhanced images based on how they fare in subsequent medical image processing tasks, like segmentation [**?**]. While quality metrics do not equate to diagnostic precision, medical professionals undeniably favor high-quality imagery for precise diagnosis. Beyond just machine-based quality metrics, prior understanding from pretrained segmentation models can also aid in optimizing medical SR models, akin to existing perceptual loss techniques.

While SR has seen significant advancements, current CNN-based models are not without challenges. As the depth of these networks increases, they demand considerable computational resources and memory, making them unsuitable for deployment on compact devices like smartphones. Additionally, CNNs typically focus on localized image areas due to the restrictive kernel sizes of their convolutional processes. This means they struggle with efficiently modeling long-range dependencies within images. Thus, for optimal network performance, it is crucial to consider both localized and broader image information. Efficiency in terms of speed and memory usage is crucial for deploying SR models for images on devices with limited resources. Recent innovations leverage techniques like distillation and aggregation, making extensive use of channel splitting and merging to maximize the utilization of available hierarchical features. On the other hand, sequential network processes, which avoid regularly revisiting previous states and adding extra nodes, can help minimize memory usage and decrease the time taken, making them more efficient. Recent developments in efficient SR aim to decrease the number of parameters

and FLOPs [**?**]. They enhance the feature strength by optimizing feature usage with intricate layer connection methods. However, such complex structures may not always result in faster performance and can pose challenges when trying to implement them on devices with limited resources. Based on this approach, we designed a lightweight network structure primarily by layering multiple optimized convolution and activation layers while reducing the dependency on feature fusion. We introduced a unique sequential channel attention mechanism, where each pixel is allocated a significance based on its local and global contexts to emphasize high-frequency details. Moreover, we modified the residual block for SR and introduced a superior residual block to further hasten the network's inference. In light of this, we review the leading efficient SR model, RFDN [**?**], aiming to strike a better balance between image reconstruction quality and inference speed. Initially, we reevaluate the effectiveness of various elements within the residual feature distillation block introduced by RFDN [**?**]. We recognize that, although feature distillation considerably cuts down on the number of parameters and boosts overall performance, it is not entirely optimized for hardware, thereby affecting RFDN [**?**] processing speed. To address this, we introduce a new model, the DRFDCAN. This model is designed to simplify the network's architecture while preserving its efficiency and capability. The proposed work recognizes the limitations entrenched within the contemporary CNN-based SR models. These challenges encompass formidable computational demands, escalating memory requirements, and the intricate puzzle of effectively modeling long-range dependencies within images. Efficiency emerges as a central tenet of our investigation. We underscore the vital importance of efficiency, particularly concerning speed and memory utilization, by emphasizing that such efficiency is not merely a desirable attribute but an absolute imperative, especially given the prevalent use of resource-constrained devices such as smartphones. The proposed method thoughtfully evaluates recent innovations in the realm of SR. These innovations, including techniques like distillation, aggregation, channel splitting, and merging, represent a concerted effort to unlock the full potential of the hierarchical features within the SR domain. A notable facet of our research method revolves around reducing model parameters and floating-point operations (FLOPs). This reduction is achieved through intricate layer connection methods and judicious feature usage, aligning perfectly with their overarching pursuit of efficiency. The emergence of DRFDCAN stands as a crowning achievement of this scientific endeavor. The model streamlines network architecture while preserving efficiency and elevating the SR process to new heights.

The proposed work leads to pioneering strides by enhancing model compactness and expediting the inference process without compromising image reconstruction quality. We also introduced a specialized channel attention block designed to amplify high-frequency features, leveraged a novel residual-within-residual network approach for enhanced processing, and designed an innovative feature extractor that is adept at capturing intricate edge and texture details. This study is further distinguished by its comprehensive examination of the elements influencing the processing speed and memory utilization within SR models.

The main highlights of our work include:

- We reassess the efficiency of RFDN [**?**] and delve into what slows it down. Introducing our innovative network, the DRFDCAN, we have been able to improve model compactness and speed up inference without compromising on the accuracy of SR restoration.
- We suggest a channel attention block tailored for high-frequency features to amplify these details, and we introduce a residual-within-residual network approach that leverages residual learning for quicker processing and reduced memory usage.
- By examining the intermediate features extracted using the feature extractor of the contrastive loss, we found that features from the initial layers play a vital role in models focused on PSNR. This observation led us to design a unique feature extractor that captures more detailed information on edges and textures.

- We thoroughly examine the elements that affect the processing speed and memory usage of SR models.

2. Related Work

Medical image SR, an area that is rapidly gaining traction in research circles, focuses on improving the clarity and resolution of medical images, making them more valuable for clinical analyses. Initial approaches relied on traditional techniques such as interpolation and example-based learning. However, with the advent of deep neural networks, more sophisticated methods have emerged. These not only focus on enhancing the image quality but also aim to balance computational efficiency. Contemporary strategies employ adversarial training, ViTs, and lightweight models to achieve superior outcomes. This section delves into various advancements in this domain, highlighting both their innovations and limitations.

2.1. Introduction to Image SR Techniques

Before the widespread use of deep neural networks, initial approaches to medical image SR largely depended on interpolation, reconstruction, and example-based learning techniques [?]. While these methods used multiple frames [?] and reference sections [?] to reconstruct HR images, their effectiveness was limited. This was mainly due to their constrained ability to represent details and the insufficient information provided by the training data. While SRCNN [?] necessitates an initial up-sampling step to preprocess LR images, this means that most subsequent processes happen in a high-dimensional space, leading to increased computational demands. To address this, ref. [?] introduced the Efficient Sub-Pixel Convolutional Neural Network (ESPCN). By positioning the up-sampling layer towards the end of the model, it ensures that feature extraction takes place in a lower-dimensional space, greatly reducing computation and spatial complexity. Furthering this [?] expanded the network depth and employed the residual structure. This demonstrated that deepening the network could enhance its output. Since these developments, researchers have consistently advanced algorithm performance using various intricate network designs, including but not limited to residual learning [?], dense connections [?], and attention mechanisms [?].

Adversarial training, alongside perceptual loss, as described by [?], has become common in single-slice SR, largely due to the significant computational demands of 3D operations, as noted by [?]. Ref. [?] presented a deep learning method combining transformers and generative adversarial networks (T-GANs), which is proposed to enhance the SR reconstruction of medical images. By integrating transformers into the GAN framework, the model achieves superior texture extraction and focuses on critical areas through global image matching. The training involves a multi-task loss function composed of content loss, adversarial loss, and adversarial feature loss. Ref. [?] presents the fine perceptive generative adversarial networks (FP-GANs) to enhance the resolution of MR images. The FP-GANs tackle time and hardware limitations by employing a divide-and-conquer strategy, separately processing the low-frequency and high-frequency components of MR images. Through wavelet decomposition, the MR image is split into global approximation and anatomical texture sub-bands, with each GAN focusing on its specific sub-band. The generator uses multiple residual-in-residual dense blocks for improved feature extraction and a texture-enhancing module to balance global and detailed features. In ref. [?] the study introduces a constrained CycleGAN. This model operates on unpaired images from different ultrasound probes and integrates two new loss functions—identical loss and correlation coefficient loss—to enhance structural consistency and backscattering patterns. Instead of relying on post-processed images, CycleGAN directly uses envelope data from beamformed radio frequency signals.

ViTs enhance various medical image processing tasks, such as reconstruction, as shown by [?]. Ref. [?] introduced the L-former, which is a lightweight transformer designed for generating realistic medical images, outperforming traditional GANs. Unlike

typical transformer models, the L-former efficiently combines transformers for LR features with CNNs for HR output, ensuring lower computational costs. Ref. [?] addresses challenges in multicontrast SR techniques for MRI. While multicontrast MRI SR can leverage complementary information from different imaging contrasts for higher-quality images, current methods fall short in two main areas: capturing long-range dependencies that are crucial for complex MR images and fully utilizing multicontrast features at various scales. To tackle these challenges, the paper introduces McMRSR++, a novel SR network that employs transformers for long-range dependency modeling and features a multiscale matching and aggregation method to harness multicontrast features effectively. Ref. [?] presents MRI-Net, a deep learning model based on U-Net architecture, designed to upgrade low-resolution brain MRI scans to a higher resolution. This advancement could enhance medical diagnosis and reduce the cost of high-resolution MRI imaging by improving lower-resolution images. The network shows superiority at a 3×3 down-sampling index, but its performance on different scales or subtler down-sampling, which is often required in real-world applications, is not discussed.

