

Image Denoising Using Generative Adversarial Network by Recursive Residual Group

Maysaa A. Ulkareem Naser ^{1*}, Abbas H. Hassin Al-Asadi ²

^{1,2} Department of Computer Information Systems, University of Basrah, Basrah, Iraq

¹ Department of Computer Science, Kufa University, Kufa, Iraq

Email: ¹ maysaa.naser@uobasrah.edu.iq, ² abbas.hassin@uobasrah.edu.iq

*Corresponding Author

Abstract—Cardiac magnetic resonance imaging (CMR) is a vital tool for noninvasively assessing heart shape and function, offering exceptional spatial and temporal resolution alongside superior soft tissue contrast. However, CMR images often suffer from noise and artifacts due to cardiac and respiratory motion or patient movement impacting diagnostic accuracy. While real-time noise suppression can mitigate these issues, it comes at a high computational and financial cost. This paper introduces a method that includes a complete way to clean up medical images by using a new Denoising Generative Adversarial Network (D-GAN). The D-GAN architecture incorporates a recursive residual group-based generator and a discriminator inspired by PatchGAN. The recursive residual group-based generator and the Selective Kernel Feature Fusion (SKFF) mechanism are part of a new D-GAN architecture that makes denoising work better. A PatchGAN-based discriminator designed to improve adversarial training dynamics and texture modeling for medical images. These innovations offer improved feature refinement and texture modeling, enhancing the denoising of cardiac MRI images. allows the model to get a doubling context of local and global, informational, and hierarchical developed features located in the generator. Our technique outperforms other methods in terms of PSNR and SSIM. With scores of 0.837, 0.911, and 0.971 for noise levels of 0.3, 0.2, and 0.1, and PSNR scores of 29.48 dB, 32.58 dB, and 37.85 dB, the results show that the D-GAN method is better than other methods.

Keywords—Generative Adversarial Network; Image Denoising; Cardiac Magnetic Resonance Imaging; Recursive Residual; Deep learning.

I. INTRODUCTION

One of the most significant and beneficial imaging modalities for the noninvasive evaluation of the shape and function of the heart is cardiac magnetic resonance imaging (CMR). The ECG gating wrong action that the patient was exposed to arrhythmia and inability to hold his breath for long caused artifacts and noises during the acquisition of CMR images, such as Gaussian noise, Rician noise, and spike noise [1]-[5], which will greatly affect the cardiovascular image of the patient's diagnosis [6]-[11]. The stopping noise in time is handled well through the process acquisition, but the expensive demands and requirements of the hardware are very costly. In recent years, deep learning has emerged as a powerful tool in image processing, providing cost-effective solutions for tasks like denoising, super-resolution, and intelligent image analysis [12]-[41]. The concept of image denoising is considered a classical subject in computer image scope and an essential part of the actual image processing apps [42-47]. Image denoising used to be done in a lot of

different ways, including non-local self-similarity [48], sparse representation [49], Markov random field (MRF) [50], and many more. However, the primary problem with the prior method is its high optimization time [51].

Chen et al. [52] came up with the idea of a trainable nonlinear response diffusion (TNRD) version. This version uses a feed-ahead deep network structure and a set number of gradient descent inference steps to finish. A multi-layer perceptron (MLP) [53] has been shown to work well in the field of image denoising, where machines are becoming smarter and deeper learning methods are being developed. Yang et al. came up with a new way to clean up CT images by using the Wasserstein distance and perceptual similarity in a Generative Adversarial Network (GAN) [54]. The framework, which is based on denoising output capabilities in a predefined characteristic area, uses real picture functions to efficiently suppress noise.

The previous method's disadvantages that cannot be ignored are that their reliance on education is precise fashions for wonderful noise levels, and the effect of denoising on noise images of different levels is limited. Zhang et al. suggested a new way of thinking by looking at image denoising as a basic discriminant learning task [55]. They focused on the idea of using a feed-forward denoising convolutional neural network (DnCNN) to separate noise from the noisy image. By integrating batch normalization and residual mastering, this method streamlines the education procedure and enhances denoising talent. Deep learning, especially inside the realm of artificial intelligence, transcends those barriers via hastily and autonomously denoising medical images. In this paper To make denoising work better and use less computing power in medical imaging, we present a model for denoising a generative adversarial network (D-GAN) hybrid design. It includes our generator model made by enhancin MIRNeT [56] and a new discriminator that is based on the Patch GAN approach. In the field of scientific image diagnostics processing, using low-resolution (LR) images with less detail is challenging because they don't have enough texture information, which could make it harder to accurately diagnose heart problems [57].

As a result, an LR cardiac image needs to be changed into HR images of free-respiration long-axis and short-axis CMR images with very high quality and resolution. Harris and Goodman et al. highlighted super-resolution [58], [59], and emphasized the sizeable mapping correlation between LR



images and HR images. By utilizing deep knowledge of methodologies to establish these mapping relationships through large-scale image training, we can structure true HR images using LR images. Dong et al. were the first to make software with a deep learning version to handle the super-resolution mission [60]. They used a three-layer convolutional neural network (CNN) [61] to figure out how the complex mapping dynamics between LR and HR images work. The incorporation of Mean Squared Error (MSE) as a loss function ensures the inclusion of high-quality images.

Still, because the MSE is a loss function, the image's input resolution may mean that there isn't enough high-frequency texture information. This is especially important when working with very high image resolutions. Ledig et al. introduced Generative Adversarial Networks (GAN) as a device for image super-resolution [62]. The mapping relationship between the LR and HR images is figured out by looking at how the generator and discriminator work together. A new perceptual loss function is also added to improve the details of image textures.

However, the original training of GAN is unstable; it is easy to introduce non-existent features into the generated image, leading to a low accuracy in image reconstruction. This persistent problem has posed a giant hurdle that GAN has been facing. Arjovsky et al. introduced the concept of WGAN to address the instability issues associated with traditional GAN training [63]. If you use an almost certain discriminator to improve the made tool and lower the Wasserstein distance, the made image distribution will gradually match that of the real image. Gulrajani and his team conducted this study. This new technology called WGAN has made things better, but the first versions of it had problems, like using bad samples and having trouble getting the models to agree [64]. As a result, changes were made to WGAN. For example, an excessive learning of charge gradient penalty was added to each pattern to speed up convergence, and the Adam optimizer was used to improve overall performance and the quality of the outputs.

Ran et al. introduced a method called RED-WGAN; this method relied on the Wasserstein generative adversarial network to decrease the noise in the brain MR image [65]. The network generator has shown the residual automatic coding network as a way to look for similar structures between adjacent slices. While successful in noise discounting, this approach does not particularly address image reconstruction at a magnified scale. Sood et al. use the SRGAN method on MR images of the prostate to get an impressive $8\times$ fine-decision reconstruction. Finding the best MOS, which is what this method gives you, is very subjective, and the PSNR/SSIM results are not what you were hoping for [66]. A cascaded generative disagreement network was created by Han et al. to give each cell in the microscope image a high level of improvement. This method focuses solely on content loss and adversarial loss functions. The DnSRGAN [67] says that it first cleans up the heart image to a good standard, then enlarges it by 4 times, and finally uses gradient penalty and more WGAN loss to make the reconstruction more accurate. A lot more advanced image denoising is available in GANs. However, the more advanced image denoising that was available before had some

problems, like mode collapse and instability during training. To name these issues, we introduce a new Denoising Generative Adversarial Network (D-GAN) that developed feature refinement and computational efficiency. This paper presents a D-GAN architecture with a recursive residual group-based generator and a discriminator that is based on Patch GAN. These innovations offer improved feature refinement and texture modeling, enhancing the denoising of cardiac MRI images.

1. A new D-GAN architecture that combines advanced feature refinement with the recursive residual group-based generator and the SKFF mechanism makes denoising work better. .
2. A PatchGAN-based Discriminator designed to improve adversarial training dynamics and texture modeling for medical images.
3. Meticulous refinement of the Resblock architecture through iterative, exceptional tuning and structural optimizations ensures seamless integration in the overarching generator framework.
4. A thorough evaluation of our D-GAN on CMR datasets, demonstrating superior denoising performance compared to conventional models.

II. RELATED WORK

A. Denoising Convolutional Neural Network

Convolutional neural networks (CNNs), a promising alternative for medical image denoising, have emerged. By using supervised CNN models, it is easy to get back images that are free of noise from noisy observations by looking at the redundancy that is present in large datasets [68]. Zhang et al. [55] proposed a denoising convolutional neural network (DnCNN) [69] as a means to enhance overall task performance. To avoid problems with gradient dispersion that come up when the network layer is deepened, DnCNN doesn't learn the noise image directly. Instead, it trains the network using the L2 norm [70] of the noise output as the loss function. DnCNN specializes in residual learning. The BN layer and residual learning are used together by the network to make the design work better and get rid of noise from images without specific noise. Zhao et al. [71] are based on the SRGAN architecture. Feed-forward denoising convolutional neural networks (DnCNNs) use a gradient penalty (GP) method to deal with the discriminator gradient vanishing problem. They also use a feed-forward denoising neural network to clean up the CMR image before it is sent in.

B. Generative Adversarial Networks

Goodfellow et al. [72] delivered a groundbreaking concept of an adversarial network for image making, revolutionizing the landscape of deep learning with its profound effect. This novel idea provides for image processing by network. Scott et al. [73] proposed a new model based on the generative adversarial networks concept. The author suggests that GANs can produce the best work in image processing for text design. Liu et al. [74] applied a study to produce handwriting features through the usage of multiscale and multi-class conditions to GANS. Currently,

the best design for image-processing apps in deep learning is CNN. DCGAN [75], which is considered one of the most important attempts at merging CNN with GAN, and Yan et al. [76]. Use DCGAN to do image superresolution, denoising, and deconvolution. Although DCGAN has a sturdy structure, it does not provide a super solution for the problem to diminish it in the source for the training stability of GAN. During the DCGAN training process, however, it remains important to balance the G and D differences from DCGAN [77]-[85]. WGAN carefulness [63] is a big step toward making GAN training more stable by using a single loss function that is centered on the Wasserstein Distance. The unique idea behind WGAN is to improve training methods by using Wasserstein distance. To maintain Lipschitz continuity in the network, we must clip the weights within a specific range. This base helps make GANs more stable, and WGAN-GP [86] strengthens its position as a major step forward in the process of removing noise from neural networks. As more research and development is done in denoising and neural networks, this paper presents a new version of image denoising that is expected to provide competitive and promising results in this rapidly changing field of study.

