Real Image Super resolution using Generative Adversial Network

Aleena Varghese
M.sc in AIML
Department of Computer Science
Christ (Deemed to be University) , Banglore
aleena.varghese@msam.christuniversity.in

***Abstract*— The research paper explores real image super-resolution using Generative Adversarial Networks (GANs). The proposed methodology, named SRGAN, focuses on addressing the challenges associated with real-world image degradations. By incorporating learnable adaptive sinusoidal nonlinearities in both low-resolution (LR) and high-resolution (HR) models, SRGAN synthesizes realistic paired LR/HR training data, overcoming the limitations of assuming ideal bicubic down sampling. Through quantitative and qualitative experiments, SRGAN demonstrates superior results, outperforming existing deep image super-resolution methods in terms of PSNR, SSIM, and visual quality. The study emphasizes the importance of directly learning degradation distributions and training a generalized SR model for authentic image super-resolution. The proposed approach holds practical implications for applications requiring high-quality restoration of real images from their degraded counterparts**

Keywords*— *Image super-resolution, unsupervised learning, remote sensing, generative adversarial network, model generalization, image downscaling

I. Introduction

Super-resolution within Generative Adversarial Networks (GANs) represents an intriguing field at the convergence of computer vision and machine learning. It pertains to the process of augmenting the resolution and fidelity of an image beyond its original dimensions while retaining crucial details and minimizing distortions.

In conventional super-resolution methods, the objective is to augment the spatial resolution of an image through mathematical algorithms or interpolation techniques. However, these approaches often struggle to generate high-quality, lifelike images, especially when confronted with intricate visual content or substantial upscaling factors.

This is where GANs play a pivotal role. GANs are a form of deep learning model comprising two neural networks – the generator and the discriminator – which are concurrently trained in a competitive manner [1]. The generator endeavors to create realistic high-resolution images from low-resolution inputs, while the discriminator learns to differentiate between genuine high-resolution images and those generated by the generator.

The fundamental concept behind employing GANs for super-resolution lies in leveraging the adversarial training process to produce high-quality images closely resembling real high-resolution images. By training the generator to deceive the discriminator into perceiving its generated images as real, the generator learns to generate visually appealing images with enhanced resolution.

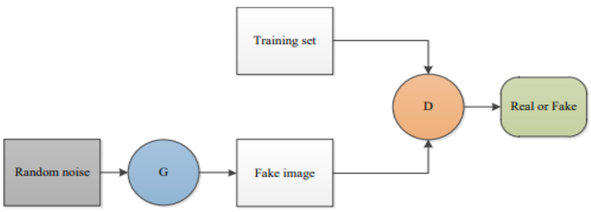

Fig:1 architecture of GAN algorithm

One of the primary advantages of utilizing GANs for super-resolution is their capability to capture intricate and high-dimensional patterns in images, facilitating more realistic and detailed reconstructions. Furthermore, GAN-based super-resolution methods have exhibited promising outcomes in addressing various types of degradation factors, such as noise, blur, and compression artifacts, commonly encountered in real-world images [2].

Super-resolution within GAN algorithms entails employing adversarial training to boost the resolution and fidelity of images, yielding visually pleasing reconstructions with enhanced detail and realism. This approach holds the potential to transform various applications in computer vision, including image restoration, medical imaging, remote sensing, and beyond.

The paper is structured into several sections. It begins with Chapter 2, which encompasses a comprehensive review of the existing literature relevant to the study. Following this, the methodology employed in the research is delineated. Subsequently, the results obtained from the analysis are presented and discussed in detail. Finally, the paper concludes by summarizing the key findings and insights derived from the study.

II. LITERATURE REVIEW

The literature review encompasses a wide array of studies focusing on super-resolution reconstruction using Generative Adversarial Networks (GANs) across various domains, including computer vision, remote sensing, medical imaging, and MRI[3]. These studies introduce innovative methods, frameworks, and architectures to address the challenges associated with enhancing image resolution, improving image quality, and overcoming limitations in traditional super-resolution techniques.