2.2. Lightweight Models for Medical Image SR

However, the enhanced reconstruction achieved by deepening the network also leads to a notable rise in computational demands and processing time [?]. This poses a limitation for the practical application of SR. As a result, many research efforts have been directed towards developing lightweight SR algorithms to overcome this obstacle. Ref. [?] introduced an efficient ViT that is tailored for SISR in medical imaging, incorporating residual dense connections and local feature fusion. While ViTs have excelled in various computer vision tasks, they face challenges in low-level medical image processing. The proposed model overcomes these challenges and also integrates a perceptual loss based on medical image segmentation to enhance image quality. While the method achieves the best PSNR scores in six modalities, it does not dominate in all seven, suggesting room for further improvement or adaptability challenges in certain modalities. This is achieved by utilizing hierarchical features from residual branching and integrating a spatial attention mechanism within the residual blocks. This is used to improve the resolution of chest CT images used in diagnosing COVID-19, presented in [?]. Recognizing that most existing SR algorithms, optimized for natural images, are not apt for medical images and tend to consume more computational resources by increasing the network depth, the study introduces a novel method: the residual feature attentional fusion network (RFAFN). The RFAFN employs a contextual feature extraction block (CFEB) for efficient feature extraction, a feature-weighted cascading strategy (FWCS) that selectively fuses detailed information and a global hierarchical feature fusion strategy (GHFFS) for better feature aggregation. The presented method boosts the reconstruction quality by increasing the network depth, making it resource-intensive and not ideal for machines with limited capacity.

Medical image SR has recently garnered significant attention from the research community. For instance, a proposed IMDN [?] addresses these challenges, which belong to computing limitations with cascaded information multidistillation blocks (IMDB) that extract and fuse hierarchical features. A contrast-aware channel attention mechanism evaluates feature importance. Additionally, an adaptive cropping strategy (ACS) enables the processing of real images of any size, allowing block-wise SR with a consistent and trained model. The presented method fails to effectively super-resolve images at any arbitrary scale factor, limiting their practical application. Another good example is the LCRCA presented in ref. [?], a lightweight skip-concatenated residual channel attention network for image SR. While deep neural networks produce images that are closer to the original HR versions, their extensive and sometimes unnecessary structures and parameters become burdensome. The LCRCA is designed for high-quality SR in environments with limited computational resources. It incorporates a deep residual block (DRB) for precise residual information, a new channel attention mechanism called statistical channel attention (SCA) to amplify the DRB's feature maps, and a skip concatenation (SC) technique for efficient

information flow. The proposed LCRCA, though efficient, still trades off some details for its lightweight nature. Existing super-resolution methods often struggle with preserving high-frequency details, leading to blurred features or introducing artifacts and structural deformations. RMISR-BL [**?**] addresses these issues through a pyramidal feature multidistillation network. Key components include a multidistilliation block that combines pyramidal convolution and shallow residual blocks, a two-branch super-resolution network for optimizing visual quality, and the use of contextual loss and L1 loss in a gradient map branch to enhance visual perception. The introduction of pyramidal feature multidistillation and other advanced techniques resulted in increased computational demands. This limits the method's real-time applicability in resource-constrained environments and on less powerful hardware.

3. Materials and Methods

In Section **??**, we presented information about used datasets and metrics. In Section **??**, we introduce the proposed DRFDCAN, where we use a novel approach, known as a "residual in residual" (RIR). The architecture is introduced for constructing exceptionally deep networks that are composed of multiple residual groups, each interconnected with lengthy skip connections. Within each residual group, there are several residual blocks linked by short skip connections. Additionally, a channel attention mechanism is suggested to dynamically adjust channel-specific features by accounting for relationships among different channels. In the next section, Section **??**, we summarize the baseline RFDN [**?**] (Figure **??**) with its shallow residual block and shortcuts, which we try to overcome.
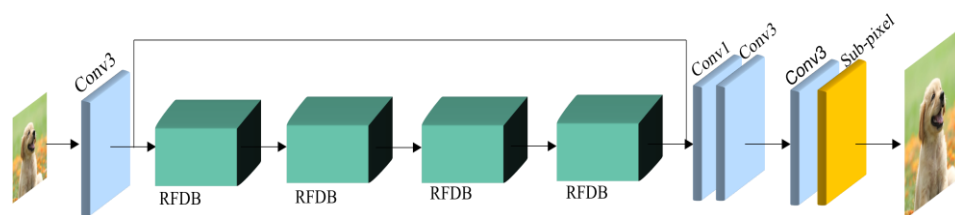

Figure 1. Architecture of baseline model RFDN [**?**].

3.1. Dataset-Specific Performance Analysis and Metrics

This section analyzes the performance of the DRFDCAN. Here, the discussion will focus on the unique challenges and outcomes associated with each dataset. For the OASIS dataset [**?**], the discussion might revolve around the model's ability to reconstruct high-fidelity images from single-modality MRIs, considering factors such as the presence of age-related atrophy or lesions. In the context of the BraTS dataset [**?**], the efficacy of the DRFDCAN in dealing with multimodal data is scrutinized. The discussion could highlight the model's capacity to enhance the contrast and details from different MR sequences, which are crucial for the accurate delineation of tumor boundaries. With the ACDC dataset, the focus might shift to how the DRFDCAN handles the dynamic range of cardiac images and whether it can effectively capture the nuances of cardiac cycles, including the changes in the heart's structure during diastole and systole. For the COVID dataset, the section would likely explore the model's effectiveness in enhancing CT images that are vital for detecting and assessing the progression of COVID-19-related lung anomalies.

In this study, we utilize four renowned medical image datasets to assess the SR capabilities and resilience of our proposed approach, aiming to closely replicate real-world clinical settings. Our testing grounds include the OASIS dataset [**?**], offering single-modality brain MR images; the BraTS dataset [**?**], which provides multimodal brain MR images; the ACDC dataset [**?**], containing cardiac MR images; and the COVID dataset [**?**], which presents chest CT scans.

To gauge the quality of the reconstructed images, we employ two metrics: the PSNR and the structural similarity index (SSIM). The PSNR serves as an indicator of the rela-

tionship between the highest possible signal and the background noise, acting as a quality assessment index based on its sensitivity to errors.

$$PSNR = 10 \times log_{10}\left(\frac{(2^n - 1)^2}{MSE}\right) \qquad (1)$$

On the other hand, the SSIM quantifies the resemblance between two images, considering their luminance, contrast, and structural attributes.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{\left(\mu_2^x + \mu_2^y + C_1\right)\left(\sigma_2^x + \sigma_2^y + C_2\right)} \qquad (2)$$

To comprehensively assess our model's computational demands, consistent with numerous studies [**?**], we determine the Multi-Adds of our model using a predetermined HR image size of 512 × 512.

