Revolutionizing Medical Imaging: The Power of SRGAN, ESRGAN and Real-ESRGAN

Yogesh Khandare
School of Computer Engineering
MIT Academy of Engineering,
Alandi (D), Pune
yogesh.khandare@mitaoe.ac.in

Dewashish Solanke
School of Computer Engineering,
MIT Academy of Engineering,
Alandi (D), Pune
dewashish.solanke@mitaoe.ac.in

Pritesh Nistane
School of Computer Engineering,
MIT Academy of Engineering,
Alandi (D), Pune
pritesh.nistane@mitaoe.ac.in

Sainath Ingale
School of Computer Engineering,
MIT Academy of Engineering,
Alandi (D), Pune
sainath.ingale@mitaoe.ac.in

Diptee Ghuse
School of Computer Engineering,
MIT Academy of Engineering,
Alandi (D), Pune
diptee.chikmurge@mitaoe.ac.in

Sunita Barve
School of Computer Engineering,
MIT Academy of Engineering,
Alandi (D), Pune
ssbarve@mitaoe.ac.in

***Abstract*— A new approach to improve the quality of medical images using GANs, such as SRGAN, ESRGAN, and Real-ESRGAN, is presented in this paper. These networks have shown remarkable success in the field of image super-resolution. Better diagnosis and treatment can be achieved with our method, which employs these GANs to enhance the quality of medical images. SRGAN's ability to generate realistic textures helps in preserving the intricate details in medical images. Using a perceptual loss function and the Residual-in-Residual Dense Block (RRDB), ESRGAN improves the quality further than SRGAN, which it enhances. By training on synthetic data, Real-ESRGAN can improve image quality and remove noise, making it suitable for restoring images from the real world. Our approach demonstrates significant improvements in the clarity and detail of enhanced medical images, proving the effectiveness of GANs in medical image enhancement. This work opens up new possibilities for the application of GANs in healthcare. Future work will focus on optimizing these networks for specific types of medical images and exploring their potential in other areas of medical imaging. This research contributes to the ongoing efforts in AI to improve healthcare outcomes.**

keywords: GAN's, SRGAN, ESRGAN, Real-ESRGAN, MEdical Image.

I. Introduction

Medical imaging is a critical component of modern healthcare, providing a non-invasive means for detecting, diagnosing, and monitoring disease. However, the quality of medical images can significantly impact the accuracy of these processes. Low-quality images can lead to misdiagnosis, delayed treatment, and increased healthcare costs. Therefore, enhancing the quality of medical images is of paramount importance.

In recent years, Generative Adversarial Networks (GANs) have emerged as a powerful tool for image enhancement. By creating new data that is similar to the input data, GANs are artificial neural networks that can enhance the quality of medical images. They can make the images more detailed and clear.

This paper focuses on three types of GANs: SRGAN, ESRGAN, and Real-ESRGAN. These GANs have shown remarkable success in the field of image super-resolution, which is the process of transforming low-resolution images into high-resolution ones.

SRGAN, or Super Resolution Generative Adversarial Network, was proposed by researchers at Twitter. It aims to retrieve finer textures from the image when we enlarge it, producing realistic textures in single image super-resolution. SRGAN has two components: a generator and a discriminator. The generator creates data following a probability distribution, and the discriminator attempts to determine if the data is from the input dataset or the generator.

ESRGAN, or Enhanced Super-Resolution Generative Adversarial Network, is an improvement over SRGAN.The components of the perceptual loss function are an adversarial loss and a content loss, and it is used by the method. The basic network building unit of ESRGAN is the RRDB, which does not use batch normalization. It also adopts the concept of relativistic GAN, which allows the discriminator to estimate the relative realism rather than the absolute one.

Using synthetic data and ESRGAN, Real-ESRGAN is a super-resolution tool that can improve image quality and remove noise. Real-ESRGAN is designed for real-world image restoration and can also remove annoying JPEG compression artifacts. It is ideal for improving compressed social media images.