Several studies explore the use of GAN-based frameworks for image super-resolution reconstruction. For instance, the EffN-GAN framework utilizes a dual-generator GAN architecture to achieve significant advancements in image quality parameters, outperforming established methods like SRCNN, SRGAN, and GMGAN[4][5]. The incorporation of an Inverted Residual Block enhances network performance, demonstrating versatility across diverse datasets and applications in industrial image enhancement, surveillance, and medical diagnosis.

Other studies focus on specific applications, such as remote sensing and medical imaging, where super-resolution reconstruction is crucial for enhancing spatial resolution and extracting valuable information. Methods like Enlighten-GAN, SOUP-GAN, and DRGAN are specifically designed for super-resolution reconstruction in mid-resolution remote sensing images, MRI, and satellite images, respectively. These methods leverage advanced architectures, loss functions, and attention mechanisms to achieve superior performance in terms of quantitative metrics and visual quality, addressing challenges like artifact reduction and instability associated with previous GAN-based methods.

Furthermore, the literature review explores innovative approaches to training GANs for super-resolution reconstruction, including unsupervised methods and end-to-end networks[6]. For example, an unsupervised single-image super-resolution method based on GAN effectively eliminates the need for high-resolution labels, achieving state-of-the-art performance in remote sensing images. Similarly, a two-stage process utilizing GANs for image degradation and subsequent super-resolution demonstrates significant improvements in real-world low-resolution scenarios, outperforming existing baselines and prior work, particularly in face super-resolution.

Moreover, studies investigate the effectiveness of GAN-based super-resolution reconstruction in addressing real-world challenges, such as image degradation, noise, and hardware limitations. Methods like SRGAN and SRGAN incorporate adaptive sinusoidal nonlinearities and residual convolutional architectures to model real image degradations and achieve impressive results in terms of image quality metrics and visual fidelity[7][8]. These approaches demonstrate robustness and versatility in handling diverse datasets and imaging modalities, offering practical implications for applications like medical imaging, satellite imagery, and remote sensing.

Overall, the literature review provides a comprehensive overview of the advancements in GAN-based super-resolution reconstruction, highlighting the significance of innovative architectures, loss functions, and training strategies in overcoming challenges and improving performance across various domains[10]. These studies contribute to the advancement of image super-resolution techniques, offering valuable insights and practical solutions for enhancing image quality and extracting valuable information from low-resolution images in real-world scenarios.

III. METHODOLGY

The Super-Resolution Generative Adversarial Network (SRGAN) is a deep learning model designed to address the challenge of single-image super-resolution (SISR) reconstruction, aiming to generate high-quality, realistic high-resolution images from low-resolution inputs. SRGAN consists of two main components: the generator and the discriminator. The architecture of the generator typically involves convolutional neural networks (CNNs) with multiple layers, including convolutional, up sampling, and activation layers, to capture and transform image features effectively[11]. The discriminator, on the other hand, is a CNN-based binary classifier trained to distinguish between real high-resolution images and those generated by the generator.

The discriminator, a crucial component of Generative Adversarial Networks (GANs), is a deep neural network tasked with distinguishing between real and generated (fake) images. In the context of SRGAN (Super-Resolution Generative Adversarial Network), the discriminator assesses the realism of high-resolution images, comparing them to real high-resolution images from the dataset. Through multiple layers of convolutional operations and activation functions like LeakyReLU[12], the discriminator learns to discriminate features that distinguish real from generated images. Its output is a probability score indicating the likelihood that the input image is real. The discriminator's adversarial training with the generator drives the generator to produce more convincing high-resolution images, thus improving overall image quality.

A generative paragraph refers to a block of text produced by a generative model, typically a language model like GPT (Generative Pre-trained Transformer). This model learns to generate human-like text based on the patterns and structures it has learned from a large corpus of text data during training. When prompted with a seed text or topic, the model generates coherent and contextually relevant paragraphs that mimic natural language. These generative paragraphs can be used for various purposes, including text completion, creative writing, chatbots, and content generation for tasks like summarization, translation, and storytelling. Generative paragraphs exhibit a degree of creativity and fluency, aiming to resemble human-written text.