3.2. Network Architecture

The architecture of our proposed model is illustrated in Figure **??**b. The main building model in DRFDCAN includes three parts: at first, there are three residual channel attention network blocks (RCABs), convolution layers with 3 × 3 and 1 × 1 kernel sizes, and ESA. I_{LR} represents the input LR image, and I_{SR} denotes the output of DRFDCAN (Figure **??**). At the beginning, the network extracts the coarse features using convolution layer 3 × 3, as follows:

$$F_0 = C_{ext}(I_{LR}) \qquad (3)$$

where C_{ext} represents the convolution layer operation, and F_0 denotes the extracted feature maps.
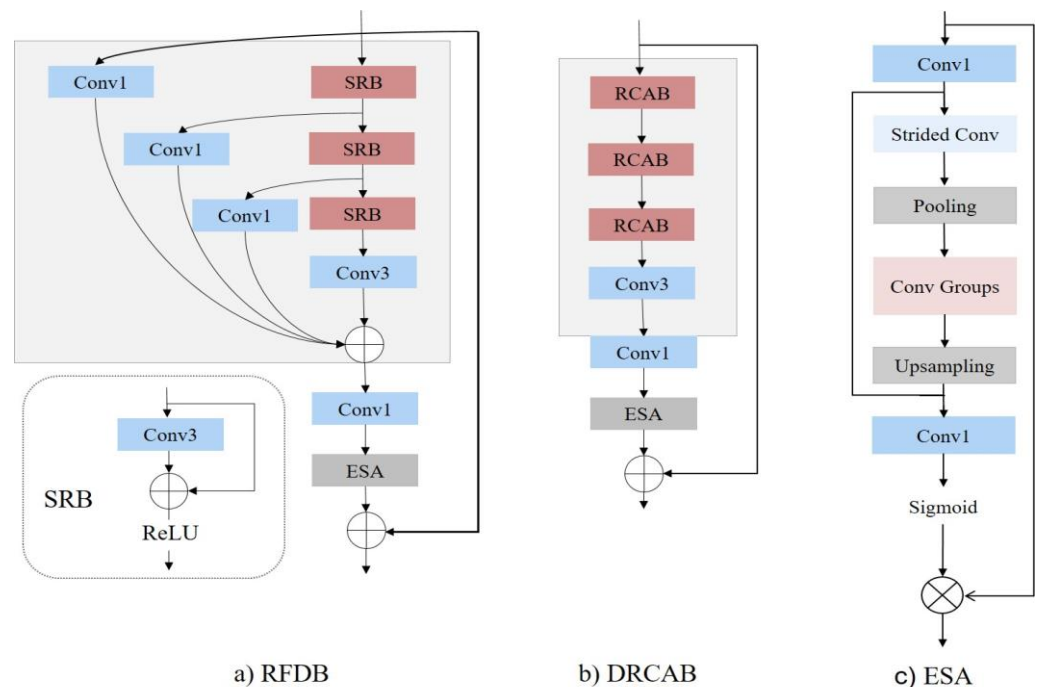

a) RFDB   b) DRCAB   c) ESA

Figure 2. (**a**) RFDB represents the residual feature distillation block. (**b**) DRCAB stands for the residual local feature block. (**c**) ESA is an abbreviation for Enhanced Spatial Attention. In both (**a**,**b**), the number of channels in the output feature maps is indicated adjacent to each layer.

To improve the performance of the baseline, we modify the channel attention block of the deep residual channel attention networks (RCANs). Into the RCAN, the RIR structure

is introduced, which comprises G residual groups (RG) and long skip connections (LSC). Each RG is further composed of B RCAB, which is interconnected by short skip connections (SSC) (Figure **??**). We change the RIR connection architecture with a new one, where instead of LSC, we use SSC. That RIR architecture enables the training of extremely deep CNNs, exceeding 400 layers, for high-performance image super-resolution. In the proposed model, we try to avoid using extremely networked layers, because this approach makes the model too complex for computation. In our model, by using SSC, we make the model more lightweight, and the process can be expressed by:

$$F_n = H_{RCAB}^n \left(H_{RCAB}^{n-1} \left(\ldots H_{RCAB}^0 (F_0) \right) \right); \tag{4}$$

where H_{RCAB} denotes the function of RG and F_n is the n-th layer of RCAB. In Figure **??**, channel attention (CA) has been integrated into RB. Two 3×3 convolution layers in RB extract the smooth feature from LR, after which the obtained information becomes the input for CA. The input feature is divided into channel attention and spatial attention layers, and both layers extract coarser features and concatenate the final result. The output is a reconstructed I_{SR} super-resolution image.

$$I_{SR} = F_{rec}((F_{smooth}(F_n) + F_0)) \tag{5}$$

where F_{rec} function represents the reconstruction module, which is composed of a single 3×3 convolution layer and a non-parametric sub-pixel operation. Additionally, F_{smooth} corresponds to a 3×3 convolution operation.
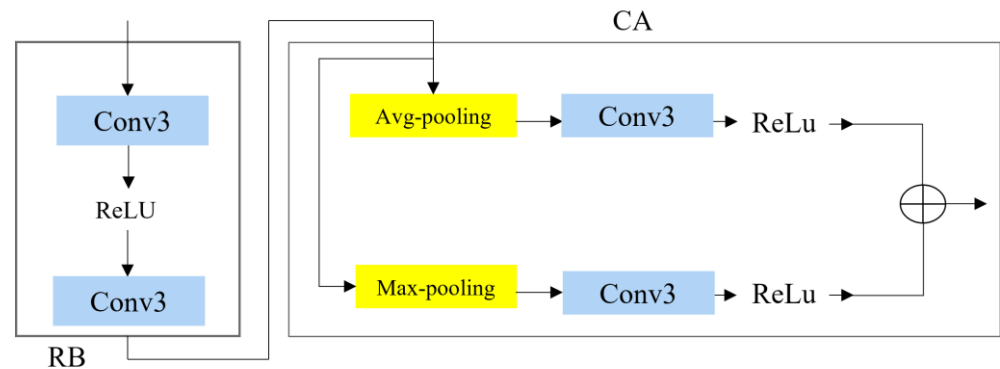

Figure 3. Illustration of RCAB block.
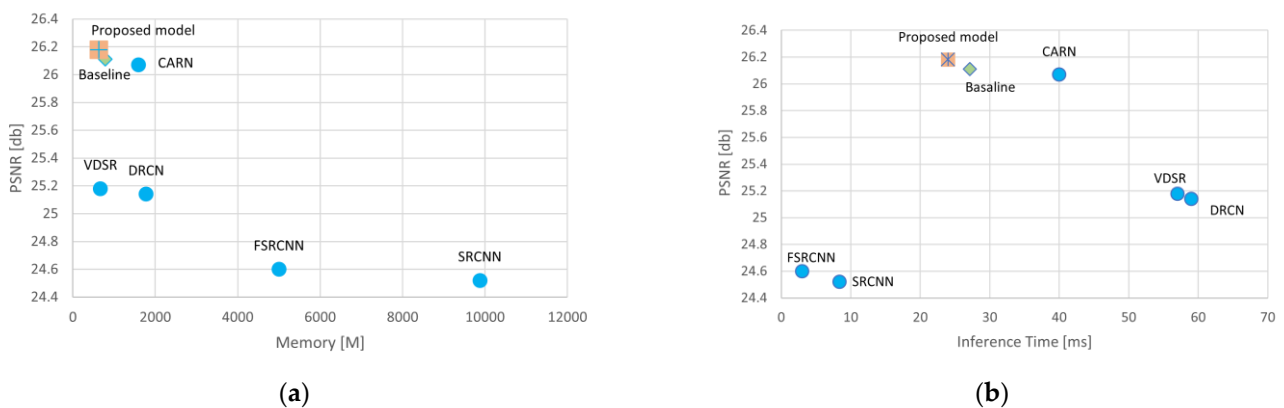

(**a**)

(**b**)

Figure 4. Illustration of PSNR, inference time, and memory consumption of SISR models on the dataset OASIS [**?**] for scale factor 4×. (**a**) present memory conceptions of baseline model with others. (**b**) present inference time conceptions of baseline model with others.

A standard loss function of the proposed model is as follows:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^{N} ||F_{DRFDCAN}(I_{LR}^n) - I_{HR}^n||_1 \tag{6}$$

where $F_{DRFDCAN}$ denotes the function of the DRFDCAN, θ represents the adjustable parameter of the proposed model, and $||.||_1$ denotes the L1 regularization. I_{LR} and I_{HR} represent input LR and coresponding HR images.

