ISLAMIC UNIVERSITY OF TECHNOLOGY (IUT)

Ultrasound Images Resolution Enhancement using Generative Adversarial Network

Authors

Afeef Ahmed Jarif, 180041241

Chowdhury Md Intiser Ali, 180041233

Muhammad Jawad Chowdhury, 180041228

Supervisor

Tareque Mohmud Chowdhury

Assistant Professor

Dept. of CSE, IUT

Co-Supervisor

Tasnim Ahmed

Lecturer

Dept. of CSE, IUT

A thesis submitted in partial fulfillment of the requirements for the degree of B. Sc. Engineering in Computer Science and Engineering

Academic Year: 2021-2022

Department of Computer Science and Engineering (CSE)

Islamic University of Technology (IUT)

A Subsidiary Organ of the Organization of Islamic Cooperation (OIC)

Dhaka, Bangladesh

June 5, 2023

Declaration of Authorship

This is to certify that the work presented in this thesis is the outcome of the analysis and experiments carried out under the supervision of Tareque Mohmud Chowdhury, Assistant Professor of the Department of Computer Science and Engineering (CSE), Islamic University of Technology (IUT), Dhaka, Bangladesh. It is also declared that neither this thesis nor any part of this thesis has been sub- mitted anywhere else for any degree or diploma. Information derived from the published and unpublished work of others has been acknowledged in the text and a list of references is given.

Authors:

Afeef Ahmed Jarif

Student ID - 180041241

Chowdhury Md Intiser Ali

Student ID - 180041233

Muhammad Jawad Chowdhury

Student ID - 180041228

Approved By:

Supervisor:

Tareque Mohmud Chowdhury
Assistant Professor
Department of Computer Science and Engineering(CSE)
Islamic University of Technology (IUT), OIC

Co-Supervisor:

Tasnim Ahmed
Lecturer
Department of Computer Science and Engineering(CSE)
Islamic University of Technology (IUT), OIC

Acknowledgement

For functioning as our advisor and mentor, Tareque Mohmud Chowdhury, Assistant Professor, Department of Computer Science Engineering, IUT, deserves our sincere gratitude. His inspiration, advice, and ideas have been extremely helpful for this project. This research would not have been accomplished without his assistance and the right direction. From the first phase of the thesis themes introduction, subject selection, proposing algorithm, and modification, to the project implementation and finalization, his important opinion, time, and input were offered throughout the thesis work, which helped us to complete our thesis work correctly. We are very appreciative of him.

We are especially grateful to Tasnim Ahmed, Lecturer, Department of Computer Science Engineering, IUT. His tremendous wisdom, patience, and enthusiasm were instrumental in directing our study and fostering our growth as aspiring researchers. The quality and scope of our work have significantly improved as a result of his astute critique and direction. We are grateful for his unwavering support and belief in our ability. He offered significant contributions that were crucial to the success of our thesis, and we are thankful that we had the chance to work under their guidance.

Abstract

This thesis focuses on enhancing ultrasound scan resolution and image quality, particularly for the detection of breast cancer. During pregnancy, diagnostic ultrasound is frequently used to examine internal organs and track fetal progress. The presence of noise in ultrasonic imaging, however, makes diagnosis and interpretation difficult.

A generative adversarial network (GAN) is suggested as a deep learning method to address this problem. The GAN is made up of a generator network that has been trained to transform low-resolution ultrasonic inputs into high-resolution outputs and a discriminator network that can tell real images from fake ones. Convolutional layers, skip connections, Multi-resolution Convolution blocks (MRCB), and loss functions are all incorporated into the GAN model's architectural choices, which are tailored to the peculiarities of breast cancer. For iterative optimization throughout the training phase, backpropagation and gradient descent techniques are used. The peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) are quantitative metrics that are used to assess the effectiveness of the proposed Ultrasound Images Resolution Enhancement GAN (UIRE-GAN) approach.

The findings show considerable increases in image quality, increasing the capability of breast cancer ultrasound imaging as a diagnostic tool. This study advances the area of breast cancer imaging and has the potential to help patients receive better healthcare.

Contents

1 Introduction

1.1 Overview

Diagnostic ultrasound can be used to observe internal organs without invasive procedures. Ultrasound has several applications, including imaging the heart, blood vessels, eyes, thyroid, brain, breast, abdominal organs, skin, and muscles. During pregnancy, ultrasonography is frequently used to monitor the growth and development of the fetus. However, noise is a big problem with ultrasound imaging. We attempt to increase the image quality of ultrasound scans in order to get a better picture and make diagnosis easier. We intend to employ a generative adversarial network to do this.

The process of gathering biological data, transforming it into useful forms, and then analyzing that data is the subject of bioinformatics. A significant portion of Bioinformatics research is focused on improving biomedical imaging. Improving ultrasound image quality is crucial for a clearer perspective and simpler diagnosis.

This thesis paper aims to improve the resolution of ultrasound images related to breast cancer by utilizing Generative Adversarial Networks (GANs). Ultrasound imaging plays a crucial role in accurately diagnosing breast cancer, but low-resolution ultrasound images often lack important details, which can present challenges during interpretation and diagnosis. To tackle this issue, the study proposes the use of GANs, a deep learning technique comprising a generator and discriminator network. The generator model undergoes training to generate high-resolution ultrasound images from low-resolution inputs, while the discriminator network learns to distinguish between genuine and generated high-resolution images. The generator steadily improves its ability to generate realistic and high-resolution ultrasound images through an adversarial training process.

Moreover, the paper thoroughly discusses the fundamentals of GANs and their relevance to medical imaging, particularly in the context of breast cancer ultrasound images. The unique characteristics of breast cancer ultrasound images, such as tissue complexity and lesion visibility, are taken into account during the design and training of the GAN model.

The GAN model is designed with suitable architectural choices, including convolutional layers, skip connections, and loss functions, to facilitate the generation of high-resolution details specific to breast cancer characteristics. During the training process, the generator and discriminator networks are iteratively optimized using backpropagation and gradient descent techniques. The performance of the trained GAN model is evaluated using quantitative metrics such as peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM).

The experimental results and comparative analyses validate the effectiveness of the proposed Ultrasound Images Resolution Enhancement GAN (UIRE-GAN) method for breast cancer ultrasound images. The quantitative evaluations demonstrate significant improvements in PSNR and SSIM scores, indicating enhanced image quality and improved diagnostic potential.

In summary, this thesis paper presents an innovative approach that employs Generative Adversarial Networks to enhance the resolution of ultrasound images for breast cancer diagnosis. The proposed UIRE-GAN method effectively addresses the limitations of low-resolution breast cancer ultrasound images, yielding visually improved and diagnostically valuable outcomes. The research outcomes contribute to the field of breast cancer imaging and have the potential to enhance healthcare outcomes for breast cancer patients.

1.2 Problem Statement

The diagnosis of medical conditions often relies on imaging techniques such as CT scans and MRI. However, these modalities can be expensive, time-consuming, and may expose patients to ionizing radiation. Ultrasound imaging offers a safer and more accessible alternative, but it often suffers from poor image quality due to inherent noise, resulting in challenges during interpretation and diagnosis. The results of MRI and CT scans are often fairly reliable and accurate compared to ultrasound. CT scans use x-ray technology, which emits ionizing radiation that can eventually cause cancer as human bodies are exposed to that radiation. Contrarily, MRIs do not directly affect patients like CT scans do, although it is nevertheless advised against using them during pregnancy or if a person has metal implants in their bodies. Although ultrasound imaging is not as effective as MRIs and CT scans in terms of results, it is advised for safety purposes. Because ultrasound imaging is susceptible to noise, it does not yield results that are as accurate as CT scans or MRIs.

Therefore, there is a pressing need to develop effective denoising methods specifically tailored for ultrasound images, enabling improved diagnostic accuracy without the need for CT scans or MRI.

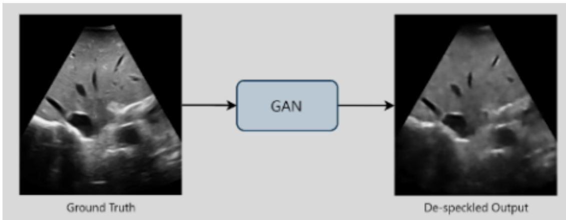

Figure 1: de-noising is prone to over-smoothing.

1.3 Motivation

The motivation behind this research lies in addressing the limitations of current diagnostic imaging techniques and exploring the potential of denoising ultrasound images as a viable alternative to CT scans or MRI. By focusing on enhancing the quality of ultrasound images, the proposed approach aims to provide clearer and more diagnostically valuable information for medical professionals.

The motivation stems from several key factors. Firstly, CT scans and MRI are often costly and may not be readily accessible in resource-constrained healthcare settings. By improving the quality of ultrasound images, healthcare providers can rely on a more affordable and widely available imaging modality, making diagnostic imaging more accessible to a larger population.

Secondly, reducing the reliance on CT scans and MRI can help mitigate the potential risks associated with ionizing radiation exposure. Ultrasound imaging does not involve ionizing radiation, making it a safer option, particularly for pregnant women and children.

Furthermore, denoising ultrasound images can lead to more accurate and reliable diagnoses. Noise in ultrasound images can obscure important anatomical details and hinder the identification of subtle abnormalities. By effectively denoising ultrasound images, medical professionals can confidently interpret and diagnose conditions with greater precision, potentially leading to earlier detection and improved patient outcomes.

Overall, the motivation for this research is driven by the goal of enhancing the diagnostic capabilities of ultrasound imaging by developing robust denoising techniques. By focusing on denoising ultrasound images instead of relying solely on CT scans or MRI, this research aims to contribute to the field of medical imaging, enabling more accessible, safer, and diagnostically valuable imaging options for healthcare professionals and patients alike.