III. METHODOLOGY

This part describes the steps that are taken in a planned way to turn raw medical images into high-fidelity copies, which solves the common problem of noise that can hide important diagnostic information. The sections that follow explain why each choice was made, how the generator and discriminator models are put together, and how the training process works in more detail. Our approach is a plan for how to use GANs to their full potential in medical imaging. Our ultimate goal is to help make diagnostic tools in healthcare practices more accurate and reliable. Fig. 1 serves as our architectural cornerstone, providing a succinct visual representation of the D-GAN overarching structure. We designed our generator architecture to seamlessly fuse input images with denoised outputs, encapsulating the essence of our holistic denoising approach.

The sections that follow break down the D-GAN's parts and give a thorough explanation of our generator, explaining their main functions and roles, starting with the basic Residual Block (Res Block) and moving on to the more advanced Dual Attention Unit (DAU) and Multi-Scale

Residual Block (MRB). We meticulously engineer each element to refine feature maps and elevate denoising performance. As we move through the Recursive Residual Group (RRG), our architecture is a hierarchical refinement of features that shows a full understanding of the complexities of medical data while carefully balancing the need for fast computing.

Along with our generator model, we show how good our Patch GAN-based Discriminator Model is at telling the difference between things. This model is set to improve the competitive training dynamics that are necessary for high-fidelity denoising results. Furthermore, we present empirical validation through a comprehensive analysis of experimental results.

A. Generator Model

We developed an architectural framework that revolves around the Recursive Residual Group (RRG). This iterative implementation endows our generator with formidable capabilities tailored specifically for the denoising of medical images. At the heart of our framework lies a final convolutional layer, responsible for generating the denoised output. A residual connection seamlessly fuses this output with the input image, symbolizing the holistic representation of the denoising process. See Fig. 2, an overview of our generator model. The nuanced architectural details of our generator.

Residual block is the cornerstone of our generator architecture, inspired by the leading ResNet framework known for its effectiveness in various computer vision tasks [87]. Each Res Block has two convolutional layers that are carefully designed to pull out complex and detailed information from the input image. This is followed by activation of a Leaky Rectified Linear Unit (Leaky ReLU). We use this activation because it can add robust non-linearity while protecting the deep learning network from the risk of vanishing gradients. Leaky ReLU is excellent at avoiding the saturation problem, where a neuron can get “stuck” at one of its extreme values. However, the vanishing gradient problem is still a challenge with leaky ReLU. A way to address this in really deep networks is to connect each layer to several other layers.

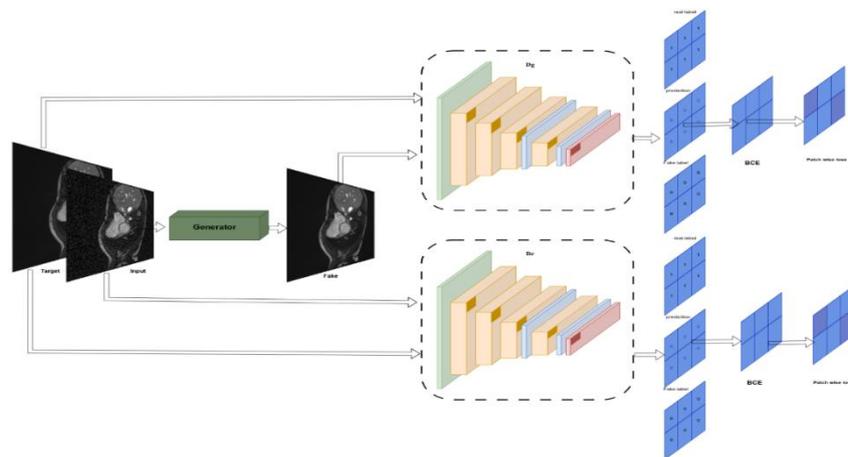


Fig. 1. The proposed network, D-GAN, features a novel architectural cornerstone that provides a succinct visual representation of the D-GAN overarching structure, which learns enriched feature representations for denoising and enhancement

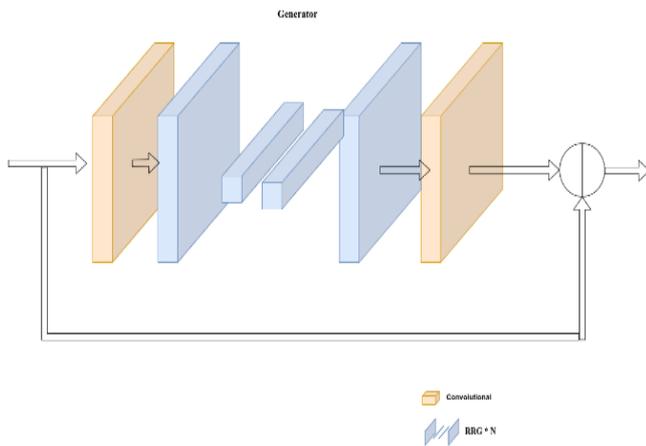


Fig. 2. Our generator model highlights the architectural design. The multi-scale residual block (MRB) is at the center of the generator. Its main branch is responsible for keeping high-resolution representations that are accurate in space across the whole network, and its set of parallel branches work together to provide better contextualized features. Through selective kernel feature fusion (SKFF), it's also possible for information to flow between parallel streams. This lets high-resolution features be combined with low-resolution features and vice versa

This allows for the gradient to skip some layers that would otherwise dilute it. This helps with robust learning and feature expression. The block has the ability to store and amplify fundamental information from the input source, all while extracting additional features. We have carefully improved our Res Block architecture by trying different things and making architectural changes. This makes sure that it works perfectly with the generator framework. See Fig. 3. This illustrative representation encapsulates the essence of the Res Block. Connection reaffirms its pivotal role in feature extraction and conservation within our generator model.

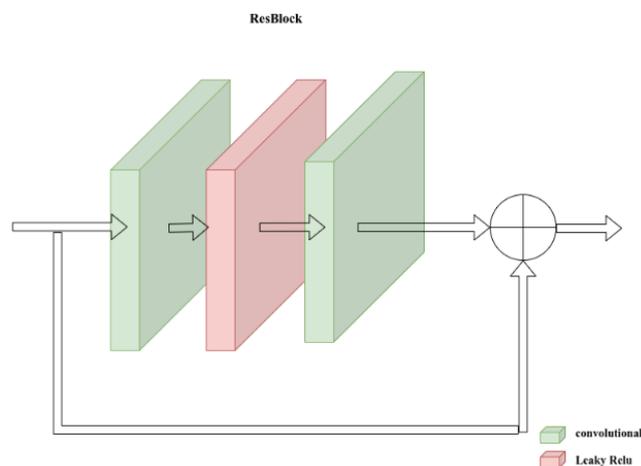


Fig. 3. Res Block is the cornerstone of our generator architecture

Selective Kernel Feature Fusion (SKFF): takes into account the adaptive neuronal capacity of neurons in the visual cortex, which means that these neurons can change their receptive field in response to outside stimuli [88]. Such CNN adaptability could be enhanced by introducing a nonlinear process in which self-attention abilities are used to fuse features acquired at different resolutions. This method, called SKFF, lets the network change receptive fields on the fly and improve feature representation by using two main operations, fuse and select, as seen in Fig. 4. The Fuse

operation amalgamates multi-scale features from parallel convolution streams using element-wise summation, denoted $L = L_1 + L_2 + L_3$. Its architecture is characterized by the implementation of one of the new residual connections, a connection that bypasses one or more convolutional layers. As a result, the as $L = L_1 + L_2 + L_3$. Following this, instead of employing global average pooling (GAP), we utilize global maximum pooling (GMP) to compute channel-channel statistics, generating a 3D output tensor L with dimensions $H \times W \times 3$ [89]. Channel-channel statistics g are then derived from L , represented as $g \in R^{1 \times 1 \times C}$. Subsequently, a dense feature S_0 with dimensions, $1 \times 1 \times r$ is generated using a channel-downscaling convolution layer, where $r = \frac{C}{4}$ for all experiments. The Select operator performs the SoftMax function to the feature descriptors v_1, v_2 , and v_3 , yielding attention activations s_1, s_2 , and s_3 , respectively. These activations adaptively recalibrate multi-scale feature maps L_1, L_2 , and L_3 . The recalibrated features are aggregated using the following equation: $U = s_1 \cdot L_1 + s_2 \cdot L_2 + s_3 \cdot L_3$. Here, U represents the aggregated feature map. Receptive fields are constantly adjusted by the SKFF module using fuse and select techniques, effectively integrating multi-scale features and enhancing feature representation. Notably, SKFF employs approximately 6 times fewer parameters compared to concatenation-based aggregation while producing superior results, as demonstrated in the experiments in Section 4.

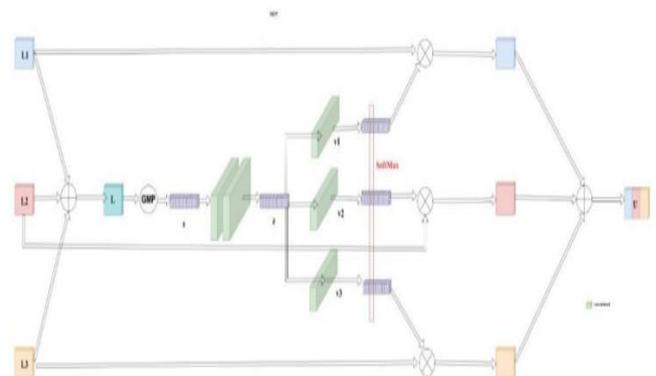


Fig. 4. This technique is known as selective kernel feature fusion (SKFF). It operates on features from multiple convolutional streams and performs aggregation based on self-attention

Dual Attention Unit (DAU): The DAU is an important part of our medical image denoising generator architecture; its job is to improve feature extraction within convolutional streams [90]. It works with the SKFF block to combine data from different resolutions and makes it easier for data to flow within a feature tensor along both spatial and channel dimensions. The DAU is very important for selectively recalibrating features and improving the model's ability to tell the difference between things by making informative features louder and noise and irrelevant information quieter. This recalibration process is facilitated by the integration of channel attention and spatial attention processes.

