Progressive Upscaling Super Image Resolution

MINOR PROJECT REPORT

Submitted in partial fulfilment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING

by

Aman Kumar Karn	Ankit Khandelwal	Gaurav Singh
En. No. 00596302720	En. No. 00796302720	En. No. 01596302720

Guided by

Ms. Swati Malik
Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
MAHARAJA SURAJMAL INSTITUTE OF TECHNOLOGY
(AFFILIATED TO GURU GOBIND SINGH INDRAPRASTHA UNIVERSITY, DELHI) DELHI – 110058

Dec 2023

CANDIDATE'S DECLARATION

It is hereby certified that the work which is being presented in the B. Tech Minor Project Report entitled **"Progressive Upscaling Super Image Resolution"** in partial fulfilment of the requirements for the award of the degree of **Bachelor of Technology** and submitted in the **Department of Computer Science & Engineering** of **MAHARAJA SURAJMAL INSTITUTE OF TECHNOLOGY, New Delhi (Affiliated to Guru Gobind Singh Indraprastha University, Delhi)** is an authentic record of our own work carried out during a period from **Aug2023 to Dec 2023** under the guidance of **Ms. Swati Malik, Assistant Professor.**

The matter presented in the B. Tech. Major Project Report has not been submitted by me for the award of any other degree of this or any other Institute.

Aman Kr. Karan	**Ankit Khandelwal**	**Gaurav Singh**
00596302720	**00796302720**	**01596302720**

CERTIFICATE

This is to certify that the above statement made by the candidate is correct to the best of my knowledge. They are permitted to appear in the External Minor Project Examination

(Ms. Swati Malik)
Assistant Professor,CSE
Mentor

(Dr. Gunjan Beniwal)
Assistant Professor,CSE
Project Coordinator

Dr. Nishtha Jatana
HOD, CSE

ACKNOWLEDGEMENT

We express our deep gratitude to **Ms. Swati Malik**, Assistant Professor, Department of Computer Science &Engineering for his valuable guidance and suggestion throughout Our project work. We are thankful to **Dr. Gunjan Beniwal,** Project Coordinator, for their valuable guidance.

We would like to extend my sincere thanks to **Head of the Department, Dr. Nishtha Jatana**for his time to time suggestions to complete my project work.

Aman Kr. Karan	**Ankit Khandelwal**	**Gaurav Singh**
00596302720	**00796302720**	**01596302720**

Table of Contents

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

SRCNN: Super-Resolution Convolutional Neural Network

GAN: Generative Adversarial Network

SRGAN: Super-Resolution Generative Adversarial Network

ESRGAN: Enhanced Super-Resolution Generative Adversarial Network

ProGAN: Progressive Growing of GANs

PSNR: Peak Signal-to-Noise Ratio

ABSTRACT

Recently, it has been demonstrated that deep neural networks can significantly improve the performance of single image super-resolution (SISR). Numerous studies have concentrated on raising the quantitative quality of super-resolved (SR) images. However, these methods that target PSNR maximization usually produce blurred images at large upscaling factor. The introduction of generative adversarial networks (GANs) can mitigate this issue and show impressive results with synthetic high-frequency textures. Nevertheless, these GAN-based approaches always have a tendency to add fake textures and even artifacts to make the SR image of visually higher-resolution.

This report presents a comparison of various super image resolution algorithms. The algorithms are evaluated based on their performance in enhancing the resolution of images. The algorithms compared include the algorithm that contain SRCNN at 1st stage (2x stage), SRGAN at 2nd Stage and ESRGAN at final stage,and algorithm with ProGAN at each stage. The comparison is based on value of PSNR, value of mean square error and quality of image. The results show that while SRCNN served as a foundational step, SRGAN and ESRGAN exhibited notable improvements, with ESRGAN capturing finer details. ProGAN, however, showcased a more consistent and coherent refinement.The paper concludes with a discussion of the strengths and weaknesses of each algorithm and recommendations for future research

CHAPTER 1

INTRODUCTION

This paper presents a comparison of various super image resolution algorithms. The algorithms are evaluated based on their performance in enhancing the resolution of images. The algorithms compared include the algorithm that contain SRCNN at 1^{st} stage (2x stage), SRGAN at 2^{nd} Stage and ESRGAN at final stage . The comparison is based on value of PSNR, value of mean square error and quality of image. The results show thatwhile SRCNN served as a foundational step, SRGAN and ESRGAN exhibited notable improvements, with ESRGAN capturing finer details. ProGAN, however, showcased a more consistent and coherent refinement.The paper concludes with a discussion of the strengths and weaknesses of each algorithm and recommendations for future research.

In recent years, several algorithms have been proposed for super image resolution. These algorithms can be broadly classified into two categories: traditional methods and deep learning methods. Traditional methods include bicubic interpolation, Lanczos resampling, and nearest neighbor interpolation. Deep learning methods include generative adversarial networks (GANs), super-resolution generative adversarial networks (SRGANs), and enhanced super-resolution generative adversarial networks (ESRGANs)[1].

Image super-resolution is a process to recover an image of high-resolution (HR) from a low-resolution (LR) image [2]. In simple terms, it can also be referred to as image interpolation, scaling, upsampling, zooming, or enlarging [2]. The purpose of image super-resolution is to obtain a high pixel density and refined details from low-resolution (LR) image(s) that cannot be seen with the naked eye. Image super-resolution was very useful in applications that required recognition or detection purposes.

In the past, many image processing-based techniques have been used in image super-resolution before deep learning-based methods were started. Li et al. [5] and Nasrollahi et al. [6] classified image super-resolution methods into three different groups, namely, interpolation-based method, reconstruction-based method, and learning-based method. Several interpolation-based methods such as linear [7], bilinear [8], or bicubic [9,10] interpolations can be found in image super-resolution applications. These methods are simple, but the high-frequency details of the image are not restored and therefore more sophisticated insight may be required to recover the image [5,6].

The reconstruction-based method includes the sharpening of edge details [11], regularization [12,13], and deconvolution [14] techniques. Researchers also used these techniques for image reconstruction. The learning-based method has an advantage over the interpolation-based method by restoring the missing high-frequency details through a relationship established between the LR image and the HR image.

Conventional upscaling methods, in their haste to increase resolution, often stumble upon pitfalls that result in a noticeable degradation of image quality. These pitfalls manifest as blurriness, pixelation, and an overall loss of visual fidelity.

The core concept underpinning this project is the incremental enhancement of resolution. Rather than a single, abrupt transformation, progressive upscaling divides the process into smaller, more manageable increments. In each of these steps, the algorithm meticulously analyzes and enhances a portion of the image, with a relentless focus on preserving the fine details that might otherwise be lost.

Preserving the original characteristics of the image is a nuanced challenge. The algorithm must delicately balance the enhancement of resolution while safeguarding color fidelity, edge sharpness, and the overall aesthetic. Each type of content, whether it's a photograph, illustration brings its own set of characteristics and challenges that the algorithm must adapt to.

We hope that this study will provide valuable insights into the performance of various super image resolution algorithms and help researchers and practitioners in selecting the most appropriate algorithm for their specific needs.

CHAPTER2

LITERATURE REVIEW

The field of image super-resolution has witnessed significant advancements, driven by the continuous pursuit of achieving higher visual fidelity and enhanced details in reconstructed images. In this literature review, we delve into key datasets used for training and evaluation, followed by an exploration of influential methodologies, including the fully progressive approach and the perception-oriented network.

A Fully Progressive Approach to Single-Image Super-Resolution [3]

This paper proposes a novel single-image super-resolution (SISR) technique that employs a progressive approach to enhance image details at various scales. The proposed method, dubbed ProSR, consists of multiple sub-networks that progressively refine the low-resolution (LR) input image to produce a high-resolution (HR) output image. Each sub-network focuses on specific aspects of the image, such as restoring high-frequency details and preserving structural information. This progressive approach enables ProSR to achieve impressive resolution improvements, outperforming state-of-the-art SISR methods in terms of both quantitative metrics and visual quality.

