Project Phase- II Report On

Single Image Super Resolution using GAN

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

Bachelor of Technology

in

Computer Science and Engineering

By

Abel John Mathew (U2003003)
Ashwin Saji (U2003220)
Athif Ahamed (U2003050)
Didin Shibu (U2003069)

Under the guidance of

Dr. Jincy J Fernandez

Department of Computer Science and Engineering
Rajagiri School of Engineering & Technology (Autonomous)
(Parent University: APJ Abdul Kalam Technological University)
Rajagiri Valley, Kakkanad, Kochi, 682039
April 2024

CERTIFICATE

This is to certify that the project report entitled ***"Single Image Super Resolution using GAN"*** *is a bonafide record of the work done by* ***Abel John Mathew (U2003003)***, ***Ashwin Saji (U2003220)***, ***Athif Ahamed (U2003050)*** *and* ***Didin Shibu (U2003069)*** *submitted to the Rajagiri School of Engineering & Technology (RSET) (Autonomous) in partial fulfillment of the requirements for the award of the degree of Bachelor of Technology (B. Tech.) in Computer Science and Engineering during the academic year 2023-2024.*

Ms. Jincy J Fernandez
Project Guide
Assistant Professor
Dept. of CSE
RSET

Dr. Sminu Izudheen
Project Coordinator
Professor
Dept. of CSE
RSET

Dr. Preetha K G
Head of the Department
Professor
Dept. of CSE
RSET

ACKNOWLEDGMENT

We wish to express our sincere gratitude towards **Prof.(Dr). P. S. Sreejith**, Principal of RSET, and **Dr. Preetha K.G**, Head of the Department of Computer Science & Engineering for providing us with the opportunity to undertake our project, ”Single Image Super Resolution using GAN”.

We are highly indebted to our project coordinator, **Dr. Sminu Izudheen**, Professor, Department of Computer Science for her valuable support.

It is indeed our pleasure and a moment of satisfaction for us to express our sincere gratitude to our project guide **Dr. Jincy J Fernandez** for her patience and all the priceless advice and wisdom she has shared with us.

Last but not the least, We would like to express our sincere gratitude towards all other teachers and friends for their continuous support and constructive ideas.

Abel John Mathew

Ashwin Saji

Athif Ahamed

Didin Shibu

Abstract

Single Image Super-Resolution (SISR) stands as a cutting-edge solution to the ever rising challenge of enhancing the quality of low-resolution images. This project uses advanced deep learning techniques, particularly Generative Adversarial Networks (GANs), to achieve excellent results. The significance of SISR is highlighted by its rising impact across diverse industries. In the realm of healthcare, the project brings a notable advantage by providing enhanced resolution in medical imaging, contributing to more accurate diagnostics. The advantages extend to security and surveillance applications, where the project outperforms existing methods by delivering sharper, more detailed images, thereby improving object recognition and tracking capabilities.

The proposed SISR project introduces a change by utilizing deep learning to capture complex mappings from low to high resolution, in contrast to conventional interpolation-based approaches. The project's effectiveness is also evident in the field of remote sensing, where it provides enhanced spatial resolution in satellite imagery to enable accurate environmental monitoring. This project's innovative design and training techniques provide a notable edge over previous efforts, guaranteeing the production of high-resolution photographs.

Beyond the said applications, the project advances the state of the art in multimedia experiences by meeting the rising demand for high-quality visuals. This project represents a significant leap forward in SISR, showcasing advantages over existing works through its innovative methodologies, superior performance, and potentials across a wide range of real-world scenarios.

Contents

List of Abbreviations

1. RSI - Remote Sensing Images
2. SISR - Single Image Super Resolution
3. MISR - Multi Image Super Resolution
4. GAN - Generative Adverserial Networks
5. RBAN - Residual Balanced Attention Network
6. UNet - U-Shaped Network
7. SISR - Single Image Super Resolution
8. LR - Low Resolution
9. HR - High Resolution
10. SR - Super Resolution
11. PSNR - Peak Signal-to-Noise Ratio
12. SSIM - Structure Similarity Index
13. IFS - Independent Feature Similarity
14. LPIPS - Learned Perceptual Image Patch Similarity
15. NIQE - Natural Image Quality Evaluator
16. ENIQA - Entropy-based No-reference Image Quality Assessment
17. AID - Aerial Images Dataset
18. UCMERCED - University of California Merced
19. SRCNN - Super Resolution using Convolutional Neural Networks
20. VDSR - Very Deep Super Resolution
21. EDSR - Enhanced Deep Residual Networks for Single Image SR
22. SRGAN - Super-Resolution Generative Adversarial Networks
23. DRLN - Densely Residual Laplacian Super-resolution
24. BAM - Balanced Attention Module
25. VGG - Visual Geometry Group
26. GMAC - Giga Multiply-Add Operations per Second
27. GFLOP - Giga Floating-Point Operations per Second

List of Figures

List of Tables

Chapter 1

Introduction

One of the most important and exciting subfields of computer vision and image processing is Single Image Super-Resolution (SISR). Enhancing a low-resolution image's spatial resolution in order to create a higher-resolution version while maintaining crucial information is the main objective of SISR. This method is especially useful for a number of applications, such as digital photography, satellite imagery[1], surveillance, and medical imaging[2].
Obtaining high-resolution photographs can be difficult or resource-intensive in many real-world situations. In order to overcome this constraint, SISR approaches use complex algorithms to estimate the missing pixel information and produce images with greater clarity and finer features.
SISR approaches performance has greatly increased due to recent developments in deep learning. Even with these developments, research is still ongoing to address issues including managing various degradation factors, lowering computing complexity, and guaranteeing real-time performance. Single Image Super-Resolution remains a dynamic and evolving topic with important implications for several applications across various industries as the need for high-quality visual content grows.

1.1 Background

The background of the project encompasses the ever-present challenge of dealing with low-resolution images, a rising issue across various domains.Present-day photos obtained from various sources, including medical imaging equipment, surveillance cameras, and remote sensing satellites, frequently have spatial resolution issues. These constraints make it more difficult to identify minute details and extract important information from the visual data. Conventional techniques for upscaling low-resolution photos have not been sufficient to extract fine details and have produced visually poor results. This shortcoming is especially noticeable in domains like medical diagnostics, where the accuracy and clarity of images have a direct bearing on the diagnosis and treatment choices made thereafter. The importance of Single Image Super-Resolution (SISR) in this context cannot be overstated. SISR is a ground-breaking remedy that goes beyond the drawbacks of conventional methods. Utilizing cutting-edge deep learning methods like Generative Adversarial Networks (GANs), SISR has the potential to completely change the image processing industry. The project's background extends to remote sensing applications,where enhanced spatial resolution satellite and aerial imagery can transform urban planning, disaster response, and environmental monitoring. The shortcomings of existing scenarios highlight how urgent it is to have a technology like SISR in order to satisfy the growing demand for high-quality pictures. SISR is becoming a center of attention for innovation and research in computer vision and image processing as a result of its rising relevance in solving real-world problems and revolutionizing the image enhancement industry.

1.2 Problem Definition

Enhancing image quality in low-resolution scenarios through advanced high-resolution reconstruction techniques.

1.3 Scope and Motivation

The scope of SISR spreads over a wide range of industries, as a response to dealing with low resolution images. In the medical field, SISR holds the potential to revolutionize medical imaging, enabling doctors in making more accurate and faster medical diagnoses. Beyond health care SISR plays a major role in surveillance and security, where the ability to identify and analyze the fine details in low-resolution images is critical for threat detection. Moreover the the applications of SISR also proved to be of great importance in the field of remote sensing, where satelite imagery can be improved to support more precise and accurate analysis of the data. As technology advances the scope of SISR keeps on expanding from scientific research to multimedia applications where the demand of better quality images are on the rise.

The motivation behind SISR lies in its potential across various industries and applications. In healthcare, the motivation rises from the need to enhance the clarity of medical images, thereby improving diagnostic accuracy and patient care. In the context of security and surveillance, the motivation is to improve the effectiveness of surveillance system by providing sharper, more detailed images for better object recognition and tracking. Additionally the motivation extends to the field of remote sensing, where improvement of satelite imagery supports better decision making. Beyond all these, the capability of SISR lies in its ability to redefine visual experiences in multimedia, catering to the the growing demand for high quality images and videos. As a result I can say that the motivation behind SISR lies in its potential to address real world challenges.

1.4 Objectives

- Develop an ML-based single image super-resolution application that would enhance the clarity of low resolution images.
- Ensure that the important details and features in the image are preserved and not distorted after the upscaling process.
- Train the network to produce accurate higher resolution image from the input that is a lower resolution image.
- Conduct extensive experiments on benchmark datasets to evaluate and analyze the performance of models.
- Develop a user-friendly interface for applying the SISR model to input images and getting the outputs.

1.5 Challenges

The challenges faced include adversarial attacks, Trade-off Between Realism and Generalization, Complexity of Real-World Degradations.
Developing a Single Image Super-Resolution (SISR) project using Generative Adversarial Networks (GANs) presents several challenges. First and foremost is the quality and quantity of the dataset, where obtaining a diverse and representative set of high-resolution images can be limited. Additionally, computational resources pose a significant challenge

due to the resource-intensive nature of GANs, impacting both training time and hardware requirements. Mode collapse, a phenomenon where the generator produces limited variations, is another issue that needs careful consideration. Hyperparameter tuning, involving aspects like learning rates and batch sizes, is crucial but intricate. Selecting appropriate loss functions, balancing between content and adversarial loss, influences the perceptual accuracy of generated images. Addressing artifacts introduced during the upscaling process is essential, as is the definition of reliable evaluation metrics for assessing super-resolved image quality. Achieving stability in adversarial training, generalization to unseen data, and real-time inference are additional challenges. Ethical considerations, ensuring compliance with norms, and the transferability of the model across different scenarios are also vital aspects of the development process. Combining advanced techniques, staying abreast of the latest research, and engaging in iterative development are essential to overcoming these multifaceted challenges.

1.6 Assumptions

Firstly assume that there is an abundant availability of a diverse and representative dataset of high resolution images for training the machine learning models, then I assume that the model can handle common artifacts introduced during the super-resolution process, such as noise or blurriness. Also assume the SRGAN model can be seamlessly integrated into existing systems if required.

1.7 Societal / Industrial Relevance

Medical Field: SISR can improve the quality of diagnostic images [2] which can help the doctors in identifying even small details which result in improved diagnosis of diseases.
Surveillance and Security: Improved video surveillance using SISR can improve the ability to identify and analyze object's which leads to better security for the public.
Remote Sensing: Large areas are captured by remote sensing satellites at different spatial resolutions. With SISR, satellite imagery can be improved[1], giving users a clearer, more detailed perspective of cities, landscapes, and other environmental elements.
Digital Photography and Entertainment: SISR can improve the quality of digital photographs, hence providing the users with more details and quality in their photographs.

1.8 Organization of the Report

The report is organized into seven distinct chapters, each contributing to a comprehensive exploration of the development of a Single Image Super-Resolution system utilizing Generative Adversarial Networks (GANs).

Chapter 1 :- This chapter serves as the introduction, providing a clear overview of the project's background and motivation, setting the foundation for subsequent research.