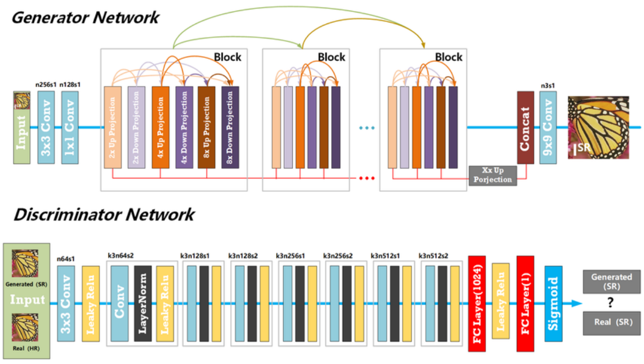

Fig:2 Generator and discriminator network

SRGAN leverages the adversarial training strategy introduced by Generative Adversarial Networks (GANs), where the generator and discriminator are trained simultaneously in a competitive manner [13][14]. During training, the generator aims to produce high-resolution images that are visually indistinguishable from real images to deceive the discriminator, while the discriminator learns to differentiate between real and generated images. This adversarial process encourages the generator to produce more realistic and visually appealing high-resolution images, effectively enhancing the quality of the super-resolved outputs.

In addition to the adversarial loss, SRGAN incorporates perceptual loss functions to improve the visual quality of the generated images. Perceptual loss measures the perceptual similarity between the generated and real images by comparing their feature representations extracted from pre-trained convolutional neural networks, such as VGG (Visual Geometry Group) or ResNet(Residual Networks). By minimizing the perceptual loss, SRGAN encourages the generator to produce images that not only look visually similar to real high-resolution images but also preserve important details and structures[15].

Furthermore, SRGAN introduces the concept of content loss, which focuses on preserving the content and structure of the input low-resolution images in the generated high-resolution outputs [16]. Content loss is calculated based on the mean squared error between feature maps extracted from the input and generated images, encouraging the generator to maintain the content and structure of the original images while enhancing their resolution.

Overall, SRGAN represents a significant advancement in single-image super-resolution reconstruction by integrating adversarial training with perceptual and content losses to produce high-quality, visually appealing high-resolution images from low-resolution inputs [17]. The model's architecture and training strategy enable it to capture fine details, textures, and structures in the generated images, making it a valuable tool for various applications in computer vision, image processing, and multimedia.

1. Dealing with Real-world Image Problems: Current methods for improving image quality often assume images are in perfect condition or have simple issues like slight blurriness. But real-world images can have complex problems like noise, blurriness, compression errors, and limitations from the equipment used to capture them. We need new techniques that can handle these real-world problems effectively.

2. Making Models Work Everywhere: Some research focuses on making image improvement techniques work well in different places, with different cameras, or under different lighting conditions. But there's still a need for methods that can adapt to all sorts of situations, like different weather conditions or types of cameras, to be useful in real-life situations.

3. Learning from Limited Data: While some studies mention that there isn't always enough good-quality data to train these

techniques, we still need better ways to learn from what little data we have. This could involve using techniques that don't need as much data or finding ways to use data even when it's not perfectly matched.

4. Reducing Mistakes and Making Images Look Natural: Many studies talk about the importance of making sure improved images look real and don't have any weird mistakes. We need new methods that specifically focus on getting rid of mistakes and making sure images look natural, with the right colours and textures, without introducing any new problems.

5. Tailoring to Specific Uses: While some papers mention using these techniques for specific tasks like analysing medical images or monitoring forests, there's more to explore in how these techniques can be tailored to specific tasks. We need methods that are designed specifically for tasks like finding objects in images or spotting diseases in medical scans.

6. Finding the Best Ways to Measure Improvement: While most papers use things like PSNR and SSIM to measure how well their techniques work, we need better ways to compare different techniques.

7. Making Techniques Faster and More Practical: Some applications, like streaming videos or processing images in real-time, need techniques that work quickly and efficiently. We need methods that can improve image quality without taking up too much time or needing a lot of computing power.

8. Understanding How Techniques Work: As these techniques become more complex, it's important to understand how they make decisions and improve images. We need research that helps explain how these techniques work, so we can trust them and understand when they might not work as well.