Progressive Perception-Oriented Network for Single Image Super-Resolution[4]

This research introduces a novel SISR architecture, called PPON, that focuses on perceptual quality. PPON comprises multiple stages, each geared towards extracting and enhancing different levels of image features. The first stage focuses on minimizing pixel-wise error, while the second stage utilizes the features extracted by the previous stage to refine structural information. The final stage employs fine structure features distilled by the second phase to produce more realistic results. This progressive architecture, coupled with a perceptual loss function, enables PPON to achieve superior visual quality compared to existing SISR methods.

Enhanced Deep Residual Networks for Single Image Super-Resolution (Wang et al., 2018)[15]

This paper proposes an enhanced deep residual network (EDSR) architecture for SISR. EDSR introduces residual connections, densely connected convolutional layers, and a recursive residual block design, enabling it to achieve significant improvements in both PSNR and SSIM metrics.

Multi-Scale Residual Network for Single Image Super-Resolution (Zhang et al., 2018)[16]

This paper presents a multi-scale residual network (MSR) architecture for SISR. MSR employs a multi-scale approach, using multiple residual networks at different scales to capture image details at varying frequencies. This multi-scale design enables MSR to achieve superior results, particularly for high-resolution upscaling tasks.

Attention-Based Deep Fusion Network for Single Image Super-Resolution (Zhang et al., 2020)[17]

This paper proposes an attention-based deep fusion network (AFNet) for SISR. AFNet utilizes an attention mechanism to selectively fuse features from different scales, allowing it to focus on more relevant features and enhance the overall quality of the super-resolved image. AFNet demonstrates superior performance compared to previous SISR methods in terms of both PSNR and SSIM metrics.

CHAPTER 3

RESEARCH METHODOLOGY

3.1 Algorithm Used

3.1.1 SRCNN (Super resolution convolutional neural network)

The Super-Resolution Convolutional Neural Network (SRCNN) is a groundbreaking deep learning architecture designed to address the challenge of image super-resolution. Image super-resolution involves enhancing the resolution of an image, generating a higher-resolution version from a lower-resolution input. SRCNN, introduced by Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang in 2014, has been instrumental in advancing the field of computer vision.

SRCNN leverages the power of convolutional neural networks (CNNs), a class of deep learning models proven effective in various image processing tasks. The network is trained end-to-end to learn the mapping between low-resolution and high-resolution image patches. The architecture consists of three main layers: the patch extraction layer, non-linear mapping layer, and reconstruction layer.

In the patch extraction layer, the SRCNN takes as input low-resolution image patches. These patches are extracted from the input image using a sliding window approach. The non-linear mapping layer, which follows, is the core of the network. It employs multiple convolutional and rectified linear unit (ReLU) layers to learn complex mappings between low and high-resolution patches. The non-linear mapping layer captures intricate features and relationships within the image data, enabling the network to understand the underlying structures.

The final reconstruction layer then transforms the high-dimensional feature maps into the desired high-resolution output. The entire network is trained using pairs of low and high-resolution images to optimize the model parameters through a process known as supervised learning.

SRCNN has demonstrated superior performance compared to traditional interpolation methods in terms of reconstructing high-resolution details and preserving image quality. Its success has inspired further research in the field of deep learning-based image super-resolution, leading to the development of more sophisticated architectures and techniques. SRCNN's impact extends beyond image super-resolution, influencing advancements in various computer vision applications, including image recognition, object detection, and image synthesis.

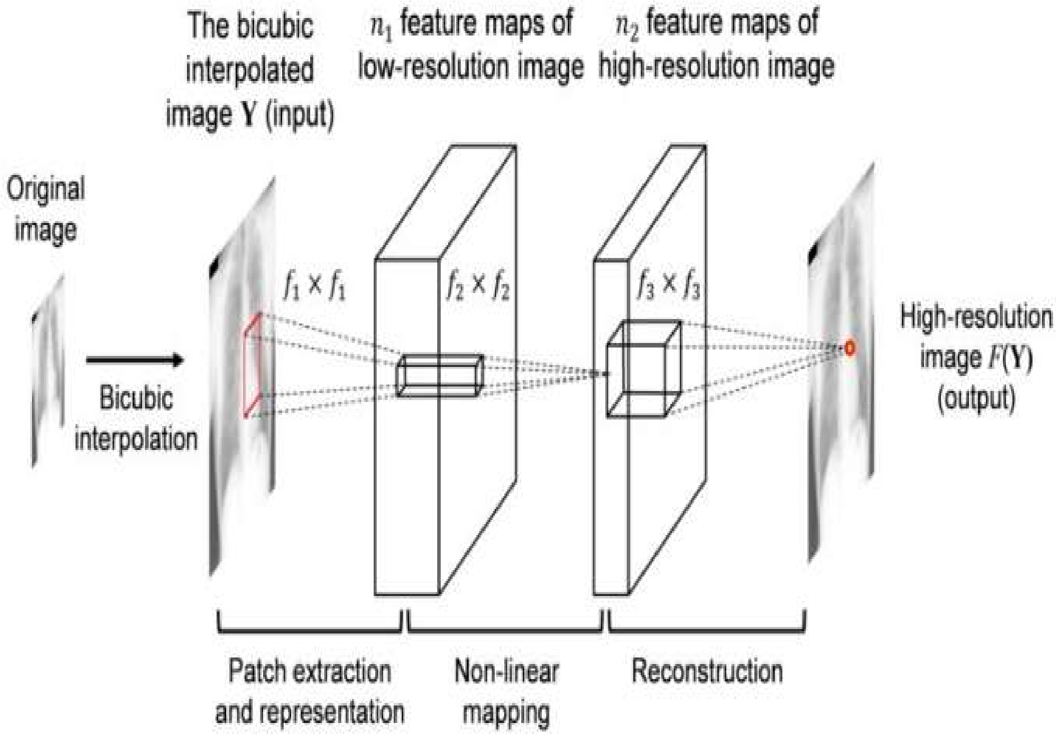

Fig 3.1 Working of SRCNN

3.1.2 GAN (Generative Adversial Network)

Generative Adversarial Networks, or GANs, represent a revolutionary paradigm in machine learning, introduced by Ian Goodfellow and his colleagues in 2014. A GAN comprises two neural networks, a generator, and a discriminator, engaged in a dynamic adversarial training process. This innovative architecture has demonstrated remarkable capabilities in generating realistic and high-quality synthetic data, making it a cornerstone in the field of generative modeling.

The generator's primary role in a GAN is to create synthetic data instances that resemble real data. It starts with random noise and progressively refines its output through the layers of the network, attempting to generate data that is indistinguishable from authentic samples. The discriminator, on the other hand, evaluates these generated samples along with real data, discerning between the two. As training progresses, the generator refines its ability to create more realistic data, while the discriminator improves its discriminatory skills.

The adversarial nature of the GAN training process is characterized by a continuous feedback loop. The generator strives to produce more convincing data to deceive the discriminator, and conversely, the discriminator aims to become more adept at distinguishing real from generated data. This competition between the two networks results in a dynamic equilibrium, where the generator continuously refines its outputs, approaching a point where it becomes challenging for the discriminator to differentiate between real and synthetic data.

GANs find applications in a diverse range of fields, such as image and video synthesis, style transfer, and even text-to-image generation. The ability of GANs to generate data that closely mirrors real-world distributions has spurred innovation in the creation of deepfake images and videos, raising ethical concerns about the potential misuse of this technology.

Despite their success, GANs present challenges such as mode collapse, training instability, and the potential generation of biased outputs. Ongoing research is focused on addressing these issues to unlock the full potential of GANs and further advance the capabilities of generative models in capturing and reproducing complex data distributions.