Channel Attention (CA) Branch: The DAU CA branch utilizes the inter-channel relationships of convolutional feature maps. Mathematically, given a feature map M with dimensions $H \times W \times C$, the squeeze operation aggregates spatial information through global average pooling, resulting

in a feature descriptor d of dimensions $1 \times 1 \times C$. This descriptor is processed further using two convolutional layers, and then a sigmoid activation is applied, yielding activation maps, d , of the same dimensions. Ultimately, the output of the CA branch is obtained by element-wise multiplication of the original feature map M with the activations, d ., effectively emphasizing informative channels while suppressing noise.

Spatial Attention (SA) Branch: At the same time, the SA branch exploits spatial dependencies within convolutional features. The global pooling and maximum pooling procedures are applied individually along the channel dimensions of the input feature map M , generating feature maps. \bar{f} of dimensions with dimensions $H \times W \times 2$. These maps are concatenated and processed through convolutional layers followed by sigmoid activation, resulting in a spatial attention map with dimensions $H \times W \times 1$. This map is then used to modify the original feature map M , selectively enhancing regions of interest while suppressing noise. Fig. 5. visually illustrates the operational principles and structure of DAU. The DAU makes noise reduction better by recalibrating certain features along spatial and channel dimensions. This makes it easier to tell the difference between features.

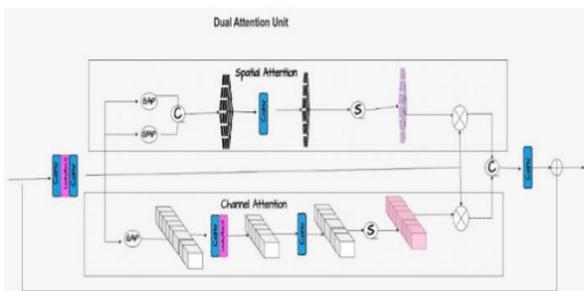


Fig. 5. Dual Attention Unit (DAU) incorporating spatial and channel attention mechanisms

Multi-Scale Residual Block (MRB): At the center of the complexities of our generator design lies the multi-scale residual block. We carefully created this basic building block to improve feature maps at different sizes. By combining our two systems, we were able to make medical images more detailed. These mechanisms are the Dual Attention Unit and the Selective Kernel Feature Fusion. In Fig. 6, the MRB is shown to be part of the generator's complex structure. It coordinates the improvement of the feature map on different levels, creating a complete idea for feature amplification. By fully utilizing the DAU mechanism, the MRB enhances the model's dual attention by prioritizing specific details from the feature space. At the same time, adding the SKFF mechanism gives the MRB the ability to do nuanced selective fusion, which means it can pick up both small details in the local area and the big picture. The MRB with DAU in the focus area improves the selectivity of basic classifiers, which lets fine-detail treatment be preferred and makes the image better. As a result, SKFF stands for smart knowledge forget-free mechanism, which allows MRB to perform the exact fusion of styles, i.e., mixing local accents and global contextual cues. The combination of DAU and SKFF mechanisms in the MRB represents a complex way to improve features that goes beyond common approaches. By using dual attention and

selective fusion together, MRB increases the generator's unique flexibility, making it easier to improve the complex patterns found in medical images. So, MRB is a new way to develop medical images that are less noisy by using generative adversarial networks.

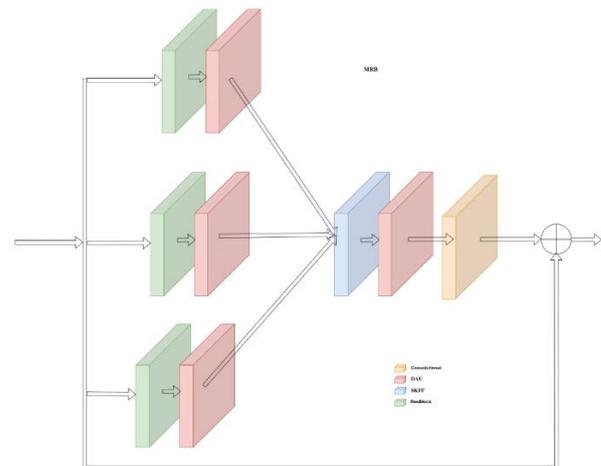


Fig. 6. Update The architecture of the MRB block lies at the heart of the complexities of our generator design, specifically the multi-scale residual block. By fully using the DAU mechanism, the MRB improves the model's dual attention aspect, which lets details from the feature space be prioritized in a selective way. At the same time, adding the SKFF mechanism gives the MRB the ability to do nuanced selective fusion, which means it can pick up both small details in the local area and the big picture

In our generator architecture, the Recursive Residual Group (RRG) is like a key player in a complicated tapestry. It makes sure that the integration of many Multi-Scale Residual Blocks (MRBs) works well together. This well-planned integration is carefully made to create a hierarchical refinement of features, which is the most important thing to do to capture the subtleties of medical images while also making the computer work efficiently. Fig. 7 serves as a vivid exposition of the recursive application of MRBs within the RRG. As the strategic leader, the RRG is in charge of deploying MRBs in stages, which leads to a gradual improvement of features. This recursive paradigm gives the model the ability to recognize and combine complex patterns in a hierarchical way, which helps us get a full picture of the medical data that lies beneath. In addition to helping to improve features, using multiple MRBs in the RRG makes the model much better at removing noise from medical images.

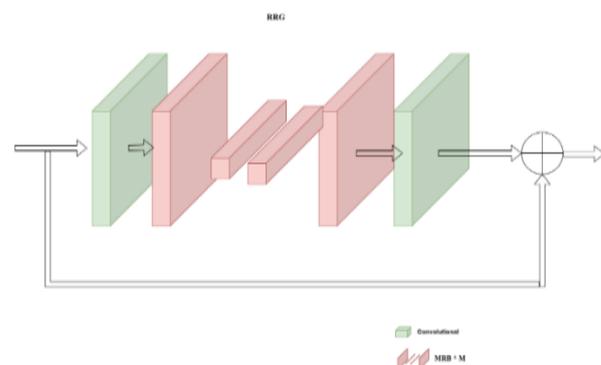


Fig. 7. The Recursive Residual Group (RRG) architecture plays a key role in coordinating the joining of several Multi-Scale Residual Blocks (MRBs). We meticulously craft this orchestrated integration to cultivate a hierarchical refinement of features

B. Discriminator Model

In the context of image generation problems, conventional loss functions such as L2 and L1 often result in blurry outputs, lacking high-frequency crispness. However, these losses effectively capture low frequencies, motivating the exploration of alternative frameworks that leverage this capability. Our discriminator architecture is based on this observation. It was inspired by the patch GAN approach [91], which models high-frequency structure and relies on an L1 term to make sure low-frequency correctness. We design our discriminator architecture, known as patch GAN, to penalize structures at local image patch scales. The goal of this specialized discriminator is to decide whether each $N \times N$ patch in an image is real or fake. To do this, it averages each response and uses convolutional processes across the whole image. Down-sampling blocks in the discriminator gradually make the input's spatial dimensions smaller. This lets the model focus on local features. Fig. 8 illustrates this down-sampling process.

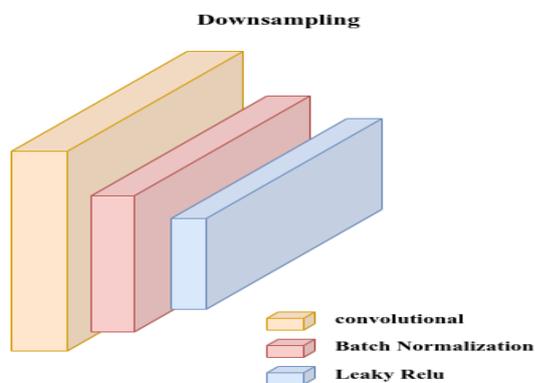


Fig. 8. Down-sampling blocks in the discriminator gradually make the input's spatial dimensions smaller. This lets the model focus on local features

There is batch normalization, leaky rectified linear unit (Leaky ReLU) activation, and then each down-sampling block is a convolutional layer. The Patch-GAN discriminator architecture includes multiple up-sampling blocks, with the number of filters increasing in each successive block. Before the last convolutional layer, we use zero-padding operations to increase the field of view and improve the ability to tell the difference. This layer produces an output tensor with dimensions (batch size, 30, 30, 1), where each 30×30 patch corresponds to a 70×70 portion of the input image. Our Patch-GAN discriminator also uses a Markovian approach, which means that the image is modeled as a Markov random field. This assumption assumes independence between pixels separated by more than a patch diameter, consistent with well-established principles in texture modeling and style, as shown in Fig. 9. In Experimental Section 4 of this paper, we demonstrate the effectiveness of our Patch-GAN discriminator, showing its ability to produce high-quality results even with smaller patch sizes. This scalability feature enables faster training and application on randomly large images, making it a versatile tool for various image processing tasks. The summary features of our discriminator model:

- Using down-sampling blocks to gradually reduce the number of spatial dimensions helps the discriminator focus on local features.

- Composite Input, The discriminator receives two input images, real (target) and generated (fake), concatenated along the channel axis for realistic assessment and classification.
- Dual-Input Design lets the discriminator accurately judge the authenticity of generated images, which makes adversarial training more effective.
- Real and Generated Assessments, The discriminator loss assesses contributions from both real and generated data, evaluating the model's ability to discern authentic images and identify generated (fake) pictures.