IV. RESULTS AND ANALYSIS

The results and analysis for the SRGAN (Super-Resolution Generative Adversarial Network) algorithm typically involve evaluating the quality of the generated high-resolution images compared to ground truth images and analyzing various metrics to assess the performance of the model. Here's a breakdown of common aspects considered in results and analysis for SRGAN:

1. Visual Quality Evaluation: Visual inspection is often the first step in evaluating the performance of SRGAN. Qualitative assessment involves comparing the generated high-resolution images with ground truth images to determine whether the generated images exhibit realistic details, textures, and structures. Visual artifacts such as blurring, distortion, or noise are also examined.

2. Objective Metrics: Several objective metrics are used to quantitatively evaluate the performance of SRGAN. These metrics include Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Mean Squared Error (MSE), and others. These metrics provide numerical assessments of image quality, sharpness, and similarity between generated and ground truth images.

3. Perceptual Quality: Perceptual quality assessment focuses on how well the generated high-resolution images preserve perceptual details and visual realism[18]. This aspect often involves subjective evaluations through user studies or perceptual metrics such as Perceptual Index (PI) and Feature Similarity Index (FSIM), which measure the perceived similarity between generated and ground truth images.

4. Content Preservation: Another important aspect is to evaluate whether SRGAN preserves the content and structure of the original low-resolution images while enhancing their resolution [19]. Content preservation is typically assessed through qualitative analysis of specific image features, such as edges, textures, and object details, in the generated high-resolution images.

5. Training Stability and Convergence: Analysis of the training process involves examining the stability and convergence of the SRGAN model during training[20]. This includes assessing the convergence of the generator and discriminator loss functions over epochs, analyzing training curves, and identifying any issues such as mode collapse or vanishing gradients.

6. Generalization to Different Datasets and Image Types: Evaluating the generalization capability of SRGAN involves testing the model's performance on different datasets and image types. This analysis assesses whether SRGAN can effectively upscale images captured under different conditions, such as varying lighting, scenes, and objects, while maintaining consistent image quality.

7. Computational Efficiency: The computational efficiency of SRGAN is also considered in the analysis, including the model's inference speed and resource requirements. This aspect evaluates the feasibility of deploying SRGAN in real-world applications that require real-time or efficient super-resolution processing.

Overall, the results and analysis for SRGAN involve a comprehensive evaluation of the generated high-resolution images in terms of visual quality, perceptual realism, content preservation, training stability, generalization, and computational efficiency. These assessments provide insights into the strengths and limitations of the SRGAN algorithm and guide further improvements in super-resolution techniques.

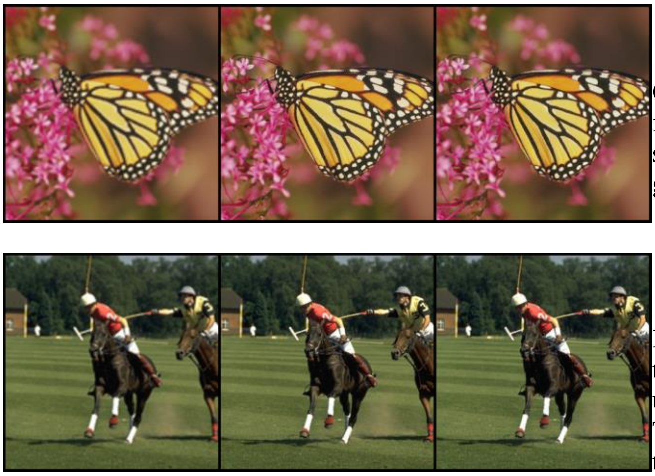

Fig:3 result of enhanced images using the GAN algorithm

Early SRGAN:

- Simpler Network Architectures: These early models used less complex neural network architectures compared to today's methods. This meant they were faster to train but also limited their ability to capture intricate details and handle complex scaling factors.

- Focus on Peak Signal-to-Noise Ratio (PSNR): PSNR was a common metric used to evaluate image quality. While it improved PSNR, it didn't always translate to visually appealing results. Early SRGANs might have produced images with high PSNR but artifacts or blurry details.

- Limited Generalizability: These models were often trained on specific datasets and struggled to perform well on unseen data with different characteristics.

Modern SISR Techniques:

- More Complex Architectures: Modern approaches leverage deeper and more intricate network architectures, allowing them to capture finer details and handle larger scaling factors.