3.3. Rethinking the RFDB

Being inspired by the RFDN [**?**], we rethink its main block, the residual feature distillation block (RFDB). In Figure **??**, the RFDB initially follows a progressive feature refinement and distillation strategy, and it employs a 1×1 convolution to reduce the number of channels. Finally, it incorporates an ESA layer and establishes a residual connection. In detail, the feature refinement and distillation process comprise multiple steps. For each stage, the RFDB utilizes a refinement module (RM) composed of a shallow residual block (SRB) to enhance the extracted features. Additionally, it employs a distillation module (DM), represented as a single 1×1 convolution layer, to distill the features. When provided with input features F_{in}, the overall structure can be described as follows:

$$\begin{aligned} F_{d\_1}F_{c\_1} &= DM_1(F_{in}), RM_1(F_{in}) \\ F_{d\_2}F_{c\_2} &= DM_2(F_{c_1}), RM_2(F_{c_1}) \\ F_{d\_3}F_{c\_3} &= DM_3(F_{c\_2}), RM_3(F_{c\_2}) \\ F_{d\_4} &= DM_4(F_{c\_3}) \end{aligned} \tag{7}$$

In this equation, DM_n and RM_n refer to the n-th distillation and refinement modules, respectively. F_{d_n} represents the features distilled in the n-th step, and F_{c_n} represents the features refined in the n-th step, which will undergo further processing by subsequent layers. Finally, all the distilled features generated from previous distillation steps are concatenated together as follows:

$$F_{distillation} = Concat\left(F_{d_1}F_{d_2}F_{d_3}F_{d_4}\right) \tag{8}$$

The concatenation operation goes along the channel dimension.

In total, the main block of the RFDN [**?**], the RFDB, uses split division to reach two feature extraction connections: a refinement layer, which obtains coarse features, and a distillation layer where several convolution layers of 1×1 extract more information and decrease the number of parameters. Though the RFDB has managed to reduce the number of parameters, FLOPs, memory consumption, and runtime, SRB shows a high number of attributes and low performance. Based on our observations, memory consumption and PSNR stand out as critical attributes for IoT devices and lightweight models. Enhancing these factors not only boosts the model's performance but also optimizes its evaluation time, making it more suitable for real-time applications (Figure **??**). The graph depicts a selection of image processing models, each represented as a point indicating their respective performance in terms of PSNR and inference time. A clear trade-off is visible: as the inference time decreases, moving towards the left, the PSNR tends to increase, moving upwards, suggesting that quicker models also improve image quality. The proposed model stands out, achieving the highest PSNR in the least inference time, positioning it as the most efficient option among those shown. The baseline provides a point of reference, sitting in the middle of the graph, while the CARN model showcases a balance of speed and quality. Other models like the VDSR and DRCN are clustered to the right, indicating longer inference times but also delivering high-quality images as reflected by their PSNR values. In contrast, the FSRCNN and SRCNN [**?**] are the slowest and produce lower-quality images, signaling a potential area for improvement. (Figure **??**b).

Residual channel attention block: The RCAB is the main block of the proposed model, which can smoothly extract more features and reduce attributes and runtime while main-

taining PSNR. As Figure **??**b demonstrates, we remove the distillation layers to focus further on the refinement ones. In this way, we avoid making the model deeper, while using an RIR structure. The process can be described as

$$
\begin{aligned}
F_{refinement\_1} &= RM_1(F_{in}) \\
F_{refinement\_2} &= RM_2\left(F_{refinement\_1}\right) \\
F_{refinement\_3} &= RM_3\left(F_{refinement\_2}\right)
\end{aligned}
\tag{9}
$$

where F_{in} is given input features which feed first the RM_n refinement layer, and $F_{refinement\_n}$ represents n-th extracted feature. Each refinement layer illustrates the n-th RCAB that works using an RIR structure of connection. Two 3×3 convolution layers in RB obtain more information to feed CA. Utilizing two pooling methods, max pooling and average pooling aids with overcoming information loss issues. Both outputs from the pooling are concatenated with each other to pass to the next refinement layer (Figure **??**).

$$
F_{refinement\_1} = RM_1(F_{in}) = F_{CA\_1}\left(F_{M_p}(F_{RB\_1}(F_{in})) + F_{A_p}(F_{RB\_1}(F_{in}))\right)
\tag{10}
$$

F_{RB_n} and F_{CA_n} represent the n-th RIR connection of the n-th refinement layer. F_{M_p} and F_{A_p} denote max and avarage poolings. A final output of RCAB feeds the 1×1 convolution layer, and then the ESA block. Next comes concatenation of the final output and F_{in}, as shown in Figure **??**b:

$$
F_{final\_output} = F_{in} + F_{refinement\_3}
$$

4. Experimental Results

In this section, the performance evaluation of the DRFDCAN is presented. The assessment is meticulously designed to measure the model's super-resolution capabilities and robustness across different clinical imaging scenarios. This evaluation is critical to ensuring that the model can be reliably deployed in real-world medical settings, where the demand for high-fidelity image reconstruction is coupled with the need for rapid diagnostic insights. The section will begin by detailing the evaluation metrics, such as PSNR, SSIM, to provide a quantitative analysis of the SR images against the ground truth. Additionally, qualitative assessments may also be incorporated, wherein expert radiologists visually inspect the SR images for clinical usability.

To ensure the robustness and generalizability of the model, tests are conducted across various datasets:

The OASIS dataset [**?**] for single-modality brain MR images,
The BraTS dataset [**?**] for multimodal brain MR images,
The ACDC dataset [**?**] for cardiac MR images, and
The COVID dataset [**?**] for chest CT scans.

The datasets are chosen to reflect the diversity of imaging modalities and the complexities inherent in different anatomical regions, pathologies, and imaging conditions.

4.1. Implementation Detail

The proposed method and other leading-edge models are tested using PyTorch. We carried out all tests using an Nvidia Quadro RTX 8000 GPU Corporation in the U.S. For data augmentation, we used random rotations at angles of $90°$, $180°$, and $270°$, along with horizontal flipping. We utilized a batch size of 64 and employed the Adam optimizer with parameters β1 set to 0.9, β2 at 0.99, and ε at 10^{-8} for training our model. We started with a learning rate of 5×10^{-4}, which we reduced by half every 50,000 iterations, continuing for a total of 300,000 iterations.

4.2. Comparison with State-of-the-Art Methods

To showcase the effectiveness of our method, we set it against several top-tier lightweight super-resolution networks such as SRCNN [**?**], RDST [**?**], RFAFN [**?**], IMDN [**?**], RFDN [**?**], RMISR-BL [**?**], and LCRCA [**?**]. For each of these networks, we utilize the source codes that are officially released by their respective authors and retrain them using the identical dataset and training specifications that were used for the proposed mode in our study.

In this analytical segment in Table **??** we scrutinize various state-of-the-art models and their performance in the domain of super-resolution, as manifested through distinct metrics on a particular dataset. The proposed model emerges as a clear victor in terms of image quality, boasting the highest PSNR at 29.51 dB.

Table 1. Illustrate comparison of lightweight models based on OASIS dataset.

Model	PSNR [db]	SSIM	Memory [M]	Inference Time [ms]
		×**4**		
RFDN [?]	29.11	82.20%	788.19	28.03
SRCNN [?]	27.55	80.42%	2387.93	36.58
RDST [?]	29.17	82.32%	790.25	29.74
RFAFN [?]	29.22	83.01%	780.18	28.41
IMDN [?]	29.12	82.23%	856.02	28.07
LCRCA [?]	29.16	82.36%	895.36	28.45
MRI-Net [?]	29.20	81.57%	1258.81	29.72
Proposed model	**29.51**	**91.56%**	**630**	**24.07**

This metric indicates its superior ability to reconstruct images that closely resemble the originals. In a comparative assessment with the proposed model, other models manifest a noticeable lag in performance. Specifically, when observing the PSNR values, SRCNN [**?**] records a difference of 1.96 dB less than the proposed model. Similarly, RFDN [**?**], RDST [**?**], RFAFN [**?**], IMDN [**?**], and LCRCA [**?**] trail behind by 0.40 dB, 0.34 dB, 0.29 dB, 0.39 dB, and 0.35 dB, respectively. MRI-Net [**?**] stands in the middle ground in terms of PSNR, suggesting that while its image enhancement quality is good, it is not the best among the compared models. In terms of SSIM, MRI-Net [**?**] is closer to the lower end, which may imply that while it can enhance the resolution, it might not preserve structural features as well as the top-performing models. MRI-Net [**?**] has a relatively high memory consumption, which could be a drawback for deploying it in environments with limited memory resources. The inference time of MRI-Net [**?**] is competitive, although not the fastest, which suggests that it could still be used in near-real-time applications.