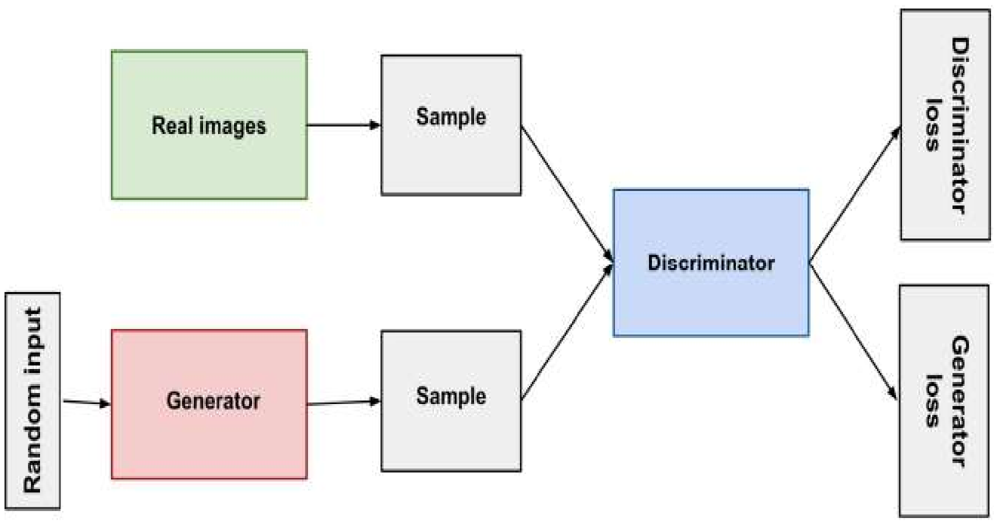

Fig 3.2 Working of GAN

3.1.3 SRGAN (Super-Resolution Generative Adversial Network)

The Super-Resolution Generative Adversarial Network (SRGAN) is an advanced deep learning model designed for single-image super-resolution. Introduced by Christian Ledig and his team in 2017, SRGAN represents a significant leap in generating high-quality, realistic, and perceptually enhanced high-resolution images from lower-resolution inputs.

SRGAN leverages the power of Generative Adversarial Networks (GANs) to produce more visually pleasing super-resolved images compared to traditional methods. The architecture consists of a generator and a discriminator, each playing a distinctive role in the training process. The generator, often referred to as the super-resolution network, is responsible for transforming low-resolution images into high-resolution counterparts. Meanwhile, the discriminator assesses the generated images, distinguishing between super-resolved images and authentic high-resolution images. This adversarial interplay refines the generator's ability to produce more authentic and visually appealing results.

What sets SRGAN apart is its use of perceptual loss, incorporating features from pre-trained deep convolutional neural networks, such as VGG-19. By introducing perceptual loss into the objective function, SRGAN focuses not only on pixel-wise accuracy but also on capturing high-level semantic features. This results in super-resolved images that are not only sharper but also maintain more realistic textures and structures.

The training of SRGAN involves a delicate balance between the adversarial loss, ensuring that the generated images are indistinguishable from authentic high-resolution images, and the perceptual loss, emphasizing the preservation of perceptual quality. This dual-approach training significantly improves the visual quality of super-resolved images compared to earlier methods that primarily relied on mean squared error.

SRGAN has found applications in various domains, including medical imaging, satellite imagery, and enhancing the visual quality of low-resolution videos. Despite its success, ongoing research continues to refine and extend SRGAN, addressing challenges like fine-tuning, generalization to diverse datasets, and mitigating artifacts in the super-resolved images to further enhance its practical utility.

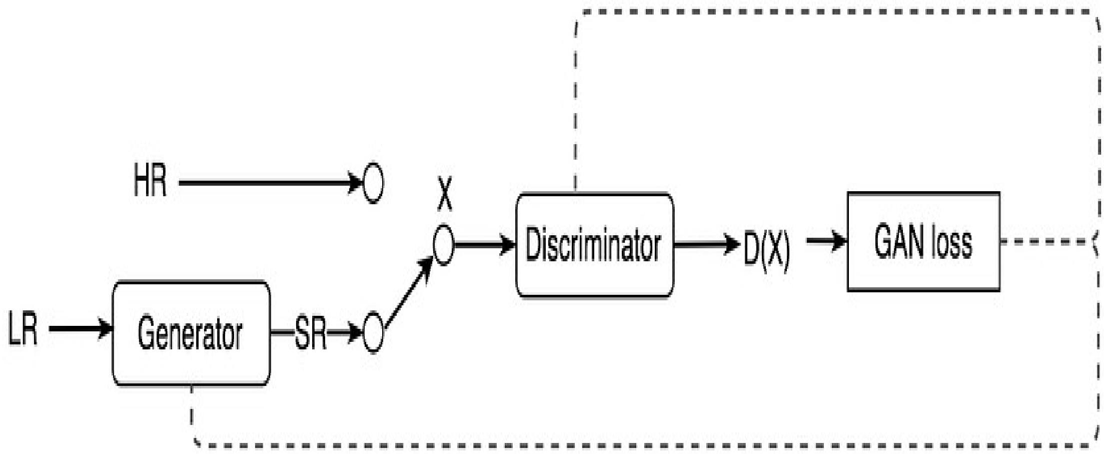

Fig 3.3 SRGAN Architecture

3.1.4 ESRGAN (Generative Adversarial Network)

Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) represents a significant advancement in the realm of image super-resolution, building upon the foundations laid by its predecessor, SRGAN. Developed by Xintao Wang and his collaborators in 2018, ESRGAN introduces novel techniques to further enhance the quality and perceptual fidelity of super-resolved images.

At its core, ESRGAN employs a Generative Adversarial Network (GAN) architecture, featuring a generator and a discriminator. The generator is responsible for transforming low-resolution images into high-resolution counterparts, while the discriminator evaluates the generated images, discerning between authentic high-resolution images and those produced by the generator. This adversarial training process encourages the generator to produce more realistic and visually pleasing super-resolved images.

What sets ESRGAN apart is its use of a novel architecture and advanced training strategies. ESRGAN replaces the original SRGAN's generator with a more sophisticated network architecture based on Residual-in-Residual Dense Blocks

(RRDB). These blocks enhance the model's ability to capture intricate details and features in the super-resolution process, leading to sharper and more realistic results.

ESRGAN also introduces the concept of perceptual loss, leveraging a pre-trained deep neural network to measure the perceptual similarity between the generated and target images. This perceptual loss, often derived from features extracted by networks like VGG-19, helps guide the generator in producing images that not only have high pixel-wise accuracy but also closely resemble the perceptual qualities of authentic high-resolution images.

The impact of ESRGAN extends to various applications, including image upscaling, medical imaging, and improving the visual quality of low-resolution videos. Researchers and practitioners continue to explore ways to refine ESRGAN and address challenges, such as mitigating artifacts, enhancing generalization across diverse datasets, and improving computational efficiency, to ensure its continued effectiveness and applicability in real-world scenarios.

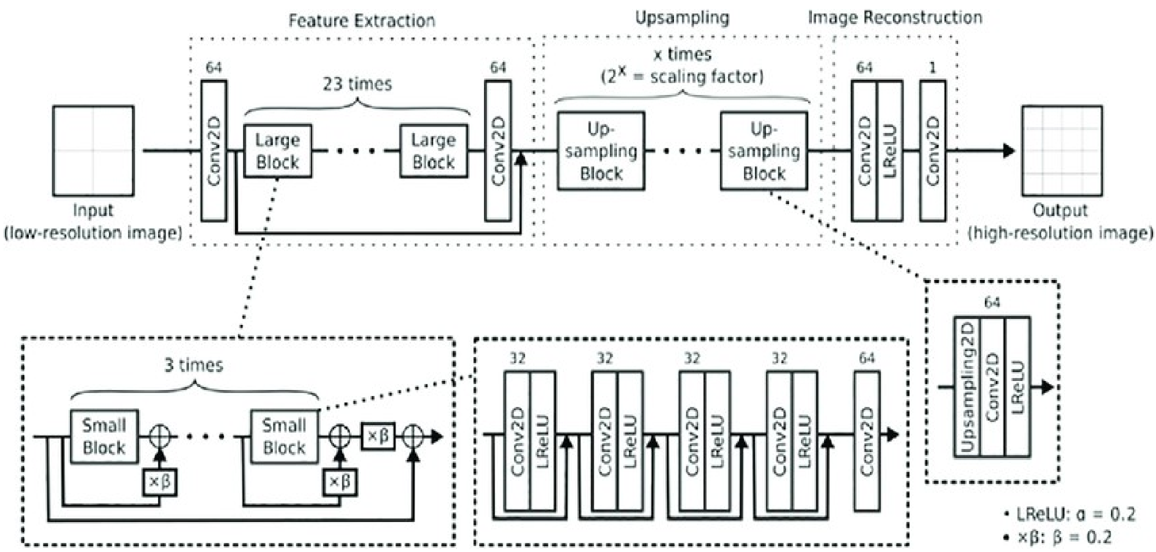

Fig 3.4 ESRGAN Architecture

3.1.5 ProGAN (Generative Adversarial Network)

Progressive Growing of GANs (ProGAN) is an innovative deep learning architecture introduced by Tero Karras and his colleagues in 2017. ProGAN addresses challenges in training Generative Adversarial Networks (GANs) by gradually increasing the resolution during the training process, enabling the generation of high-quality, high-resolution images.