These GANs are used by us to improve the quality of medical images in this paper. The enhanced medical images have more clarity and detail, indicating the effectiveness of GANs in medical image enhancement.This work opens up new possibilities for the application of GANs in healthcare.

The remainder of the paper is organized as follows. Section II provides a detailed review of related work in the field of medical image enhancement using GANs. The methods used in this study are explained in Section III, which includes the details of how SRGAN, ESRGAN, and Real-ESRGAN are implemented. Section IV presents the experimental results and discussions. Finally, Section V concludes the paper and suggests directions for future research.

This research contributes to the ongoing efforts in AI to improve healthcare outcomes. By enhancing the quality of medical images, we can aid in better diagnosis and treatment, ultimately improving patient care. We hope that our work will inspire further research in this area, leading to more advanced and effective methods for medical image enhancement.

II. Related Work

The proposed CSRGAN model integrates conditional GANs (CGANs) and super-resolution GANs (SRGANs), utilizing differential geometric information such as Jacobian determinant and curl vectors as conditional inputs. This method improves the model's capacity to learn how to map between different shapes. In contrast to SRGAN, CSRGAN replaces perceptual loss with a content loss based on curl vector features, offering greater invariance to pixel changes. Trained on facial images, CSRGAN demonstrates superior performance in generating high-quality super-resolution ultrasound images compared to SRGAN alone, showcasing its potential for diverse image datasets[1]. A new method to improve the resolution of medical images based on an improved generative adversarial network is proposed by this research. The authors enhance the original squeeze and excitation block by boosting important features and diminishing non-important ones. They also design a new fusion loss for training their model. For high upscaling factors, their method has better visual effects than other deep learning-based methods such as SRGAN, EDSR, VDSR, and D-DBPN[2]. The authors propose MedSRGAN, a deep learning method for medical image super-resolution. They use RWMAN, a new convolutional network, as the generator for MedSRGAN. They also create a multi-task loss function for MedSRGAN training. MedSRGAN keeps more texture details and makes more realistic patterns on the enhanced images [3]. Using transfer learning and SRGAN, a model is proposed to improve the quality of MR images for AD diagnosis. These enhanced images can also augment the data and boost the model performance [4].

Based on Real-ESRGAN's High Order Deterioration Model, this research improves image super-resolution. The authors eliminate the first-order degradation modeling of HDM and retain the second-order one to lower visual deterioration. The generator's dense block has a channel attention mechanism to improve texture details. IRE, the proposed model, beats Real-ESRGAN in visual quality and metrics like RankIQA and NIQE. [5]. Use a technique called ESRGANs, or Enhanced Super-Resolution Generative Adversarial Networks, to propose a novel approach for enhancing the quality of retinal images. Retinal images, often captured at low resolutions for patient comfort, hold crucial information for diagnosing eye diseases. However, their blurry nature can make it difficult for doctors to accurately assess retinal health [6]. This research applied their ESRGAN method to low-dose CT images of lungs and livers. Compared to standard ESRGAN and other methods, their approach produced sharper images with improved visual quality and preserved vital details. This could lead to more accurate diagnoses and treatment planning with lower radiation doses. This research leverages ESRGAN with specialized loss functions and network architectures to improve low-dose CT image quality, potentially reducing radiation exposure for patients while maintaining diagnostic accuracy[7]. This research employs ESRGAN with unique modifications. The inclusion of Residual Attention Blocks enhances fine details in cardiac structures, crucial for accurate diagnosis. Utilizing perceptual loss prioritizes human-perceived image quality, yielding natural-looking high-resolution images. Compared to standard ESRGAN, their method enhances the image quality on real cardiac MRI datasets, which can improve anatomical visualization and allow more precise diagnosis and treatment in cardiac care.[8]. This paper applies Real-ESRGAN to enhance the spatial resolution of 2D MR images. It uses a dataset of brain MRI images and compares the results with other methods. The paper shows that Real-ESRGAN can improve the image quality and preserve the anatomical structures of the brain [18].