Chapter 2 :- In this chapter a thorough literature survey is conducted, offering a comprehensive review of existing Super Resolution methods.

Chapter 3 :- This chapter is dedicated to detailing the hardware and software requirements essential for the project.

Chapter 4 :- The system architecture is elucidated, delving into the specifics of various steps involved in the development process.

Chapter 5 :- This chapter describes in detail how the GAN based model was implemented, showcasing relevant architecture diagrams and code snippets. The chapter also includes wireframes created for the project interface as well as the database design.

Chapter 6 :- In this chapter, the results of the project are formulated in the form of graphs and tables, providing qualitative analysis of findings achieved throughout this project.

Chapter 7 :- This is the concluding chapter that summarizes key findings and contributions and provides recommendations for further investigation.

Chapter 2

Literature Survey

2.1 Existing Methods

2.1.1 SRCNN:

In 2016, Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang pioneered the Super-Resolution Convolutional Neural Network (SRCNN) [3]. Their groundbreaking approach, detailed in "Image Super-Resolution Using Deep Convolutional Networks", showcased the power of deep learning in achieving state-of-the-art results for single-image super-resolution. SRCNN's architecture as shown in Figure 2.1, though relatively simple, effectively captures complex relationships within image patches. However, challenges include its reliance on fixed-size inputs, potentially limiting adaptability to varying resolutions and posing challenges in generalizing across diverse image types. This architecture begins with a low-resolution input image, which is divided into overlapping patches to effectively capture local information. The subsequent convolutional layers perform feature extraction, learning hierarchical representations from low to high-level features. Rectified Linear Unit (ReLU) activation functions introduce non-linearity to the model, enhancing its ability to capture complex relationships in the data. Notably, SRCNN is characterized by a relatively shallow network depth, consisting of a few convolutional layers. The final convolutional layer is dedicated to reconstructing the high-resolution image.

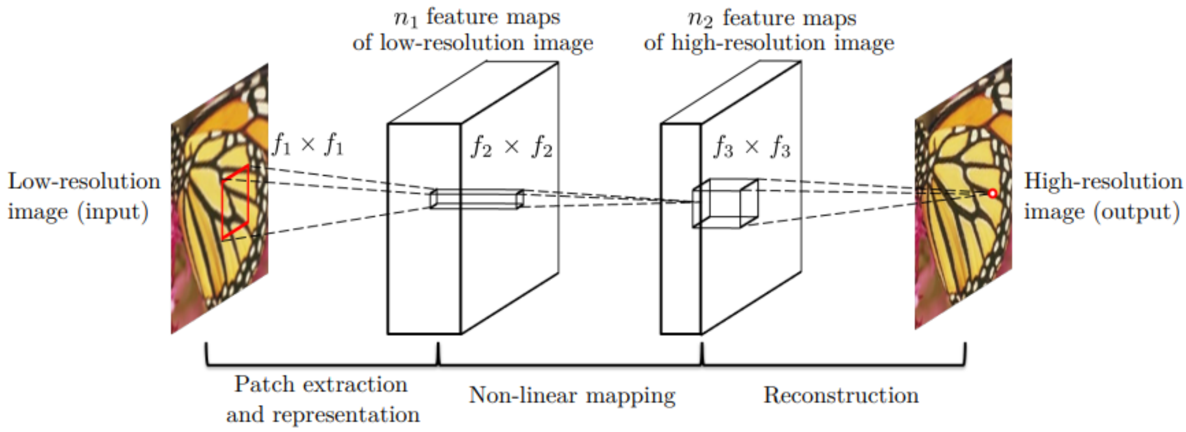


Figure 2.1: Super-Resolution Convolutional Neural Network (SRCNN)

2.1.2 VDSR:

Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee introduced the Very Deep Super-Resolution (VDSR) model in 2016. Outlined in "Accurate Image Super-Resolution Using Very Deep Convolutional Networks"[4], The VDSR network structure contains a total of d layers, where all layers except the first and last have the same specifications. Each layer consists of 64 channels of size $3 \times 3 \times 64$, with a channel operating over 64 channels (feature maps) in a 3×3 spatial domain as shown in Figure 2.2. The first layer handles the input image, and the last layer, which is responsible for image reproduction, consists of a single channel of size $3 \times 3 \times 64$. The network takes as input an interpolated low-resolution image (adjusted to the required size) and predicts the finer details of the image. A challenge with using a very deep network to predict dense results is reducing the size of the feature map after each convolution operation. To overcome this challenge, zero padding is used before convolution to keep all map objects, including the output image, consistent. The null pad is an effective tool to maintain accurate predictions even for pixels near the image boundary. This approach differs from many other methods that can limit the image that can be obtained, especially when dealing with large demanding environmental areas.

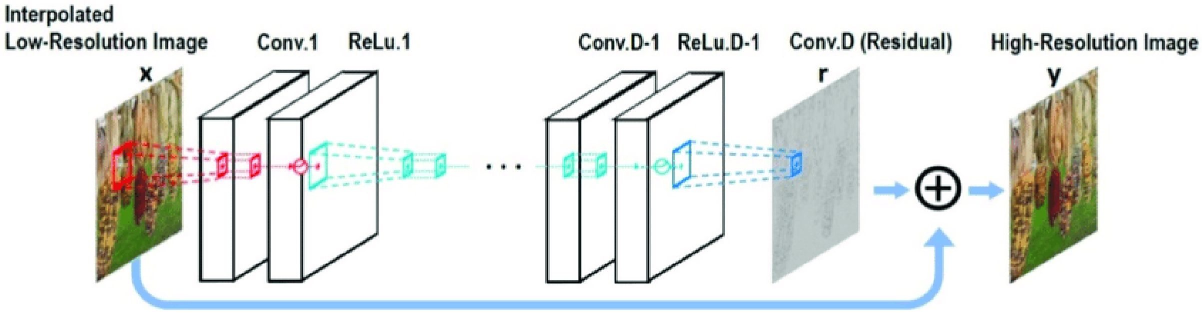


Figure 2.2: Very Deep Convolutional Networks (VDSR)

2.1.3 SRGAN

In 2017, the paper-Realistic Single Image Super-Resolution Using a Generative Adversarial Network"[5] by Christian Ledig et al. introduced the Super-Resolution Generative Adversarial Network (SRGAN), a transformative approach to single-image super-resolution. SRGAN integrates a generator, SRResNet, with a discriminator within a GAN (generative adversarial network) framework as shown in Figure 2.3. Notably, SRGAN goes past standard super-resolution ways by adding adversarial training and perceptual loss to the procedure. This special fusion tends to not really the pixel-wise errors but also leads to perceptual quality which are of high-resolution images that are highly realistic.

The adversarial preparation step in the SRGAN plays a key role of enhancing the realism of the generated images. But this creative approach also presents its contributors with challenges. Training GANs can be intrinsically difficult, and issues such as mode collapse represent potential trouble for the approach. It is crucial to maintain a delicate balance between generator and discriminatifier, and fine-tuning hyperparameters becomes decisive ensuring stable and efficient training of SRGAN. Of course, there are suprisingly several other challenges in designing and training GANs but despite these as an approach, SRGAN tackles an important advancement in the quest for photo-realistic single-image super-resolution. SRGAN uses both content loss and adversarial loss during the training of the generator. Content loss calculates the perceptual variance between the produced and ground truth high-resolution images. On the other hand, adversarial loss motivates the generator to create images that closely resemble authentic high-resolution images which can fool the discriminator when evaluated it.

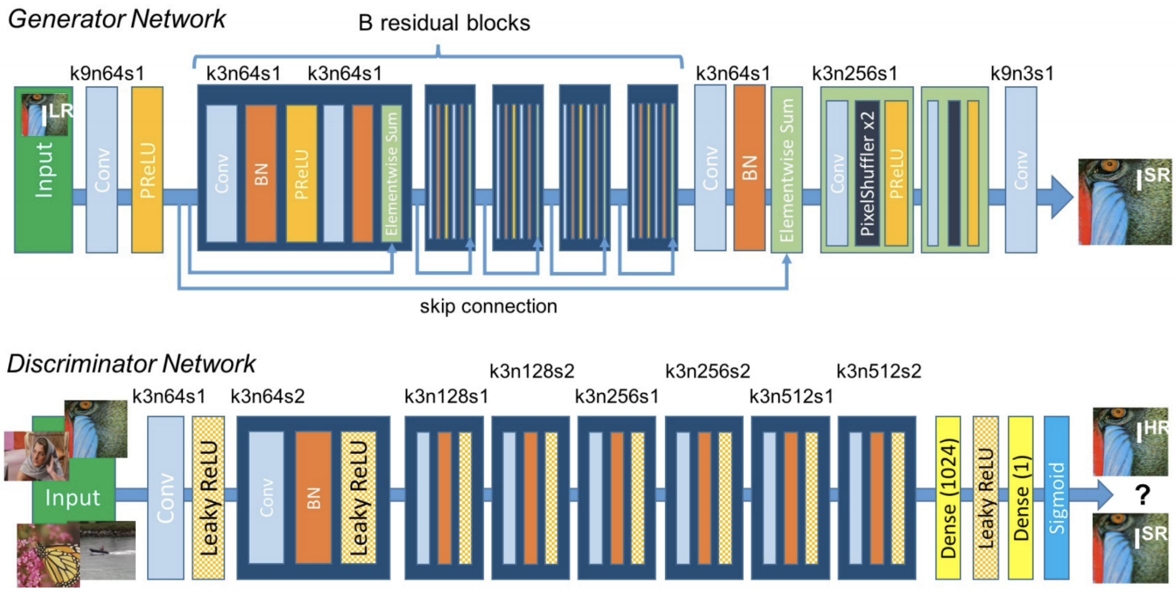


Figure 2.3: Single Image Super-Resolution Using a Generative Adversarial Network (SR-GAN)

The gaps faced in the current SISR methods like SRGAN are that it doesn't do well for images which have degradations like atmospheric propogation, lens blur, noise etc. These are normal degradations which are seen among most images but SRGAN is trained on a high qual-ity dataset hence it doesn't help in real world images. SRGAN also produces a lot of artifacts and inconsistencies because it is trained in a large dataset. Another problem faced in SISR models is that Although the goal of SISR techniques is to enhance images, maintaining realism and naturalness is a challenge. Edited images can sometimes lack natural texture and detail, leading to blurry results. Another problem faced in SISR is that Quantifying the subjective quality of enhanced images is challenging. Traditional metrics such as PSNR and SSIM do not always correlate well with human perception, making it difficult to objectively assess the quality of images produced.

2.2 Literature Papers

2.2.1 SISR of RSI

This paper [1] takes a multifaceted approach to capturing and addressing the intricacies of real-world degradation. The conventional models for blur often fall short in representing the complex factors contributing to degradation in Remote Sensing Images (RSIs). To overcome this limitation, the study transcends simplistic blur models and introduces a modified KernelGAN framework as shown in Figure 2.4. This sophisticated tool is specifically tailored to scrutinize real RSIs, extracting precise blur kernels that intricately account for atmospheric turbulence, sensor noise, and compression artifacts. Blur kernel estimation with KernelGAN. The Discriminator tries to find the difference between the real patches and the ones generated by the Generator, like this the Generator learns to downsample the image effectively to an extent where it can fool the Discriminator.