Figure **??** highlights the importance of optimizing for both performance and efficiency. In the context of medical imaging, where these models might be used, the graphs suggest that the most valuable models are those that can produce high-quality images quickly and with less demand on memory resources, allowing for fast and efficient diagnosis without compromising on the clarity of the images needed by medical professionals.

These discrepancies underscore a relative shortcoming in image clarity for these models when positioned against the proposed model. Turning our attention to SSIM, a critical metric that assesses structural and visual congruency with the original HR images, the proposed model again takes the lead with an exceptional 91.56%. Conversely, RFAFN [**?**] trails with an SSIM of 83.01%, hinting at a certain level of structural deviation in its outputs. From a resource-efficiency standpoint, the proposed model demonstrates admirable frugality, consuming a mere 630M of memory. This is strikingly lower than its counterparts, marking it as a viable choice for scenarios with stringent memory constraints. At the other end of the spectrum, SRCNN [**?**] consumes a substantial 2387.93 M, making it the most memory-intensive model amongst those evaluated.
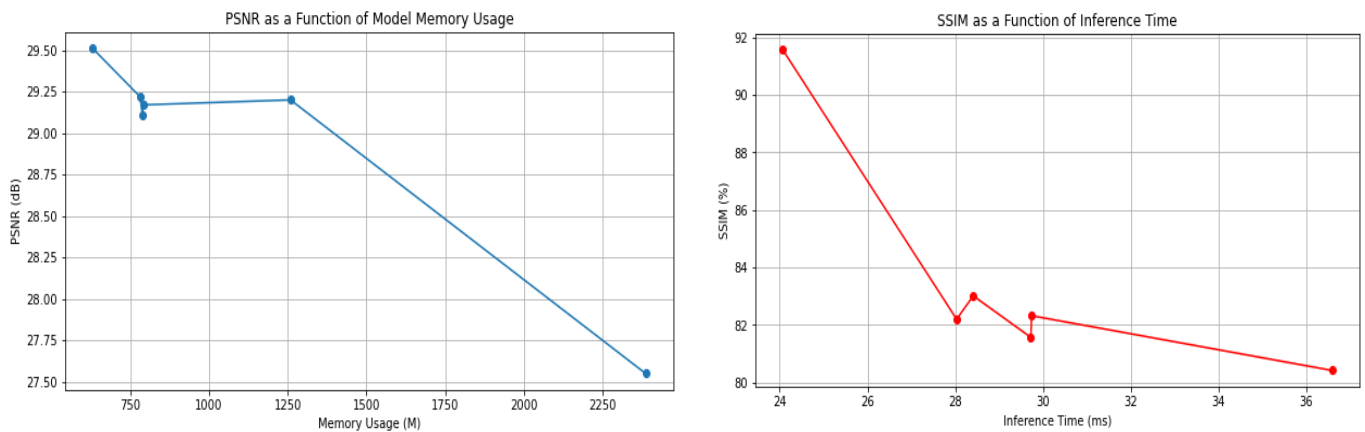
Figure 5. In the provided graphs, we have two plots, each representing the relationship between model performance and computational requirements, one for the PSNR against model memory usage, and the other for SSIM against inference time.

In the area of real-time performance, the proposed model clinches another win with the swiftest inference time of 24.07 ms. This rapidity indicates its suitability for applications demanding instant results. In contrast, SRCNN [**?**] registers the lengthiest inference time of 36.58 ms, which may limit its utility in time-sensitive scenarios. Drawing parallels with other models, the likes of RFAFN [**?**], RDST [**?**], and LCRCA [**?**] offer commendable metrics, showcasing near neck-to-neck performances in PSNR and SSIM. Their memory consumption and inference times are also competitive, situating them as strong contenders in the super-resolution landscape. However, placing all these models under an overarching lens, the proposed model emerges as a paradigm of balance, harmonizing impeccable image quality with resource and time efficiency. While models like the SRCNN [**?**] have created the foundation for super-resolution, they now seem to be eclipsed by the newer, more optimized entrants in the arena.

From Table **??**, the superior performance of the proposed model is unequivocal, especially at the ×2 magnification, where it consistently registers the highest PSNR and SSIM values across datasets such as BraTS in Figure **??**, OASIS in Figure **??**, ACDC in Figure **??**, and COVID-CT in Figure **??**. Notably, while the proposed model remains preeminent at ×2 magnification, the competitive landscape becomes more nuanced at the ×4 magnification. In this context, RFAFN [**?**] achieves the highest PSNR of 30.45, but the proposed model retains its supremacy in SSIM with a score of 85.12%.
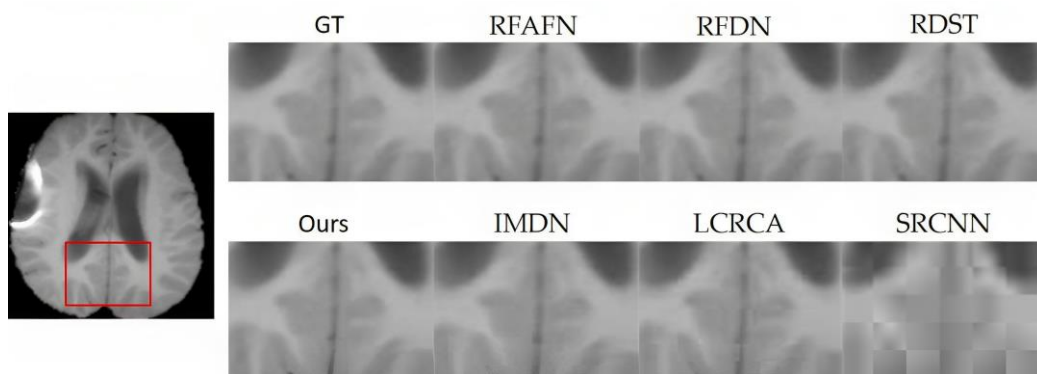

Figure 6. SR outcomes for a randomly selected slice from the BraTS (BraTS: https://www.med.upenn.edu/cbica/brats2020/data.html, accessed on 20 May 2023) testing subset are shown.

Table 2. The table provides a comprehensive assessment of various super-resolution models evaluated over different datasets and magnifications using the PSNR and SSIM metrics.