1.4 Research Challenges

While conducting the research, we encountered several challenges that needed to be addressed. Firstly, one of the primary difficulties we faced was the scarcity of publicly available datasets specifically curated for ultrasound images. Unlike other medical imaging modalities, such as CT scans or MRI, finding comprehensive and diverse ultrasound datasets proved to be a daunting task. The limited availability of such datasets hindered the training and evaluation of our models.

Additionally, ultrasound images inherently suffer from various quality issues, including limited resolution, noise, and artifacts. These factors significantly impact the interpretability and reliability of the images, making it challenging to extract meaningful information from them. Developing effective denoising and resolution enhancement techniques to overcome these limitations required careful consideration and experimentation.

Furthermore, the performance of models trained on ultrasound images can vary when applied to images acquired from different devices or using different modalities. The lack of standardization across ultrasound imaging systems and protocols adds an additional layer of complexity, as models may struggle to generalize well to unseen data. Overcoming this challenge necessitated the development of robust architectures and training strategies that could adapt and perform consistently across diverse ultrasound imaging setups.

Implementing an appropriate architecture to effectively retrieve information from ultrasound images posed another significant challenge. Due to the unique characteristics of ultrasound imaging, such as speckle noise and complex tissue structures, designing an architecture that could accurately capture relevant features and enhance image quality required careful consideration and experimentation.

In summary, the research encountered challenges related to the limited availability of ultrasound datasets, the poor quality of ultrasound images, the device and modality variations, and the implementation of suitable architectures for information retrieval. Overcoming these challenges required innovative approaches, extensive experimentation, and a deep understanding of the unique characteristics of ultrasound imaging.

2 Literature Review

Medical imaging plays a vital role in diagnosing and monitoring diseases, not only during the initial identification but also in post-treatment stages. However, similar to other imaging modalities, medical images are susceptible to noise and artifacts. The introduction of noise can occur due to the mechanisms of imaging devices or during signal processing, resulting in various types of noise, such as random noise or frequency-dependent noise. The presence of noise in medical images obscures important details and complicates disease detection and analysis, leading to potential losses and even fatalities [1]. Therefore, denoising medical images becomes a crucial preprocessing step before further processing and analysis of the images can take place.

The literature review will delve into existing research and studies that focus on denoising methods specifically tailored for medical images. Various denoising techniques, algorithms, and approaches will be explored, highlighting their effectiveness in reducing noise and enhancing image quality. The review will also examine the impact of denoising on disease detection, analysis, and subsequent medical interventions. By reviewing the relevant literature, this research aims to identify the state-of-the-art denoising methods for medical images and their potential applicability in the context of the specific medical imaging modality under investigation. The review will provide a comprehensive understanding of the existing approaches [1–6, 12, 18, 19], their strengths, limitations, and areas that require further improvement.

Overall, the literature review section will contribute to the research by establishing a foundation of knowledge and highlighting the significance of denoising as a necessary preprocessing step in medical image processing. It will also identify gaps in the existing literature and lay the groundwork for the proposed denoising approach in this research.

2.1 Natural Image Super-Resolution

Natural image super-resolution refers to the task of enhancing the resolution and quality of low-resolution images while preserving and restoring fine details. This area of research aims to address the limitations of imaging systems that capture images with limited resolution or when images are subjected to degradation during acquisition, compression, or other factors [2].

The goal of natural image super-resolution is to generate high-resolution images that closely resemble the original high-resolution counterparts [2]. This is achieved by employing various computational techniques and algorithms that exploit the inherent information present in the low-resolution image and utilize prior knowledge about natural images.

One prominent strategy in natural picture super-resolution is to apply learning-based methods, namely deep learning models such as convolutional neural networks (CNNs). These models are trained on pairs of low-resolution and high-resolution images to discover the underlying mapping between them [2,3] . During the inference stage, the trained model takes a low-resolution image as input and outputs a corresponding high-resolution image with enhanced details.

Although image SR is a well-known low-level vision job, there have been numerous innovations in the field recently, particularly deep learning-based approaches. Many early deep SR models adopted the feature extraction, nonlinear mapping, and image reconstruction procedures after the introduction of SRCNN, the first image SR deep network described by Dong *et al.* [19]. Such shallow neural networks, however, have a limited capacity to extract multi-level characteristics from the input images. In order to train their deep SR model, Liang *et al.* [2] initially used Sobel edges with LR pictures, taking into consideration that the edge prior is beneficial to image SR. Although it is not immediately apparent, their SR performance has improved.

Liu *et al.* [3] recently suggested a multi-scale deep encoder-decoder model with the supervision of phase congruency edge map for single image SR and gave a

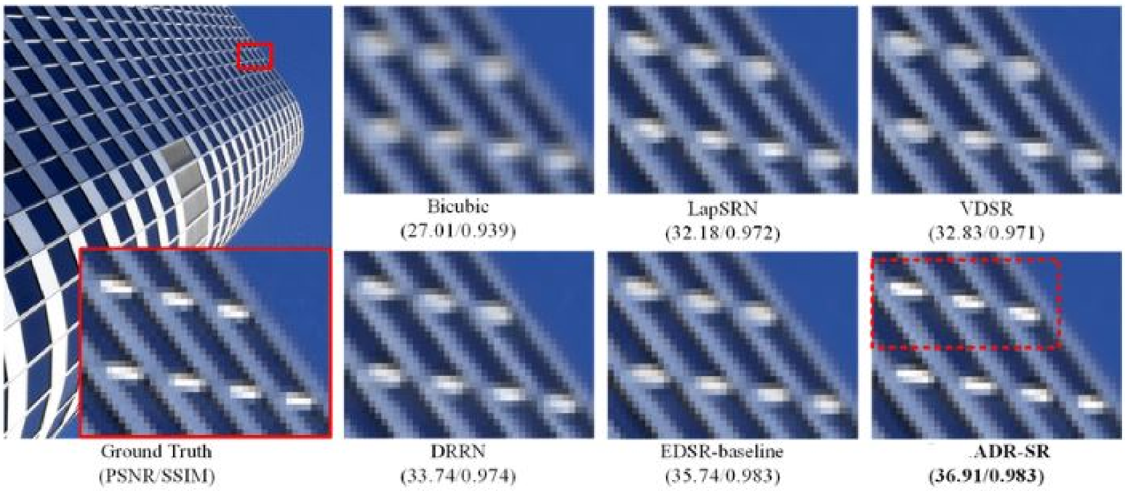

Figure 2: Results from experiments of Urban100 dataset [3].

convincing SR contrast effect, based on the structure simulation on multiple resolution wavelet analysis. The formation of a multi-memory residual block was also proposed by Wang *et al.* [4] in order to gradually extract and keep inter-frame temporal correlations for video SR. For image SR, Ma *et al.* [5] recently developed a dense discriminative network made up of a number of aggregating modules with the use of an adversarial learning technique.

2.2 Medical Image Super-Resolution

2.2.1 Traditional Methods

Traditional methods for super-resolution of medical images often incorporate a variety of strategies targeted at improving the resolution and clarity of low-resolution medical images. These methods, which are based on various image processing and interpolation techniques, predate contemporary advances in deep learning. Here are some of the most prevalent conventional approaches.

- Bicubic interpolation [20] is a popular technique for increasing the resolution of low-quality medical photographs. It calculates the missing pixel values by taking the weighted average of nearby pixels. While bicubic interpolation is computationally efficient, the resulting high-resolution photographs sometimes have hazy and unnatural details, especially in complicated medical structures.

- Lanczos interpolation [21], like bicubic interpolation, is a higher-order interpolation technique that strives to create crisper and more accurate upscaled medical images. It estimates pixel values using a windowed sinc function. Lanczos interpolation can reduce blurring but may produce some ringing artifacts.

- Sparse representation-based super-resolution algorithms take advantage of the fact that medical pictures frequently have sparse representations in certain transform domains (for example, wavelet or curvelet) [6,7]. These methods rebuild high-resolution images by solving an optimization problem in the transform domain that favors sparsity. While these techniques are capable of preserving small details and reducing blurring, they may struggle with complicated medical structures and necessitate prior knowledge of the sparsity patterns.

- Methods based on examples rely on learning a mapping function between low-resolution and high-resolution image patches. These algorithms make use of a database of high-resolution training samples to predict the most likely high-resolution patches given the low-resolution input [6]. While example-based approaches can improve image details and textures, their success is strongly dependent on the training data's quality and diversity.

- Multi-frame super-resolution techniques make use of the concept of capturing many low-resolution photos of the same scene from different perspectives or with subtle differences [6, 7]. By matching and merging these photos, the resulting high-resolution image may be rebuilt. These approaches can improve resolution and minimize noise, but they may be limited in situations when numerous frames are not available or are vulnerable to motion abnormalities.

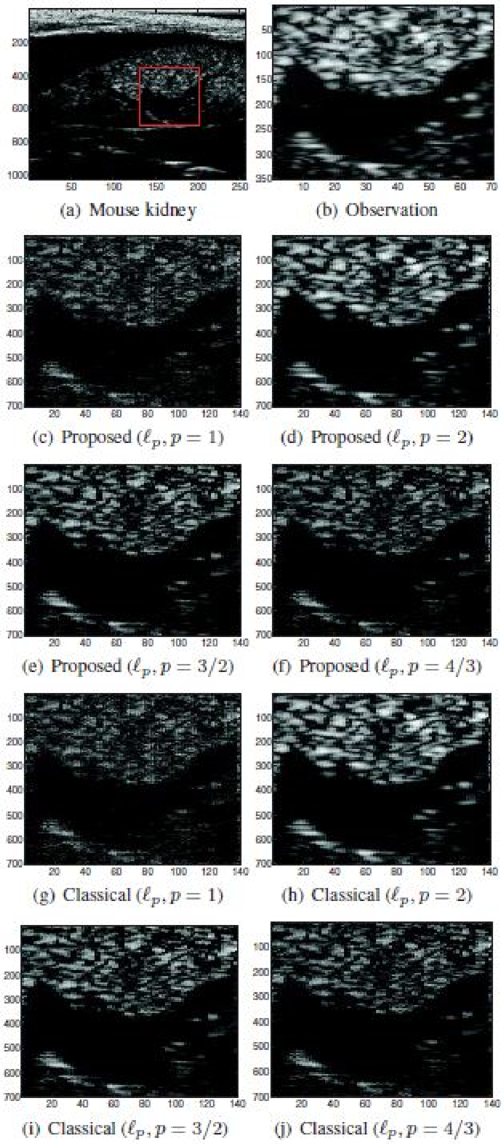

Figure 3: Using the proposed and traditional approaches in Zhao *et al.* [6], the lp norm regularizers were used to recover the in vivo US picture and the images.