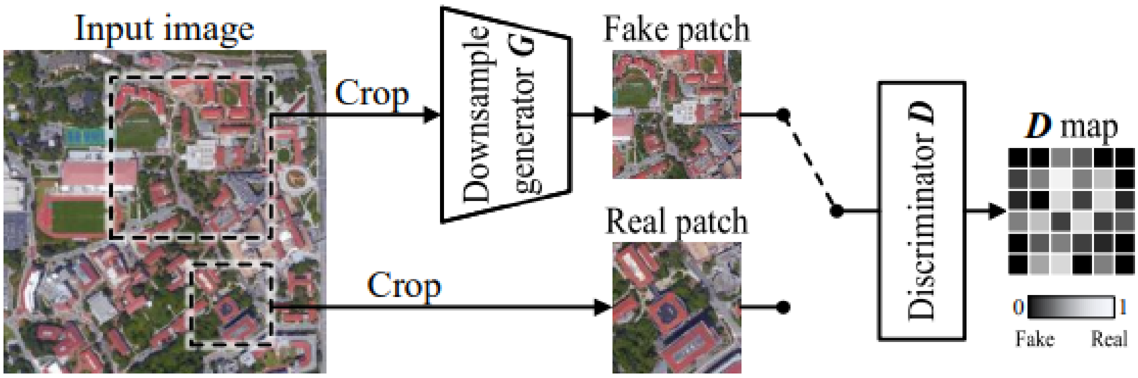


Figure 2.4: KernelGAN

To add on to the realistic degradation of the dataset, random noise patches are calculated and added on to the image as shown in Equation 2.1. A threshold 'v' is found out by the author from images like watery surfaces or bare land.

$$\sigma(\mathbf{n}_i) < v, \tag{2.1}$$

Algorithm 1 Realistic data pairs generation

Input: Real RSIs set $\mathcal{X}$
Output: Realistic image pairs $\{\mathbf{I}_{\mathrm{LR}}, \mathbf{I}_{\mathrm{HR}}\}$

Initialize kernel dataset $\mathcal{K} = \emptyset$
Initialize noise dataset $\mathcal{N} = \emptyset$
for all $\mathbf{I}_{\mathrm{real}}$ that $\mathbf{I}_{\mathrm{real}} \in \mathcal{X}$ **do**
    Estimate $\mathbf{k}$ from $\mathbf{I}_{\mathrm{real}}$ by solving Equation (4) and add $\mathbf{k}$ to $\mathcal{K}$
    Crop $\mathbf{n}$ from $\mathbf{I}_{\mathrm{real}}$
    if $\mathbf{n}$ meets Equation (5) **then**
        Add $\mathbf{n}$ to $\mathcal{N}$
    end if
end for
for all $\mathbf{I}_{\mathrm{real}}$ that $\mathbf{I}_{\mathrm{real}} \in \mathcal{X}$ **do**
    Generate $\mathbf{I}_{\mathrm{HR}}$ using Equation (2)
    Randomly select $\mathbf{k}_i \in \mathcal{K}, \mathbf{n}_j \in \mathcal{N}$
    Generate $\mathbf{I}_{\mathrm{LR}}$ using Equation (3)
end for
return $\{\mathbf{I}_{\mathrm{LR}}, \mathbf{I}_{\mathrm{HR}}\}$

Figure 2.5: Algorithm for the creation of HR-LR pair

A crucial first step towards guaranteeing that super-resolution models can function well in real-world scenarios is the creation of realistic training data. The paper uses a careful algorithm to generate paired datasets of Low-Resolution (LR) and High-Resolution (HR) images by leveraging the predicted blur kernels, noise patches, and genuine RSIs. The super-resolution model is trained on this artificial dataset, offering a multitude of scenarios which imitate real-world degradation difficulties. By convolving estimated kernels and patches with real remote sensing images to create low-high resolution pairings, the model learns atmospheric turbulence, sensor noise, and other degradation inducing factors and their nuances as shown in Figure 2.5.

The main part of this framework is the Residual Balanced Attention Network (RBAN). It is a novel architecture designed for Single Image Super-Resolution (SISR) tasks of RSIs. The RBAN architecture is made up of three key components: shallow feature extraction, deep feature extraction with Residual Balanced Attention Groups (RBAG), and a Balanced Attention Module (BAM) for reconstructing the image which involves upsampling and convolutional layers. The attention-driven feature extraction is the main feature of RBAN, with RBAGs processing features across multiple scales to gather information and preserve important details essential for accurately reconstructing fine textures. The addition of BAM brings in a smart way of deciding how much importance each type of information should get in super-resolution tasks. It focuses more on the information that is most helpful for making the images look better. To compliment the Generator which is the RBAN architecture, the paper introduces a modified UNet discriminator within a Generative Adversarial Network (GAN) framework as shown in Figure 2.6. The discriminator acts like a critic, carefully checking how realistic each pixel looks and giving useful feedback to the RBAN while it learns. By using a Relativistic GAN (RGAN) framework, the RBAN is guided to create super-resolution images that not only look good but also seem very real when it comes to representing remote sensing scenes.

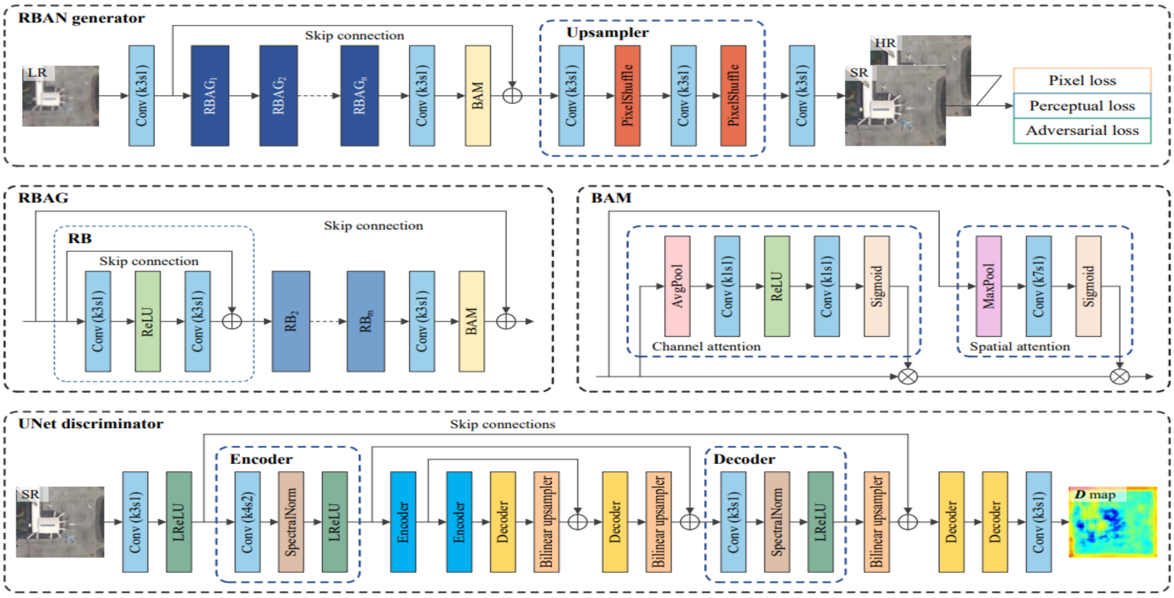


Figure 2.6: RBAN-UNet Architechture

Working together in a Collaborative training with the GAN framework results in a shared learning experience. The UNet discriminator plays an active role in evaluating how accurate the images are. Through this adversarial learning algorithm, the super-resolution results are fine-tuned to be more in line with the complexity found in real-world remote sensing imagery. The discriminator ability to provide feedback at the pixel level ensures that local textures are accurately represented, increasing the fidelity of the overall super-resolution process. The loss function and optimization used are basednon a comprehensive approach. The goal of the training program is to minimize a combination of three different types of loss: pixel loss, perceptual loss, and adversarial loss.

Pixel loss is The main component, measuring the pixel-wise difference between the calculated super-resolution image and the ground truth High-resolution (HR) image. Perceptual loss takes the training a step further by capturing higher-level image quality.

This is achieved by comparing features extracted from the pre-trained trained VGG network, which is a well-known image classification network capable of extracting high level semantic information from images. Finally, the addition of adverserial loss guides the RBAN to generate images that is very similar to the ground truth but also can effectively fool the UNet discriminator. As the training process unfolds, the RBAN model and the UNet discriminator together undergoes meticulous optimization to enhance its Super-Resolution (SR) performance on remote sensing images.

The collaboration between RBAN and GAN frameworks allows for a comprehensive learning experience, where the model adapts itself to real-world degradation complexities, while continuously refining its output to match the nuances of the actual RSI. In the final phase of the study, the effectiveness of the trained RBAN model is evaluated on existing standards.

This validation step is important in finding out the real-world application of the model and its ability to produce high-quality super-resolution reconstructions of real remote sensing data. Paving the way for various applications such as overlay mapping, disaster assessment, and environmental monitoring In summary, this research presents a comprehensive and sophisticated approach to address the challenges of super-resolution in remote sensing imagery the inspection process. From capturing real-world damage through advanced blur kernel estimation and noise collection to generate a synthetic realistic training data and using attention blocks for feature extraction. This approach is very comprehensive and innovative.

2.2.2 Real-ESRGAN

Real-ESRGAN revolutionizes image super-resolution with an inventive methodology and advanced architecture, adept at accurately simulating and rectifying the complexities of real-world image degradations. At its core, the model adopts a high-order degradation modeling process as shown in Figure 2.7, surpassing conventional first-order models. This extension allows Real-ESRGAN to emulate a diverse range of intricate degradation processes observed in real-world images, resulting in a more flexible and genuine representation of degradation phenomena.

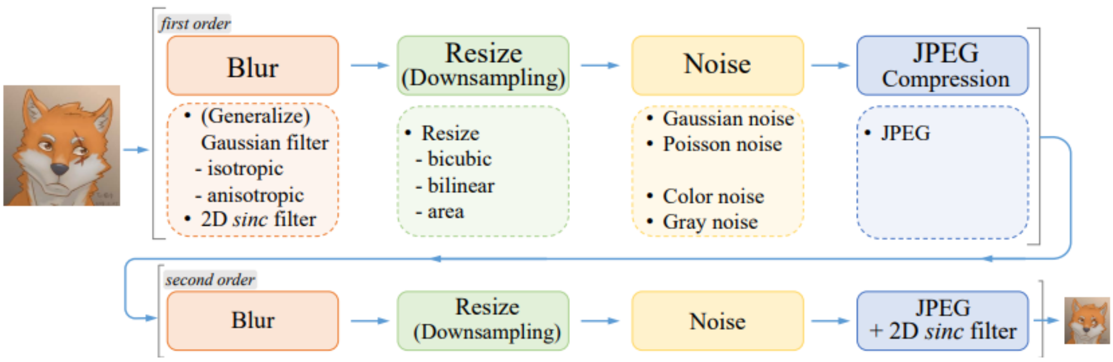


Figure 2.7: Second order Degradation model for practical and realistic degradations

A standout feature of Real-ESRGAN involves the incorporation of sinc filters into the synthesis process as shown in Figure 2.8. These filters play a pivotal role in reproducing common ringing and overshoot artifacts prevalent in degraded images. By seamlessly integrating these filters, Real-ESRGAN effectively eliminates artifacts during image restoration, leading to a significant improvement in visual sharpness in the restored images. The model's nuanced approach to artifact management is a fundamental aspect of its strategy, recognizing and mitigating visual distortions commonly encountered in real-world degraded images. In the Real-ESRGAN+ variant, a strategic enhancement is introduced to the training process by integrating sharpened ground-truth images. This variation aims to strike a delicate balance between achieving sharpness and suppressing overshoot artifacts frequently observed in real-world scenarios. By training the model with these sharpened ground-truth images, Real-ESRGAN+ aims to attain an optimized equilibrium, overcoming challenges related to overshoot artifacts and bolstering the model's overall effectiveness in producing high-quality super-resolved images.