	BraTS		ACDC		COVID-CT	
			×2			
Mode	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RFDN [?]	33.65	91.89%	32.84	91.24%	36.87	93.58%
SRCNN [?]	30.45	90.20%	32.58	92.15%	35.89	92.65%
RDST [?]	33.89	92.27%	33.65	93.41%	36.74	94.87%
RFAFN [?]	33.56	92.84%	**35.28**	**94.86%**	37.46	95.45%
RMISR-BL [?]	29.46	89.54%	30.08	90.23%	32.89	90.78%
LCRCA [?]	32.56	92.11%	31.02	90.84%	35.48	91.99%
MRI-Net [?]	31.89	91.57%	31.58	90.99%	32.63	91.14%
Proposed model	**34.20**	**94.68%**	35.15	94.78%	**37.89**	**96.08%**
			×4			
	BraTS		ACDC		COVID-CT	
Model	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RFDN [?]	30.11	83.55%	29.21	82.35%	31.87	83.99%
SRCNN [?]	28.53	80.89%	27.55	80.24%	28.78	81.24%
RDST [?]	30.11	79.58%	27.05	79.23%	33.69	87.99%
RFAFN [?]	30.45	81.25%	29.24	80.99%	32.14	89.88%
RMISR-BL [?]	30.27	92.68%	27.52	80.41%	31.88	89.12%
LCRCA [?]	30.21	81.75%	29.12	82.84%	30.84	82.13%
MRI-Net [?]	29.38	81.00%	27.15	79.89%	29.56	82.17%
Proposed model	**31.55**	**85.12%**	**30.74**	**84.26%**	**34.87**	**89.86%**

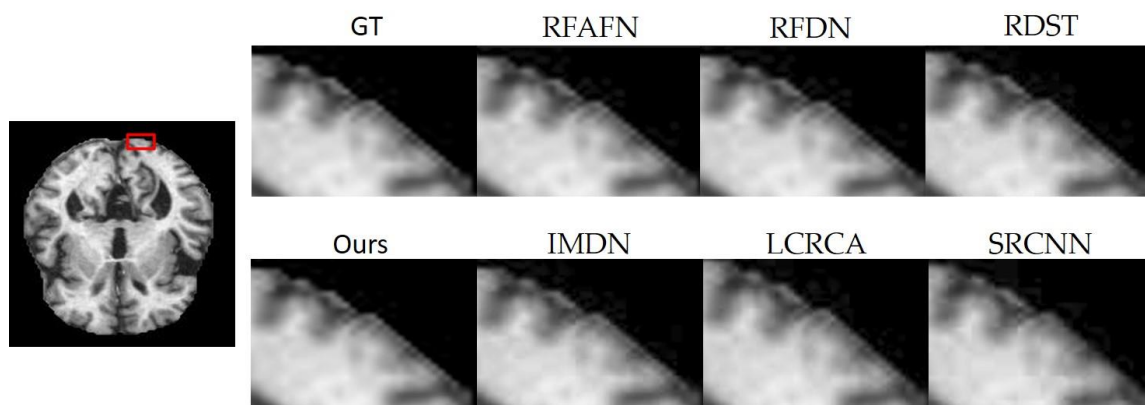**Figure 7.** A comparison of the proposed method and state-of-the-art techniques for the ×4 super-resolution task using the OASIS (OASIS: https://www.oasis-brains.org/, accessed on 10 June 2023) dataset is presented.
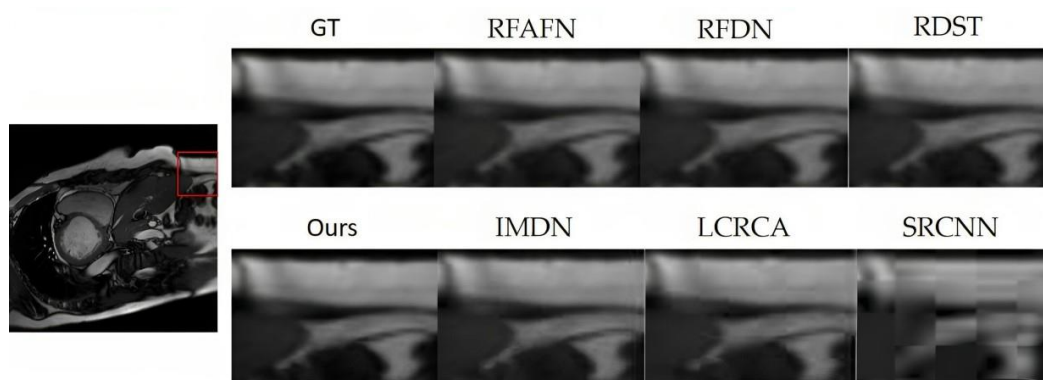
Figure 8. SR outcomes for a randomly chosen slice from the ACDC (ACDC: https://www.creatis.insalyon.fr/Challenge/acdc/databases.html, accessed on 25 June 2023) testing dataset are presented.
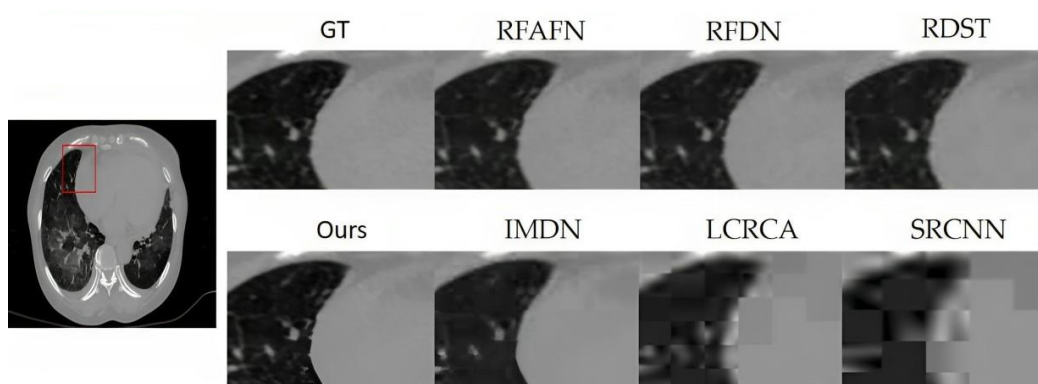
Figure 9. SR findings for a randomly selected slice from the COVID-CT (COVID-19 CT: https://zenodo.org/record/3757476, accessed on 2 June 2023) testing dataset are displayed.

At the ×4 magnification, certain models, especially the foundational super-resolution model SRCNN [**?**], encounter challenges. MRI-Net [**?**] performance drops noticeably at a ×4 magnification for both metrics, indicating that its SR capabilities are more constrained with cardiac MR images when the upscaling task is more challenging. The decline in SRCNN [**?**] performance at this magnification underscores the advancements that contemporary models have achieved in methodology and optimization. Nonetheless, in the COVID-CT dataset at a ×4 magnification, the proposed model reasserts its preeminence with an unmatched PSNR of 34.87 and SSIM of 89.86%. This superior performance accentuates its adaptability and robustness across varying datasets and magnifications. RDST [**?**], RMISR-BL [**?**], and LCRCA [**?**] demonstrate variable performances in different datasets, but none consistently outperform the proposed model. While pioneering models like SRCNN [**?**] have established foundational benchmarks, they appear to be surpassed by more advanced and refined successors that adeptly balance quality and efficiency. In this array of state-of-the-art models, the proposed model distinctly emerges as the exemplar, representing the pinnacle of current super-resolution achievements.

The above analysis underscores the brisk advancements in the super-resolution realm, with contemporary models, particularly the proposed one, leading the charge by offering both superior quality and efficiency.

5. Conclusions

Super-resolution techniques, especially in the domain of medical imaging, have undergone substantial advancements. The criticality of image quality in medical diagnostics necessitates innovations that balance clarity with computational efficiency. Traditional CNN-based SR models, despite their capabilities, grapple with computational constraints, particularly on devices with limited resources. Analyzing the performance of the state-of-

the-art RFDN model [**?**] brought to light certain inefficiencies, especially concerning processing speed in real-world deployments. In response to these challenges, our DRFDCAN has emerged as a promising alternative. It successfully simplifies the network structure, ensuring faster inference while either maintaining or even enhancing image reconstruction quality. The inclusion of a dedicated channel attention block for high-frequency features signifies the model's adeptness at capturing intricate details. Furthermore, the integration of a more sophisticated residual block underscores our commitment to improving the processing speed without compromising the integrity of the image enhancement. A notable revelation from our research is the pronounced significance of the initial layer features in models that prioritize PSNR. Harnessing this knowledge, we have crafted a distinctive feature extractor, tailored to distill intricate edge and texture details more effectively. The interpretability of deep learning models like DRFDCAN, crucial for clinical trust, remains an area that is ripe for exploration. Adversarial robustness is another concern, as the integrity of medical diagnoses must be safeguarded against potential manipulations. The integration of SR technologies into existing medical workflows, adherence to ethical considerations, and regulatory compliance also pose significant challenges that must be met with deliberate and focused research efforts.