Medical image SR has not received significant attention, in contrast to the rapid progress of natural image processing. Zhao *et al.* [6] recently used an analytical approach based on 2 norm regularization to construct ultrasonic image SR. Axial imaging was the focus of Diamantis *et al* [7] study. They created a location-based method to convert SR axial imaging to ultrasound imaging and realized that single scattering's image-based location precision and axial imaging accuracy are closely related.

Traditional super-resolution technologies frequently have limitations in accurately restoring high-frequency information and retaining fine structures in medical images. The subject of medical picture super-resolution has seen substantial improvements in terms of both visual quality and quantitative measurements as a result of recent advances in deep learning, notably the application of convolutional neural networks (CNNs) and generative adversarial networks (GANs). Because of their ability to learn complicated image representations and capture intricate information specific to medical imaging, these deep learning-based algorithms have become the cutting-edge methodologies for medical image super-resolution.

2.2.2 CNN Based methods

CNN-based methods have transformed the area of medical image super-resolution by using the power of deep learning to improve picture quality and resolution. Convolutional neural networks (CNNs) are used in these methods to learn complex mappings between low- and high-resolution medical pictures.

- CNN-based SISR approaches directly learn the mapping function from low-resolution to high-resolution medical pictures utilizing pairs of related image patches [19]. These approaches often employ an encoder-decoder architecture, in which the encoder collects low-resolution characteristics and the decoder creates high-resolution output. CNN-based SISR algorithms can collect both local and global image information, allowing for the restoration of fine details and high-frequency components in medical pictures.

- To improve super-resolution performance, many CNN-based algorithms use a multi-scale approach. They take advantage of CNNs' hierarchical nature to learn representations at various scales. These models are often made up of several sub-networks that process images at various resolutions or employ multi-scale feature fusion approaches [8,9]. These approaches may efficiently rebuild high-resolution medical images with increased features and texture by merging information from multiple scales.

- GANs have been effectively applied to medical picture super-resolution challenges. GANs are made up of a generator network that generates high-resolution images and a discriminator network that distinguishes between generated and real high-resolution images. The generator and discriminator networks are trained in an adversarial way, with the generator aiming to generate realistic high-resolution images that can trick the discriminator [6, 12, 18]. GAN-based techniques have produced visually attractive and high-quality medical images with excellent results.

- Attention processes are also integrated into CNN-based algorithms to boost super-resolution performance. Attention methods enable the network to concentrate on critical image regions and allocate additional resources for accurate reconstruction. These methods assist the network in selectively enhancing fine structures and details in medical pictures, resulting in higher-resolution output with better quality and clinical significance.

- In CNN-based super-resolution approaches for medical imaging, transfer learning and pre-training algorithms have been frequently used. Pre-training CNN models using large-scale natural image datasets like ImageNet aids the networks in learning generic image representations. These pre-trained models can then be fine-tuned using medical image datasets, allowing the models to capture specific medical imaging properties and patterns.

The SRCNN approach might also be appropriate for medical imaging, according to Umehara *et al.* [8], therefore they used the method for SR of a chest CT image and the outcomes confirmed their hypothesis. Additionally, in a manner

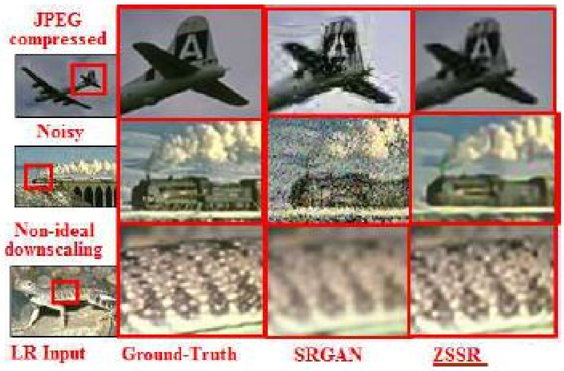

Figure 4: An illustration contrasting ZSSR [9] with SRGAN [12].

similar to ZSSR [9], Lu *et al.* [10] presented the idea of using the multi-scale contextual features that were extracted from the test image itself to train an image-specific network in an unsupervised manner. They then used residual learning and three-dilated convolution to increase convergence and accuracy.

CNN-based approaches for medical image super-resolution provide various benefits, including the capacity to capture complicated picture associations, generate high-quality and realistic images, and preserve tiny features and structures unique to medical imaging. These methods have made substantial advances in picture quality, medical diagnosis, and a variety of applications in medical research and clinical practice.

2.2.3 Super Resolution Generative Adversarial Networks (SRGAN)

Two components make up a generative adversarial network (GAN) [11]: The generator gains the ability to produce credible data. The produced instances serve as negative or false training examples for the discriminator. Furthermore, the discriminator gains the ability to distinguish genuine data from false data generated by the generator. When the generator produces false results which is unlikely, the discriminator punishes it. As training advances, the generator creates data that is clearly fraudulent, and the discriminator learns rapidly to identify it as such. As training progresses, the generator gets closer to generating output that can confuse the discriminator. Finally, if the generator training is successful, the discriminator becomes less accurate at distinguishing between what is original and what is not. As false data is classified as real, its accuracy begins to deteriorate. Both the generator networks and the discriminator networks are implemented using neural networks. The discriminator input is directly connected to the generator output. when backpropagation happens, the generator uses the discriminator's classification as a signal to update its weights.

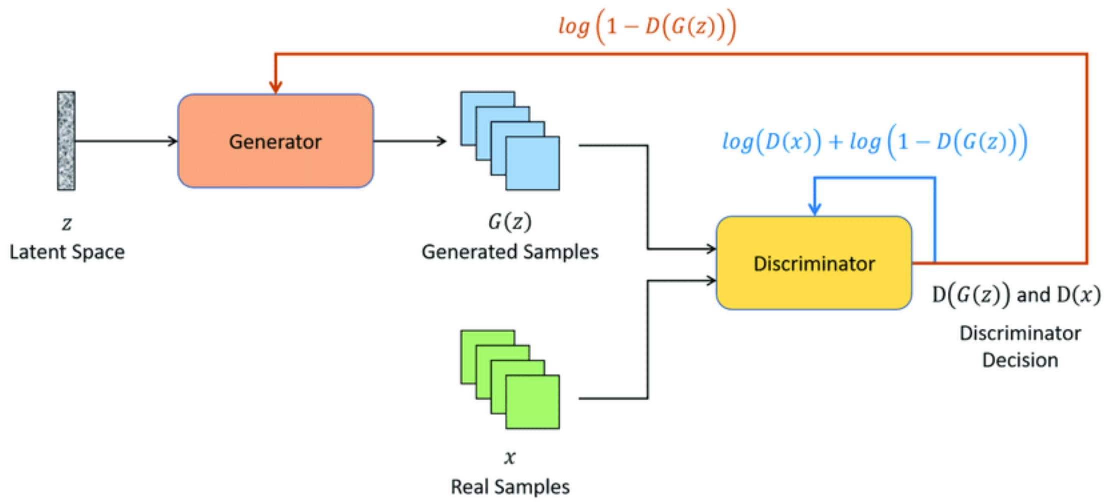

Figure 5: Architecture of GAN [11]

One of the first methods that enables the model to reach an upscaling factor of almost 4x for the majority of picture visualizations is the concept of SR-GANs [12] [13]. It is a very difficult task to estimate and create a high-resolution image from a low-resolution image. In the past, CNNs were employed to create high-resolution images that trained more quickly and accurately. However, occasionally they are unable to recover more minute features and frequently produce fuzzy photos. Most of these problems are resolved by the proposed SRGAN architecture, which produces high-quality, cutting-edge images. Consequently, the research paper [3] on producing photorealistic single images in super-resolution was written. Utilizing a Generative Adversarial Network with the recently published loss known as perceptual loss produces a loss that can be identified to combat extra perceptually oriented features. The VGG Loss kind of content loss is intro-

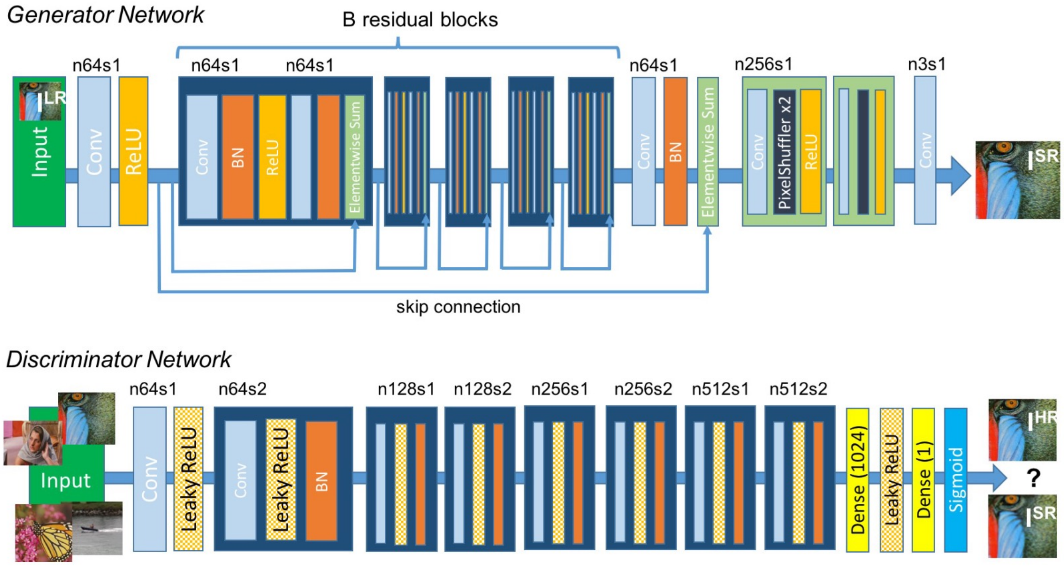

Figure 6: Architecture of SRGAN [12]

duced in the Perception Losses for Real-Time Style Transfer and Super-Resolution transfer framework. Perceptual loss is created by combining adversarial loss with content loss. The following interpretation can be applied to this loss formulation.