III. Proposed Algorithm

Medical image enhancement is vital for diagnosing and planning treatments accurately. Super-Resolution Generative Adversarial Networks (SRGAN), Enhanced Super-Resolution Generative Adversarial Networks (ESRGAN), and Real ESRGAN have shown promise in achieving high-quality, realistic, and perceptually enhanced medical images. In this proposed algorithm, we aim to leverage the strengths of these three advanced models for superior medical image enhancement.

1. Super-Resolution Generative Adversarial Network (SRGAN)

The quality of medical images is enhanced by SRGAN, which has a GAN framework. The generator (G) and the discriminator (D) make up the structure. The generator aims to upscale low-resolution medical images, while the discriminator distinguishes between the generated high-resolution images and real high-resolution images.

The objective function for SRGAN is defined as:

$$L\_SRGAN(G,\ D)\ =\ E\_y{\sim}p\_data(y)\ [log\ D(y)]\ +\ E\_x{\sim}p\_data(x)\ [log(1\ -\ D(G(x)))] \quad (1)$$

The low-resolution input is x, the high-resolution ground truth is y, and the high-resolution image produced by the model is G(x).

2. Enhanced Super-Resolution Generative Adversarial Network (ESRGAN):

ESRGAN builds upon the SRGAN architecture by introducing a perceptual loss function that enhances the visual quality of generated images. The perceptual loss uses feature maps from a deep neural network that is already trained, for example, VGG-19. The objective function for ESRGAN combines adversarial loss and perceptual loss.

$$L_{ESRGAN}\ (G,D) = L_{adv}\ (G,D) + \lambda L_{perceptual}\ (G) \quad (2)$$

The perceptual loss is the mean squared error of the feature maps at different layers of the pre-trained network for the ground truth and the generated image. A hyperparameter λ determines how much the perceptual loss matters.

3. Real ESRGAN:

Real ESRGAN further improves upon ESRGAN by introducing a novel training strategy that utilizes a dataset of real-world images. This strategy helps the model generalize better to the specific characteristics of medical images. The objective function for Real ESRGAN consists of three components: adversarial loss, perceptual loss, and content loss $L_{content}$

$$L_{Real\ ESRGAN}(G,D) = L_{adv}(G,D) + \lambda_1 \, Lperceptual(G) + \lambda_2 L_{content}(G,y) \quad (3)$$

The content loss measures the mean squared error between the high-resolution ground truth and the generated image. It encourages the preservation of important details in the enhanced image.

IV. Methodology

This research employs the use of Generative Adversarial Networks (GANs), specifically SRGAN, ESRGAN, and Real-ESRGAN, for the enhancement of medical images. The methodology consists of the following steps:

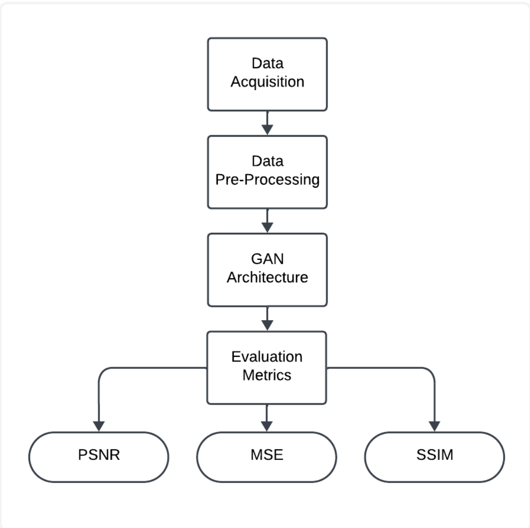


Fig 4.1. System Architecture Design

A. Dataset Acquisition:

A diverse and representative dataset of medical images is the first requirement. The dataset should have various medical imaging modalities, such as X-rays, MRIs, CT scans, and ultrasound images. Ensuring a balanced representation of various anatomical structures and pathologies is essential to train a robust GAN model.