The key feature of ProGAN is its progressive growth strategy. Unlike traditional GANs that train on a fixed resolution from the beginning, ProGAN starts with a low resolution and incrementally adds layers to both the generator and discriminator as training progresses. This incremental growth allows the model to learn coarse details before refining the finer details of the image. The approach not only enhances the stability of training but also facilitates the generation of high-resolution images with more realistic details.

ProGAN introduces a unique training scheme where lower-resolution images are initially generated, and then, as the model matures, additional layers are progressively added to increase the output resolution. This gradual growth mitigates issues such as mode collapse and training instability, which are common challenges in training GANs.

The generator in ProGAN employs a block-wise structure where each block corresponds to a specific resolution. As the training advances, new blocks are added, allowing the generator to learn increasingly detailed features. The discriminator, in turn, also undergoes a similar incremental growth, becoming more adept at distinguishing real images from generated ones at higher resolutions.

One of the notable advantages of ProGAN is its ability to generate high-fidelity images across various resolutions, making it particularly suitable for tasks like image synthesis and style transfer. ProGAN has been influential in the development of subsequent GAN architectures and has set the stage for further innovations in the field of generative modeling.

The success of ProGAN highlights the importance of progressive training strategies in mitigating the challenges associated with GANs, demonstrating how a carefully designed growth mechanism can lead to more stable and effective training for high-resolution image generation.

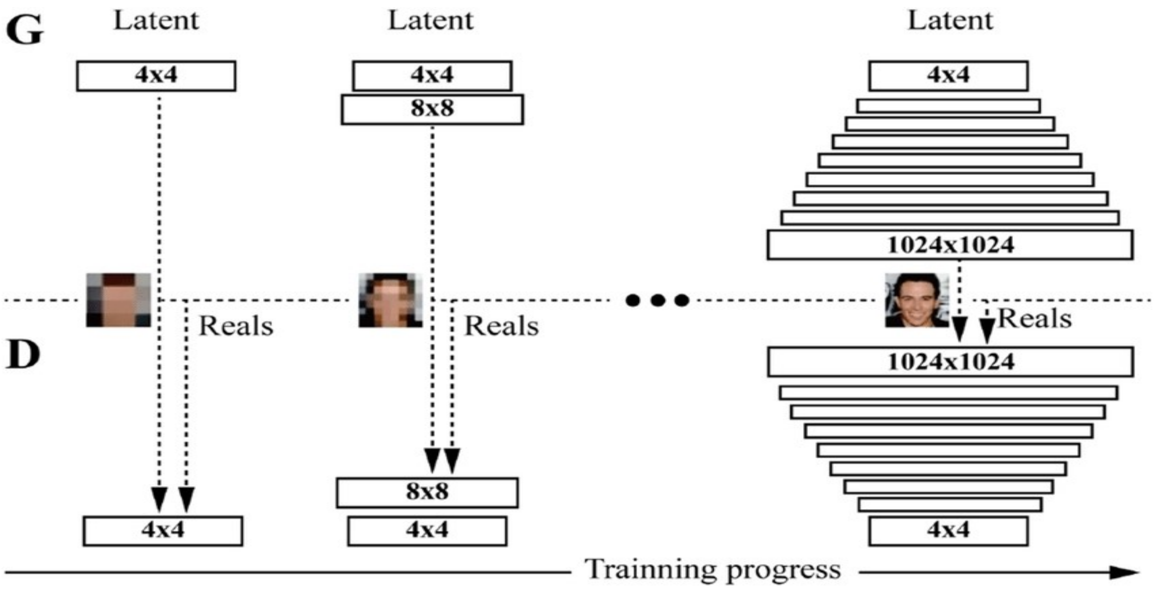

Fig 3.5 Working of ProGAN

Table 2.1 Comparison SRCNN, GAN, SRGAN, ESRGAN, and ProGAN

Feature	SRCNN	GAN	SRGAN	ESRGAN	ProGAN
Type	Convolutional Neural Network (CNN)	Generative Adversarial Network (GAN)	GAN-based Super-Resolution (SR)	GAN-based SR	Progressive Growing of GANs (PGGAN)
Architecture	Three-layer CNN	Generator and discriminator	Generator and discriminator	Generator and discriminator	Multi-scale discriminator
Loss function	Mean Squared Error (MSE)	Binary cross-entropy	MSE, perceptual loss	MSE, perceptual loss	Wasserstein distance, perceptual loss
Performance	Good for low-resolution images	Excellent for high-resolution images	High-quality SR	State-of-the-art SR	State-of-the-art image generation
Limitations	Prone to over-smoothing	Can produce artifacts	Can be unstable during training	Can be computationally expensive	Can be sensitive to hyperparameters

Applications	Image enhancement, videoupscaling	Image generation, image editing	Image super-resolution	Image super-resolution, image generation	Image generation, image editing

3.2 DATASETS

DIV2K Dataset: The DIV2K dataset emerges as a pivotal resource in the realm of single-image super-resolution. Comprising 800 high-definition images for training, each accompanied by downscaled versions at 2x, 3x, and 4x resolutions, it provides a robust foundation for evaluating the effectiveness of upscaling techniques. Additionally, the inclusion of validation and test data, with low-res versions for feedback and final evaluation, respectively, ensures a comprehensive assessment of model performance.[18]

Flickr30k Dataset: Focusing on image description benchmarking, the Flickr30k dataset and its Entities subset contribute significantly to the exploration of textual entity localization. With 244k coreference chains and 276k annotated bounding boxes, it sets a new benchmark in this domain. The baseline, featuring image-text embeddings, object detectors, color classification, and a bias towards larger objects, establishes a strong foundation for evaluating complex models, particularly in the context of image-sentence retrieval.[19]

3.3 SOFTWARES

Google Colab

Google Colab, short for Colaboratory, is a free, cloud-based platform provided by Google that allows you to write and execute Python code in a web-based interactive environment. It is particularly popular among researchers, students, and data scientists for its ease of use and the ability to run code on Google's powerful cloud servers, often with free access to GPUs (Graphics Processing Units) and TPUs (Tensor Processing Units).

Key features of Google Colab include:

1. Free Access to GPUs and TPUs: Colab provides free access to GPU and TPU resources, which can significantly speed up the training of machine learning models.

2. Jupyter Notebook Integration: Colab is built on top of Jupyter Notebooks, which allows you to create and share documents that contain live code, equations, visualizations, and narrative text.

3. Cloud Storage Integration: Colab integrates with Google Drive, allowing you to save and share your Colab notebooks directly on Google Drive.

4. Pre-installed Libraries: Colab comes pre-installed with many popular Python libraries and frameworks, making it easy to get started with tasks like data analysis, machine learning, and deep learning.

5. Collaboration Features: You can share your Colab notebooks just like you would with Google Docs or Sheets. This makes it easy to collaborate with others in real-time.

Google Colab notebooks are similar to Jupyter Notebooks, allowing you to run code in cells and see the output interactively. You can also add text cells to provide explanations, instructions, or comments.