The architecture of Real-ESRGAN prominently features a U-Net discriminator with spec-

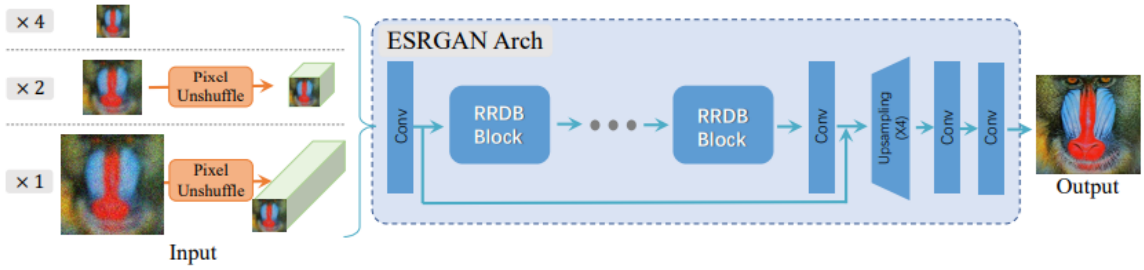


Figure 2.8: Architechture of Real-ESRGAN

tral normalization. This discriminator architecture plays a pivotal role in the adversarial training process, utilizing spectral normalization to enhance its ability to distinguish real images from complex training outputs. This strategic discriminator architecture contributes to the overall effectiveness of the adversarial training approach, enabling the generator to produce more realistic images closely aligned with the characteristics of authentic high-resolution images. Real-ESRGAN further incorporates a second-order degradation model, a decision made through empirical considerations to strike a balance between simplicity and effectiveness in high-order degradation modeling. This second-order degradation process provides the model with better control and reflection of the influence of degradations on the Real-ESRNet during the super-resolution process, enhancing the model's adaptability and overall performance.

The generator network, denoted as Real-ESRNet, serves as a central component responsible for transforming low-resolution input images into high-resolution outputs. Within this network, high-order degradation modeling, sinc filters, and the second-order degradation model are effectively implemented. Real-ESRNet is designed not only to restore image sharpness but also to focus on the realistic restoration of textures in real-world samples. The model showcases an ability to restore intricate textures, such as those found in brick, mountain, and tree surfaces, while actively avoiding the introduction of unnatural textures observed in other methods. In visual comparisons with several leading methods, including ESRGAN, DAN, CDC, RealSR, and BSRGAN, Real-ESRGAN emerges as a superior performer. The model excels in both artifact removal and the restoration of intricate texture details in real-world images, demonstrating its versatility and effectiveness in addressing the complexities of image super-resolution. Real-ESRGAN, with its comprehensive methodology and innovative architecture, stands at the forefront of advancements in the field, providing a robust solution for generating high-quality, realistic, and visually appealing super-resolved images.

2.2.3 GAN in Visual Synthesis: Algorithms and Applications

Visual synthesis in computer science refers to the use of computational methods to generate visual content such as images or videos, for various tasks such as realistic rendering, procedural generation, image manipulation, creating immersive experiences in virtual environments, etc. Visual Synthesis methods have several applications in the fields of Entertainment and Media, Computer Aided Design, Simulation and training, Medical imaging and even Computer vision. One of the most effective and powerful methodologies for achieving visual synthesis is by using GANs or Generative Adversarial Networks. The paper 'Generative Adversarial Networks for Image and Video Synthesis: Algorithms and Applications' by Ming Yu Liu *et al.* [6] provides an overview of GANs, highlighting it's different variations along with a detailed explanation on algorithms and applications in visual synthesis.

The two most prominent variations of GAN that are widely used in visual synthesis tasks are unconditional and conditional GAN. In unconditional GANs, the generator is defined to be a mapping function that converts a noise signal z to a generated output $G(z)$, where $z \in Z$, and Z is usually a random Gaussian variable [6]. The discriminator tells apart real images x from the training data set D and fake images from G. On the other hand, in conditional GANs the generator takes an extra input y as the control signal, which could be another image, text, or a categorical label. The discriminator tells apart real from fake by utilizing the data in y [6]. Fig 2.9 illustrates the differences between unconditional and conditional GANs. In both scenarios, the joint utilization of the discriminator and authentic training data shapes an objective function for image synthesis. This data-centric approach in defining the objective function serves as a potent tool for tackling various computer vision tasks such as Image-to-Image translation, Image inpainting, and Neural rendering, among others.

Image-to-Image translation

Image translation is a visual synthesis technique where the main task is to map an image from one visual domain to another. An example of such could be creating photos from

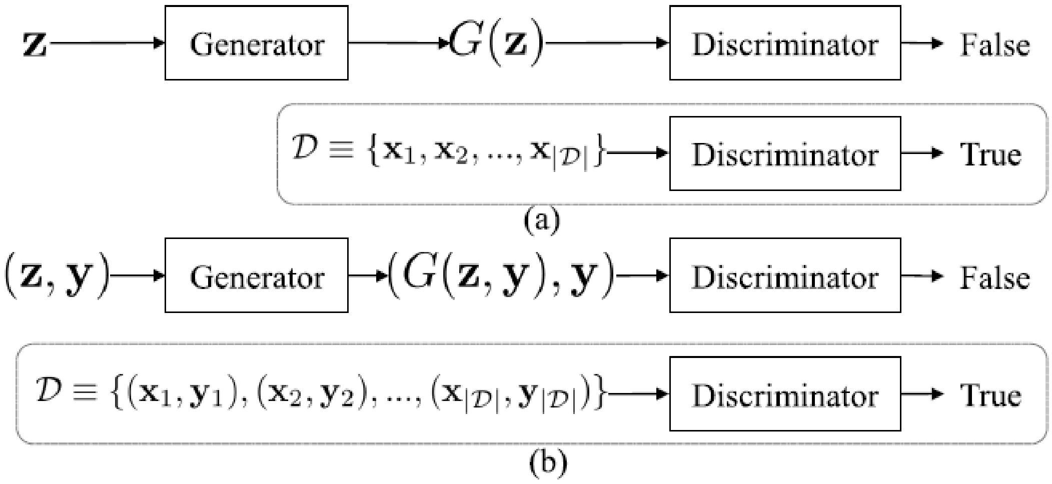


Figure 2.9: (a)Unconditional GANs (b)Conditional GANs

labels, converting a picture taken from one season to another, etc. It can make use of supervised learning, where sample pairs of training images are available, or unsupervised learning, where such training data is unavailable and there are only two independent sets of images. One of the widely used supervised image-to-image translation techniques is Pix2Pix GAN [7]. Pix2Pix is a conditional GAN framework that makes use of the pixelwise $l1$ loss between the ground truth and generated image. It also makes use of a patch-wise discriminator called PatchGAN that discriminates each local image patch rather than taking the image as whole. Pix2Pix architecture thus decreases the workload of the discriminator for improved performance because it requires far less model capacity to discriminate local patches than whole images.

In unsupervised image translation the prominent technique used is SPAGAN or Spatial Attention GAN [8], which is a variation of CycleGAN, an architecture that makes use of two generators and discriminators. Here, one generator-discriminator pair is used for translation from visual domain A to B, while the other is used to translate from B to A. One identifying feature of SPA-GAN is it's use of spatial attention maps that highlight the most discriminative regions between the source and target domains. The architecture of SPA-GAN is lightweight and does not introduce any additional attention networks or supervision. SPA-GAN's generator adopts an encoder-decoder architecture.

Image inpainting

Image inpainting, in computer-science is the method of filling in missing pixels of an image such that the result is realistic and indistinguishable from the ground truth. Inpainting algorithms can be used to remove distracting objects or retouch undesired regions in photos and can be further extended to other tasks, including image uncropping, stitching, rotation, recomposition, retargeting, compression, harmonization, super-resolution and more.

Traditional methods to perform image inpainting include Patch-match which copy background patches based on metrics such as euclidean distance to paste them onto missing regions. Recently deep learning networks such as GANs have shown incredible results in this process. GANs have enabled remarkably realistic and contextually aware reconstructions of missing regions.

Context encoders [9] were the first GAN based approach in image inpainting. It is an image inpainting convolutional neural network (CNN) that is trained to generate the contents of an arbitrary image region conditioned on its surroundings.They are typically trained in an unsupervised manner, using a dataset of images with masked regions. The goal of the training process is to learn a representation of the image that can be used to reconstruct the missing pixels in a way that is both semantically and visually plausible. Context encoders are effective for image inpainting because they are able to learn both the appearance and the semantics of visual structures. This means that they can not only fill in missing pixels with colors and textures that are consistent with the surrounding image, but they can also generate objects and scenes that make sense in the context of the image.

Another prominent strategy used is semantic inpainting, which is a technique that can handle varying mask shapes. The algorithm used here is called PEPSI or Parallel Extended decoder Path for Semantic Inpainting network [10]. This architecture is designed to address the limitations of traditional coarse-to-fine networks in terms of computational efficiency. It reduces convolution operations by employing a shared encoding network and a parallel decoding network featuring coarse and inpainting pathways. In this design, the coarse path generates an initial inpainting outcome, facilitating training of the encoding network to anticipate features for the Context Aggregation Module (CAM). Concurrently, the inpainting path refines features reconstructed by the CAM, yielding a superior-quality inpainting result.

Neural rendering

Neural rendering is a rapidly evolving field at the intersection of computer graphics and machine learning. It utilizes deep generative models to perform tasks such as creating hyper-realistic images from novel viewpoints and semantically manipulating the characteristics of an image. Neural rendering techniques often surpass the capabilities of traditional rendering techniques.

The most effective strategy when performing neural rendering using GANs is to use 3-D to 2-D transform as part of the network training.These operations enable to model the geometry and appearance of the scene in the feature space. The general pipeline of this methodology is given in Fig 2.10.