The burgeoning fields of transfer learning and few-shot learning offer promising methods for DRFDCAN to quickly adapt to new tasks, potentially revolutionizing its application to rarer diseases or novel imaging modalities. These approaches could significantly diminish the data requirements for training high-performance SR models, a particularly pertinent advantage in the medical domain, where data availability can be limited due to privacy concerns and regulatory constraints. The proposed DRFDCAN model stands as a significant advancement in medical SR, but fully harnessing its potential will require navigating the complex interplay of technical limitations, ethical considerations, and emerging research directions. Future work should focus on expanding the generalizability and robustness of the model, exploring the ethical landscape of artificial intelligence in healthcare, and ensuring compliance with evolving regulatory standards. The insights from our study thus not only pave the way for the next generation of SR models but also establish a comprehensive roadmap for the responsible and effective integration of these technologies into medical diagnostics.

The evolution of SR, as showcased by our work, signifies a paradigm shift towards models like DRFDCAN. These models prioritize both computational efficiency and unparalleled image reconstruction, which is particularly pivotal in the medical field, where diagnostic accuracy hinges on image clarity. The insights gleaned from our study pave the way for future research, emphasizing the importance of balancing the trifecta of speed, accuracy, and computational demands in the ever-evolving realm of medical super-resolution.

6. Discussion

The development and application of the DRFDCAN holds promise for future diagnostic facilities through several significant outcomes. Firstly, the model's ability to enhance image clarity could revolutionize the precision of diagnostics. This clarity allows for the detection of minute pathological details, which is critical in diagnosing diseases at an early stage when they are most treatable. Moreover, the computational efficiency of DRFDCAN could enable its integration into diagnostic workflows without the need for extensive hardware upgrades. Such efficiency supports the use of SR in a broader range of clinical settings, including those with limited resources. The ability to process images quickly also means that results can be obtained faster, which is crucial in acute medical situations where time is of the essence. If DRFDCAN proves to be adaptable to different imaging modalities, it would be a significant boon, offering a unified SR solution across diverse diagnostic platforms. This flexibility would ensure that high-resolution imaging is not a specialized service but a standard offering, improving diagnostic services across the board. Furthermore, in imaging modalities where radiation is a concern, DRFDCAN could

contribute to patient safety. By allowing for lower radiation doses while still maintaining image quality, the model can help in minimizing the risk associated with repeated exposure.

The proposed model's prowess in image reconstruction is highlighted by its superior PSNR values, which surpass those of established models such as SRCNN [**?**], RFDN, RDST [**?**], RFAFN [**?**], IMDN [**?**], and LCRCA [**?**]. For instance, SRCNN [**?**], an earlier convolutional-neural-network-based model, lags significantly, with a PSNR difference of 1.96 dB compared with the proposed model. This considerable gap underscores the advancements made by the proposed model in producing high-fidelity images. Similar patterns emerge with the SSIM metric, where the proposed model maintains a lead with an impressive 91.56%, indicating its exceptional capability to preserve the structural integrity and texture of the enhanced images. This is particularly noteworthy in comparison to RFAFN [**?**], which exhibits an SSIM of 83.01%, pointing to potential compromises in structural fidelity. From a computational resource standpoint, the proposed model demonstrates a remarkable efficiency. It consumes only 630 M of memory, which is significantly less than the SRCNN [**?**] hefty 2387.93 M. This makes the proposed model not only more suitable for environments with limited hardware capabilities but also aligns with the increasing need for energy-efficient AI models, thereby contributing to sustainable computing practices in healthcare. The ability of the DRFDCAN to maintain its performance across various planes and types of images speaks to its robustness and potential for widespread clinical application. As medical imaging technology evolves, such an SR tool could become integral to the development of automated diagnostic systems, enhancing not only image resolution but also facilitating the integration of artificial intelligence in diagnosis, ultimately leading to improved patient outcomes.

Author Contributions: Conceptualization, S.U. and S.M. (Sevara Mardieva); methodology, S.U. and S.M. (Sevara Mardieva); software, S.U., and S.M. (Sevara Mardieva); validation, S.U. and S.A.; formal analysis, S.A. and T.W.; investigation, S.A. and T.W.; resources, S.U. and S.M. (Shakhnoza Muksimova); data curation, S.U. and S.M. (Shakhnoza Muksimova); writing—original draft preparation, S.U. and S.M. (Sevara Mardieva); writing—review and editing, S.A. and T.W.; visualization, S.M. (Sevara Mardieva) and S.M. (Shakhnoza Muksimova); supervision, S.A. and T.W.; project administration, S.A. and T.W.; funding acquisition, T.W. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Open Access.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

Definition	Abbreviation
SR	Super-resolution
CNN	Convolutional neural networks
RFDN	Residual Feature Distillation Network
DRFDCAN	Deep Residual Feature Distillation Channel Attention Network
PSNR	Peak signal-to-noise ratio
MRI	Magnetic Resonance Imaging
CT	Computed Tomography
CAD	Computer-Aided Diagnosis
HR	High-resolution
LR	Low-resolution

SISR	Single Image Super-Resolution
GANs	Generative adversarial networks
ViTs	Vision transformers
RIR	Residual in residual
DRCAB	Residual local feature block
ESA	Enhanced Spatial Attention
RCABs	Residual channel attention network block
RCAN	Residual channel attention networks
LSC	Long skip connections
SSC	Short skip connections
CA	Channel attention
RM	Refinement module
SRB	Shallow residual block
DM	Distillation module
SSIM	Structural similarity index

References

1. Zhang, Y.; Zhang, K.; Chen, Z.; Li, Y.; Timofte, R.; Zhang, J.; Zhang, K.; Peng, R.; Ma, Y.; Jia, L.; et al. NTIRE 2023 challenge on image super-resolution (x4): Methods and results. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023; pp. 1864–1883.
2. Umirzakova, S.; Ahmad, S.; Mardieva, S.; Muksimova, S.; Whangbo, T.K. Deep learning-driven diagnosis: A multi-task approach for segmenting stroke and Bell's palsy. *Pattern Recognit.* **2023**, *144*, 109866. [CrossRef]
3. Kaur, P.; Singh, R.K. A review on optimization techniques for medical image analysis. *Concurr. Comput. Pract. Exp.* **2023**, *35*, 7443. [CrossRef]
4. Gendy, G.; He, G.; Sabor, N. Lightweight image super-resolution based on deep learning: State-of-the-art and future directions. *Inf. Fusion* **2023**, *94*, 284–310. [CrossRef]
5. Umirzakova, S.; Ahmad, S.; Khan, L.U.; Whangbo, T. Medical Image Super-Resolution for Smart Healthcare Applications: A Comprehensive Survey. *Inf. Fusion* **2023**, *103*, 102075. [CrossRef]
6. Wei, C.; Ren, S.; Guo, K.; Hu, H.; Liang, J. High-resolution Swin transformer for automatic medical image segmentation. *Sensors* **2023**, *23*, 3420. [CrossRef] [PubMed]
7. Azad, R.; Kazerouni, A.; Heidari, M.; Aghdam, E.K.; Molaei, A.; Jia, Y.; Jose, A.; Roy, R.; Merhof, D. Advances in medical image analysis with vision transformers: A comprehensive review. *arXiv* **2023**, arXiv:2301.03505. [CrossRef] [PubMed]
8. Yan, Y.; Liu, R.; Chen, H.; Zhang, L.; Zhang, Q. CCT-Unet: A U-shaped Network based on Convolution Coupled Transformer for Segmentation of Peripheral and Transition Zones in Prostate MRI. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 4341–4351. [CrossRef] [PubMed]
9. Kastryulin, S.; Zakirov, J.; Pezzotti, N.; Dylov, D.V. Image quality assessment for magnetic resonance imaging. *IEEE Access* **2023**, *11*, 14154–14168. [CrossRef]
10. Kelkar, V.A.; Gotsis, D.S.; Brooks, F.J.; Prabhat, K.C.; Myers, K.J.; Zeng, R.; Anastasio, M.A. Assessing the ability of generative adversarial networks to learn canonical medical image statistics. *IEEE Trans. Med. Imaging* **2023**, *42*, 1799–1808. [CrossRef] [PubMed]
11. Güven, S.A.; Talu, M.F. Brain MRI high resolution image creation and segmentation with the new GAN method. *Biomed. Signal Process. Control* **2023**, *80*, 104246.
12. Liu, K.; Yu, H.; Zhang, M.; Zhao, L.; Wang, X.; Liu, S.; Li, H.; Yang, K. A Lightweight Low-dose PET Image Super-resolution Reconstruction Method based on Convolutional Neural Network. *Curr. Med. Imaging* **2023**, *19*, 1427–1435. [CrossRef] [PubMed]
13. Liu, J.; Tang, J.; Wu, G. Residual feature distillation network for lightweight image super-resolution. In Proceedings of the Computer Vision–ECCV 2020 Workshops, Glasgow, UK, 23–28 August 2020; Springer International Publishing: Cham, Switzerland, 2020; Volume 16, pp. 41–55.
14. Isaac, J.S.; Kulkarni, R. Super resolution techniques for medical image processing. In In Proceedings of the 2015 International Conference on Technologies for Sustainable Development IEEE, Mumbai, India, 4–6 February 2015; pp. 1–6.
15. Kaji, S.; Kida, S. Overview of image-to-image translation by use of deep neural networks: Denoising, super-resolution, modality conversion, and reconstruction in medical imaging. *Radiol. Phys. Technol.* **2019**, *12*, 235–248. [CrossRef] [PubMed]
16. Greenspan, H. Super-resolution in medical imaging. *Comput. J.* **2009**, *52*, 43–63. [CrossRef]
17. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image super-resolution using deep convolutional networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 295–307. [CrossRef]
18. Ruan, H.; Tan, Z.; Chen, L.; Wan, W.; Cao, J. Efficient sub-pixel convolutional neural network for terahertz image super-resolution. *Opt. Lett.* **2022**, *47*, 3115–3118. [CrossRef]
19. Xie, C.; Zhu, H.; Fei, Y. Deep coordinate attention network for single image super-resolution. *IET Image Process.* **2022**, *16*, 273–284. [CrossRef]