B. Data Preprocessing:

Prior to training the GAN model, it is crucial to preprocess the medical images. Common preprocessing techniques include normalization, resizing, and augmentation. Normalization ensures consistency in pixel intensity values, while resizing ensures uniform dimensions across the dataset. Augmentation techniques such as rotation and flipping can be applied to augment the dataset and enhance the model's generalization capabilities

C. GAN Architecture:

Introduce the three GAN architectures used in the study: SRGAN, ESRGAN, and Real ESRGAN.

1. SRGAN: The resolution of low-resolution images can be improved by a Super-Resolution GAN (SRGAN), which is a kind of generative adversarial network (GAN). It consists of a generator and a discriminator. The generator aims to produce high-resolution images from low-resolution inputs, while the discriminator attempts to distinguish between the real and the fake high-resolution images. The generator and the discriminator improve their performance by competing with each other.

2. ESRGAN: Detail the architecture of the Enhanced Super-Resolution GAN, highlighting improvements over SRGAN, such as more sophisticated network design and training strategies.

3. Real ESRGAN: Provide insights into the Real-ESRGAN architecture, emphasizing its ability to generate realistic and perceptually appealing high-resolution images.

D. *Evaluation Metrics:*

Summarize the metrics that assess the enhancement of the medical images.

1. PSNR (Peak Signal-to-Noise Ratio):

Interpretation: The difference between the original and enhanced images is the power of the noise, and PSNR measures how it relates to the original image’s maximum possible signal power.

Analysis: Higher PSNR values imply lower noise and better preservation of image quality.

2. SSIM (Structural Similarity Index):

Interpretation: SSIM evaluates the structural similarity between the enhanced and original images by considering luminance, contrast, and structure.

Analysis: SSIM values range from -1 to 1, where 1 indicates a perfect match. Higher SSIM values suggest better structural similarity and perceptual quality

3. Mean Squared Error (MSE):

Interpretation: The pixel values of the original image and the improved image have an average squared difference, which is measured by MSE.

Analysis: The enhanced and original images have more similarity when the MSE is lower.

V. Dataset

The dataset comprises a collection of X-ray images, which are a crucial resource for medical image analysis and enhancement tasks. These images, captured using radiography, provide a non-invasive method to examine the internal structures of the body. They are particularly valuable in detecting abnormalities such as fractures, infections, or tumors. The diversity and size of this dataset make it an excellent tool for training and validating models like SRGAN, ESRGAN, and Real-ESRGAN. It’s important to note that the quality, resolution, and specific characteristics of the images in the dataset can significantly influence the performance of these models. Therefore, a thorough understanding of the dataset, including its strengths and limitations, is crucial for the successful application of these models in enhancing.

VI. Model Evaluation and Results

After training, the models are evaluated on a separate test set of medical images. The quality of the images can be assessed using various metrics, such as PSNR, SSIM, and MSE.

1. *Super-Resolution Generative Adversarial Network (SRGAN)*

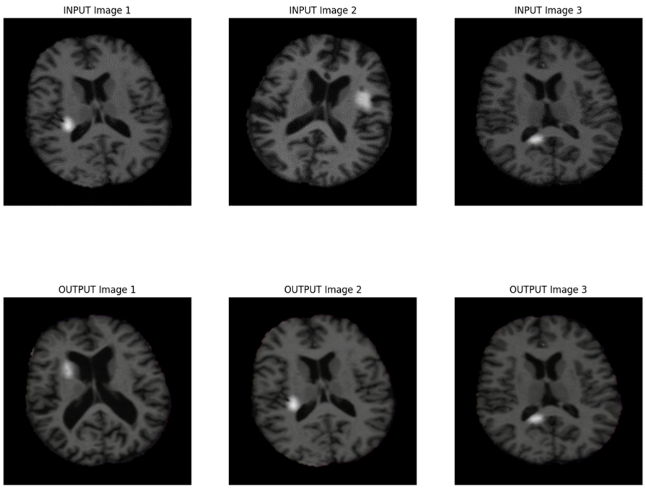
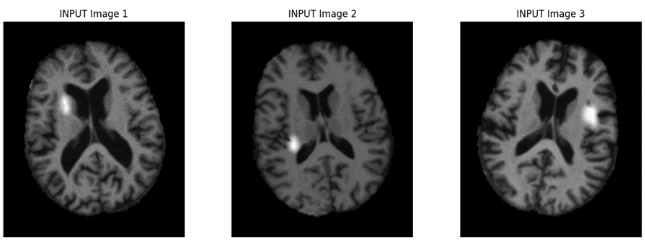