Remember that the resources provided by Google Colab are not unlimited, and the availability of GPU and TPU resources is subject to demand. If you're working on resource-intensive tasks, be aware of the limitations and consider alternatives if necessary.

Jupyter Notebook

Jupyter Notebooks are a popular open-source tool for interactive computing. They allow you to create and share documents that contain live code, equations, visualizations, and narrative text. Jupyter supports multiple programming languages, but it is most commonly used with Python.

Key features and concepts related to Jupyter Notebooks:

1. Cells: The basic unit of content in a Jupyter Notebook is the cell. There are two main types of cells: code cells and markdown cells.

- Code Cells: These are used to write and execute code. You can run a code cell by clicking the "Run" button, pressing Shift+Enter, or using other keyboard shortcuts.

- Markdown Cells: These contain formatted text written in Markdown, a lightweight markup language. You can use markdown cells to add explanations, documentation, and other text to your notebook.

2. Interactive Output: When you run a code cell, the output is displayed directly below the cell. This allows you to see the results of your code immediately, which is especially useful for data analysis and exploration.

3. Kernel: A Jupyter Notebook runs a kernel, which is a computational engine that executes the code in the notebook. The kernel can be restarted, cleared, or interrupted, and it keeps track of the variables in memory.

4. Extensions: Jupyter Notebooks support extensions that can add functionality, themes, and shortcuts. You can customize your Jupyter environment to suit your preferences.

5. Exporting and Sharing: Jupyter Notebooks can be exported to various formats, including HTML, PDF, and slideshows. You can also share your notebooks with others by sending them the `.ipynb` file or by using platforms like GitHub.

6. Integration with Data Science Libraries: Jupyter Notebooks are widely used in data science and machine learning. They integrate well with popular libraries such as NumPy, Pandas, Matplotlib, and Scikit-learn.

Jupyter Notebooks provide an interactive and versatile environment for data analysis, research, and education. They are widely used in the data science and machine learning communities.

3.4 LIBRARIES

Tensorflow[26]

TensorFlow, an open-source machine learning library from the Google Brain team, has emerged as a cornerstone in the field, offering a robust set of tools and libraries for diverse machine learning applications. Whether tackling neural network models, natural language processing, or computer vision, TensorFlow provides a flexible architecture with both high-level APIs, like Keras, and low-level APIs, catering to the needs of beginners and experienced developers alike. TensorFlow 2.x, a significant evolution of the framework, has enhanced the user experience with eager execution enabled by default, simplifying the model-building process. Its seamless integration with Keras and support for a multitude of platforms, including CPUs, GPUs, TPUs, and mobile devices, underline TensorFlow's versatility, making it suitable for various research and production scenarios. With TensorBoard for visualization and extensive community support, TensorFlow stands as a comprehensive ecosystem empowering developers to create, train, and deploy machine learning models efficiently.

TensorFlow's influence extends beyond its technical capabilities; it fosters a vibrant community of researchers and developers. The official documentation and tutorials serve as valuable resources for both beginners and seasoned practitioners, while community forums provide platforms for knowledge sharing and issue resolution. TensorFlow Hub, a repository of pre-trained models, further facilitates collaborative learning by enabling the reuse of model components for diverse tasks through transfer learning. This collaborative ethos, combined with the framework's technical prowess, solidifies TensorFlow's role as a pivotal force in the ever-evolving landscape of machine learning and artificial intelligence.

Cv2[27]

`cv2` is the Python interface for OpenCV, a comprehensive open-source library designed for computer vision and machine learning applications. Leveraging `cv2`, developers can seamlessly integrate OpenCV's diverse functionalities into Python scripts, making it a popular choice for tasks like image processing, object detection, and video analysis. Whether reading and manipulating images, applying complex algorithms, or interfacing with cameras, the `cv2` module serves as a bridge between Python and the extensive capabilities of OpenCV, facilitating the development of applications ranging from computer vision research to real-world projects.

One of the strengths of `cv2` lies in its efficiency and versatility. From basic image loading and color space conversions to advanced features like contour detection and face recognition, the library's wide array of functions empowers developers to address various computer vision challenges with ease. As a widely adopted tool in the fields of artificial intelligence and robotics, `cv2` continues to play a pivotal role in advancing applications that rely on visual data processing.

Numpy[28]

NumPy, short for Numerical Python, is a foundational library for numerical computing in Python. Its core functionality revolves around the efficient handling of arrays and matrices, providing a versatile and high-performance platform for mathematical operations. NumPy introduces the `ndarray` (n-dimensional array) data structure, which not only enables efficient storage and manipulation of large datasets but also facilitates vectorized operations, eliminating the need for explicit loops. This feature makes NumPy an essential tool for scientific computing, machine learning, and data analysis, as it significantly enhances computational efficiency and code readability.

In addition to its array-centric capabilities, NumPy offers a rich collection of mathematical functions for operations such as linear algebra, Fourier analysis, and statistical computations. Its seamless integration with other scientific computing libraries in the Python ecosystem, such as SciPy and Matplotlib, enhances its utility in a broader context. Whether working with large datasets in data science applications or implementing complex mathematical algorithms in scientific research, the versatility and performance of NumPy have established it as a fundamental building block for a wide range of numerical computing tasks in the Python programming language.

Pytorch[29]

PyTorch, commonly referred to as "torch," is a widely-used open-source machine learning library developed by Facebook. Renowned for its dynamic computational graph, PyTorch provides a flexible and intuitive platform for building and training deep learning models. One of its key strengths lies in its dynamic computation graph, allowing for more natural expression of complex models and dynamic changes during training. This feature contrasts with static computation graphs employed by some other deep learning frameworks, enabling easier debugging and experimentation. PyTorch's dynamic nature, coupled with its seamless integration with Python, has contributed to its popularity among researchers and practitioners in the machine learning community.

PyTorch offers a comprehensive ecosystem that includes tools and modules for various aspects of deep learning, such as neural network construction, optimization, and deployment. The library supports dynamic neural networks through its `torch.nn` module, and the `torch.autograd` module facilitates automatic differentiation, a crucial component for training deep learning models. Additionally, PyTorch has gained traction in the research community due to its backing by Facebook AI Research (FAIR), fostering innovation and advancements in the field. With its user-friendly design, extensive documentation, and active community, PyTorch continues to be a preferred choice for both beginners and experienced practitioners in the rapidly evolving landscape of deep learning.

Glob[30]

The `glob` module in Python is a versatile utility within the standard library that facilitates the searching and retrieval of files based on specified patterns. The `glob.glob()` function, in particular, enables users to obtain a list of file paths that match a given pattern, employing wildcard characters such as '*' to represent any sequence of characters in filenames. This functionality is especially useful for tasks involving batch processing, data loading, or any scenario where a collection of files needs to be accessed programmatically. The module's straightforward interface and compatibility with various pattern-matching scenarios make it a valuable tool for file manipulation in diverse Python applications.

An additional feature of the `glob` module is its support for more complex patterns and recursive searching. By incorporating double asterisks `` in the pattern, users can perform searches that extend into subdirectories, allowing for the retrieval of files in a nested directory structure. This flexibility makes `glob` well-suited for projects where files are organized hierarchically, providing a streamlined approach to file handling in Python scripts.