20. Muhammad, W.; Bhutto, Z.; Masroor, S.; Shaikh, M.H.; Shah, J.; Hussain, A. IRMIRS: Inception-ResNet-Based Network for MRI Image Super-Resolution. *CMES-Comput. Model. Eng. Sci.* **2023**, *136*, 1121–1142. [CrossRef]

21. Prasad Gunasekaran, K. Ultra Sharp: Study of Single Image Super Resolution using Residual Dense Network. *arXiv* **2023**, arXiv:2304.10870.

22. Li, X.; Li, M.; Yan, P.; Li, G.; Jiang, Y.; Luo, H.; Yin, S. Deep learning attention mechanism in medical image analysis: Basics and beyonds. *Int. J. Netw. Dyn. Intell.* **2023**, *2*, 93–116. [CrossRef]

23. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In Proceedings of the Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; Springer International Publishing: Berlin/Heidelberg, Germany, 2016; pp. 694–711.

24. Yang, Z.; Yang, H.; Yin, Z. Super-resolution reconstruction for the three-dimensional turbulence flows with a back-projection network. *Phys. Fluids* **2023**, *35*, 055123.

25. Du, W.; Tian, S. Transformer and GAN-Based Super-Resolution Reconstruction Network for Medical Images. *Tsinghua Sci. Technol.* **2023**, *29*, 197–206. [CrossRef]

26. You, S.; Lei, B.; Wang, S.; Chui, C.K.; Cheung, A.C.; Liu, Y.; Gan, M.; Wu, G.; Shen, Y. Fine perceptive gans for brain mr image super-resolution in wavelet domain. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *34*, 8802–8814. [CrossRef] [PubMed]

27. Sun, X.; Li, H.; Lee, W.N. Constrained CycleGAN for effective generation of ultrasound sector images of improved spatial resolution. *Phys. Med. Biol.* **2023**, *68*, 125007. [CrossRef] [PubMed]

28. Guo, P.; Mei, Y.; Zhou, J.; Jiang, S.; Patel, V.M. ReconFormer: Accelerated MRI reconstruction using recurrent transformer. *IEEE Trans. Med. Imaging* **2023**, 1. [CrossRef]

29. Zheng, T.; Oda, H.; Hayashi, Y.; Nakamura, S.; Mori, M.; Takabatake, H.; Natori, H.; Oda, M.; Mori, K. L-former: A lightweight transformer for realistic medical image generation and its application to super-resolution. In *Medical Imaging Image Processing*; SPIE: St. Bellingham, WA, USA, 2023; Volume 12464, pp. 245–250.

30. Lyu, J.; Li, G.; Wang, C.; Cai, Q.; Dou, Q.; Zhang, D.; Qin, J. Multicontrast MRI Super-Resolution via Transformer-Empowered Multiscale Contextual Matching and Aggregation. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, 1–11. [CrossRef]

31. Kalluvila, A. Super-Resolution of Brain MRI via U-Net Architecture. *Int. J. Adv. Comput. Sci. Appl.* **2023**, *14*, 26–31. [CrossRef]

32. Zhu, J.; Yang, G.; Lio, P. A residual dense vision transformer for medical image super-resolution with segmentation-based perceptual loss fine-tuning. *arXiv* **2023**, arXiv:2302.11184.

33. Yang, K.; Zhao, L.; Wang, X.; Zhang, M.; Xue, L.; Liu, S.; Liu, K. Residual Feature Attentional Fusion Network for Lightweight Chest CT Image Super-Resolution. *Comput. Mater. Contin.* **2023**, *75*, 5159–5176. [CrossRef]

34. Hui, Z.; Gao, X.; Yang, Y.; Wang, X. Lightweight image super-resolution with information multi-distillation network. In Proceedings of the 27th ACM International Conference on Multimedia, Nice, France, 21–25 October 2019; pp. 2024–2032.

35. Peng, C.; Shu, P.; Huang, X.; Fu, Z.; Li, X. LCRCA: Image super-resolution using lightweight concatenated residual channel attention networks. *Appl. Intell.* **2022**, *52*, 10045–10059. [CrossRef]

36. Ren, S.; Guo, K.; Ma, J.; Zhu, F.; Hu, B.; Zhou, H. Realistic medical image super-resolution with pyramidal feature multi-distillation networks for intelligent healthcare systems. *Neural Comput. Appl.* **2023**, *35*, 22781–22796. [CrossRef]

37. Marcus, D.S.; Wang, T.H.; Parker, J.; Csernansky, J.G.; Morris, J.C.; Buckner, R.L. Open Access Series of Imaging Studies (OASIS): Cross-sectional MRI data in young, middle aged, nondemented, and demented older adults. *J. Cogn. Neurosci.* **2007**, *19*, 1498–1507. [CrossRef] [PubMed]

38. Menze, B.H.; Jakab, A.; Bauer, S.; Kalpathy-Cramer, J.; Farahani, K.; Kirby, J.; Burren, Y.; Porz, N.; Slotboom, J.; Wiest, R.; et al. The multimodal brain tumor image segmentation benchmark (BRATS). *IEEE Trans. Med. Imaging* **2014**, *34*, 1993–2024. [CrossRef] [PubMed]

39. Bernard, O.; Lalande, A.; Zotti, C.; Cervenansky, F.; Yang, X.; Heng, P.A.; Cetin, I.; Lekadir, K.; Camara, O.; Ballester, M.A.G.; et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: Is the problem solved? *IEEE Trans. Med. Imaging* **2018**, *37*, 2514–2525. [CrossRef] [PubMed]

40. Ma, J.; Wang, Y.; An, X.; Ge, C.; Yu, Z.; Chen, J.; Zhu, Q.; Dong, G.; He, J.; He, Z.; et al. Toward data-efficient learning: A benchmark for COVID-19 CT lung and infection segmentation. *Med. Phys.* **2021**, *48*, 1197–1210. [PubMed]