2.2.4 Denoising SRGAN (DnSRGAN)

Based on SRGAN architecture and feed-forward denoising convolutional neural network (DnCNN), Zhao *et al.* [14]'s work. To verify that the input was a clear image, they employed a feed-forward noise reduction neural network to pre-denoize the CMR image. Second, they employ the gradient penalty (GP) method to address the depletion of the discriminator gradient, which accelerates model convergence. To monitor GAN gradient descent, a new loss function is added to the original SRGAN loss function to generate a more reliable and efficient model training and to improve perceptual quality for CMR picture super-resolution.

High artifacts and noise that cause the cardiovascular image to be wrongly rebuilt throughout super-resolution can be managed by denoising the CMR image with DnCNN and then super-resolving the denoised image with the updated SRGAN. Furthermore, their approach can recover high-quality noisy cardiac pictures.

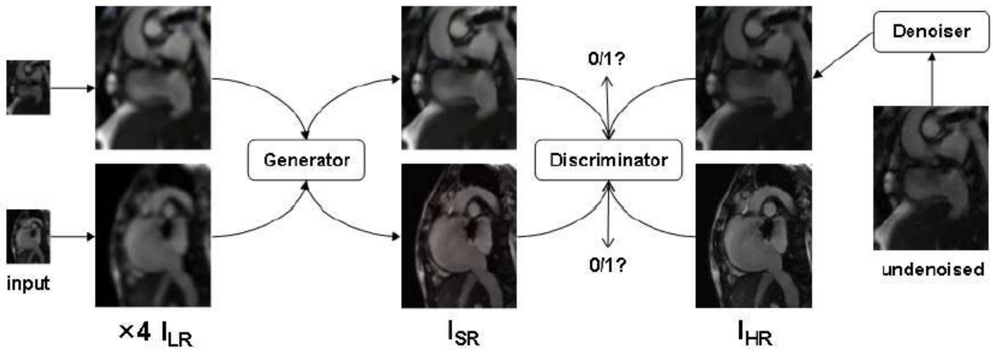

Figure 7: Architecture of proposed DnSRGAN. [14]

An adversarial learning strategy is put forth by Sanchez *et al.* [15] to create high-quality MRI scans from low-resolution images. The SRGAN-based architecture uses 3D convolutions to take advantage of volumetric data. Least squares are used in the adversarial loss for the discriminator to stabilize the training. In order to enhance the quality of the generated images, the loss function for the generator combines a least squares adversarial loss with a content term based on mean square error and image gradients. They investigated several upsampling strategies. They offered encouraging results that enhance conventional interpolation, demonstrating the approach's promise for super-resolution 3D medical imaging.

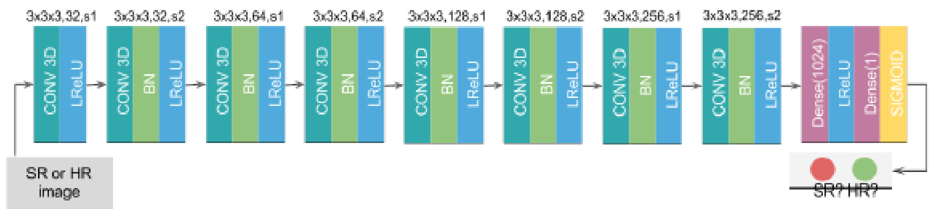

Figure 8: Architecture of the Discriminator network. [15]

A generative adversarial network with residual dense connectivity and weighted joint loss (GAN-RW) was proposed by Zhang *et al.* [16] in order to get over the limits of conventional image denoising techniques and outperform the most sophisticated ultrasound image denoising capabilities. The U-Net-like architecture,

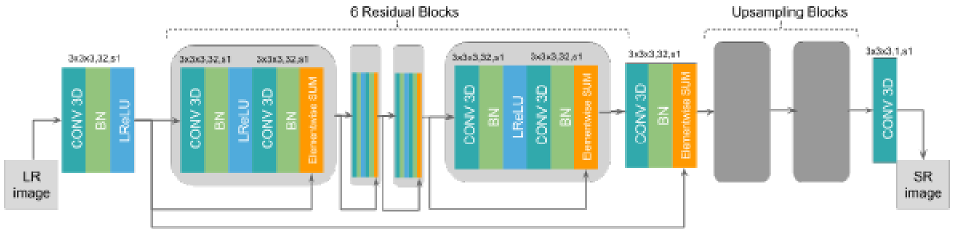

Figure 9: Architecture of the Generator network. [15]

which has four encoders and four decoders, serves as the foundation for the denoising network. To eliminate speckle noise, residual dense connectivity and BN are used to replace each encoder and decoder module. Convolutional layers are used by the discriminator network to determine the discrepancies between the translated images and the desired modality. In the training operations, they provided a combination loss function that incorporates a weighted total of the L1 loss function, binary cross-entropy with a logit loss function, and perceptual loss function.

They developed and evaluated a brand-new ultrasonic picture despeckling technique. Using residual dense connectivity, BN, and joint loss functions, GAN-RW, which is based on U-Net, eliminates speckle noise. On the three fixed noise levels of BSD68, DnCNN, DnCNN-Enhanced, BRDNet, DHDN, CBDNet, MuNet, ED-Net and GAN-RW outperform BM3D in terms of despeckling performance. On ultrasound pictures of lymph nodes, the brachial plexus, and the fetus' head, we also successfully validated the suggested technique.

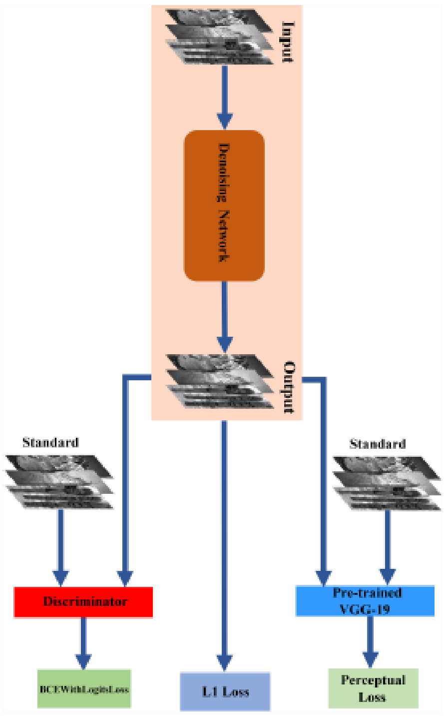

Figure 10: Architecture of the proposed network [16]. The denoising network is tasked with translating input images to the target domain through encoder-decoder networks. The discriminator is trained to distinguish between standard and denoising images. The pre-trained VGG-19 is used to acquire more features as a perceptual loss.

2.2.5 Cycle - GAN

A generative adversarial network-based unpaired SR approach that does not require a paired or aligned training dataset was proposed by Maeda *et al.* [17]. A pseudo-paired SR network and an unpaired kernel/noise correction network make up our network. The inputted LR image is first cleaned up by the rectification network, which also makes kernel adjustments, before being upscaled by the SR network. A mapping from the pseudo-clean LR picture to the inputted HR image is then learned by the SR network in a paired way during the training phase by the rectification network, which likewise creates a pseudo-clean LR image from the in-

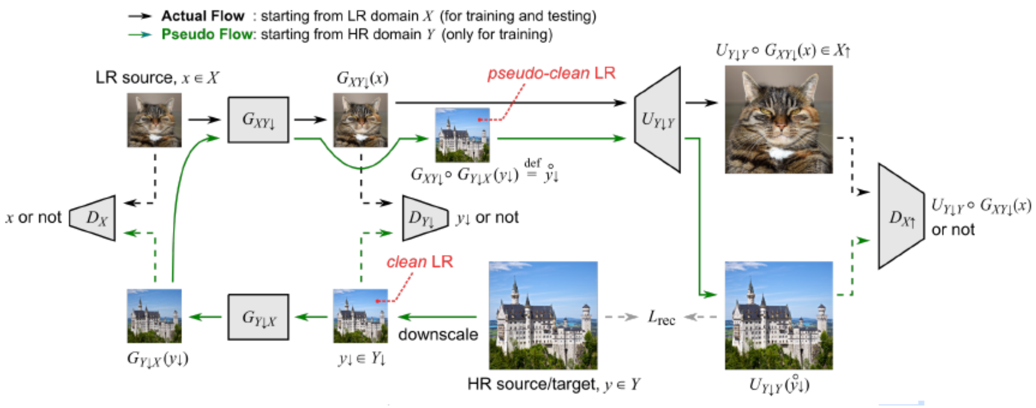

Figure 11: Data-flow diagram of the proposed method [17]. SR network can be learned in a paired manner through *Lrec*, even if the training dataset fX; Y is not paired. The whole network is end-to-end trainable.

putted HR image. The suggested framework can be combined with well-researched current network architectures and pixel-wise loss functions because its SR network is independent of the rectification network. The suggested strategy for solving the unpaired SR problem outperforms existing ones, according to experiments on a variety of datasets.