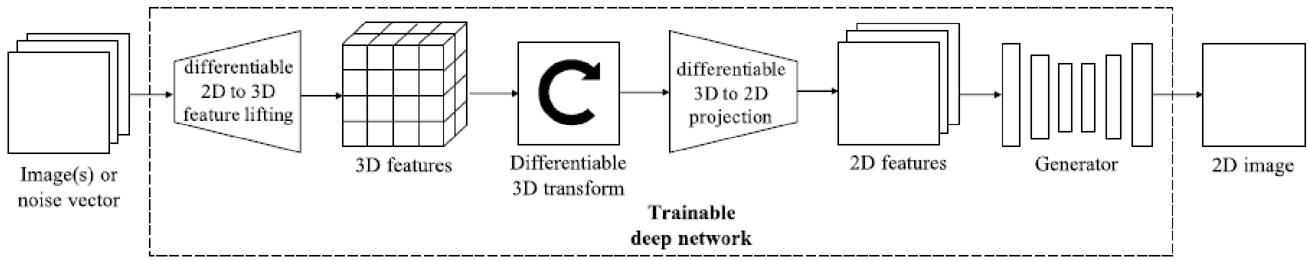


Figure 2.10: 3-D to 2-D transformation as part of neural rendering

There are many advantages of learning a 3-D representation and modeling the process of image projection and formation such as the ability to reason in 3-D, control the pose, and produce a series of consistent views of a scene. One of the prominent methodologies used here is HoloGAN or Holographic GAN [11].

HoloGAN is a cutting-edge neural rendering technique that pushes the boundaries of unsupervised 3D reconstruction from natural images. It leverages the power of deep learning to extract implicit scene representations directly from photographs, paving the way for novel view synthesis and scene understanding applications using disentangled representations. The working of HoloGAN is based on unconditional GANs. Here, a strong learning bias about the 3D world is introduced into the generator architecture which learns 3D representations from 2D images without labels. HoloGAN specifically generates images by learning 3D representations in the form of a voxel grid, and renders it realistically such that it fools the discriminator. Perspective control is achieved by directly applying rigid-body 3D transformations to the learnt 3D features. In other words, the images created by the generator are a perspective-dependent mapping from a learnt 3D representation to the 2D image space. This is different from other GANs which learn to map a noise vector z directly to 2D features to generate images.

Another hybrid strategy for neural rendering is EG3D or Efficient Geometry-aware 3D Generative Adversarial Network [12]. It is a tri-plane-based 3D GAN framework that enables high-quality geometry-aware image synthesis. It is hybrid in the sense that it combines both explicit and implicit representations of 3D geometry. This means it uses a mesh (explicit) to define the overall shape and an implicit function (implicit) to capture finer details. The 3D GAN framework consists of various components, including a pose-conditioned StyleGAN2-based feature generator coupled with a mapping network, a tri-plane 3D representation paired with a lightweight feature decoder, a neural volume renderer, a super-resolution module, and a pose-conditioned StyleGAN2 discriminator featuring dual discrimination [12]. This structure effectively separates feature generation from neural rendering, allowing for the utilization of a potent StyleGAN2 generator for broad 3D scene generalization. Furthermore, the streamlined 3D tri-plane representation proves both expressive and efficient, facilitating real-time high-quality 3D-aware view synthesis.

2.2.4 MWDCNN

Three integral components of MWDCNN[13] make up the methodology, each actively contributing to the comprehensive framework as shown in Figure 2.11. The first component—Dynamic Convolutional Block (DCB)—utilizes a dynamic convolution mechanism; its primary function is the dynamic adjustment of multiple convolutions' parameters as shown in Figure 2.12. This strategic tweaking aims at achieving an equilibrium: on one hand it optimizes denoising performance – on the other, it ensures efficient management of computational costs. Strategically integrating discriminative learning with signal processing techniques, specifically wavelet transformation, constitutes the second component: the Wavelet Transform and Enhancement Block (WEB). This amalgamation endeavors to adeptly quash noise–a critical aspect in facilitating intricate detail recovery within image denoising processes. Ultimately refining features derived from preceding stages hinges on a pivotal role played by our third element–the Residual Block (RB). Enhancing denoising effects and contributing to the reconstruction of clean images: this is the primary function of our refinement process, achieved through improved residual dense architectures. They have designed a robust and versatile framework–a structured integration of these components–to rigorously address denoising challenges while maintaining adherence to academic standards.

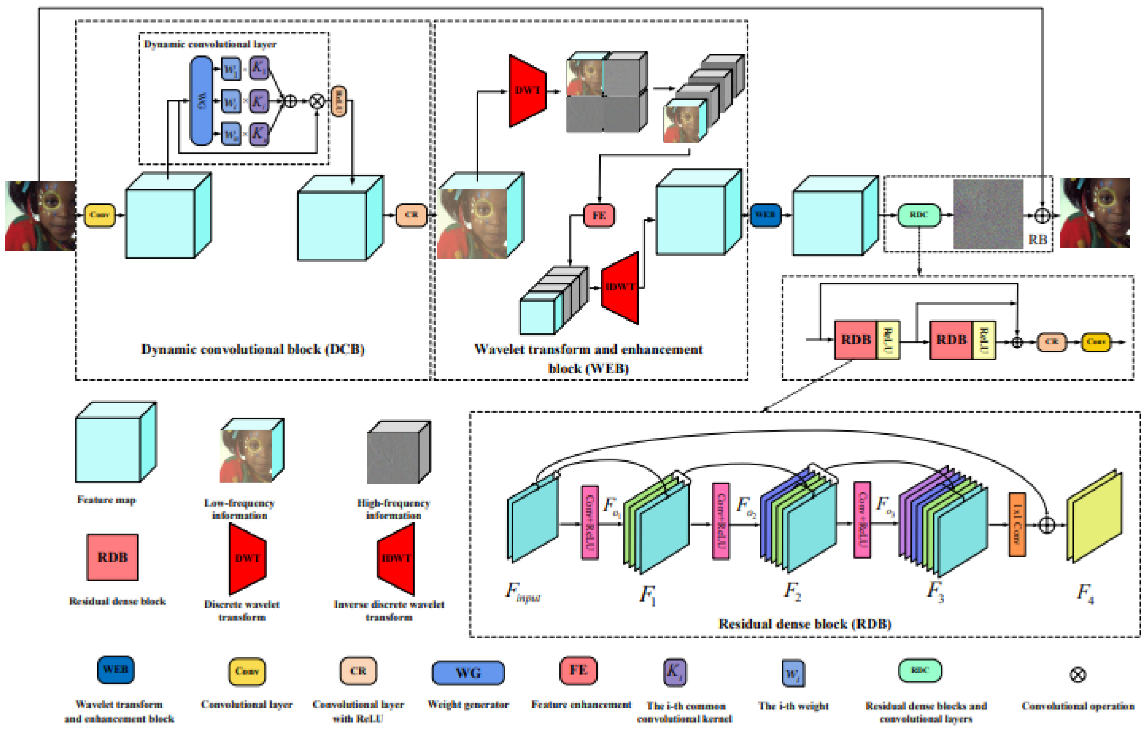


Figure 2.11: MWDCNN Network Architecture

The Dynamic Convolution Block (DCB) comprises five layers: a convolutional layer, a 3-layer dynamic convolution, and another convolutional layer as shown in Figure 2.12. The DCB initiates by employing weighted summation (WG); this allows it to procure four weights–guiding four parallel convolutional kernels in an adjusted fashion based on weight distribution. Following that step – the outcomes of this weighted-convolution operation serve as inputs for a subsequent 5×5 convoluted-layer process. The third step involves the linear fusion of: (i) the output from DCB's initial convolutional layer, and (ii) the result of dynamic convolution's second operation; this amalgamation incorporates bespoke parameters–specifically tailored for diverse noisy images. Two stacked wavelet

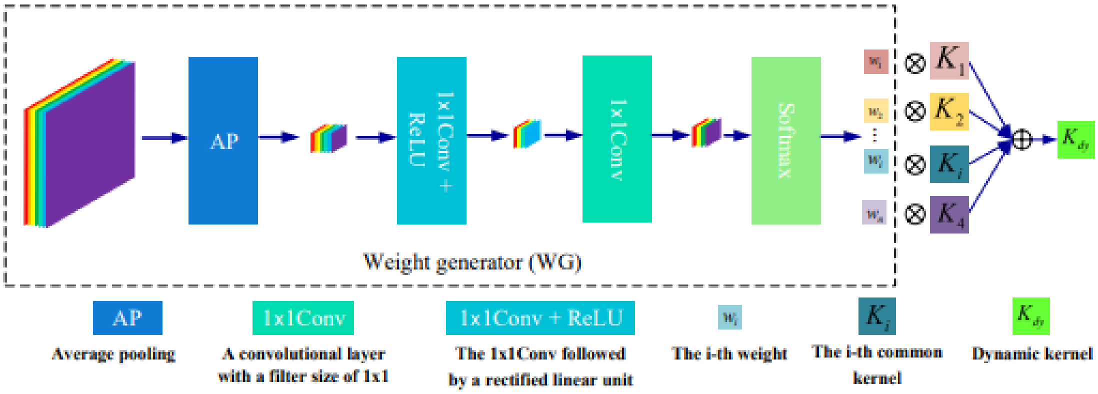


Figure 2.12: DCB Weight Generator

transform and enhancement blocks structure the Wavelet Transform and Enhancement Block (WEB), each block comprising four layers. First, it applies a Discrete Wavelet Transform (DWT) to convert linearly constructed data into four distinct frequency features. Next, leveraging a structural network guides signal processing techniques through the Feature Enhancement (FE) mechanism that also operates as a 4-layer Residual Dense Block (RDB). The process endeavors to quell noise and salvage complex details in the image denoising procedure. In accomplishing this, it culminates with an application of the Inverse Discrete Wavelet Transform (IDWT) – a step that reconverts frequency features into their original form: linear structural information.