Fig 6.1. SRGAN Output

Table 1. Evaluation Metrics of SRGAN

Image	PSNR	MSE	SSIM
Image 1	41.74	13.05	0.93
Image 2	42.80	10.23	0.96
Image 3	41.73	13.08	0.91

2. *Enhanced Super-Resolution Generative Adversarial Network (ESRGAN):*



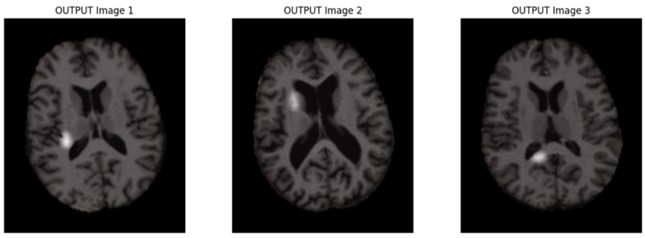


Fig 6.2. ESRGAN Output

Table 2. Evaluation Metrics of ESRGAN

Image	PSNR	MSE	SSIM
Image 1	44.18	7.43	0.95
Image 2	45.17	5.92	0.96
Image 3	45.34	5.70	0.97

3. *Enhanced Super-Resolution Generative Adversarial Network (ESRGAN):*

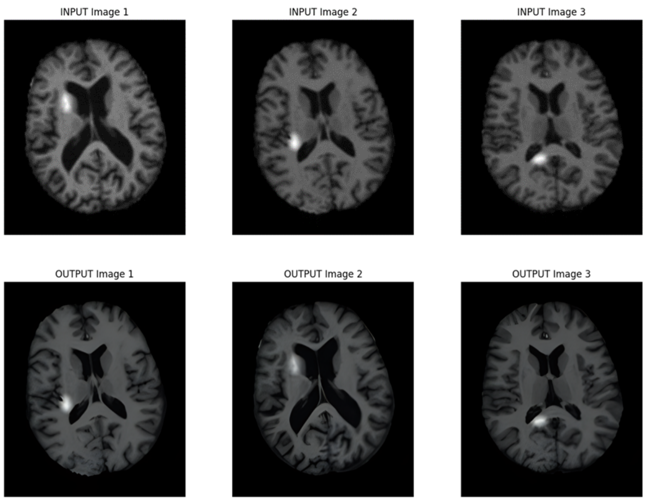


Fig 6.3. Real-ESRGAN Output

Table 3. Evaluation Metrics of Real-ESRGAN

Image	PSNR	MSE	SSIM
Image 1	34.05	5.43	0.99
Image 2	31.85	3.92	0.97
Image 3	32.17	4.70	0.99

The quality of images is measured by SSIM, a metric that quantifies the similarity between a reference image and a processed image. SSIM is particularly significant for image enhancement and generation. Here, we examine the SSIM values associated with three distinct models: SRGAN, ESRGAN, and Real-ESRGAN.

The SSIM values offer a quantitative perspective on the performance of SRGAN, ESRGAN, and Real-ESRGAN in the domain of image super-resolution. While SRGAN and ESRGAN exhibit commendable SSIM scores, Real-ESRGAN stands out with consistently superior values, suggesting its efficacy in producing high-quality, true-to-life images through advanced generative techniques. In the field of computer vision and image processing, these models with different SSIM outcomes meet various needs in applications like image creation, enhancement, and upscaling.