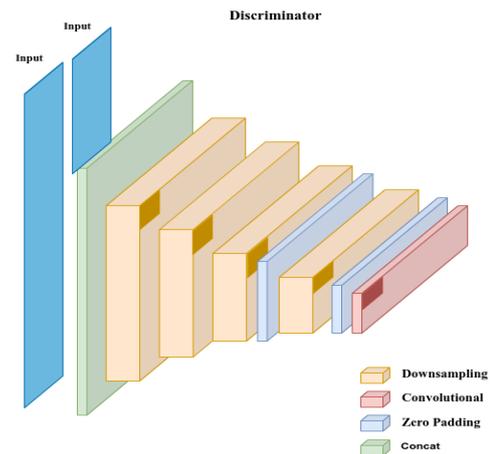


Fig. 9. The discriminator architecture, known as Patch-GAN, aims to penalize structure at local image patch scales

IV. RESULTS AND DISCUSSION

A. Data Preparation

In the planning stage for our medical image denoising model, we do a lot of careful work to make sure that the training dataset is reliable and useful. We follow strict ethical rules and put patient privacy and data security first when using the ACDC (Automated Cardiac Diagnosis Challenge) dataset from the University Hospital of Dijon [92]. To train and develop our proposed denoising models, we split the dataset into 70% (101 images) for training, 15% (22 images) for validation, and 15% (22 images) for testing. Before training. This dataset, encompassing diverse cardiac pathologies and physiological parameters, serves as the foundation for our denoising model. The precise annotations made by clinical experts help us set a reliable ground truth for our denoising task by choosing clean images from the ACDC training dataset. To simulate real-world noise patterns in medical images, it becomes imperative to introduce noise into the clean images. We can emulate various noise characteristics prevalent in medical imaging by adopting a Gaussian-distributed random noise approach. In empirical evaluation, we obtained occurrences of Gaussian distributions with noise factors of 0.5, 0.1, 0.15, 0.2, and 0.3 for medical imaging: small-scale, medium-scale, and big-scale noises, respectively, which allow us to cover the range of realistic noises in medical imaging. The noise factors are conventionally set out to account for different levels of noise, with low noise being represented by a value of 0.5, 0.1, or 0.15; moderate noise by 0.2; and high noise by a value of 0.3. This method allows our net to learn all the noise spectrum

from the medical images with the maximum possible frequencies. The calibrated noise level keeps the balance between image quality and noise level magnitude. A minimum noise factor of 0.2 was chosen on purpose as the best value for the system after testing. Fig. 10 is an example of the effect of noise addition, where a reference medical image is displayed next to its noisy variants with 0.2 as a noise factor. This image shows that the noise addition process has two goals: it adds real medical image noise to the training data and creates ground truth for the training noise reduction model.

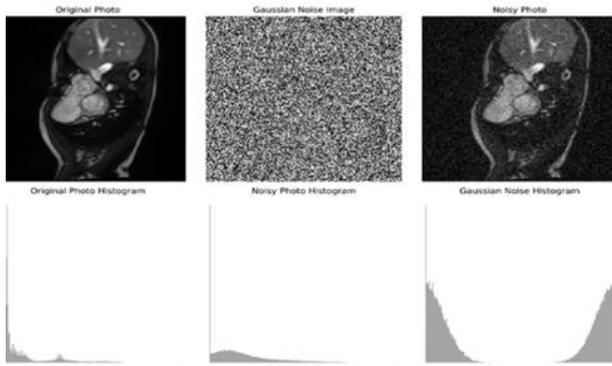


Fig. 10. The example illustrates the impact of noise addition, showcasing a clean medical image alongside its noisy counterparts generated with noise factors of 0.2

In addition to noise, protecting pixel values within a valid range is critical. The clipping process prevents defects caused by out-of-range pixel values. This meticulous attention to detail ensures that noisy images are not only realistic but also aids in effective noise reduction training. Batch processing achieves efficient data set management by dividing the data set into smaller batches. For both training and testing sets, it optimizes memory usage, facilitates parallel computation, and ensures optimized model updates based on multiple samples. We can process image data efficiently by converting image matrices to TensorFlow, which simplifies integration with our deep learning model.

B. Experimental Procedure

After curating a dataset consisting of images blurred by three different Gaussian noise distributions ($\sigma = 0.1$, $\sigma = 0.2$, and $\sigma = 0.3$), we embark on training two distinct models: the D-GAN model (Generative Adversarial Network (GAN) with Patch GAN discriminator) and our generator model only (see in Fig. 2) in the network was performed. Here was the experimental part that helped us evaluate the proposed method of denoising the given medical images with different noise types.

1) Comparison Training Setup for D-GAN Model and Generator-Only Networks

We will explain in more detail how to train both the GAN and generator-only networks below, including the different loss function formulations, optimization strategies, and parameter configurations. The goal of training both models was to show that the GAN model is better than the generator network at the task of medical image denoising. To show how well the GAN framework works at making high-quality denoised images, we will compare their training methods and rate how well they can remove noise from images.

• Training Setup for (GAN) Network

Optimization Strategy: SGDM can improve overall performance, make training more stable, and speed up convergence when used in denoising GANs. This makes it a good choice for difficult tasks like image denoising. The GAN network updates simultaneously during training, enabling the improved discriminator to interact with the improved generator. The stochastic gradient descent with momentum (SGDM) is the essential optimization algorithm for enhancing the model performance. The algorithm can correctly update and change parameters during the learning process to eventually minimize the difference between real and generated images.

Batch Size and Epochs: The correctly chosen batch size of 32 balances computational efficiency and other essential aspects of the training phase. The GAN models go through a careful training process for the duration of 30 epochs, making them well-prepared for the required parameter tuning and the eventual convergence to the optimal denoising performance.

Learning Rate and Momentum: The learning rate is carefully set to 0.0002 for the discriminator and 0.0001 for the generator. This allows parameters to be updated slowly and reduces oscillations during training. Furthermore, we adopt momentum values of 0.9 for the generator and 0.5 for the discriminator to incorporate historical gradients and enhance convergence stability.

Loss Function Formulation: The generator loss function has two parts: adversarial loss (GAN loss) and mean absolute error (L1 loss). These parts work together to make sure that the images that are generated are realistic and correct in terms of their structure. The GAN loss measures how well the model can create images that can't be told apart from real data, and the L1 loss measures how different pixels in the generated and target images are from each other. The total generator loss seamlessly integrates both losses, embodying a configurable weighting parameter (λ). We tested a range of values (50, 100, 200) to identify the optimal balance for specific datasets and tasks. When the set value of high $\lambda = 200$ is too high, the model overly prioritizes pixel accuracy, leading to less realistic images. Low conversely = 50; a lower value shifts the focus toward the GAN loss, resulting in more realistic images but sacrificing detail and fidelity to the actual data. The right $\lambda=100$ requires careful tuning and can significantly impact image quality.

$$gan\ loss = BinaryCrossEntropy (tf.ones\ like(D(G(input\ image))), D(G(input\ image))) \quad (1)$$

Complementing this, the L1 loss ($l_1\ loss$) evaluates the pixel-wise dissimilarity between the generated and target images:

$$l_1\ loss = MeanAbsoluteError(target, G(input\ image)) \quad (2)$$

$$total\ gen\ loss = gan\ loss + LAMBDA \times l_1\ loss \quad (3)$$

Notably, LAMBDA=100 is a judiciously chosen hyperparameter

The Discriminator Loss assumes an adversarial training shake with two parts: positive real data assessment and

straightforward fake data assessment. The real loss evaluates the discriminator's ability to discern authentic images.'

$$real\ loss = BinaryCrossEntropy(tf.ones\ like(D(real\ image)), D(real\ image)) \quad (4)$$

Conversely, the generated loss (generated loss) quantifies the discriminator's proficiency in identifying generated (fake) images:

$$generated\ loss = BinaryCrossEntropy(tf.zeros\ like(D(G(input\ image))), D(G(input\ image))) \quad (5)$$

The collective discriminator loss (total_disc_loss) encapsulates both real and generated losses:

$$total\ disc\ loss = real\ loss + generated\ loss \quad (6)$$

Training Execution: During training iterations, the generator produces denoised images, which are investigated by the discriminator. The calculated losses for the discriminator and the generator tell backpropagation and gradient descent how to change the parameters. This process of iterative refinement goes on for the set number of epochs. The training progress is carefully tracked, and full summaries are saved so they can be viewed and analyzed.

- *Training Setup for Generator-only Network*

Loss Function Utilization: Unlike the GAN network, the generator-only network undergoes training using the Charbonnier loss function. The generator can learn sophisticated strategies for noise reduction thanks to this loss function, which stands out for its robustness to outliers. The Charbonnier Loss:

$$Loss(I, I^*) = \sqrt{||I - I^*||^2 + \epsilon^2} \quad (7)$$

where I^* denotes the ground-truth image, and ϵ is a constant that we empirically set to 10^{-3} for all the experiments

Optimization and Training Execution: The generator-only network optimization approach and training methodology are essentially the same as those of the GAN network, but they are limited to the generator parameters. Over the set number of epochs, we use stochastic gradient descent with momentum to reduce the generator loss over and over again.

- *Evaluation Metrics*

In evaluating the denoising performance of the Generative Adversarial Network (GAN), two primary metrics were employed: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). PSNR is a number that measures the quality of denoising by comparing the strongest signal power to the strongest corrupting noise power. Higher values mean better denoising. Computed in decibels (dB), PSNR provides a straightforward measure of image fidelity. SSIM assesses structural similarity between denoised and ground truth images, considering luminance, contrast, and structure. SSIM values range from -1 to 1, with 1 representing perfect similarity and 0 representing no similarity, embodying perceptual aspects of human vision. The process of evaluating involves using the trained GAN model to clean up noisy medical images, calculating the

PSNR and SSIM between the cleaned up and original images, and collecting metrics across the dataset to get a general idea of how well the denoising worked. Higher PSNR and SSIM values signify superior denoising quality and structural similarity. These metrics are numerical ways to rate how well denoising works. They make it easier to compare different methods and help make GAN-based denoising algorithms better, which makes the results of medical image denoising more reliable. The Results:

The SSIM (Structural Similarity Index) serves as a crucial metric, measuring the similarities between denoised images and a ground truth image. According to the SSIM indicator, the results of the GAN approach are 0.837, 0.911, and 0.971 for noise levels of 0.3, 0.2, and 0.1, respectively. Moreover, the generator network, considering the same noise level in Fig. 11, obtained an SSIM score of 0.833, 0.891, and 0.934, respectively. This means that GAN usually gets better results than the generator network when it comes to noise levels, and it keeps impressing with its ability to keep image structure and fidelity better.