The implementation employs a 10-layer enhanced residual dense architecture for The Residual Block (RB) as shown in Figure 2.13. Two combinations of 4-layer Residual Dense Blocks (RDBs) and Rectified Linear Units (ReLU), serve as the initial components to intensify feature refinement in image denoising. Addressing long-term dependency concerns, it utilizes two residual learning operations: one each from every RDB- ReLU couplet; another across all three— RDB, ReLU, and WEB — thereby enhancing memory abilities throughout shallow-to-deep layers during the process of image denoising. A convolutional layer, followed by ReLU, actively prevents over-enhancement and refines these features. In the final step: a second convolutional layer reconstructs noisy mappings; simultaneously–through a residual learning operation–it uses obtained noisy mappings to construct clean images from provided noisy images. Contents

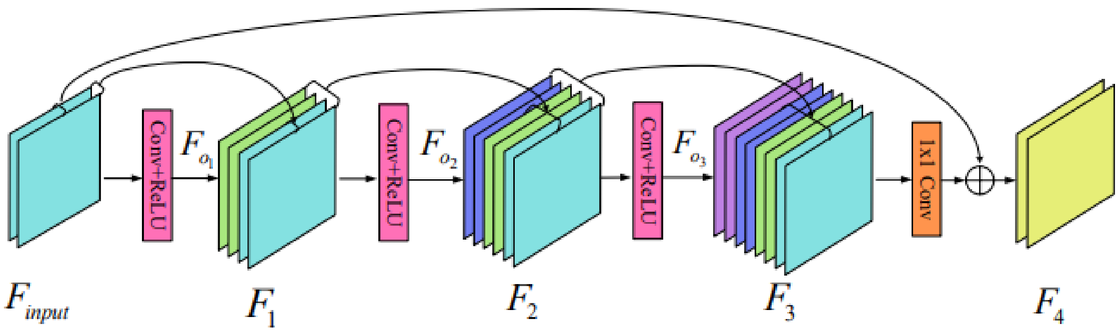


Figure 2.13: Residual Block

2.2.5 Smart brain tumor diagnosis system utilizing DCNN

In the study "Smart brain tumor diagnosis system utilizing deep convolutional neural networks "[2] , a brain tumor is diagnosed using a contemporary method by applying deep learning techniques. For the accurate and trustworthy evaluation of brain tumors using MRI, the system makes use of CNN and the EfficientNetV2 model, which has been optimized using the Ranger optimizer. This work explores the application of a CNN-based algorithm to help radiologists and doctors distinguish between cancerous and normal tissues on high-definition MRI scans with many gray levels. Pre-processing, data production, CNN-based tumor detection, and diagnosis are all included in the suggested system. The work uses contemporary optimizers and exper- imental evaluation of various CNN designs to determine which is best suited for brain tumor diagnosis. According to the study, a very high test sensitivity (¿99for the detection of the main tumoral entities, which include pituitary tumors, gliomas, and meningiomas. Convolutional Neural Networks (CNNs) and the EfficientNetV2 model, enhanced with the Ranger optimizer and extensive pre-processing, are used in the paper "Smart brain tumor diagnosis system utilizing deep convolutional neural networks" to achieve accurate and dependable brain tumor classification based on MRI images. Pre-processing, data production, CNN framework for tumor identification, and diagnosis are the four primary phases of the technique. In the pre-processing stage, the image datasets are cropped and filtered to remove image impairments such as noise, blur, and brightness. The next stage involves data augmentation and balancing the dataset during training. In the CNN framework stage, the EfficientNetV2 model is used for feature extraction, and the Ranger optimizer is applied to improve the learning stability and capability of processing images with varying quality. Finally, in the diagnosis stage, the system classifies the tumor type based on the extracted features. Through experimental comparisons with different CNN architectures and contempo- rary optimizers, the efficacy of the proposed method is assessed, revealing the superiority of the EfficientNetV2 + Ranger model in achieving high test accuracy rates. Specifically, the model achieves 99meningioma, glioma, and pituitary tumors. The technique offers a dependable and ef- fective method of tumor categorization based on MRI imaging data, advancing AI-based diagnostic systems for brain cancers.

2.3 Summary and Gaps Identified

Paper Title	**Advantages**	**Disadvantages**
Single-image super resolution of remote sensing images with real-world degradation modeling, Remote Sensing, vol. 14, no. 12, p. 2895, 2022, J. Zhang et al	1. Residual learning to learn the residual mapping between the LR and HR images which improves the reconstruction quality 2.Balanced attention modules (BAMs) to selectively enhance important features and suppress irrelevant ones	1.It requires a large amount of training data to achieve good performance. 2.This method may be computationally expensive.
"Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure in Synthetic Data," Proceedings of the IEEE/CVF Methodology International Conference on Computer Vision, 2021, X. Wang	1. It offers a flexible and realistic degradation process. 2. effective in removing artifacts and restoring texture details in real-world images.	1. It may produce twisted lines and unpleasant artifacts. 2.It struggles with out of distribution complex degradations.

Paper Title	Advantages	Disadvantages
Generative Adversarial Networks for Visual Synthesis: Algorithms and Applications, PROCEEDINGS OF THE IEEE,Vol. 109, No. 5, May 2021—,MING-YU LIU et al.	1. GAN based methods show exceptional accuracy in generating content 2. Highly realistic image generation	1. Difficulty in training 2. Limited control
Multi-stage image denoising with the wavelet transform. Pattern Recognition, vol. 134, p. 109050, 2023.	1. Adaptability 2. Tradeoff between denoising and computational cost, 3. Improved denoising	1. It Lacks Specific Wavelet transform, 2. Complexity and Computational Requirement
Smart brain tumor diagnosis system utilizing deep convolutional neural networks, Multimedia Tools and Applications,pp.1-27,2023, Springer,Y. Anagun	1. The model was trained using the cutting-edge optimizer Ranger, which improved the stability and convergence of the network and reduced the variance in the process.	2. Training deep learning models can be computationally intensive and require specialized hardware, such as GPUs, to achieve reasonable training times.

Table 2.1: Summary of Scientific Papers

2.3.1 Gaps Identified

1. **Real-world degradation modeling**: SISR methods were primarily developed and assessed on synthetic datasets, emphasizing the necessity for more authentic degradation models that accurately capture the complexities seen in real-world images.

2. **Generalization to diverse datasets**: Achieving consistent performance across a variety of datasets and image content proved challenging. Some techniques excelled with specific image characteristics but struggled to generalize across a broader

spectrum.

3. **Handling different scaling factors**: Despite a focus on upscaling by specific factors (e.g., 2x, 4x), adapting to various scaling factors, particularly arbitrary ones, remained a persistent challenge.

4. **Addressing large upscaling factors**: Dealing with substantial upscaling factors (e.g., 8x or more) in super-resolving images posed difficulties, prompting a need for methods capable of effectively managing such scenarios.

5. **Computational efficiency**: Certain SISR techniques were computationally demanding, limiting their practical application in real-time settings. The challenge was to develop algorithms that are efficient without compromising quality.

Chapter 3

Hardware and Software Requirements

3.0.1 Hardware Requirements

- RAM - 8 GB (minimum)
- Processor - Intel i5(7th gen) (minimum)
- System type - 64 bit operating system
- System architecture - x86 based
- GPU VRAM - 4GB

3.0.2 Software Requirements

- Tensor Flow
- Google Colab
- Python
- Operating System: Windows/Linux
- Visual Studio
- OpenCV
- Pytorch
- NumPy
- HTML
- CSS
- JavaScript

3.1 Functional Requirements (Numbered List/ Description in Use Case Model)

The different functional requirements of this project are:

1. User authentication and authorization:
    - The system should support user registration and login functionality
    - Different user roles such as regular users, administrators, etc should have appropriate access levels.
2. Input image upload
    - Users should be able to upload single images for super-resolution processing.
    - Supported image formats and size limits should be specified.
3. Super-resolution processing:
    - The system should utilize a GAN-based method for single-image super-resolution.
    - Specify the desired level of upscaling (e.g., 2x, 4x) based on GAN model capabilities.
4. Real-Time Processing:
    - The system should provide near-real-time super-resolution processing for uploaded images.
    - Specify the expected processing time for different image sizes.
5. Output Image Download:
    - Users should be able to download the super-resolved images.
    - Output images should be available in standard formats such as JPEG, PNG, etc.
6. Quality Metrics:

- The system should provide quantitative measures of image quality, such as PSNR (Peak Signal-to-Noise Ratio) or SSIM (Structural Similarity Index), for both input and output images.

7. Web Interface:
   - The web app should have an intuitive and user-friendly interface.
   - Ensure compatibility with different web browsers.
8. Error Handling:
   - Implement robust error handling for various scenarios, such as incorrect file formats, server errors, or network issues.
9. Security:
   - Ensure secure transmission and storage of user-uploaded images.
   - Protect against common web vulnerabilities, such as SQL injection and cross-site scripting.
10. Scalability
    - The system should be scalable and must handle concurrent requests.
    - Load balancing can be implemented to optimize the image processing pipeline.
11. Documentation:
    - Provide comprehensive documentation for users and developers, including a user manual and API documentation.
12. Testing:
    - Define a testing strategy, including unit testing, integration testing, and user acceptance testing.
    - Specify the criteria for evaluating the success of super-resolution processing.

Chapter 4

System Architecture

4.1 System Overview

The single image super resolution project is a web app that allows users to upscale as well as store their images in a unified platform. The application makes use of a deep-learning based neural network in order to upscale images. It also utilises a sophisticated database management system for functions such as storage and retrieval. The neural network that performs the upscaling process is trained with the help of two different modules:

1. Degradation Module

2. Adversarial training Module

The architecture diagram that describes the same is given in Fig 4.1.

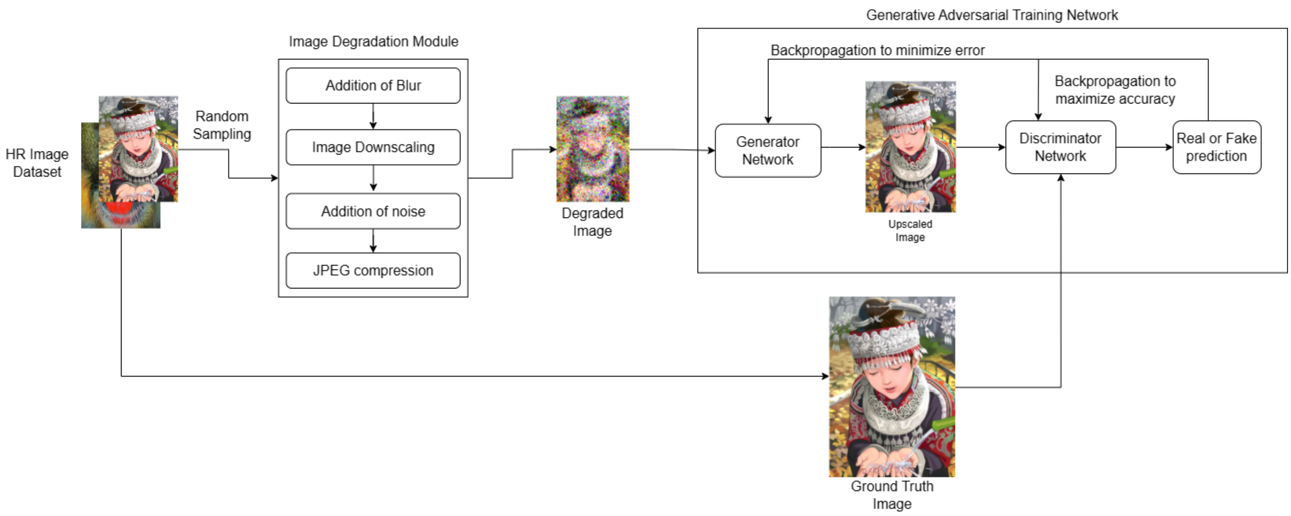


Figure 4.1: Architecture diagram of GAN based image super resolution model

4.1.1 Degradation Module

The primary function of the degradation module is to create a dataset of degraded images to train the super-resolution GAN network to achieve realistic results. The module applies four different degradations on a random image sampled from the input high-resolution dataset:

- Addition of blur
- Image downscaling
- Addition of noise
- JPEG compression

4.1.2 Generative Adversarial training Module

Images from the dataset obtained are then used as input to the generative adversarial network which performs the upscaling process by iteratively optimizing the generator and discriminator losses by comparing the upscaled degraded images with the original ground truth images, until a point is reached where the discriminator is unable to discern the difference between these pair.