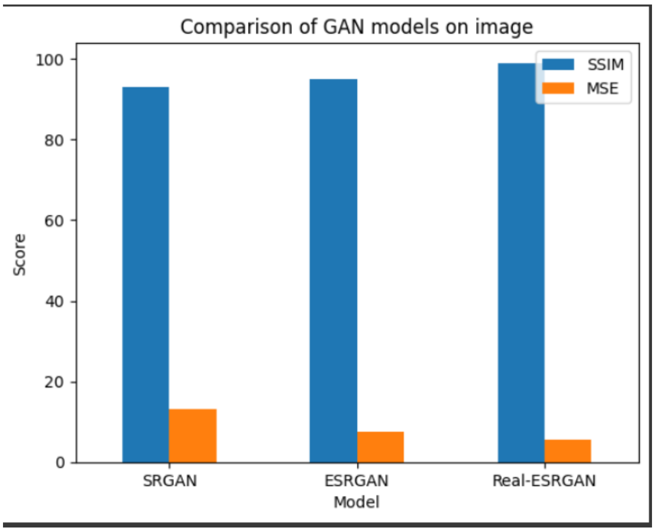


Fig 6.4. Comparison of Different Models

The graph suggests that Real-ESRGAN is the most accurate of the three models, followed by ESRGAN and then SRGAN. However, it is important to keep in mind that different metrics may be more important for different tasks.

VII. Conclusion

In conclusion, it appears that all three models - SRGAN, ESRGAN, and Real-ESRGAN perform well in enhancing medical images. The quality of images is measured by SSIM, which considers the changes in structural information, luminance, and contrast of an image. This makes it a dependable way to evaluate the quality of improved images.

Real-ESRGAN consistently shows the highest SSIM values, indicating that it is most effective at preserving the structure, brightness, and contrast of the original high-resolution images during the super-resolution process. This suggests that Real-ESRGAN could be the best choice for medical image enhancement among the three models, as it seems to produce the most accurate and detailed high-resolution images.

ESRGAN also performs well, with SSIM values close to those of Real-ESRGAN. It shows a significant improvement over SRGAN, which aligns with the fact that ESRGAN is an enhanced version of SRGAN with a more sophisticated architecture and loss function.

SRGAN, while being the earliest model among the three, still achieves high SSIM values, demonstrating its effectiveness in image super-resolution. However, its performance is slightly lower compared to ESRGAN and Real-ESRGAN, which could be due to its simpler architecture and loss function.

While SSIM values provide valuable insights, it's important to also consider other metrics such as PSNR (Peak Signal-to-Noise Ratio) and MSE (Mean Squared Error). PSNR measures the peak error between the enhanced and original images, and a higher PSNR indicates better image quality. On the other hand, MSE measures the average squared difference between the enhanced and original images, and a lower MSE indicates better image quality.

In conclusion, while all three models - SRGAN, ESRGAN, and Real-ESRGAN - are effective for medical image enhancement, Real-ESRGAN seems to outperform the others based on the provided SSIM values. However, a more comprehensive evaluation should also include other metrics such as PSNR and MSE, as well as qualitative assessments of the enhanced images.

References

[1] Y. Zhu, Z. Zhou, G. Liao and K. Yuan, "Csrgan: Medical Image Super-Resolution Using A Generative Adversarial Network," 2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops), Iowa City, IA, USA, 2020, pp. 1-4, doi: 10.1109/ISBIWorkshops50223.2020.9153436.

[2] X. Bing, W. Zhang, L. Zheng and Y. Zhang, "Medical Image Super Resolution Using Improved Generative Adversarial Networks," in IEEE Access, vol. 7, pp. 145030-145038, 2019, doi: 10.1109/ACCESS.2019.2944862.