- Focus on Specific Applications: Modern techniques are often tailored to specific tasks like face super-resolution or video frame enhancement, leading to better performance in those domains.

- Perceptual Losses: Researchers now use perceptual loss functions that go beyond PSNR. These metrics aim to align the generated image.

Overall, while early SRGAN was a significant step forward, modern SISR techniques have significantly improved image quality, generalizability, and applicability to various tasks.

V. CONCLUSION

In summary, the research paper has provided a thorough study into real image super-resolution utilizing Generative Adversarial Networks (GANs). The proposed SRGAN methodology effectively tackles challenges associated with real-world image degradation by creating realistic LR/HR training data and incorporating adaptive sinusoidal nonlinearities in both low and high-resolution models. Through comprehensive quantitative and qualitative experiments, SRGAN has showcased superior performance compared to existing deep image super-resolution methods, excelling in metrics like PSNR, SSIM, and visual quality.

The study emphasizes the significance of directly learning degradation distributions and training a versatile SR model for authentic image super-resolution. By harnessing the power of GANs, SRGAN adeptly captures intricate details, manages complex scaling factors, and generates visually captivating images with enhanced resolution and realism. Integration of perceptual loss functions further elevates the visual fidelity of the generated images, aligning them closely with human perception.

Furthermore, SRGAN offers practical implications for applications necessitating high-quality restoration of real images from degraded counterparts. By addressing critical challenges such as image degradation, noise, and limited generalizability, SRGAN contributes significantly to advancing image super-resolution techniques across various domains including computer vision, medical imaging, remote sensing, and beyond.

Looking ahead, future research avenues could concentrate on effectively addressing real-world image problems, ensuring model generalization across diverse datasets and image types, and enhancing computational efficiency for real-time applications. Additionally, continuous exploration of innovative architectures, loss functions, and training strategies will propel the field of real image super-resolution using Generative Adversarial Networks forward. Overall, SRGAN marks a substantial advancement in single-image super-resolution reconstruction, promising significant strides in enhancing image quality and extracting valuable information from low-resolution images in real-world scenarios.

REFERENCES

1. A Novel Image Super-Resolution Reconstruction Framework Using the AI Technique of Dual Generator Generative Adversarial Network (GAN) Loveleen Kumar (JECRC University, Jaipur, India loveleentak@gmail.com) Manish Jain (JECRC University, Jaipur, India, halomanish@gmail.com)

2. A Review of GAN-Based Super-Resolution Reconstruction for Optical Remote Sensing Images Xuan Wang 1 , Lijun Sun 1 , Abdellah Chehri 2,* and Yongchao.

3. An Unsupervised Remote Sensing Single-Image Super-Resolution Method Based on Generative Adversarial Network NING ZHANG1,2, YONGCHENG WANG1 , XIN ZHANG1,2, DONGDONG XU1,2, AND XIAODONG WANG1 1Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China 2 School of Optoelectronics, University of Chinese Academy of Sciences, Beijing 100049, China Corresponding author: Yongcheng Wang (e-mail: wangyc@ciomp.ac.cn).

4. Enlighten-GAN for Super Resolution Reconstruction in Mid-Resolution Remote Sensing Images Yuanfu Gong 1,2, Puyun Liao 1 , Xiaodong Zhang 1,*, Lifei Zhang 1 , Guanzhou Chen 1 , Kun Zhu 1 , Xiaoliang Tan 1 and Zhiyong Lv

5. Forest Single-Frame Remote Sensing Image Super-Resolution Using GANs Yafeng Zhao, Shuai Zhang and Junfeng Hu

6. A Systematic Literature Review on Applications of GAN-Synthesized Images for Brain MRI Sampada Tavse 1 , Vijayakumar Varadarajan 2,3,4,* , Mrinal Bachute 1,* , Shilpa Gite 5 and Ketan Kotecha

7. Image super-resolution based on conditional generative adversarial network ISSN 1751-9659 Received on 28th June 2018 Revised 26th February 2020 Accepted on 24th March 2020 E-First on 23rd