We commonly use PSNR (Peak Signal-to-Noise Ratio) to evaluate the quality of the denoised image. The higher values of PSNR indicate better image fidelity. The GAN-based model approach got PSNR scores of 29.48 dB, 32.58 dB, and 37.85 dB for noise levels of 0.3, 0.2, and 0.1, respectively, in our tests. On the other hand, the generator network got PSNR readings of 30.01 dB, 32.85 dB, and 36.02 dB, which are the different noise levels shown in Fig. 12. The GAN delivered higher PSNR values over all kinds of noise levels. The results show that the GAN consistently produced higher PSNR values across all noise levels, indicating better signal integrity when denoising.

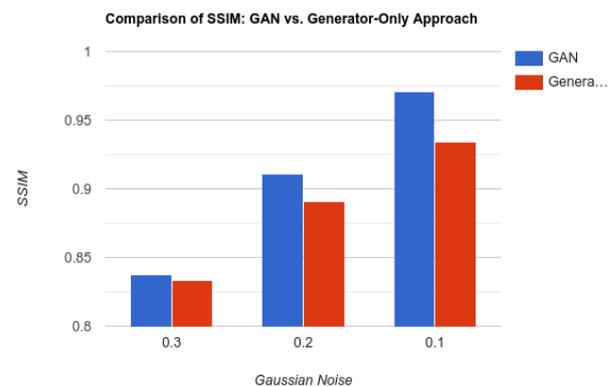


Fig. 11. Looking at the Structural Similarity Index (SSIM) scores between the images that have been cleaned up and their original images for GAN and generator networks

- *Loss Curve Figures*

We visualize the behavior and stability of the training process at a noise level of 0.2 for both the GAN and the generator network in figures of the loss function. This curve shows very clearly how generator and discriminator loss changes over epochs. It also shows how learning works and what level of optimality each network should have in this noisy situation. The loss curves act as a visual tool that helps to understand the learning process and helps monitor the fluctuations or tendencies that may impact the model's ability to clean noise.

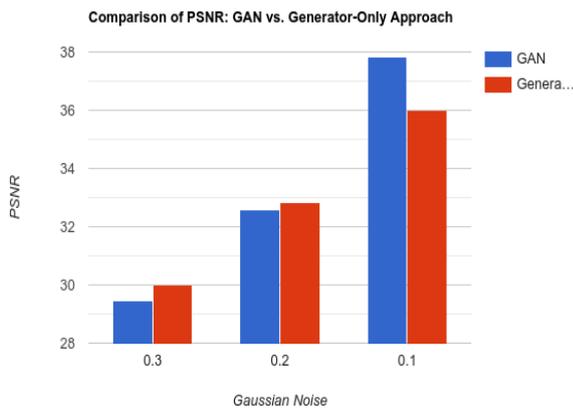


Fig. 12. We looked at the PSNR difference between the images that had been cleaned up and the original images for our model GAN and generator-only network

GAN Loss Curve: The loss curve for the GAN network illustrates the evolution of the in-traces both the decrease of generator loss and discriminator loss during the training of the network; see Fig. 13. It gives details about the interplay between the generator and discriminator, including the fact that they somehow remain closer to their optimal denoising performance. The curve shows how the loss functions change over time, showing how the model parameters are being finetuned over time and how that affects the effectiveness of denoising.

The present image in Fig. 13(A) denotes GAN training, where GAN loss is highlighted, particularly the GAN loss component. Binary cross-entropy calculates the generator's loss, assessing its "strength" from the discriminator's perspective. While the training process is going on, the generator is trying more and more to reduce this loss or error function, successfully raising the level of realism in its generated data. The optimization process is done using the Adam optimizer, with the learning rate set to $2e-4$ and the β_1 parameter set to 0.5.

The image in question is Fig. 13(B). Figure pictorially displays the evolution in discriminator loss during GAN training. The indicator emphasizes real or generated data and uses binary cross entropy to help the discriminator distinguish between the two. In the meantime, the discriminator learns to be more and more skilled in distinguishing real from fake images, which ends up making the loss of the discriminator smaller and smaller. The discriminator optimizer is being used to train this generator. It is a parameter Adam optimizer with a learning rate of $2e-4$ and a β_1 parameter of 0.5.

Image Fig. 13(C): presentation of the generator L1 loss in GAN training. The mean absolute error calculates the average distance between the generated output and the target data. In the course of training, the generator aims to shrink this loss while enhancing the similarity between the generated and target data. The optimization is accomplished with the help of the Adam optimizer, which gives $2e-4$ for a learning rate and a β_1 parameter of 0.5.

Image Fig. 13(D) is the total generator loss in GAN training, the sum of the mean absolute error and the gradient information, also known as the L1 loss. You can see how well

the generator is at tricking the discriminator by looking at the loss of GAN. The comparison criterion is L1, which shows the difference between the generated (output) data and the target data. One of the components of total generator loss is the sum of these parts with the consideration of a hyperparameter (LAMBDA). Such mistakes act not only as a guiding force in the optimization process but also as a target for minimization to improve the quality of a generator.

Generator Loss Curve: At the same time, the loss curve for the generator network shows in great detail how the generator loss function changes during training (see Fig. 14). It clarifies the optimization path of the generator as it strives to shrink the discrepancy between the generated and target images. The curve gives us useful information about how the generator network converges and stays stable, which helps us fully understand how well it can remove noise.

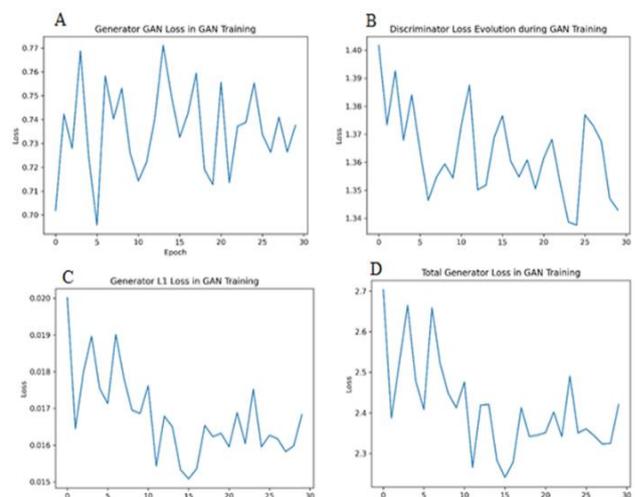


Fig. 13. The trends (A, B, C, and D) show very clearly how the generator and discriminator loss functions have changed over time and across epochs. This shows how well each model does in these situations in terms of how it learns and, finally, how well it optimizes

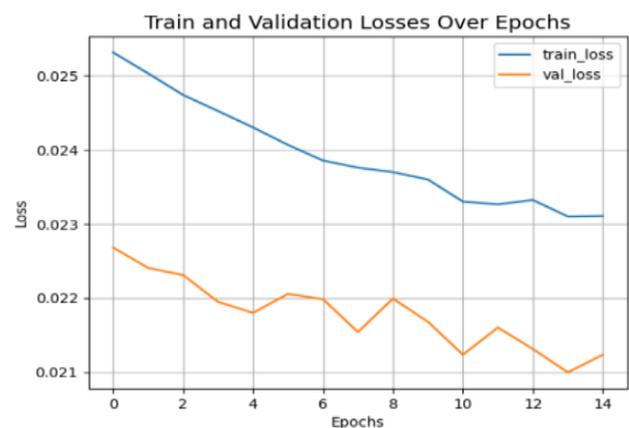


Fig. 14. The Epoch-wise Train and Validation Losses. We compute the losses using an invented custom Charbonnier Loss function, which effectively handles outliers by adding an epsilon term for cases of overfitting. The Adam optimization is used with the size set by the β_1 (0.9) and β_2 (0.999) factors. Adaptive learning rates and a schedule function (lr_schedule_fn) accomplish this

By putting these loss curves next to the quantitative evaluation metrics and sample output images, we can get a full picture of how well the GAN and generator networks

remove noise at a level of 0.2. The integration of visual representations enhances the interpretability of the results, enabling a deeper analysis of the model's training dynamics and denouncing its efficacy.

2) Comparison Network (MIRNet, D-GAN, Generator Only) Networks

With the same ACDC dataset and three different Gaussian noise distributions ($\pi = 0.1, \pi = 0.2, \text{ and } \pi = 0.3$), we trained all the models so that we could compare each network on its own (MIRNet, D-GAN, and generator only). Table I represents the results of each network. The table displays the averaged quantitative results, demonstrating that the proposed D-GAN network outperforms other networks, as indicated by the results of PSNR and SSIM. We show some examples of images that were only made by the D-GAN Recursive Residual Group-based Generator Model to give the network its power; see Fig. 15. The following examples demonstrate the effectiveness of our model approach in denoising, resulting in visually pleasing images. The sample outputs serve as representative examples of the model's performance in reducing noise and enhancing image quality across various medical imaging scenarios. By looking at these sample outputs, viewers can judge the perceptual fidelity and structural coherence of the denoised images, which gives them useful information about how well the GAN network can denoise.