CHAPTER 4

IMPLEMENTATION & RESULTS

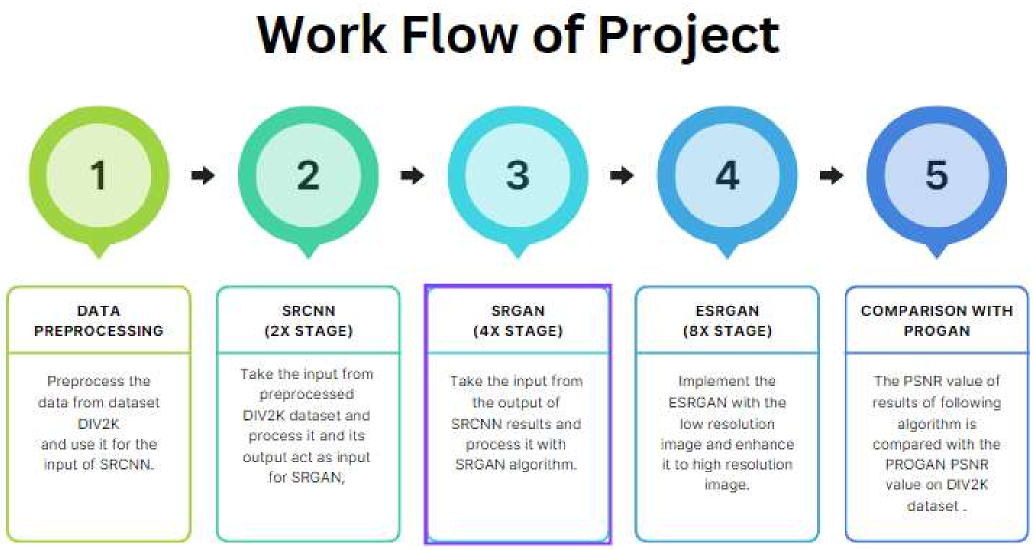

Fig 4.1 Workflow of the Project

4.1 Data Preprocessing

In the initial stage of the image super-resolution project, dataset selection played a crucial role. It was imperative to choose a dataset that comprised pairs of high-resolution and low-resolution images, as these pairs served as the foundation for training the model effectively. The existence of corresponding high and low-resolution counterparts enabled the algorithm to learn and enhance the details present in low-resolution images.

Following dataset selection, the subsequent step involved image preprocessing. To facilitate uniformity in data, pixel values were normalized to a common scale. Normalization ensured that the model processed images consistently, promoting stable and effective learning. Additionally, data augmentation techniques were employed, such as rotation and flipping, to introduce diversity into the training samples. Augmentation was essential for enhancing the model's ability to generalize well to various image conditions, thereby bolstering its performance on unseen data. This comprehensive approach to dataset selection and image preprocessing laid a solid foundation for the subsequent stages of the super-resolution task.

4.2 SRCNN (2X STAGE)

In this stage, the primary goal was to enhance image resolution through the application of deep learning techniques, a process that unfolded in several key steps. During the training phase, images underwent meticulous preprocessing using carefully selected techniques to ensure an optimal input for the subsequent deep learning model. The Super-Resolution Convolutional Neural Network (SRCNN) was identified as the most suitable model for this task and underwent training on the preprocessed dataset.

The architecture of SRCNN was designed with three fundamental layers: patch extraction, non-linear mapping, and reconstruction. These layers worked synergistically, with patch extraction capturing essential features from the low-resolution input, the non-linear mapping stage enhancing intricate image details, and the reconstruction phase generating an intermediate high-resolution image.

The ultimate output of the SRCNN model was a significantly improved version of the initial low-resolution input. This process showcased the remarkable effectiveness of the deep learning approach in the context of image resolution enhancement. By strategically combining preprocessing techniques and leveraging the three-layer architecture of SRCNN, this stage successfully demonstrated the capability of deep learning to elevate image resolution and extract finer details from lower-quality inputs.

4.3 SRGAN (4x STAGE)

In Stage 2 of the project, the focus shifted to the implementation of SRGAN (Super-Resolution Generative Adversarial Network) with the overarching objective of enhancing visual quality through adversarial training. The training phase involved preprocessing images as necessary and employing SRGAN, incorporating an adversarial loss alongside the conventional super-resolution loss. This adversarial training approach aimed to not only enhance resolution but also improve the overall visual realism of the generated images.

The architecture of SRGAN consisted of a generator responsible for super-resolution and a discriminator tasked with assessing the realism of the generated images. This adversarial setup fostered a dynamic interplay between the generator and discriminator, promoting the generation of high-resolution images that not only captured fine details but also exhibited a heightened level of visual fidelity.

The output of the SRGAN model served as a refinement of the intermediate result obtained from SRCNN. The final outcome was a visually enhanced high-resolution image, representing a significant advancement in the project's objective to improve image quality through the integration of adversarial training techniques. This stage showcased the potential of SRGAN in producing high-quality, realistic super-resolution results by leveraging the power of generative adversarial networks in the context of image enhancement.

4.4 ESRGAN (8x STAGE)

In the third stage of the project, the focus turned to the implementation of ESRGAN (Enhanced Super-Resolution Generative Adversarial Network) with the aim of further advancing resolution and refining fine details in the generated images. The training process involved meticulous preprocessing of images specifically tailored for ESRGAN training. The model was then trained using a combination of adversarial training, perceptual loss, and content loss, introducing a multi-faceted approach to enhance not only resolution but also perceptual and content quality.

ESRGAN boasted a deeper and more sophisticated architecture compared to its predecessor, SRGAN. This more intricate design allowed ESRGAN to capture and generate even finer details, pushing the boundaries of super-resolution capabilities. The architecture's depth and complexity played a pivotal role in the model's ability to discern intricate features and nuances in the images.

The output of ESRGAN represented a substantial improvement over the results obtained from SRGAN. Applied to the output of the previous stage, ESRGAN further elevated the resolution and visual quality of the images. This stage marked a significant stride in the project's overarching objective, demonstrating the effectiveness of ESRGAN in pushing the boundaries of super-resolution, capturing finer details, and producing visually stunning high-resolution images with improved quality and fidelity.

4.5 Comparison with ProGAN

In the final stage of the project, a comprehensive evaluation and comparison were conducted, considering the outputs generated by each progressive model—SRCNN, SRGAN, and ESRGAN—against the benchmark of ProGAN (Progressive Growing of GANs). The overarching goal was to assess the performance and quality of each super-resolution approach, taking into account both quantitative metrics and visual inspection.

Starting with the SRCNN model, its output served as the baseline for subsequent comparisons. While it demonstrated commendable improvements in resolution, the subsequent introduction of SRGAN elevated the visual quality by incorporating adversarial training, introducing a more realistic and detailed appearance to the generated images. The addition of perceptual and content loss in the ESRGAN model further refined these aspects.

ProGAN, known for its progressive growing architecture, was then brought into the comparison. The model's ability to generate high-quality images in a stepwise manner, progressively refining details at different resolutions, set a high standard for evaluation. ProGAN's output showcased a remarkable balance between sharpness, detail, and visual coherence, affirming its position as a formidable benchmark for comparison.

When assessing the outputs of SRCNN, SRGAN, and ESRGAN against ProGAN, distinct strengths and weaknesses emerged. While SRCNN served as a foundational step, ProGAN demonstrated superior overall quality. SRGAN and ESRGAN exhibited notable improvements over SRCNN, with ESRGAN, in particular, capturing finer details. However, ProGAN's progressive growth approach seemed to achieve a more consistent and coherent refinement across various resolutions.

4.6 RESULTS

The project aimed to enhance image resolution through a multi-stage approach employing deep learning techniques. Each stage introduced progressively advanced models to achieve superior results. The first stage focused on SRCNN, the Super-Resolution Convolutional Neural Network. The dataset selection was pivotal, requiring pairs of high and low-resolution images for effective training. During image preprocessing, pixel values were normalized to ensure consistent processing. Data augmentation techniques, such as rotation and flipping, were employed to diversify training samples.

In the training phase, the SRCNN model learned to map low-resolution images to their high-resolution counterparts. The architecture comprised three main layers: patch extraction, non-linear mapping, and reconstruction. This allowed SRCNN to effectively enhance details in low-resolution images. The output was an intermediate high-resolution image.

Fig 4.2 (a) SRCNN Input

Fig 4.2 (b) SRCNN Output

Stage two introduced SRGAN, the Super-Resolution Generative Adversarial Network. The objective was to not only improve resolution but also enhance visual quality through adversarial training. Images were preprocessed as needed, and SRGAN was trained with an adversarial loss in addition to the traditional super-resolution loss. The architecture included a generator for super-resolution and a discriminator for assessing realism. The output refined the intermediate result from SRCNN, producing a visually enhanced high-resolution image with heightened realism.