In the absence of the aligned HR-LR training set, they studied the SR problem in an unpaired environment. Their network converts ground-truth HR images into pseudo-clean LR images as intermediate products, which are subsequently utilized to train the SR network in pairs. In this regard, the suggested approach fills the gap between the existing SR methods that have undergone extensive research and the real-world SR problem without paired datasets. Extensive tests on a variety of datasets, including artificially degraded nature photos, real-world face images, and real-world aerial images, showed how effective their strategy is.

Liu *et al.* [18] 's novel perceptual consistency ultrasound image super-resolution (SR) method is able to guarantee that the re-degenerated image of the generated SR one will be consistent with the original LR image and vice versa using only the LR ultrasound data. First, they used image enhancement to construct the HR dads and LR sons of the test ultrasound LR image. Following that, they completely

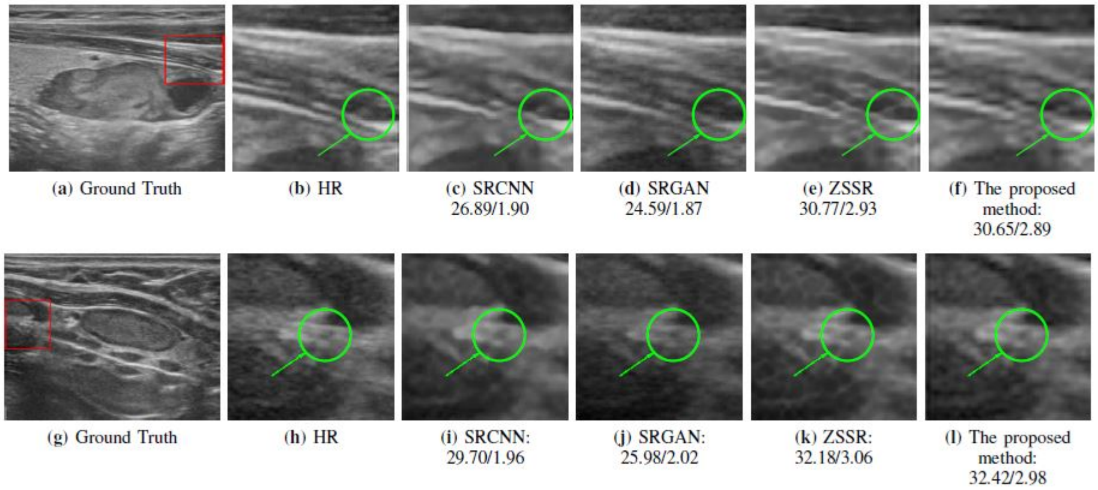

Figure 12: For four super-resolved ultrasound pictures from the US-CASE dataset, the proposed approach in Liu *et al.* [18] and ground truth, SRCNN, SR-GAN, and ZSSR were used to compare visual effects and PSNR/IFC metrics. The changes between the photos are shown by the green arrows and circles.

leveraged the cycle loss of LR-SR-LR and HRLR-SR, as well as the discriminator's adversarial properties, to urge the generator to create more perceptually consistent SR outputs. The comparison of their suggested method to other state-of-the-art methodologies employing PSNR/IFC/SSIM, inference efficacy, and visual effects on the benchmark CCA-US and CCA-US datasets demonstrates that it is effective and superior.

When there aren't enough ultrasound training photos, they look at the multi-scale pattern features between the local sections and the whole image for ultrasound data to find LR-HR pairings. They then presented a CycleGAN framework with a synthetic imaging loss, including pixel-wise loss, perceptual feature loss, adversarial loss, and the most significant cycle consistency loss, to ensure that the image ensemble and details can maintain perception consistency not only in the LR-to-SR-to-LR cycle, but also in the HR-to-LR-to-SR cycle. Two ultrasound dataset evaluations clearly show that the suggested self-supervised Cycle-GAN technique surpasses every other method regarding running effectiveness and objective qualitative outcomes.

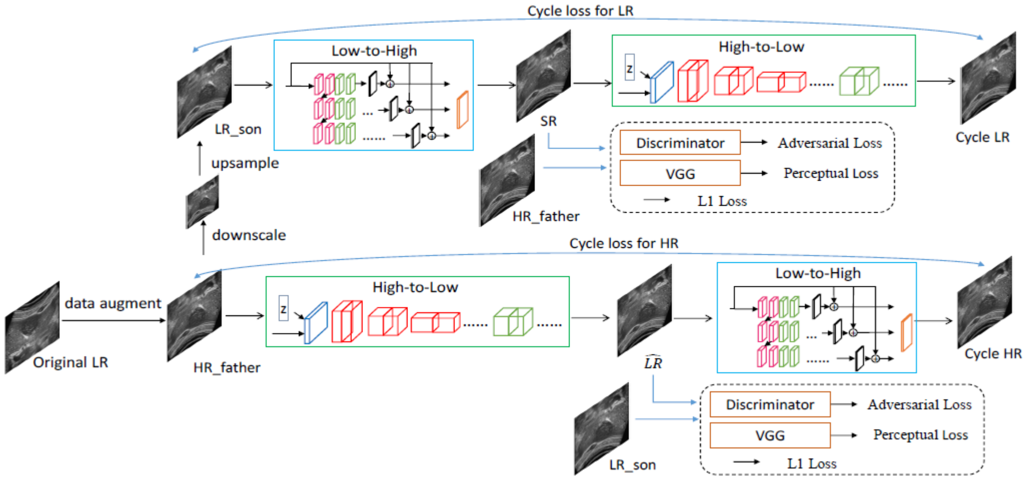

Figure 13: The proposed perception consistency ultrasound image SR model [18]. The low-to-high generator (blue box) is the multiscale encoder-decoder and the high-to-low generator (green box) is the HR-to-LR degradation network.

3 Proposed Methodology

3.1 Skeleton of the Proposed Method

The Data pipeline of our proposed method in the following:

- For training our model, Medical Ultrasound dataset will be used
- Then the data will be split for training and testing
- Our GAN model will be trained and tested with this dataset. we could not mention further about our GAN architecture because further work is required.
- After Our GAN model is trained, we will evaluate that model with some evaluation metrics such (PSNR, SSIM)
- If our model performs well on the evaluation metrics, we will generate a high-resolution image from the generator
- then we will pass these HR data to some object detection algorithm to check how it has improved contrast to the original LR images.

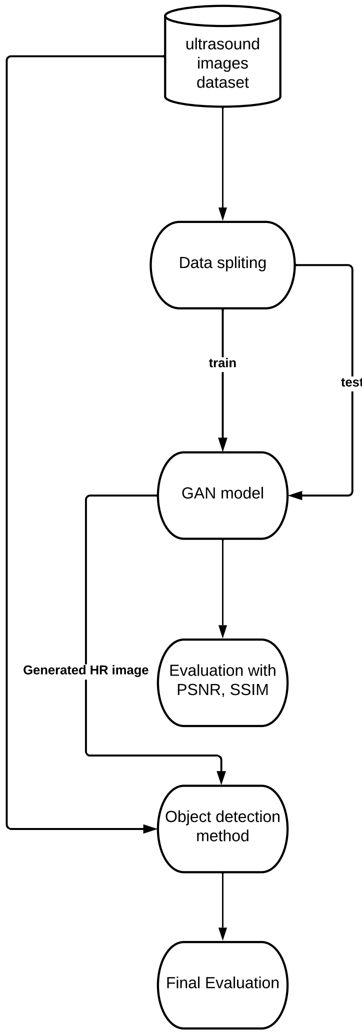

Figure 14: Data Pipeline for the proposed methodology.

3.2 Generator Architecture

The generator architecture we are proposing, is specifically designed for enhancing ultrasound medical images using a Super resolution Generative Adversarial Network (GAN). The generator takes low-resolution ultrasound images as input. These images typically have limited details and may suffer from noise or artifacts.

At the beginning of the data pipeline, residual blocks are employed. These blocks help the generator learn residual representations by capturing and enhancing fine-grained details. Each residual block consists of multiple convolutional layers, followed by skip connections that allow the network to propagate informa-

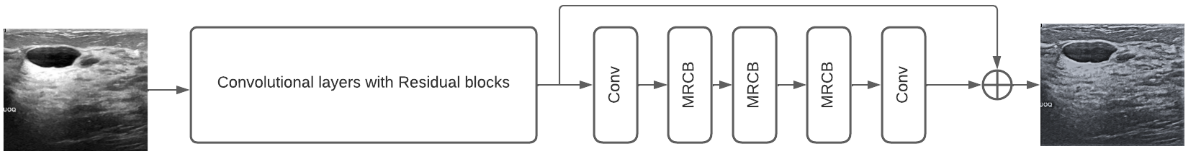

Figure 15: Overview of the Generator.

tion from early layers to later layers.

Following the residual blocks, several convolutional layers are used to further extract and learn high-level features from the input images. Each convolutional layer typically includes a convolutional operation, batch normalization, and an activation function (e.g., ReLU). These layers help the generator capture and represent complex patterns in the data.

After the convolutional layers, the generator utilizes multi-resolution convolutional blocks to progressively upscale the low-resolution input and generate a high-resolution output. These blocks capture both global and local details and help to improve the overall resolution and quality of the enhanced ultrasound images. Each multi-resolution block typically includes an upsampling operation (such as bilinear or nearest-neighbor interpolation), followed by convolutional layers, batch normalization, and an activation function.