4.1.3 Web Application

This trained generator is then integrated into a user friendly web app that is capable of providing:

- Registration for new users
- Login functionality for existing users
- Ability to upload images for upscaling process
- Ability to download images once they are enhanced
- Storage and retrieval of images as needed

4.2 Architectural Design

4.2.1 Use case diagram

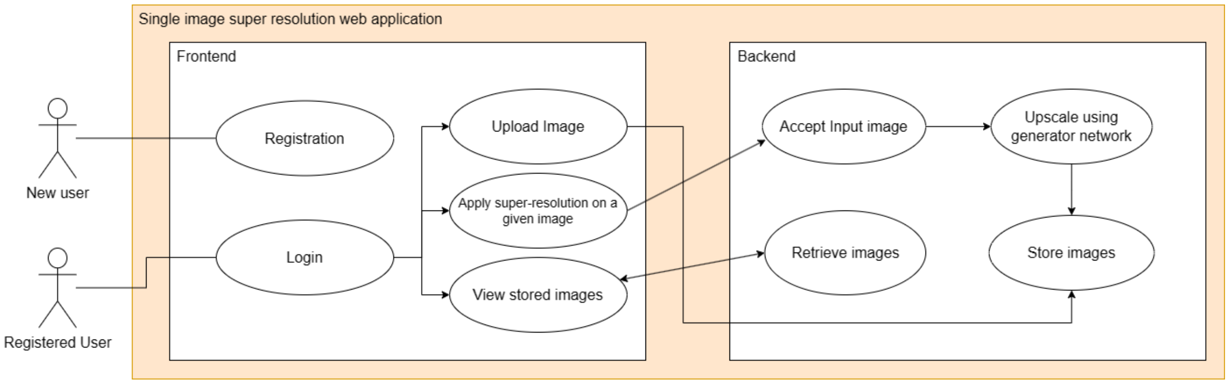


Figure 4.2: UML Use Case Diagram

4.2.2 Sequence Diagram

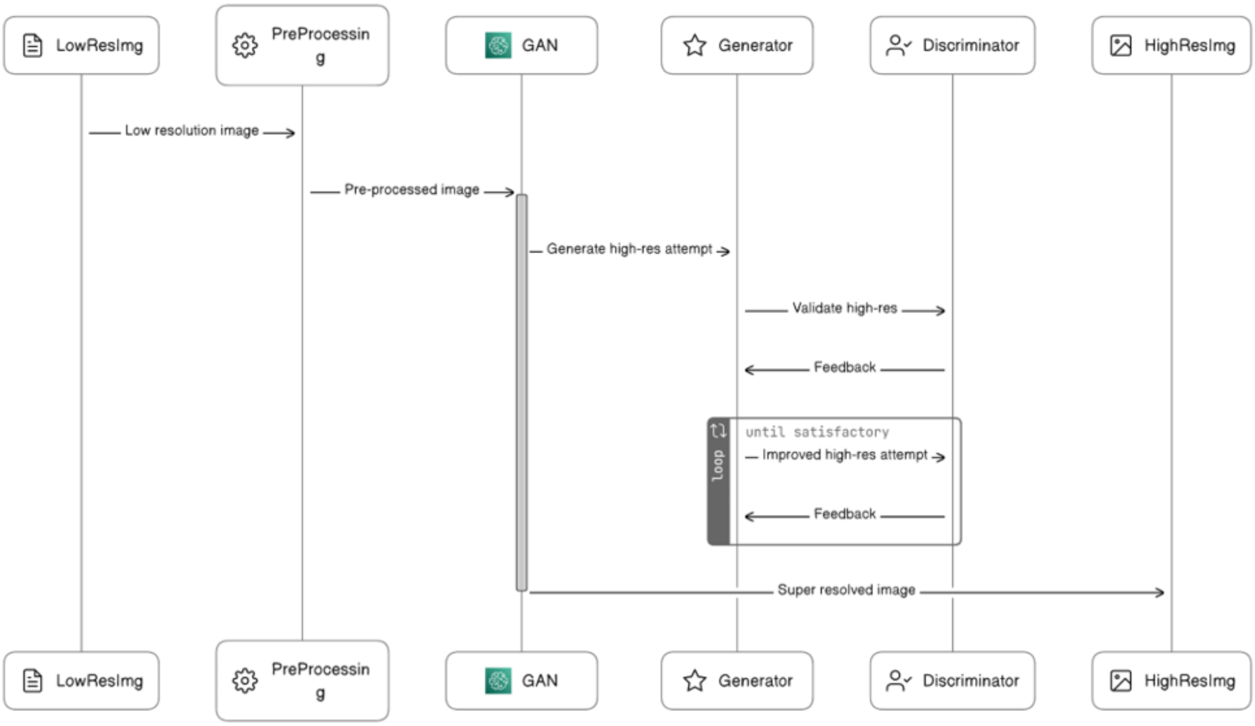


Figure 4.3: Sequence Diagram

4.3 Module Division

The different modules involved in this project are:

- Dataset collection: One or more high-resolution image datasets are collected from various sources over the internet. Module diagram is given in Fig 4.4.

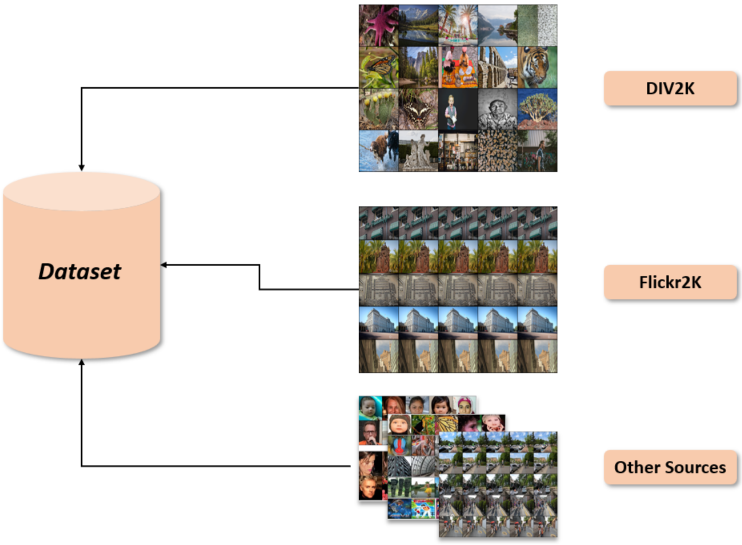


Figure 4.4: Dataset Collection Module

- Development of Image degradation model: The degradation model applies four different quality degradation filters (blur, downscaling, noise, compression) on each image in the high-resolution dataset, obtained in the previous step to create a new dataset consisting of only degraded images. Degradation module is depicted in Fig 4.5.

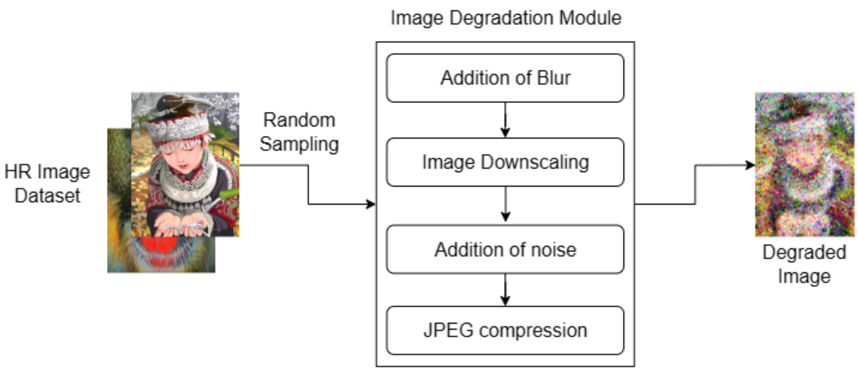


Figure 4.5: Image Degradation Module

- GAN based super resolution model: This module is responsible for the actual image upscaling process. It employs a unique adversarial training setup in order to train a generator neural network to perform this task.

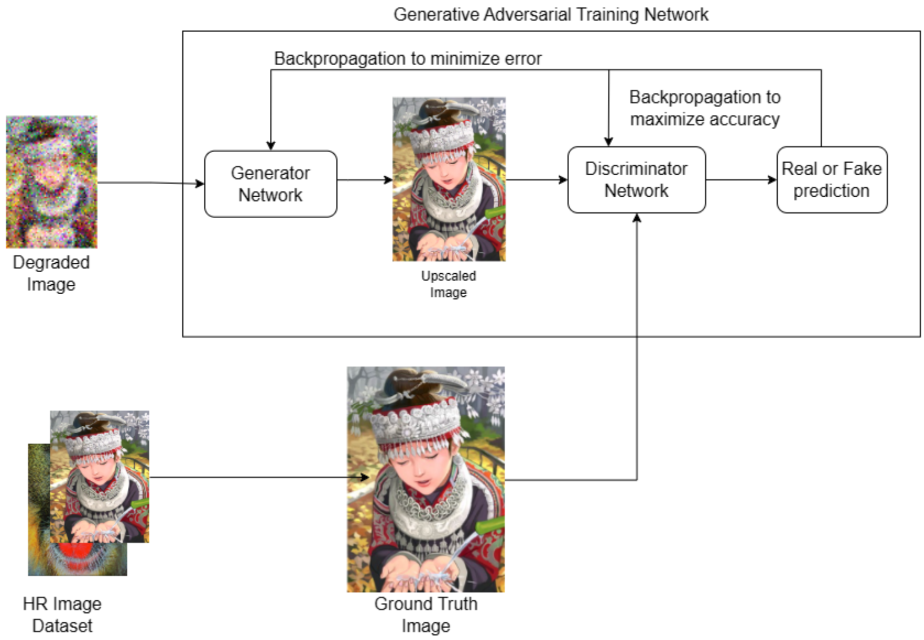


Figure 4.6: GAN based Super-Resolution Module

- Model testing and training: Once the GAN based model is developed, various tests and optimization strategies are employed in order to ensure stable performance.
- Web app integration: The trained generator obtained in the previous step is integrated into a user friendly web application that provides functionalities of registration, login, upload, upscale, download, storage, retrieval, etc.