Fig 4.3 (a) SRGAN Input

Fig 4.3 (b) SRGAN Output

The third stage incorporated ESRGAN, the Enhanced Super-Resolution Generative Adversarial Network, aiming to further enhance resolution and fine details. Images were preprocessed specifically for ESRGAN training, and the model was trained with adversarial training, perceptual loss, and content loss. ESRGAN's architecture was deeper and more sophisticated compared to SRGAN, enabling it to capture finer details.

Fig 4.4 (a) ESRGAN Input

Fig 4.4 (b) ESRGAN Output

In the final stage, the outputs from SRCNN, SRGAN, and ESRGAN were compared against ProGAN (Progressive Growing of GANs), a benchmark known for its progressive growth architecture. ProGAN's stepwise refinement across different resolutions set a high standard. The evaluation considered quantitative metrics and visual inspection, revealing ProGAN's superior overall quality. While SRCNN served as a foundational step, SRGAN and ESRGAN exhibited notable improvements, with ESRGAN capturing finer details. ProGAN, however, showcased a more consistent and coherent refinement.

TABLE 5.1 Comparison Based on PSNR Value

PSNR VALUE	PROGAN	SRCNN	SRGAN	ESRGAN
2X STAGE	35.80	34.60		
4X STAGE	30.39		42.34	
8X STAGE	27.86			38.37

This detailed project overview underscores the meticulous process of dataset selection, image preprocessing, and model training at each stage. It demonstrates the evolution from basic super-resolution techniques to advanced generative adversarial networks, with a final comparison against a state-of-the-art benchmark. The nuanced evaluation provides insights into the strengths and limitations of each approach, guiding future advancements in the field of image super-resolution through deep learning.

CHAPTER5

CONCLUSION& FUTURE SCOPE

In conclusion, this comprehensive image super-resolution project has traversed a transformative journey, showcasing the progressive refinement of models from SRCNN to SRGAN and ESRGAN. The meticulous dataset selection, image preprocessing, and innovative model architectures underscore the commitment to advancing the field of image enhancement through deep learning.

The initial stage with SRCNN laid a robust foundation by emphasizing the significance of a well-curated dataset and effective preprocessing techniques. As the project evolved, the introduction of SRGAN brought forth adversarial training, introducing a dynamic interplay between a generator and discriminator. This not only enhanced resolution but also significantly improved the visual realism of the generated images.

ESRGAN, in the third stage, further elevated the super-resolution process. Its deeper architecture, combined with adversarial training, perceptual loss, and content loss, showcased a commitment to capturing finer details and pushing the boundaries of image enhancement. The iterative refinement throughout each stage highlighted the project's dedication to continual improvement.

The final comparison against ProGAN, a benchmark in progressive growing architectures, provided nuanced insights into the relative strengths and weaknesses of each approach. While SRCNN marked a foundational step, SRGAN and ESRGAN demonstrated commendable advancements, particularly in capturing finer details. However, ProGAN maintained a superior overall quality, setting a high standard for image super-resolution.

This project's significance lies not only in its achievements in image super-resolution but also in its roadmap for future endeavors. It serves as an inspiring example of how the iterative refinement of deep learning models can lead to substantial strides in visual quality and fidelity. As the project concludes, it leaves a legacy of innovation, providing valuable insights for future research in the dynamic intersection of deep learning and image enhancement.

This research project unfolds within a meticulously defined scope that intricately combines algorithmic development, dataset preparation, quantitative evaluation, comparative analysis, and documentation. The overarching goal is to advance our understanding of single-image super-resolution through the exploration of a progressive upscaling approach and its juxtaposition with a ProGAN-based iterative method.

5.1 Algorithmic Development:

In the realm of algorithmic development, a pivotal objective is the formulation and implementation of a resilient progressive upscaling algorithm. This algorithm stands as a synthesis of advanced techniques, seamlessly integrating the power of SRCNN, SRGAN, and ESRGAN. The goal extends beyond mere amalgamation, aiming to create a robust and effective mechanism for enhancing visual content progressively. The fine-tuning of algorithmic parameters assumes a central role in this process, requiring a delicate balance between computational efficiency and the attainment of optimal visual quality. This intricate calibration ensures that the algorithm operates with efficiency, delivering enhanced visuals without undue strain on computational resources. Moreover, paramount to the success of this endeavor is the establishment of algorithmic compatibility, fostering a cohesive and seamless multi-stage enhancement process. This harmonious integration ensures that each stage complements the other, creating a unified and powerful system for elevating visual content to new levels of quality and detail.

5.2 Dataset Preparation

In the meticulous process of dataset preparation, the DIV2K dataset assumes a pivotal role as the cornerstone for training, validation, and testing phases of the algorithmic development. The use of DIV2K provides a diverse and comprehensive set of images, facilitating a robust evaluation of the progressive upscaling algorithm. To ensure the efficacy of the training process, a meticulous preprocessing step is implemented, extracting high-resolution images from the DIV2K dataset along with their corresponding low-resolution counterparts at 2x, 3x, and 4x downscaling factors. This preprocessing phase is crucial in generating a well-balanced dataset that encompasses a variety of resolutions, mirroring real-world scenarios where content may vary in quality and detail. Furthermore, the dataset undergoes a systematic division to delineate specific subsets for training, validation, and final testing. This strategic division ensures that the algorithm is trained on a diverse set of data, validated on a separate subset to fine-tune its parameters, and ultimately tested on a distinct set to assess its generalization capabilities. The combination of the DIV2K dataset and meticulous preprocessing lays a solid foundation for the algorithm's training and evaluation, setting the stage for its application in real-world scenarios where high-quality upscaling is paramount.

5.3 Evaluation Metrics:

In the realm of image processing and progressive upscaling, the implementation of robust evaluation metrics is paramount for assessing the efficacy of any developmental stage. A pivotal step involves adopting stringent criteria, with a focus on industry-standard benchmarks such as the Peak Signal-to-Noise Ratio (PSNR) and Mean Squared Error (MSE). By leveraging these metrics, a precise quantification of the fidelity and quality of the upscale images can be achieved. To ensure a systematic and comprehensive evaluation process, the establishment of a standardized evaluation framework is imperative. This framework should meticulously measure and analyze the performance of the progressive upscaling approach at each developmental stage, providing insights into its strengths and areas for improvement. Moreover, creating a benchmark for comparative analysis, particularly aligning with the ProGAN-based iterative method, allows for a contextual understanding of the proposed approach's advancements in comparison to existing state-of-the-art techniques. This holistic approach to evaluation metrics ensures not only a thorough assessment of the progressive upscaling methodology but also facilitates informed decisions regarding its optimization and potential integration into real-world applications.

5.4 Comparison with ProGAN:

In conducting a comparative analysfis with ProGAN, the first step involves retrieving and preprocessing data from a seminal research paper that employs ProGAN for iterative super-resolution. This foundational data serves as a benchmark for evaluating the proposed progressive upscaling approach. The next phase encompasses the implementation of the ProGAN-based iterative method to facilitate a direct comparison between the two methodologies. This ensures a consistent and fair evaluation, allowing for a clear understanding of the strengths and weaknesses of each approach. To ascertain their relative performance, a meticulous examination of Peak Signal-to-Noise Ratio (PSNR) and Mean Squared Error (MSE) values is conducted. These industry-standard metrics provide a quantitative basis for assessing image fidelity and quality. The in-depth evaluation aims to delineate the nuances in performance between the progressive upscaling methodology and ProGAN iterations, offering valuable insights for further refinement and optimization. This comparative analysis contributes to a comprehensive understanding of how the proposed approach aligns with or surpasses the state-of-the-art ProGAN-based iterative method in the context of super-resolution.