TABLE I. THE AVERAGE (PSNR / SSIM) MEASUREMENTS AND QUANTIFIABLE COMPARISONS ARE GENERATED UNDER GAUSSIAN NOISE WITH $\sigma = 10\%, 20\%, \text{ AND } 30\%$

Sample rate	Metrics	D-GAN	MIRNet	Generator only
10%	PSNR	37.85	35.67	36.02
	SSIM	0.971	0.9367	0.934
20%	PSNR	32.58	31.91	32.85
	SSIM	0.911	0.884	0.891
30%	PSNR	29.48	28.62	30.01
	SSIM	0.837	0.814	0.833

3) Comparison of D-GAN Network with Other Methods

On the ACDC dataset, we recreate Gaussian commotion with $\sigma = 5\%, 10\%, \text{ and } 15\%$. We compare our proposed method (D-GAN) with other state-of-the-art denoising strategies in D2S [93], counting profound pictures earlier (DIP) [94], Self2Self (S2S) [95], BM3D [96], and VBM4D [97]. For D2S and S2S, we utilize the same learning rate specified over and tune the number of cycles on our datasets. We adopt the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM) as evaluation metrics. Table II shows the quantitative data on the ACDC dataset. The proposed strategy accomplishes execution better than other strategies.

Table II displays the quantitative results of the ACDC dataset. The proposed method achieves much better performance than other methods, D2 and VBM4D. In particular, D2 does better than VBM4D for other types of noise, especially when there is a lot of it, and it gets results that are similar to VBM4D for Gaussian noise. Also, D-GAN always does better than D2S and S2S, which shows that data from other time frames can greatly improve the performance of denoising models. Single-image methods such as BM3D

and S2S use only one noisy image, so they do not have enough information to recover details corrupted by noise, resulting in an overestimation of blurring. The DIP method introduces significant structural artifacts at high noise levels. Although some details are recovered from adjacent frames, VBM4D also introduces subtle artifacts to the noisy images. There are more detailed structures that can be recovered by the D2S method from the DIP, S2S, BM3D, and VBM4D approaches. However, the D-GAN method shows that PSNR/SSIM gives better results than all of them, both in terms of the quality of the image and how well it fits our perception of it.

TABLE II. QUANTITATIVE RESULTS ON ACDC DATA SET FOR GAUSSIAN

Method	Gaussian noise		
	$\sigma = 5\%$	$\sigma = 10\%$	$\sigma = 15\%$
	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
NOISY	26.02/0.769	20.00/0.518	16.48/0.369
BM3D	32.32/0.953	28.54/0.905	26.45/0.860
VBM4D	32.54/0.957	28.96/0.911	26.88/0.863
DIP	26.95/0.875	25.55/0.815	23.48/0.718
S2S	30.41/0.942	28.45/0.912	26.90/0.880
D2S	32.16/0.960	30.26/0.936	28.22/0.887
D-GAN	38.93/0.963	37.85/0.971	34.68/0.932

This study provides good results in denoising medical images of the heart, but there is One significant limitation of GAN-based denoising methods is their handling of different noise distributions. Most current implementations primarily address Gaussian noise, which is common in MRI, but struggle with other types such as Poisson noise, often encountered in CT imaging. This disparity can lead to suboptimal performance when applied across various modalities. Additionally, GANs may not generalize well between imaging techniques, resulting in degraded image quality when moving from one medical modality to another, such as from MRI to CT. Despite the progress, it's crucial to recognize the challenges that remain within this observation. Although the performance of the ACDC dataset shows promise, the proposed method needs to be tested on a wider range of datasets to make it more reliable and useful. To create these models, we utilized the TensorFlow library together with the Python-based NN Keras toolkit. We used Nvidia GTX 1080 Ti GPUs running on a Linux cluster with 32GB of RAM for training and testing.

V. CONCLUSION

We present a Denoising Generative Adversarial network (D-GAN) architecture with a recursive residual group-based generator and a discriminator that is based on Patch GAN. These innovations offer improved feature refinement and texture modeling, enhancing the denoising of cardiac MRI images. Our technique can help mitigate the effects of patient movement, leading to more stable and consistent images and a better diagnosis. Clearer images allow for more accurate assessments of cardiac conditions. Using the PSNR experiment and SSIM as indicators, the results demonstrate that the new model effectively reduces image noise and enhances the visual perception of the image.

DATA AVAILABILITY

The ACDC (Automated Cardiac Diagnosis Challenge) dataset from the University Hospital of Dijon in this study, adhering to rigorous ethical guidelines that prioritize patient privacy and data confidentiality [92]. links to online <https://www.creatis.insa-lyon.fr/Challenge/acdc/databases.html>

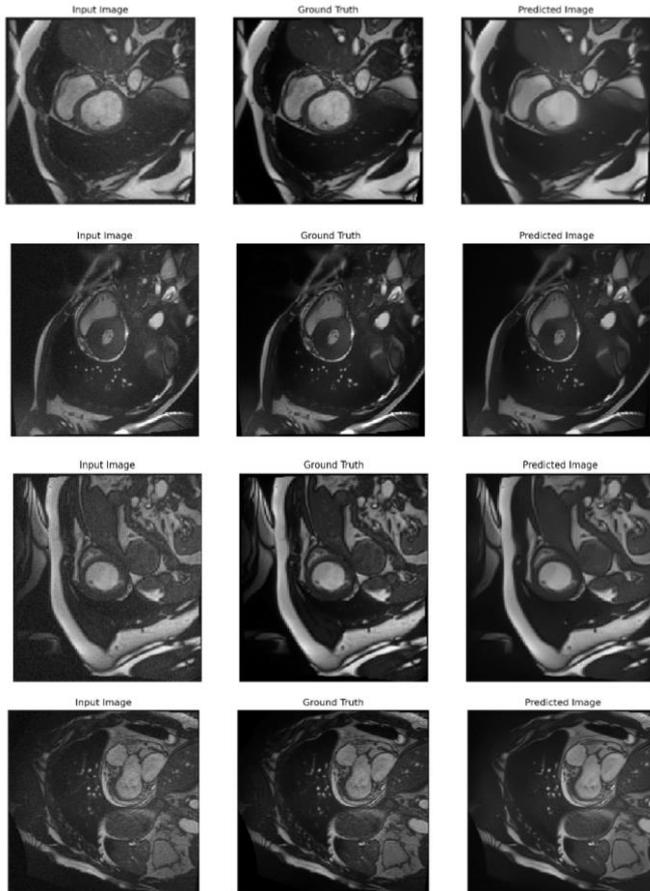


Fig. 15. Example output for images generated exclusively by the RRG-GM network. These sample outputs exemplify the denoising efficacy of our model approach, showcasing visually appealing denoised images produced by the model.

REFERENCES

- [1] A. Bustin, N. Fuin, R. M. Botnar, and C. Prieto, "Compressed-sensing to artificial intelligence-based cardiac MRI reconstruction," *Frontiers in Cardiovascular Medicine*, vol. 7, p. 17, 2020, doi: 10.3389/fcvm.2020.00017.
- [2] K. Wong, G. Fortino, and D. Abbott, "Deep learning-based cardiovascular image diagnosis: a promising challenge," *Future Generation Computer Systems*, vol. 110, p. 802-811, 2020, doi: 10.1016/j.future.2019.09.047.
- [3] G.K. Thakur, A. Thakur, S. Kulkarni, N. Khan, and S. Khan, "Deep Learning Approaches for Medical Image Analysis and Diagnosis," *Cureus*, vol. 16, no. 5, p. e59507, 2024, doi: 10.7759/cureus.59507.
- [4] T. F. Ismail, W. Strugnelli, C. Coletti, M. Božić-Iven, S. Weingärtner, K. Hammernik, T. Correia, and T. Küstner, "Cardiac MR: From theory to practice," *Front. Cardiovasc Med*, vol. 9, p. 826283, 2022, doi: 10.3389/fcvm.2022.826283.
- [5] M. A. U. Naser and A. H. H. Al-Asadi, "Hybrid Residual Blocks - Transformers Improving Cardiac MRI Segmentation in Deep Learning," *2023 4th International Conference on Communications, Information, Electronic and Energy Systems (CIEES)*, pp.1-7, 2023, doi: 10.1109/CIEES58940.2023.10378777.
- [6] M. A. Kareem *et al.*, "Predicting post-contrast information from contrast agent-free cardiac MRI using machine learning: Challenges and methods," *Front Cardiovasc Med*, vol. 9, p. 894503, 2022, doi: 10.3389/fcvm.2022.894503.
- [7] Q. Zhang *et al.*, "Toward replacing late gadolinium enhancement with artificial intelligence virtual native enhancement for gadolinium-free cardiovascular magnetic resonance tissue characterization in hypertrophic cardiomyopathy," *Circulation*, vol. 144, pp. 589-599, 2021, doi: 10.1161/CIRCULATIONAHA.121.054432.
- [8] K. Haris *et al.*, "Free-breathing fetal cardiac MRI with Doppler ultrasound gating, compressed sensing, and motion compensation," *J Magn Reson Imaging*, vol. 51, pp. 260-272, 2020, doi: 10.1002/jmri.26842.
- [9] Y. Sato and K. Ohkuma, "Verification of image quality improvement by deep learning reconstruction to 1.5 T MRI in T2-weighted images of the prostate gland," *Radiol Phys Technol*, vol. 17, pp. 756-764, 2024, doi: 10.1007/s12194-024-00819-5.
- [10] T. M. Vollbrecht *et al.*, "Deep learning denoising reconstruction for improved image quality in fetal cardiac cine MRI," *Front Cardiovasc Med*, vol. 11, p. 1323443, 2024, doi: 10.3389/fcvm.2024.1323443.
- [11] M. Tang *et al.*, "A deep learning algorithm for noise reduction in a novel respiratory-triggered single-shot phase-sensitive inversion recovery myocardial delayed enhancement cardiac MRI pulse sequence," *European Heart Journal - Cardiovascular Imaging*, vol. 24, 2023, doi: 10.1093/ehjci/jead119.372.
- [12] K. Phipps *et al.*, "Accelerated in vivo cardiac diffusion-tensor MRI using residual deep learning-based denoising in participants with obesity," *Radiology: Cardiothoracic Imaging*, vol. 3, p. e200580, 2021, doi: 10.1148/ryct.2021200580.
- [13] A. Fotaki *et al.*, "Accelerating 3D MTC-BOOST in patients with congenital heart disease using a joint multi-scale variational neural network reconstruction," *Magn Reson Imaging*, vol. 92, pp. 120-132, 2022, doi: 10.1016/j.mri.2022.06.012.
- [14] L. M. Bischoff *et al.*, "Deep learning super-resolution reconstruction for fast and motion-robust T2-weighted prostate MRI," *Radiology*, vol. 308, no. 3, 2023, doi: 10.1148/radiol.230427.
- [15] M. Tanabe *et al.*, "Feasibility of high-resolution magnetic resonance imaging of the liver using deep learning reconstruction based on the deep learning denoising technique," *Magn Reson Imaging*, vol. 80, pp. 121-126, 2021, doi: 10.1371/journal.pone.0287903.
- [16] B. Sistaninejad, H. Rasi, and P. Nayeri, "A Review Paper about Deep Learning for Medical Image Analysis," *Computational and Mathematical Methods in Medicine*, vol. 2023, no. 1, p. 7091301, 2023, doi: 10.1155/2023/7091301.
- [17] H. H. Valizadeh *et al.*, "Highly accelerated free-breathing real-time phase contrast cardiovascular MRI via complex-difference deep learning," *Magn Reson Med*, vol. 86, pp. 804-819, 2021, doi: 10.1002/mrm.28750.
- [18] J. I. Hamilton, D. Currey, S. Rajagopalan, and N. Seiberlich, "Deep learning reconstruction for cardiac magnetic resonance fingerprinting T (1) and T (2) mapping," *Magn Reson Med*, vol. 85, pp. 2127-2135, 2021, doi: 10.1002/mrm.28568.
- [19] D.yd, E. Heiberg, A.H. Aletras, and E. Hedström, "Super-resolution cine image enhancement for fetal cardiac magnetic resonance imaging," *Magn Reson Imaging*, vol. 56, no. 1, pp. 223-231, 2022, doi: 10.1002/jmri.27956.
- [20] C. M. Sandino, P. Lai, S. S. Vasanawala, and J. Y. Cheng, "Accelerating cardiac cine MRI using a deep learning-based ESPiRiT reconstruction," *Magn. Reson. Med*, vol. 85, no. 1, pp. 152-167, 2021, doi: 10.1002/mrm.2842.
- [21] A. Phaira, A. Fotakia, L. Felsnera, H. Qib, and R. M. Botnar, "A motion-corrected deep-learning reconstruction framework for accelerating whole-heart magnetic resonance imaging in patients with congenital heart disease," *Journal of Cardiovascular Magnetic Resonance*, vol. 26, no. 1, p. 101039, 2024, doi: 10.1016/j.jocmr.2024.101039.
- [22] D. Kim, M. L. Jen, L. B. Eisenmenger, and K. M. Johnson, "Accelerated 4D-flow MRI with 3-point encoding enabled by machine learning," *Mag. Reson. Med*, vol. 89, pp. 800-811, 2023, doi: 10.1002/mrm.29469.
- [23] Z. Ke *et al.*, "Deep Manifold Learning for Dynamic MR Imaging," in *IEEE Transactions on Computational Imaging*, vol. 7, pp. 1314-1327, 2021, doi: 10.1109/TCI.2021.3131564.