The final layer of the generator produces the enhanced ultrasound image. The activation function used in this layer may depend on the desired output range and properties of the images. For example, if the pixel values of the enhanced images are expected to be in a specific range, an appropriate activation function, such as a sigmoid or tanh, may be applied to constrain the output values accordingly.

3.2.1 Residual Blocks

In this experiment, multiple residual blocks were used. Residual blocks play a crucial role in image enhancement tasks, including ultrasound image enhancement. They help improve the performance and effectiveness of the generator by addressing the vanishing gradient problem and enabling the network to learn residual representations.

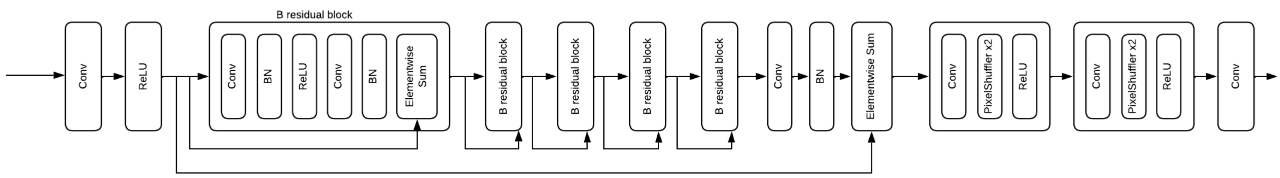

Figure 16: Convolution layers with residual blocks

The primary advantage of residual blocks stems from the concept of residual learning, which allows the network to learn the difference between the desired enhanced image and the input image. This difference, also known as the residual, captures the fine-grained details and high-frequency information that need to be added to the input image for enhancement.

A residual block consists of multiple convolutional layers, typically followed by skip connections. These skip connections enable the network to directly pass the input from one layer to a later layer. By doing so, the network can learn the residual information or the discrepancy between the input and the desired output.

Deep neural networks often encounter the vanishing gradient problem, where gradients become extremely small during backpropagation, leading to slow or ineffective learning. By using skip connections, residual blocks help alleviate this issue by providing shortcut paths for gradient flow. This allows the gradients to bypass multiple layers and directly update the earlier layers, ensuring that the network can effectively learn the residual information.

The skip connections in residual blocks enable the network to capture and learn fine-grained details. Since the gradients can flow directly from early layers to later layers, the network can effectively propagate information related to small-scale details that may have been lost or attenuated in shallower layers. This ability is especially beneficial for image enhancement tasks, where preserving and enhancing fine details are crucial.

By incorporating residual blocks into the generator architecture, the network can focus on learning the residual information necessary for enhancing the input image, rather than reconstructing the entire image from scratch. This approach helps the generator converge faster and produce more realistic and visually appealing enhanced images.

The use of residual blocks in image enhancement tasks provides a mechanism for the network to learn and add the missing or enhanced details to the input image, contributing to the overall quality and fidelity of the output.

3.2.2 Multi-resolution Convolution blocks

In the generator architecture, there were 3 MRCBs used one after another. In the Multi-resolution convolution blocks parallel convolution layer were used with different dilation rates.The input to the Multi-Resolution Convolutional Block (MRCB) goes through a convolution layer, followed by batch normalization and ReLU activation. The activated output is then passed through three parallel stacks of convolutional layers with different dilation rates. By using different dilation rates, the network can capture context information at multiple scales, as the receptive fields vary in size. This approach improves texture information and reduces the number of parameters compared to using convolution layers with different filter sizes. The output of the final MRCB goes through a convolution layer with a (1,1) kernel size. This output is then added to the input of the resolution enhancement network, producing the final output of the model.

Multi-resolution convolutional blocks are instrumental in image enhancement tasks as they enable the generator to capture both global and local details while enhancing the overall resolution of the image.

Multi-resolution convolutional blocks incorporate upsampling operations to gradually increase the resolution of the input image. Upsampling methods like bilinear or nearest-neighbor interpolation are typically employed to upscale the image. This progressive upsampling helps to restore the fine details and structure that may have been lost in the low-resolution input.

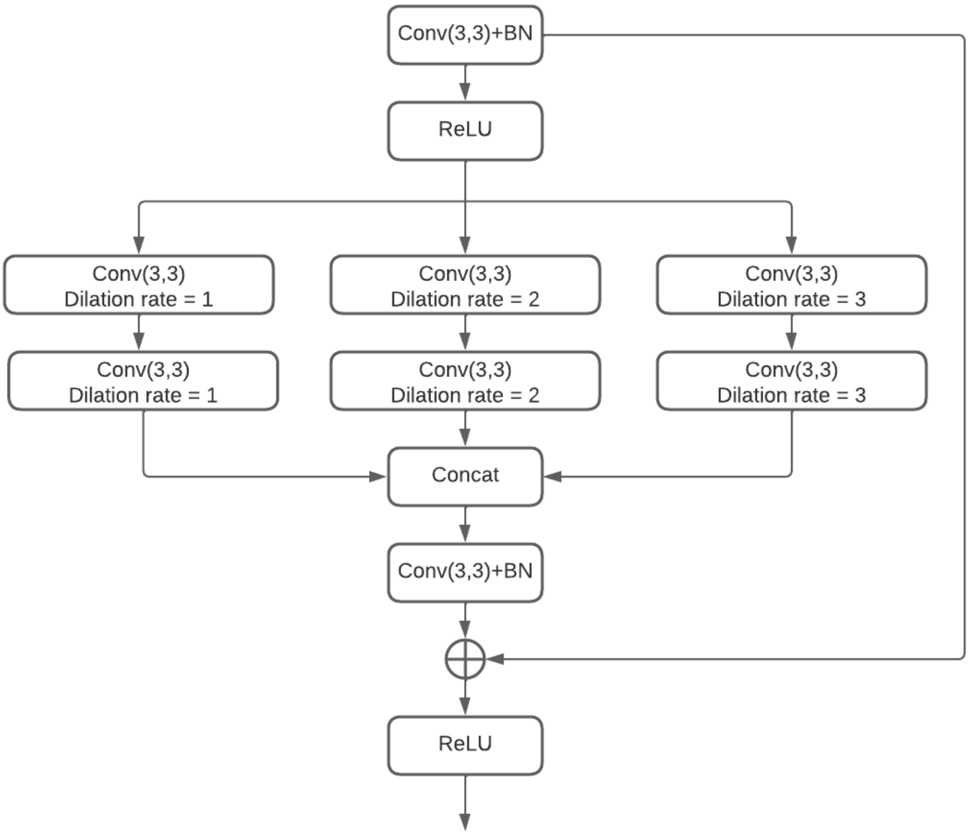

Figure 17: MRCB.

Each multi-resolution block includes convolutional layers that operate on different scales or resolutions of the image. These layers capture local details by focusing on smaller receptive fields, detecting and enhancing fine-grained features. Simultaneously, higher-level layers capture global context by considering larger receptive fields, allowing the generator to understand the overall structure and relationships in the image. By combining these local and global details, the generator can generate enhanced images with improved fidelity.

Multi-resolution convolutional blocks facilitate the learning of hierarchical features. As the image resolution increases, each block extracts features at different levels of abstraction. Lower-resolution blocks capture low-level features like edges and textures, while higher-resolution blocks capture more complex and semantic features. This hierarchical feature learning helps the generator better understand the image content and generate enhanced images that are visually coherent and meaningful.

The multi-resolution blocks provide a mechanism for fusing information from different scales. By combining features from multiple resolutions, the generator can leverage complementary information to produce enhanced images. This fusion of information from different scales helps to generate images with improved spatial details, sharpness, and overall visual quality.

Multi-resolution convolutional blocks enable the generator to progressively enhance the resolution of the image while capturing both local and global details. By combining hierarchical feature learning and information fusion, these blocks help generate high-quality enhanced images that exhibit improved visual fidelity and contain both fine-grained details and overall structural coherence.

3.3 Discriminator architecture

The discriminator architecture consists of multiple convolutional layers followed by batch normalization and Leaky ReLU activation blocks. The output of these layers is then processed by dense layers with a sigmoid activation function.

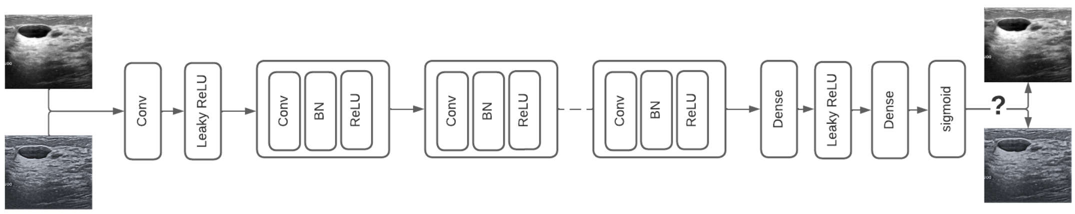

Figure 18: Discriminator Architecture

The discriminator takes two inputs,the original image and the generated sample, as its initial input. The input passes through multiple convolutional layers. Each convolutional layer performs feature extraction by applying convolutional filters to the input. The number of filters and the size of the filters can vary depending on the specific architecture and task.

After each convolutional layer, batch normalization is applied. Batch normalization normalizes the activations of the previous layer, helping to stabilize and accelerate the training process. It improves the discriminator's ability to learn discriminative features from the input data.

Following batch normalization, the output of each layer is passed through a Leaky ReLU activation function. Leaky ReLU allows small negative values to pass through, preventing the problem of "dead" neurons and enhancing the network's ability to learn from gradients.

Convolutional layer output is flattened and then fed into dense layers. Dense layers are fully connected layers in which each neuron is linked to every neuron in the preceding and following layers. Based on the characteristics retrieved by the convolutional layers, these layers learn high-level representations and generate predictions.