5.5 Visual Analysis:

Engaging in a comprehensive visual analysis is imperative for understanding the nuanced aspects of the progressive upscaling approach in comparison to ProGAN iterations. This process commences by generating visually compelling comparisons, utilizing sample images extracted from the designated test dataset. These images serve as representative examples for evaluating the perceptual quality and visual fidelity achieved by both methodologies. The qualitative assessments delve into the intricacies of the visual output, highlighting any noticeable enhancements introduced by the progressive upscaling approach and discerning potential artifacts that may arise from ProGAN iterations. By scrutinizing the visual aspects, such as sharpness, clarity, and detail retention, a holistic perspective on the strengths and limitations of each method emerges. This visual analysis not only complements quantitative metrics but also provides a more intuitive understanding of how the proposed approach and ProGAN iterations impact the overall visual appeal of the upscaled images, thereby aiding in the refinement and optimization of these super-resolution techniques.

5.6 Documentation and Reporting:

Effective documentation and reporting are integral components of presenting research findings on the progressive upscaling approach in comparison to ProGAN-based iterative methods. The process involves methodically capturing algorithmic intricacies, ensuring a detailed account of the steps involved in the implementation. This documentation extends to the meticulous preparation of datasets, elucidating the selection criteria, preprocessing steps, and any considerations taken to maintain data integrity. Additionally, the documentation encompasses the intricacies of the evaluation methodologies employed, including the rationale behind metric choices and the specific parameters used for assessment. The synthesis of results is then approached with a focus on clarity and structure, aiming to create a cogent narrative that communicates the research findings effectively. By presenting a comprehensive overview of the study, the documentation enables readers to delve into the research process, understand the decisions made, and grasp the intricacies of the progressive upscaling approach. This comprehensive reporting facilitates a deeper comprehension of the methodology's efficacy when compared to ProGAN-based iterative methods, providing valuable insights for both researchers and practitioners in the field of super-resolution techniques.

Within this comprehensive scope, the research aspires to make meaningful contributions to the landscape of single-image super-resolution, unraveling the nuanced strengths and limitations inherent in progressive upscaling strategies and offering valuable comparative insights with ProGAN-based iterative approaches.

REFERENCES

[1] “Image Super Resolution | Deep Learning for Image Super Resolution,” Analytics Vidhya, May 26, 2021. https://www.analyticsvidhya.com/blog/2021/05/deep-learning-for-image-super-resolution/

[2] Anwar, S.; Khan, S.; Barnes, N. A deep journey into super-resolution: A survey. arXiv 2019, arXiv:1904.07523. [Google Scholar] [CrossRef]

[3] Y. Wang, F. Perazzi, B. McWilliams, A. Sorkine-Hornung, O. Sorkine-Hornung, and C. Schroers, “A Fully Progressive Approach to Single-Image Super-Resolution,” Jun. 2018, doi: https://doi.org/10.1109/cvprw.2018.00131.

[4] Z. Hui, J. Li, X. Gao, and X. Wang, “Progressive perception-oriented network for single image super-resolution,” *Information Sciences*, vol. 546, pp. 769–786, Feb. 2021, doi: https://doi.org/10.1016/j.ins.2020.08.114.

[5] Li, X.; Wu, Y.; Zhang, W.; Wang, R.; Hou, F. Deep learning methods in real-time image super-resolution: A survey. J. Real-Time Image Process. 2020, 17, 1885–1909.

[6] Nasrollahi, K.; Moeslund, T.B. Super-resolution: A comprehensive survey. Mach. Vis. Appl. 2014, 25, 1423–1468.

[7] Tong, C.S.; Leung, K.T. Super-resolution reconstruction based on linear interpolation of wavelet coefficients. Multidimens. Syst. Signal Process. 2007, 18, 153–171. [Google Scholar] [CrossRef]

[8] Sun, N.; Li, H. Super Resolution Reconstruction of Images Based on Interpolation and Full Convolutional Neural Network and Application in Medical Fields. IEEE Access 2019, 7, 186470–186479. [Google Scholar] [CrossRef]

[9] Liu, J.; Gan, Z.; Zhu, X. Directional Bicubic Interpolation—A New Method of Image Super-Resolution. In 3rd International Conference on Multimedia Technology(ICMT-13); Atlantis Press: Paris, France, 2013; pp. 470–477. [Google Scholar]

[10] Kumar, G.; Singh, K. Image Super Resolution on the Basis of DWT and Bicubic Interpolation. Int. J. Comput. Appl. 2013, 65, 1–6. [Google Scholar]

[11] Dai, S.; Han, M.; Xu, W.; Wu, Y.; Gong, Y. Soft Edge Smoothness Prior for Alpha Channel Super Resolution. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8. [Google Scholar]

[12] Chang, K.; Ding, P.L.K.; Li, B. Single Image Super Resolution Using Joint Regularization. IEEE Signal Process. Lett. 2018, 25, 596–600. [Google Scholar] [CrossRef]

[13] Yu, L.; Cao, S.; He, J.; Sun, B.; Dai, F. Single-image super-resolution based on regularization with stationary gradient fidelity. In Proceedings of the 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Shanghai, China, 14–16 October 2017; pp. 1–5. [**Google Scholar**]

[14] Shan, Q.; Li, Z.; Jia, J.; Tang, C.-K. Fast image/video upsampling. *ACM Trans. Graph.* **2008**, *27*, 1–7. [**Google Scholar**] [**CrossRef**]

[15] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Jul. 2017, doi: https://doi.org/10.1109/cvprw.2017.151.

[16] Y. Liu, X. Zhang, S. Wang, S. Ma, and W. Gao, "Progressive Multi-Scale Residual Network for Single Image Super-Resolution.," Jul. 2020.

[17] Zhang, Y., Li, K., Li, Y., Zhang, L., & Wei, Y. (2020). Attention-Based Deep Fusion Network for Single Image Super-Resolution. IEEE Transactions on Image Processing, 29(12), 3568-3581

[18] @InProceedings{Timofte_2017_CVPR_Workshops,

author = {Timofte, Radu and Agustsson, Eirikur and Van Gool, Luc and Yang, Ming-Hsuan and Zhang, Lei and Lim, Bee and others},

title = {NTIRE 2017 Challenge on Single Image Super-Resolution: Methods and Results},

booktitle = {The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops},

month = {July},

year = {2017}

}

[19] P. Young, A. Lai, M. Hodosh, and J. Hockenmaier, "From image descriptions to visual denotations: New similarity metrics for semantic inference over event descriptions," in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2014, pp. 1532-1543.

[20] Timofte, Radu and Gu, Shuhang and Wu, Jiqing and Van Gool, Luc and Zhang, Lei andYang, Ming-Hsuan and Haris, Muhammad and others,"NTIRE 2018 Challenge on SingleImageSuper-Resolution",TheIEEEConferenceonComputerVisionandPatternRecognition(CVPR)Workshops,June,2018

[21] Ignatov, Andrey and Timofte, Radu and others, "PIRM challenge on perceptual imageenhancementonsmartphones",EuropeanConferenceonComputerVision(ECCV)Workshops,January,2019

[22]] Ledig, C., Theis, L., Husz´ar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., et al.: Photo-realistic single image superresolution using a generative adversarial network. In: CVPR. (2017)

[23] Jolicoeur-Martineau, A.: The relativistic discriminator: a key element missing from standard gan. arXiv preprint arXiv:1807.00734 (2018)

[24] Blau, Y., Mechrez, R., Timofte, R., Michaeli, T., Zelnik-Manor, L.: The pirm challenge on perceptual super resolution. https://www.pirm2018.org/PIRM-SR.html (2018)

[25] Dong, C., Loy, C.C., He, K., Tang, X.: Learning a deep convolutional network for image super-resolution. In: ECCV. (2014) [5.] Kim, J., Kwon Lee, J., Mu Lee,

K.: Accurate image super-resolution using very deep convolutional networks. In: CVPR. (2016)

[26] Google, “TensorFlow,” TensorFlow, 2019. https://www.tensorflow.org/

[27] “opencv-python,” PyPI, Nov. 21, 2019. https://pypi.org/project/opencv-python/

[28] Numpy, “NumPy,” Numpy.org, 2009. https://numpy.org/

[29] PyTorch, “PyTorch,” Pytorch.org, 2023. https://pytorch.org/

[30] “glob — Unix style pathname pattern expansion — Python 3.8.3rc1 documentation,” docs.python.org. https://docs.python.org/3/library/glob.html