- [24] M. Akcakaya, S. Moeller, S. Weingartner, and K. Ugurbil, "Scan-specific robust artificial-neural-networks for k-space interpolation (RAKI) reconstruction: database-free deep learning for fast imaging," *Magn Reson Med*, vol. 81, pp. 439-453, 2019, doi: 10.1002/mrm.27420.
- [25] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *IEEE Trans Med Imaging*, vol. 37, pp. 491-503, 2018, doi: 10.1109/TMI.2017.2760978.
- [26] T. Kustner *et al.*, "Deep-learning based super-resolution for 3D isotropic coronary MR angiography in less than a minute," *Magn Reson Med*, vol. 86, pp. 2837-2852, 2021, doi: 10.1002/mrm.28911.
- [27] T. A. McDonagh *et al.*, "2023 focused update of the 2021 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) with the special contribution of the Heart Failure Association (HFA) of the ESC," *European heart journal*, vol. 44, no. 37, pp. 3627-3639, 2023, doi: 10.1093/eurheartj/ehad195.
- [28] S. Ghadimi *et al.*, "Fully-automated global and segmental strain analysis of DENSE cardiovascular magnetic resonance using deep learning for segmentation and phase unwrapping," *J Cardiovasc Magn Reson*, vol. 23, p. 20, 2021, doi: 10.1186/s12968-021-00712-9.
- [29] E. Hann, *et al.*, "Deep neural network ensemble for on-the-fly quality control-driven segmentation of cardiac MRI T1 mapping," *Med Image Anal*, vol. 71, p. 102029, 2021, doi: 10.1016/j.media.2021.102029.
- [30] I. Machado *et al.*, "A Deep Learning-Based Integrated Framework for Quality-Aware Undersampled Cine Cardiac MRI Reconstruction and Analysis," *IEEE Trans BiomedEng*, vol. 71, no. 3, pp. 855-865, 2024, doi: 10.1109/TBME.2023.3321431.
- [31] N. Ajmeera and P. Kamakshi, "Sentiment analysis technique on product reviews using Inception Recurrent Convolutional Neural Network with ResNet Transfer Learning," *Smart Science*, vol. 12, no. 4, pp. 654-665, 2024, doi: 10.1080/23080477.2024.2370210.
- [32] L. T. Wang and S. Cheng, "Image super-resolution via dual-level recurrent residual networks," *Sensors*, vol. 22, p. 3058, 2022.
- [33] S. M. A. Bashir, Y. Wang, M. Khan, and Y. Niu, "A comprehensive Review of deep learning-based single image super-resolution," *PeerJ Comput. Sci.*, vol. 7, p. e621, 2021, doi: 10.7717/peerj-cs.621.
- [34] Y. Li, B. Sixou, and F. Peyrin, "A Review of the deep learning methods for medical images super resolution problems," *IRBM*, vol. 42, pp. 120-133, 2021, doi: 10.1016/j.irbm.2020.08.004.
- [35] S. Anwar and N. Barnes, "Densely Residual Laplacian Super-Resolution," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 3, pp. 1192-1204, 2022.
- [36] R. Lan *et al.*, "Cascading and enhanced residual networks for accurate single-image super-resolution," *IEEE Trans Cybern*, vol. 51, pp. 115-125, 2021, doi: 10.1109/TCYB.2019.2952710.
- [37] J. A. scanoa, M. J. Middione, A. B. Syed, C. M. Sandino, S. S. Vasanaawala, and D. B. Ennis, "Accelerated two-dimensional phase-contrast for cardiovascular MRI using deep learning-based reconstruction with complex difference estimation," *Mag. Reson. Med*, vol. 89, pp. 356-369, 2023.
- [38] H. Li, X. He, Z. Yu, and J. Luo, "Noise-robust image fusion with low-rank sparse decomposition guided by external patch prior," *Inf. Sci.*, vol. 523, pp. 14-37, 2020, doi: 10.1016/j.ins.2020.03.009.
- [39] Q. Zhang *et al.*, "Artificial intelligence for contrast-free MRI: scar assessment in myocardial infarction using deep learning-based virtual native enhancement," *Circulation*, vol. 146, no. 20, pp. 1492-1503, 2022.
- [40] S. A. Khowaja, B. N. Yahya, and S. -L. Lee, "Cascaded and Recursive ConvNets (CRCNN): An effective and flexible approach for image denoising," *Signal Processing: Image Communication*, vol. 99, p. 116420, 2021, doi: 10.1016/j.image.2021.116420.
- [41] E. Cole, J. Cheng, J. Pauly, and S. Vasanaawala, "Analysis of deep complex-valued convolutional neural networks for MRI reconstruction and phase-focused applications," *Mag. Reson. Med.*, vol. 86, pp. 1093-1109, 2021, doi: 10.1002/mrm.28733.
- [42] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei, and H. Liu, "Attention-guided CNN for image denoising," *Neural Networks*, vol. 124, pp. 117-129, 2020, doi: 10.1016/j.neunet.2019.12.024.
- [43] J. A. Oscanoa, *et al.*, "Deep Learning-Based Reconstruction for Cardiac MRI: A Review," *Bioengineering (Basel)*, vol. 10, no. 3, p. 334, 2023, doi: 10.3390/bioengineering10030334.
- [44] T. Küstner *et al.*, "CINENet: Deep learning-based 3D cardiac CINE MRI reconstruction with multi-coil complex-valued 4D spatio-temporal convolutions," *Sci Rep*, vol. 10, p. 13710, 2020, doi: 10.1038/s41598-020-70551-8.
- [45] Y. Zhang *et al.*, "Rapid 3D breath-hold MR cholangiopancreatography using deep learning-constrained compressed sensing reconstruction," *Eur Radiol*, vol. 33, no. 4, pp. 2500-2509, 2023, doi: 10.1007/s00330-022-09227-y.
- [46] D. Prokopenko, K. Hammernik, T. Roberts, D. F. A. Lloyd, D. Rueckert, and J. V. Hajnal, "The challenge of fetal cardiac MRI reconstruction using deep learning," *Springer Nature Switzerland*, pp. 64-74, 2023, doi: 10.1007/978-3-031-45544-5_6.
- [47] N. Nazir, A. Sarwar, and B. S. Saini, "Recent developments in denoising medical images using deep learning: An overview of models, techniques, and challenges," *Micron*, vol. 180, 2024, doi: 10.1016/j.micron.2024.103615.
- [48] W. Yuan, H. Liu, L. Liang, and W. Wang, "Learning the Hybrid Nonlocal Self-Similarity Prior for Image Restoration," *Mathematics*, vol. 12, p. 1412, 2024, doi: 10.3390/math12091412.
- [49] N. Qi, Y. Shi, X. Sun, W. Ding, and B. Yin, "Single image super-resolution via 2D sparse representation," *2015 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1-6, 2015.
- [50] X. Lan, S. Roth, D. Hutten ocher, and M. J. Black, "Efficient belief propagation with learned higher-order Markov random fields," in *European conference on computer vision. Berlin, Heidelberg: Springer*, vol. 2006, pp. 269-82, 2006, doi: 10.1007/11744047_21.
- [51] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, "Non-local sparse models for image restoration," *2009 IEEE 12th International Conference on Computer Vision*, pp. 2272-2279, 2009.
- [52] Y. Chen and T. Pock, "Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 6, pp. 1256-1272, 2017, doi: 10.1109/TPAMI.2016.2596743.
- [53] H. C. Burger, C. J. Schuler and S. Harmeling, "Image denoising: Can plain neural networks compete with BM3D?," *2012 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2392-2399, 2012.
- [54] Q. Yang *et al.*, "Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1348-1357, 2018.
- [55] K. Zhang *et al.*, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142-3155, 2016.
- [56] S. Zamir, A. Arora, S. Khan, M. Hayat, F. Khan, M. Yang, and L. Shao, "Learning Enriched Features for Real Image Restoration and Enhancement," in *Computer Vision-ECCV 2020: 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part XXV 16*, pp. 492-511, 2020, doi: 10.1007/978-3-030-58595-2_30.
- [57] K. K. L. Wong *et al.*, "medical image diagnostics based on computer-aided flow analysis using magnetic resonance images," *Computerized Medical Imaging and Graphics*, vol. 36, no. 7, pp. 527-541, 2012.
- [58] J. L. Harris, "Diffraction and Resolving Power," *Journal of the Optical Society of America*, vol. 54, no. 7, pp. 931-936, 1964.
- [59] J. W. Goodman, "Introduction to Fourier Optics: McGraw-Hill," *Optical Engineering*, vol. 35, no. 5, pp. 1513-1513, 1996.
- [60] C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 2, pp. 295-307, 2016.
- [61] C. Dong, C. C. Loy, and X. Tang, "Accelerating the super-resolution convolutional neural network," in *Computer Vision-ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part II 14*, pp. 391-407, 2016.
- [62] C. Ledig *et al.*, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 105-114, 2017, doi: 10.1109/CVPR.2017.19.