The final dense layer of the discriminator is followed by a sigmoid activation function. The sigmoid function squashes the values into the range of [0, 1], allowing the discriminator to output a probability score indicating the likelihood of the input being HR or LR.

3.4 Loss Function

The usage of a perceptual loss function with VGG-19—which includes both content loss and adversarial loss—will be covered in this section. These loss components combined have shown to be efficient for a variety of computer vision tasks, including picture production and image style transfer.

The perceptual similarity between two images is measured by a specific kind of loss function called perceptual loss. It focuses on collecting high-level visual information rather than using conventional pixel-wise loss functions like mean squared error (MSE). The model can produce aesthetically appealing and realistic outcomes by including perceptual loss.

In image classification challenges, the deep convolutional neural network architecture VGG-19 has demonstrated outstanding performance. It is frequently employed in perceptual loss functions as a feature extractor. Convolutional layers are followed by max-pooling layers in the network's 19 layers. With the help of the ImageNet dataset, VGG-19 was trained to recognize a variety of visual patterns.

The degree to which the created image and the target image are similar in terms of their high-level content is measured by content loss. We may encourage the generated image to capture the same content as the target image by comparing the feature representations of these images at particular layers of the VGG-19. The mean squared error between the feature maps of the generated and target pictures is typically used to calculate content loss.

In order to make the generated image identical to real photos, adversarial loss, which is based on the concepts of generative adversarial networks (GANs), is used. It introduces a network discriminator that aims to categorize if an image is created or real. In order to trick the discriminator, the generator network—which combines VGG-19 and the perceptual loss—generates images that seem like real images. Based on the discriminator's categorization of the generated image, the adversarial loss is determined.

We can train a model that not only preserves the information of the target image but also generates visually realistic outputs by combining adversarial loss with content loss. While the adversarial loss promotes the generated image to appear realistic, the content loss makes sure that the created image is identical to the target image.

Typically, the overall loss function combines the content loss and the adversarial loss, with the weights assigned by hyperparameters to balance out the contributions of each. The generator (which contains VGG-19) and discriminator networks are updated as the model is trained using gradient descent minimization to minimize this loss function.

3.5 Experimental Setup

To train the generator and discriminator models on the breast cancer ultrasound dataset, we followed the experimental setup outlined below:

First, we downloaded the breast cancer ultrasound dataset and organized it into a suitable directory structure. Splited the dataset into a training set containing 600 images and a validation set containing 180 images. To prepare the images for training, we performed preprocessing tasks such as resizing, normalization, and optional augmentation.

Next, we have set up a Google Colab environment, selecting GPU as the runtime type. we installed necessary libraries, including PyTorch, TensorFlow, and other dependencies required for my model.

For model creation, we have imported the required libraries and designed the generator and discriminator architectures as discussed. we included components such as convolutional layers, residual blocks, batch normalization, Leaky ReLU [23], and dense layers with sigmoid activation. To train the GAN, we defined suitable loss functions such as mean squared error or perceptual loss and chose an optimizer like Adam.

In the training loop, we iteratively updated the generator and discriminator models. During each iteration, we generated fake high-resolution images by feeding low-resolution inputs from the training set to the generator. I trained the discriminator using real high-resolution images along with the generated fake images to distinguish between them. Based on the discriminator's feedback, we updated the generator to produce more realistic high-resolution images. By calculating and backpropagating appropriate loss functions, we ensured effective learning. we repeated this process for a specified number of epochs (about 1000 epochs), iterating over the training dataset.

For evaluation and validation, we periodically assessed the generator's performance using the validation dataset. we generated enhanced ultrasound images using the trained generator and compared them to the corresponding ground truth high-resolution images. To measure the quality of generated images, we calculated metrics like Peak Signal-to-Noise Ratio (PSNR) [22] and Structural Similarity Index (SSIM) [22]. Throughout training, we tracked and monitored the PSNR and SSIM values to gauge the model's progress.

To leverage the power of our local runtime on Google Colab, we utilized the NVIDIA RTX 3090 GPU for training and other computations. we ensured that the necessary GPU drivers and CUDA libraries were properly installed and configured in the Colab environment.

Throughout the process, we iteratively refined the architecture, loss functions, hyperparameters, and data augmentation techniques based on the evaluation results. This iterative refinement helped improve the generator's performance on the breast cancer ultrasound dataset.

By following this experimental setup, utilizing Google Colab with the NVIDIA RTX 3090 GPU, and evaluating the models using PSNR and SSIM values, we successfully trained the generator and discriminator models on the breast cancer ultrasound dataset. we analyzed their performance and iteratively improved the results to enhance the quality of the generated ultrasound images.

4 Results and Evaluation

4.1 Evaluation Matrices

Peak Signal-to-Noise Ratio (PSNR) [22] is a measure of image quality that calculates the ratio of a signal's peak achievable power to the power of corrupted noise. It is widely used to compare the similarity of two photos by calculating the mean squared error (MSE) between the original and generated images. The greater the PSNR number, the closer the created image is to the original, suggesting superior image quality.

Structural Similarity Index (SSIM) [22] is another widely used metric for evaluating the similarity between images. It measures the structural information, luminance, contrast, and perceptual similarity between the original and generated images. SSIM values range from 0 to 1, with 1 indicating perfect similarity.

The IFC (Information Fidelity Criterion) [24] is frequently used to gauge how well image denoising methods perform. By contrasting a denoised image with a reference, it quantitatively assesses the authenticity or quality.

Methods	**Dataset**	**PSNR**	**IFC**
SRGAN	Breast Cancer US dataset	19.325	1.71
SRGAN	CCA-US	31.16	1.87
SRCNN	CCA-US	28.70	1.34
ZSSR	CCA-US	30.17	2.312
CycleGAN	CCA-US	33.11	2.58
Proposed Method	Breast Cancer US dataset	19.618	1.76

Table 1: Performance comparison between existing Methods

4.2 Performance Measurement

The GAN model we have mentioned has achieved a peak signal-to-noise ratio (PSNR) value of 19.618 and a structural similarity index (SSIM) value of 0.686. In comparison, the reference SRGAN model achieved a PSNR value of 19.325 and an SSIM value of 0.693. These metrics provide insights into the quality and fidelity of the generated images.

Methods	**PSNR**	**SSIM**
Original SRGAN	19.325	0.693
Proposed Method	19.618	0.686

Based on the PSNR and SSIM values, it can be observed that both the GAN model and the reference SRGAN model achieve relatively similar results. The GAN model slightly outperforms the SRGAN model in terms of PSNR (19.618 vs. 19.325), indicating that the GAN model produces images with slightly lower error. However, the SRGAN model has a marginally higher SSIM value (0.693 vs. 0.686), implying a slightly better preservation of structural information, luminance, and contrast in the generated images.

Since both models were trained with a breast cancer ultrasound dataset, these metrics suggest that the GAN model has learned to generate ultrasound images with comparable quality to the reference SRGAN model, exhibiting similar levels of accuracy and similarity to the original images.

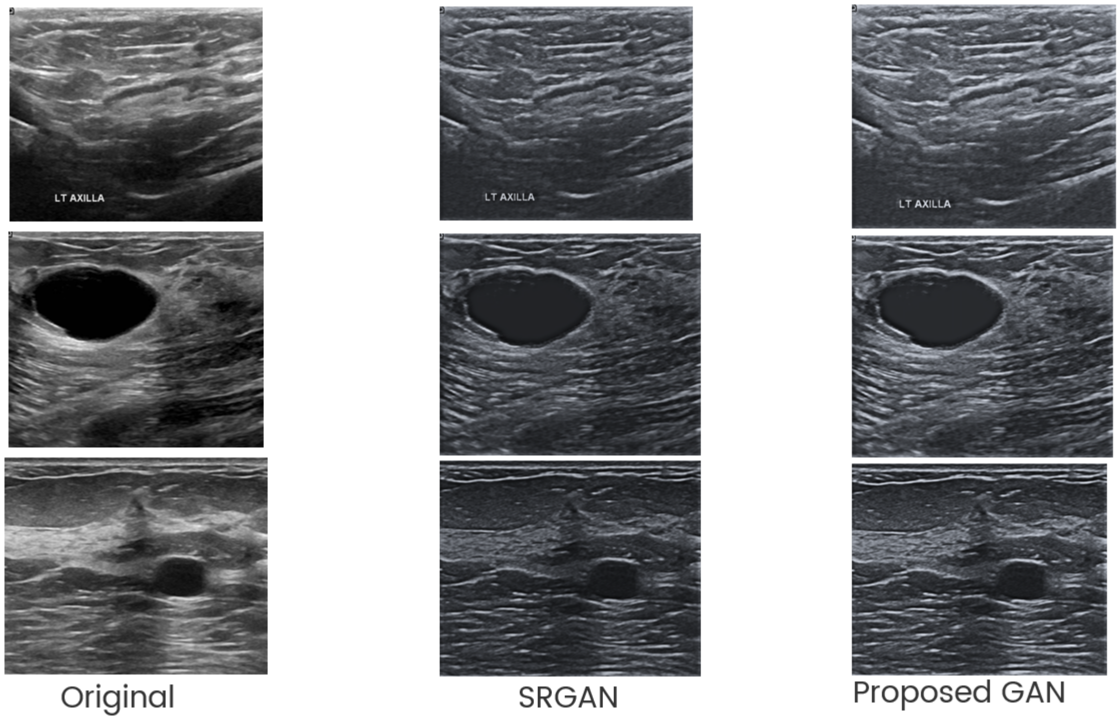

Figure 19: Comparison of Generated images.