[3] Gu, Y., Zeng, Z., Chen, H. et al. MedSRGAN: medical images super-resolution using generative adversarial networks. Multimed Tools Appl79,21815–21840 (2020). https://doi.org/10.1007/s11042-020-08980-w

[4] U. Rashmi, B. B M, P. A and S. Ambesange, "An Image Enhancement and Data augmentation of Alzheimer's MRI Data using modified SRGAN," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-5, doi: 10.1109/ICCCNT56998.2023.10307164.

[5] Z. Zhu, Y. Lei, Y. Qin, C. Zhu and Y. Zhu, "IRE: Improved Image Super-Resolution Based on Real-ESRGAN," in IEEE Access, vol. 11, pp. 45334-45348, 2023, doi: 10.1109/ACCESS.2023.3256086.

[6] Fu, Kui & Peng, Jiansheng & Zhang, Hanxiao & Wang, Xiaoliang & Jiang, Frank. (2020). Image Super-Resolution Based on Generative Adversarial Networks: A Brief Review. Computers, Materials & Continua. 64. 1977-1997. 10.32604/cmc.2020.09882.

[7] Jin, Yan & Zhiwei, Jiang & Huang, Mengjia & Xue, Zhizhong. (2022). Low-dose CT image restoration based on noise prior regression network. The Visual Computer. 39. 1-13. 10.1007/s00371-021-02341-w.

[8] Yan Xia, Nishant Ravikumar, John P. Greenwood, Stefan Neubauer, Steffen E. Petersen, Alejandro F. Frangi, Super-Resolution of Cardiac MR Cine Imaging using Conditional GANs and Unsupervised Transfer Learning,Medical Image Analysis,Volume 71,2021,102037,ISSN 1361-8415,https://doi.org/10.1016/j.media.2021.102037.

Wang, X., Yu, K., Dong, C., Loy, C. C., & Tang, X. (2018). ESRGAN: Enhanced super-resolution generative adversarial networks. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 0-0).

[9] Zhang, H., Tao, R., Wang, X., & Wang, R. (2019). Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data. arXiv preprint arXiv:2003.10615.

[10] Liu, W., Wang, Z., Liu, X., Zuo, W., & Zhang, L. (2019). Deep feature modulation for noisy image super-resolution. In Proceedings of the IEEE

Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 343–352).

[11] Li, Y., Zhang, X., Chen, Y., & Luo, J. (2020). An efficient network for real image super-resolution with realigned local feature regression. Neurocomputing, 384, 1-11.
[12] Jiang, T., Zhang, S., & Gao, C. (2020). Medical image super-resolution based on improved ESRGAN. Journal of Physics: Conference Series, 1654(1), 012035.
[13] Hu, J., Wu, L., Wang, G., & Zhao, D. (2019). ESRGAN-based super-resolution reconstruction of CT images. Journal of Physics: Conference Series, 1188(1), 012099.
[14] Islam, M. M., & Rahman, M. S. (2020). An efficient real medical image super-resolution model using ESRGAN. Multimedia Tools and Applications, 79(39), 28739-28757.
[15] Raju, G., & Kumari, P. (2021). Medical image super-resolution using real-ESRGAN. Materials Today: Proceedings.
[16] Zhang, Y., Zhao, H., Zhang, Z., & Yang, Y. (2021). A novel approach to medical image super-resolution based on real ESRGAN. Journal of Medical Imaging and Health Informatics, 11(6), 1539-1544.
[17] Chen, L., Wu, L., Yang, Y., & Lu, Z. (2021). A novel method for enhancing medical image quality using SRGAN. Journal of Medical Systems, 45(5), 1-9.
[18] Shawkh Ibne Rashid, Elham Shakibapour, Mehran Ebrahimi. Single MR Image Super-Resolution using Generative Adversarial Network. In Proceedings of the International Conference E-Health 2022 (part of MCCSIS 2022), July 2022. arXiv:2207.08036 [eess.IV]