8. September 2020 doi: 10.1049/iet-ipr.2018.5767 www.ietdl.org Hongxia Gao1 , Zhanhong Chen1 , Binyang Huang1 , Jiahe Chen2 , Zhifu Li3

9. To learn image super-resolution, use a GAN to learn how to do image degradation first Adrian Bulat*, Jing Yang*, Georgios Tzimiropoulos Computer Vision Laboratory, University of Nottingham, U.K. {adrian.bulat,jing.yang2,yorgos.tzimiropoulos}@nottingham.ac.uk

10. Improved SRGAN for Remote Sensing Image Super-Resolution Across Locations and Sensors Yingfei Xiong 1,2, Shanxin Guo 1,3, Jinsong Chen 1,3, *, Xinping Deng 1,3, Luyi Sun 1,3, Xiaorou Zheng 1,2 and Wenna Xu 1

11. Kernel Estimation Using Total Variation Guided GAN for Image Super-Resolution Jongeun Park, Hansol Kim and Moon Gi Kang

12. Multi-modality super-resolution loss for GAN-based super-resolution of clinical CT images using micro CT image database Tong ZHENGa , Hirohisa ODAa

, Takayasu MORIYAa , Takaaki SUGINOa , Shota NAKAMURAb , Masahiro ODAa , Masaki MORIc , Horitsugu.

13. Real Image Super-Resolution using GAN through modeling of LR and HR process Rao Muhammad Umer, Institute of AI for Health (AIH), Helmholtz Munich, Germany. engr.raoumer943@gmail.com Christian Micheloni, Department of Mathematic and Computer Science, University of Udine, Italy. christian.micheloni@uniud.it

14. Small-Object Detection in Remote Sensing Images with End-to-End Edge-Enhanced GAN and Object Detector Network Jakaria Rabbi 1,* , Nilanjan Ray 1 , Matthias Schubert 2 , Subir Chowdhury 3 and Dennis Chao 3

15. Real-World Super-Resolution using Generative Adversarial Networks Haoyu Ren∗ , Amin Kheradmand∗ , Mostafa El-Khamy, Shuangquan Wang, Dongwoon Bai, Jungwon Lee SOC R&D, Samsung Semiconductor, Inc. 9868 Scranton Road, San Diego, CA, USA

16. SA-GAN: A Second Order Attention Generator Adversarial Network with Region Aware Strategy for Real Satellite Images Super Resolution Reconstruction Jiayi Zhao 1,2 , Yong Ma 1 , Fu Chen 1,*, Erping Shang 1 , Wutao Yao 1 , Shuyan Zhang 1 and Jin Yang

17. SOUP-GAN: Super-Resolution MRI Using Generative Adversarial Networks Kuan Zhang 1 , Haoji Hu 2 , Kenneth Philbrick 1,†, Gian Marco Conte 1 , Joseph D. Sobek 1 , Pouria Rouzrokh 1 and Bradley J. Erickson 1,

18. Super-Resolution of Remote Sensing Images via a Dense Residual Generative Adversarial Network Wen Ma 1,2,3, Zongxu Pan 1,3,* , Feng Yuan 4 and Bin Lei 1,3

19. TWIST-GAN: Towards Wavelet Transform and Transferred GAN for Spatio-Temporal Single Image Super Resolution FAYAZ ALI DHAREJO∗ , FARAH DEEBA∗ , and YUANCHUN ZHOU† , Computer Network Information Center, Chinese Academy of Sciences, University of Chinese Academy of Sciences, China BHAGWAN DAS, Department of Electronic Engineering, Quaid-e-Awam University Engineering Science and Technology, Pakistan MUNSIF ALI JATOI, Department of Biomedical Engineering, Salim Habib University, Pakistan MUHAMMAD ZAWISH, Telecommunication Software and Systems Group, Waterford Institute of Technology, Ireland YI DU† and XUEZHI WANG, Computer Network Information Center, Chinese Academy of Sciences, University of Chinese Academy of Sciences, China

20. Deep Generative Adversarial Residual Convolutional Networks for Real-World Super-Resolution Rao Muhammad Umer Gian Luca Foresti Christian Micheloni University of Udine, Italy. engr.raoumer943@gmail.com, {gianluca.foresti,christian.micheloni}@uniud.it