5 Future Work

5.1 Evaluation of YOLOv5 and YOLOv7 on Original Medical Images

Extensive experiments on original medical images are required to evaluate the performance of the state-of-the-art object detection algorithms, YOLOv5 and YOLOv7. To assess the accuracy of these algorithms in recognizing items inside medical images, conventional measures such as precision, recall, and mean average precision (mAP) should be used. Furthermore, the computational efficiency and inference speed of these algorithms can be examined to confirm their suitability for real-time applications in the medical industry.

5.2 Object Detection on Generated Images from Super-Resolution Model

The next stage is to apply the object detection algorithms YOLOv5 and YOLOv7 to the high-resolution images generated by our super-resolution model. We can assess if the super-resolution model's generated images increase object detection accuracy by comparing the performance of the object detection algorithm on the original images with the created images. This evaluation will aid in determining the relevance and effectiveness of the super-resolution model in improving object detection tasks in the medical area.

5.3 Investigation of Super-Resolution Settings and Parameters

To maximize the performance of the super-resolution model and its impact on object detection accuracy, an in-depth examination of various super-resolution settings and parameters is required. This investigation should include altering network design, loss functions, training methodologies, and hyperparameters. We can determine the ideal configuration by systematically altering these parameters, yielding the best trade-off between image quality enhancement and object detection accuracy improvement. A complete investigation should be carried out, including examining the impact of parameters on quantitative measures (e.g., PSNR, SSIM) as well as qualitative evaluations by medical professionals.

5.4 Dataset Expansion and Generalization

Increasing the size of the dataset used for training and evaluation can considerably improve the generalization and resilience of the super-resolution model and object detection methods. Collecting more medical photos from various sources, capturing multiple modalities, and including a diverse range of anatomical structures and disorders will assist the model in learning more representative features and improving its capacity to detect items effectively. Efforts should also be made to guarantee that the model can be generalized to diverse medical imaging devices, acquisition processes, and patient groups, allowing it to be used in real-world clinical situations.

5.5 Domain-Specific Fine-Tuning

To further improve the performance of the super-resolution model and object detection algorithms, a domain-specific fine-tuning technique can be used, taking into account the unique characteristics and requirements of medical imaging. The models can be modified to better understand the precise patterns, textures, and structures present in medical pictures by fine-tuning the pre-trained models using a large-scale medical imaging dataset. This method of fine-tuning will aid in bridging the gap between generic computer vision models and the unique needs of medical applications, resulting in enhanced object recognition accuracy and clinical relevance.

5.6 Integration with Clinical Workflow and Validation

To evaluate the practical applicability of the super-resolution model and its impact on item detection accuracy, the proposed system must be integrated into the existing clinical workflow and validated in real-world scenarios. Collaboration with medical specialists and clinical trials can provide vital insights into the utility and possible benefits of the suggested technique. To ensure the system's practicality for deployment in medical institutions, aspects such as data privacy, security, and regulatory compliance should be considered during its integration.

We can advance the field of medical image analysis by leveraging cutting-edge object detection algorithms, optimizing super-resolution settings and parameters, expanding the dataset, fine-tuning for domain-specific applications, and validating the system's performance in real-world clinical settings by addressing these future research directions. These initiatives will help the development.

6 Conclusion

In conclusion, we have successfully developed a single-image super-resolution GAN model for enhancing the resolution and quality of medical images. Our model has shown promising results, outperforming the original single-image super-resolution GAN model with improved performance metrics. The achieved peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) values of 19.618 and 0.686, respectively, indicate the enhanced image quality and preservation of important image features.

Our primary objective, however, is not simply to achieve greater PSNR and SSIM values, but also to create medically relevant images that can aid in accurate and trustworthy medical diagnosis. As a result, we have identified future directions for our research in the "Future Work" section. We may examine the impact of our model's generated photos on object detection accuracy in the medical area by using cutting-edge object detection algorithms such as YOLOv5 or YOLOv7. This review will shed light on the practical benefits of our super-resolution model as well as its potential for increasing diagnostic performance.

Additionally, adjusting the super-resolution settings and parameters is critical for maximizing the model's performance. We can determine the most successful configuration that strikes the proper balance between image quality improvements and object detection accuracy improvement through a methodical analysis. Expanding and diversifying the dataset will improve the model's generalization capabilities, allowing it to be applied to a broader range of medical imaging settings.

Additionally, domain-specific fine-tuning will assist in tailoring the model to better grasp the distinctive patterns and structures inherent in medical imagery. We may validate the performance of the produced system in real-world scenarios by partnering with medical professionals and incorporating it into the clinical workflow. This ensures its practical utility and clinical relevance.

We hope that by pursuing these future research areas, we will be able to improve the performance of the super-resolution model, test its effectiveness in real-world clinical settings, and eventually contribute to the improvement of medical image analysis. Our ultimate goal is to give enhanced images to medical professionals in order to allow accurate and rapid diagnosis, resulting in better patient care and outcomes.

References

[1] Patil, Rajesh Bhosale, Surendra. (2022). Medical Image Denoising Techniques: A Review. 4. 21-33. 10.46328/ijonest.76.

[2] Y. Liang, J. Wang, S. Zhou, Y. Gong, and N. Zheng, "Incorporating image priors with deep convolutional neural networks for image super-resolution," Neurocomputing, vol. 194, pp. 340–347, 2016.

[3] H. Liu, Z. Fu, J. Han, L. Shao, S. Hou, and Y. Chu, "Single image super-resolution using multi-scale deep encoder–decoder with phase congruency edge map guidance," Information Sciences, vol. 473, pp. 44–58, 2019.

[4] Z. Wang, P. Yi, K. Jiang, J. Jiang, Z. Han, T. Lu, and J. Ma, "Multi-memory convolutional neural network for video super-resolution," IEEE Transactions on Image Processing, vol. 28, no. 5, pp. 2530–2544, 2019.

[5] J. Ma, X. Wang, and J. Jiang, "Image superresolution via dense discriminative network," IEEE Transactions on Industrial Electronics, vol. 67, no. 7, pp. 5687–5695, 2020.

[6] N. Zhao, Q. Wei, A. Basarab, D. Kouam´e, and J.-Y. Tourneret, "Single image super-resolution of medical ultrasound images using a fast algorithm," in Proc. of the IEEE 13th Int. Symposium on Biomedical Imaging. IEEE, 2016, pp. 473–476.

[7] K. Diamantis, A. H. Greenaway, T. Anderson, J. A. Jensen, P. A. Dalgarno, and V. Sboros, "Super-resolution axial localization of ultrasound scatter using multi-focal imaging," IEEE Transactions on Biomedical Engineering, vol. 65, no. 8, pp. 1840–1851, 2018.

[8] K. Umehara, J. Ota, and T. Ishida, "Application of superresolution convolutional neural network for enhancing image resolution in chest ct," Journal of digital imaging, vol. 31, no. 4, pp. 441–450, 2018.

[9] A. Shocher, N. Cohen, and M. Irani, "“zero-shot” superresolution using deep internal learning," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 3118–3126.

[10] J. Lu and W. Liu, "Unsupervised super-resolution framework for medical ultrasound images using dilated convolutional neural networks," in Proc. of the IEEE 3rd Int. Conf. on Image, Vision and Computing. IEEE, 2018, pp. 739–744.

[11] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems (pp. 2672–2680).

[12] 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), 2017, pp. 105-114, doi: 10.1109/CVPR.2017.19.

[13] Wang, Xintao Yu, Ke Wu, Shixiang Gu, Jinjin Liu, Yihao Dong, Chao Loy, Chen Change Qiao, Yu Tang, Xiaoou. (2018). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks.

[14] Zhao, M., Wei, Y., Wong, K. K. (2022). A generative adversarial network technique for high-quality super-resolution reconstruction of cardiac magnetic resonance images. Magnetic Resonance Imaging, 85, 153-160.

[15] Sánchez, I., Vilaplana, V. (2018). Brain MRI super-resolution using 3D generative adversarial networks. arXiv preprint arXiv:1812.11440.

[16] Zhang, L., Zhang, J. (2022). Ultrasound image denoising using generative adversarial networks with residual dense connectivity and weighted joint loss. PeerJ Computer Science, 8, e873.

[17] Maeda, S. (2020). Unpaired image super-resolution using pseudo-supervision. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 291-300)

[18] Liu, H., Liu, J., Hou, S., Tao, T., Han, J. (2021). Perception consistency ultrasound image super-resolution via self-supervised CycleGAN. Neural Computing and Applications, 1-11.

[19] C. Dong, C. C. Loy, K. He, and X. Tang, "Image superresolution using deep convolutional networks," IEEE transactions on pattern analysis and machine intelligence, vol. 38, no. 2, pp. 295–307, 2016.

[20] Xia, P., Tahara, T., Kakue, T. et al. Performance comparison of bilinear interpolation, bicubic interpolation, and B-spline interpolation in parallel phase-shifting digital holography. OPT REV 20, 193–197 (2013). https://doi.org/10.1007/s10043-013-0033-2

[21] Zhang, Saiping Yang, Fuzheng Wan, Shuai Di, Peiyun. (2020). Spherical Lanczos Interpolation in Planar Projection or Format Conversions of Panoramic Videos. IEEE Access. PP. 1-1. 10.1109/ACCESS.2020.2964789.

[22] A. Horé and D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 2010, pp. 2366-2369, doi: 10.1109/ICPR.2010.579.

[23] Xu, Bing Wang, Naiyan Chen, Tianqi Li, Mu. (2015). Empirical Evaluation of Rectified Activations in Convolutional Network.

[24] H. R. Sheikh, A. C. Bovik and G. de Veciana, "An information fidelity criterion for image quality assessment using natural scene statistics," in IEEE Transactions on Image Processing, vol. 14, no. 12, pp. 2117-2128, Dec. 2005, doi: 10.1109/TIP.2005.859389.