- [63] M. Arjovsky, C. Smith, and B. Léon, "Wasserstein generative adversarial networks," In *International conference on machine learning*, pp. 214–233, 2017.
- [64] G. Ishaan, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of Wasserstein GANs," in *Advances in neural information processing systems*, vol. 30, 2017.
- [65] M. Ran, J. Hu, Y. Chen, H. Chen, "Denoising of 3-D Magnetic Resonance Images Using a Residual Encoder-Decoder Wasserstein Generative Adversarial Network," *Medical image analysis*, vol. 55, pp. 165-180, 2018.
- [66] R. R. Sood and M. Rusu, "Anisotropic Super Resolution In Prostate Mri Using Super Resolution Generative Adversarial Networks," *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pp. 1688-1691, 2019, doi: 10.1109/ISBI.2019.8759237.
- [67] Z. Han, B. Wei, A. Mercado, S. Leung, and S. Li, "Spine-GAN: Semantic segmentation of multiple spinal structures," *Medical image analysis*, vol. 50, pp. 23–35, 2018, doi: 10.1016/j.media.2018.08.005.
- [68] H. Du, N. Yuan, and L. Wang, "Node2Node: Self-Supervised Cardiac Diffusion Tensor Image Denoising Method," *Applied Sciences*, vol. 13, no. 19, p. 10829, 2023, doi: 10.3390/app131910829.
- [69] S. Ghose, N. Singh, and P. Singh, "Image Denoising using Deep Learning: Convolutional Neural Network," *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 511-517, 2020, doi: 10.1109/Confluence47617.2020.9057895.
- [70] M. Rezaei, H. Yang, and C. Meinel, "Deep neural network with the l2-norm unit for brain lesions detection," in *International conference on neural information processing*, pp. 798–807, 2017, doi: 10.1007/978-3-319-70093-9_8.
- [71] M. Zhao, Y. Wei, and K. K. Wong, "A Generative Adversarial Network technique for high-quality super-resolution reconstruction of cardiac magnetic resonance images," *Magnetic Resonance Imaging*, vol. 85, pp. 153–160, 2022, doi: 10.1016/j.mri.2021.10.033.
- [72] I. Goodfellow *et al.*, "Generative adversarial networks," *Communications of the ACM*, vol. 63, no. 11, pp. 139-144, 2020, doi: 10.1145/3422622.
- [73] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative adversarial text to image synthesis," in *International conference on machine learning*, pp. 1060-1069, 2016.
- [74] J. Liu, C. Gu, J. Wang, G. Youn, and J. Kim, "Multi-scale multi-class conditional generative adversarial network for handwritten character generation," *The Journal of Supercomputing*, vol. 75, no. 4, pp. 1922–1940, 2019.
- [75] Q. Yan and W. Wang, "DCGANs for image super-resolution, denoising and deblurring," *Advances in Neural Information Processing Systems*, pp. 487–495, 2017.
- [76] K. Zhang, H. Hu, K. Philbrick, G. M. Conte, J. D. Sobek, P. Rouzrokh, and B. J. Erickson, "SOUP-GAN: Super-Resolution MRI Using Generative Adversarial Networks," *Tomography*, vol. 8, no. 2, pp. 905-919, 2022, doi: 10.3390/tomography8020073.
- [77] M. L. Salvia, *et al.*, "Deep Convolutional Generative Adversarial Networks to Enhance Artificial Intelligence in Healthcare: A Skin Cancer Application," *Sensors*, vol. 22, p. 6145, 2022, doi: 10.3390/s22166145.
- [78] J. Blarr, S. Klinder, W. V. Liebig, K. Inal, L. Kärger, and K. A. Weidenmann, "Deep convolutional generative adversarial network for generation of computed tomography images of discontinuously carbon fiber reinforced polymer microstructures," *Scientific Reports*, vol. 14, no. 1, p. 9641, 2024, doi: 10.1038/s41598-024-59252-8.
- [79] Z. Chen, Z. Zeng, H. Shen, Zheng, P. Dai, and P. Ouyang, "DN-GAN: Denoising generative adversarial networks for speckle noise reduction in optical coherence tomography images," *Biomedical Signal Processing and Control*, vol. 55, p. 101632, 2020, doi: 10.1016/j.bspc.2019.101632.
- [80] Y. Wang, S. Luo, L. Ma, and M. Huang, "RCA-GAN: An Improved Image Denoising Algorithm Based on Generative Adversarial Networks," *Electronics*, vol. 12, no. 22, p. 4595, 2023, doi: 10.3390/electronics12224595.
- [81] B. Gãšřard, M. Sangnier, U. Tanielian, "Some theoretical insights into Wasserstein GANs," *Journal of Machine Learning Research*, vol. 22, no. 119, pp. 1-45, 2021.
- [82] Y. Gao and M. K. Ng, "Wasserstein generative adversarial uncertainty quantification in physics-informed neural networks," *Journal of Computational Physics*, vol. 463, p. 111270, 2022, doi: 10.1016/j.jcp.2022.111270.
- [83] M. N. Yeasmin, M. Al Amin, T. J. Joti, Z. Aung, and M.A. Azim, "Advances of AI in image-based computer-aided diagnosis: A review," *Array*, p. 100357, 2024, doi: 10.1016/j.array.2024.100357.
- [84] G. Yang *et al.*, "DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1310-1321, 2018, doi: 10.1109/TMI.2017.2785879
- [85] Y. Li, P. Jian, and G. I. Han, "Cascaded Progressive Generative Adversarial Networks for Reconstructing Three-Dimensional Grayscale Core Images from a Single Two-Dimensional Image," *Frontiers in Physics*, vol. 10, p. 716708, 2022, doi: 10.3389/fphy.2022.716708.
- [86] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of Wasserstein gans," in *Advances in Neural Information Processing Systems*, pp. 5767–5777, 2017.
- [87] M. Shafiq and Z. Gu, "Deep residual learning for image recognition: A survey," *Applied sciences*, vol. 12, no. 18, p. 8972, 2022, doi: 10.3390/app12188972.
- [88] X. Li, W. Wang, X. Hu, and J. Yang, "Selective Kernel Networks," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 510-519, 2019.
- [89] T. Mou, S. Alam, Md. H. Rahman, G. Srivastava, M. Hasan, and M. L. Uddin "Multi-Range Sequential Learning Based Dark Image Enhancement with Color Upgradation," *Applied Sciences*, vol. 13, no. 2, p. 1034, 2023, doi: 10.3390/app13021034.
- [90] S. Woo, J. Park, J. Lee, and I. Kweon, "CBAM: Convolutional block attention module," In *Proceedings of the European conference on computer vision (ECCV)*, pp. 3-19, 2018, doi:10.1007/978-3-030-01234-2_1.
- [91] P. Isola, J. Y Zhu, T. Zhou, and A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," *Conference: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5967-5976, 2017, doi: 10.1109/CVPR.2017.632.
- [92] O. Bernard *et al.*, "Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved?," in *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2514-2525, 2018, doi: 10.1109/TMI.2018.283750.
- [93] J. Xu and E. Adalsteinsson, "Deformed2self: Self-supervised denoising for dynamic medical imaging," In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part II 24*, pp. 25-35, 2021.
- [94] V. Lempitsky, A. Vedaldi, and D. Ulyanov, "Deep Image Prior," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9446-9454, 2018.
- [95] Y. Quan, M. Chen, T. Pang, and H. Ji, "Self2self with dropout: Learning self-supervised denoising from single image," In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1890-1898, 2020.
- [96] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," in *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080-2095, 2007.
- [97] M. Maggioni, G. Boracchi, A. Foi, and K. Egiazarian, "Video Denoising, Deblocking, and Enhancement Through Separable 4-D Nonlocal Spatiotemporal Transforms," in *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 3952-3966, 2012.