4.3.1 Work breakdown

Member Name	**Modules assigned**
Abel John Mathew	Development of Generator and EfficientNet module
Ashwin Saji	Website backend, Model validation and integration.
Athif Ahamed	Dataset collection and degradation, Website Frontend
Didin Shibu	Development of Discriminator and EfficientNet module

Table 4.1: Modules assigned to each team member

4.4 Work Schedule - Gantt Chart

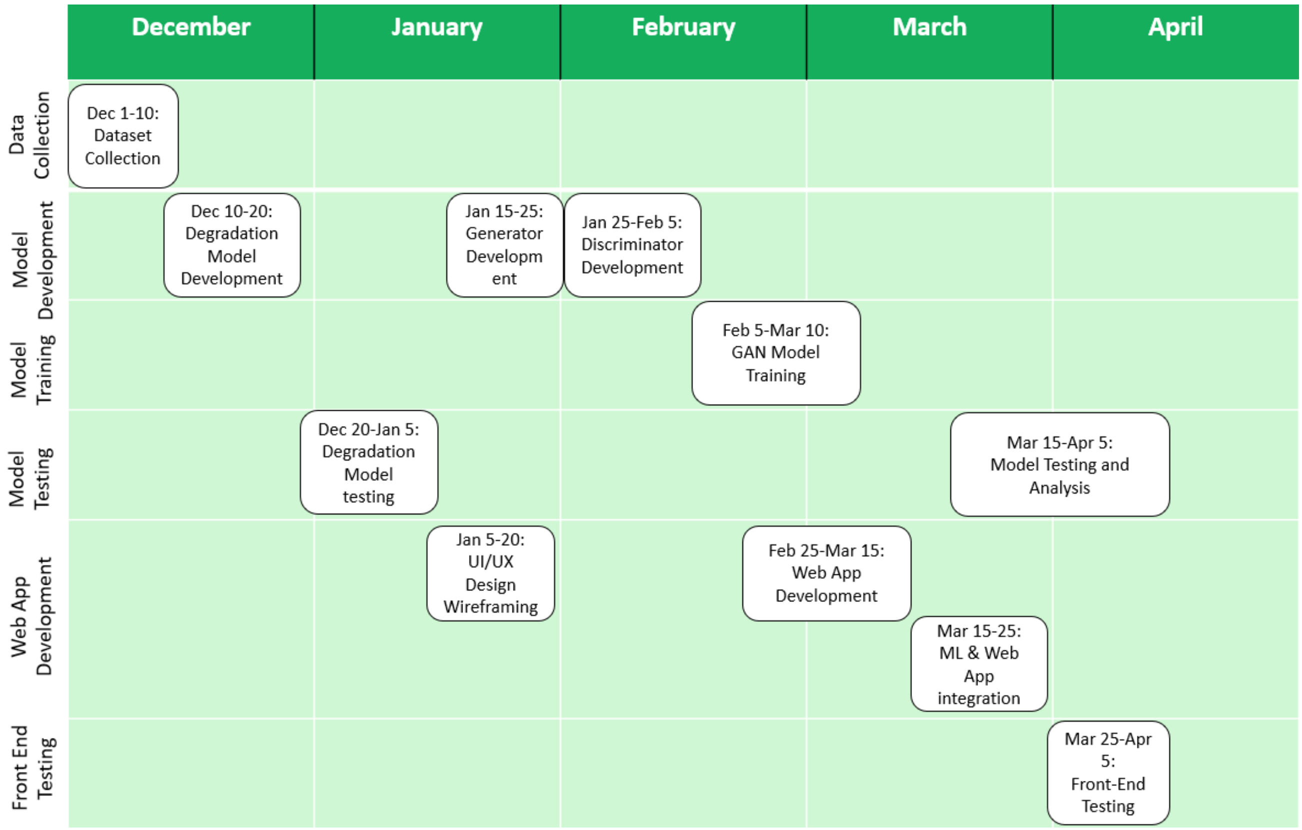


Figure 4.7: Gantt chart

4.5 Conclusion

In conclusion, this chapter provides a systematic overview of the architecture of our Single Image Super resolution application with the help of module-wise diagrams, UML diagrams and Sequence diagrams. The project modules are clearly specified along with each team member's individual module assignment. The project's roadmap is also portrayed using a gantt chart.

Chapter 5

System Implementation

5.1 Datasets Identified

The GAN model is trained using images from various popular super-resolution datasets such as

- DIV2K [14] : It consists of 800 2K images collected from various sources across the internet.
- Flickr2K [15]: It consists of 2,650 images with 2K resolution collected from Flickr, a popular image and video hosting service.
- OutdoorSceneTraining (OST) [16]: It is an outdoor scene dataset with 300 test images of outdoor scenes, and a training set of 7 categories of images with rich textures.

5.2 Proposed Methodology/Algorithms

5.2.1 Image degradation

Degradation module refers to a component or process within a super-resolution model that simulates the degradation or deterioration that occurs in images captured by real world imaging systems. This degradation typically includes factors such as noise, blur, and low resolution, which are common in images captured by cameras or other imaging devices. In the degradation module the original image goes through multiple processes that includes:

- Application of Blur

  The Blur Kernel Estimation is done using the KernelGAN [17] framework. It is used for finding realistic blur kernels with the help of a single input image. The

KernelGAN consists of a 7 layer generator which is trained on an input image and it is trained to successfully produce a blur kernel.

- Image downscaling The blurred image is downscaled to 256 x 256 and this is considered as the high resolution image.The downscaling of the image is used by the Pillow library, particularly the resize function which utilises bicubic downscaling to downscale the image.

- Noise addition

  Noise addition is done using poison and gausian noise.Gaussian noise is a type of statistical noise that follows a Gaussian (normal) distribution.Poisson noise follows a poisson distribution. Commonly found as camera sensor noise in low light conditions.

- JPEG Compression

  The image size is reduced using JPEG compression. The image size is reduced for many reasons like reduced memory and bandwidth consumption, faster processing and noise reduction.

5.2.2 GAN model

1. **Generator**

   The generator model is designed to take a low-resolution image as input and generate a high-resolution version of that image as shown in Figure 5.1. The architecture of the generator is based on the idea of residual learning and skip connections, which have been shown to improve the performance of deep neural networks in various tasks, including super-resolution.

   The skip connections and residual blocks in the generator architecture help to propagate low-level information from the input image to the output image, while the upscaling blocks increase the spatial resolution of the feature maps. This architecture allows the generator to learn a mapping from low-resolution to high-resolution images while preserving important details and features. A detailed architecture diagram of the generator is given in Figure 5.1.

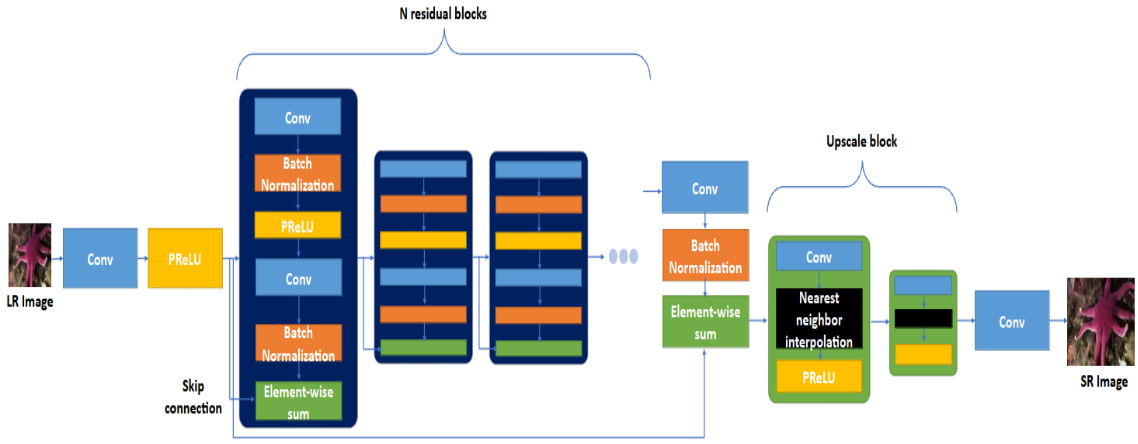


Figure 5.1: Generator Architecture

2. **Discriminator**

The discriminator model as shown in Figure 5.2 is designed to distinguish between real high-resolution images and fake (generated) high-resolution images produced by the generator model. The discriminator plays a crucial role in the training process of the GAN by providing feedback to the generator, helping it to generate more realistic and sharper images.

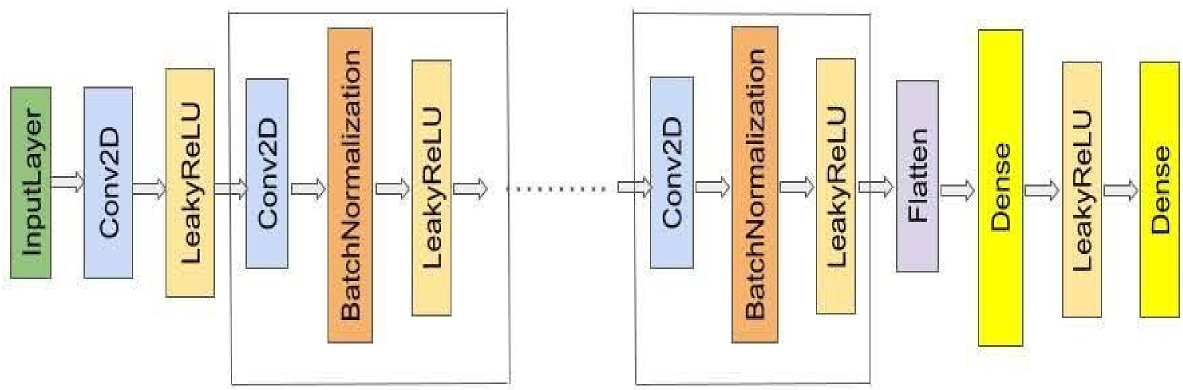


Figure 5.2: Discriminator Architecture

The discriminator model is trained in an adversarial manner, where it tries to maximize the probability of correctly classifying real images as real and generated images as fake. During training, the discriminator is first trained on batches of real and fake images, and then the generator is trained based on the discriminator's loss function called adversarial loss, to produce realistic outputs. The discriminator's architecture diagram is described in Figure 5.2.

3. **EfficientNet Feature Extractor**

   The EfficientNet feature extractor as shown in Figure 5.3 is a deep convolutional neural network architecture that is used to extract high-level features from the input images. These features are then utilized to calculate the content loss (also known as the perceptual loss) during the training of the GAN model.

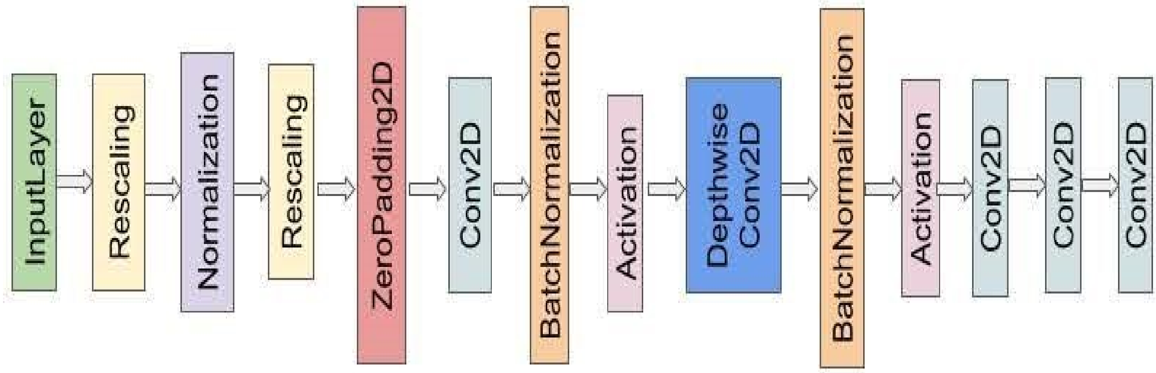


Figure 5.3: EfficientNet Feature Extractor Architecture

During the training process, the EfficientNet feature extractor is used to extract features from both the generated high-resolution images (produced by the generator) and the real high-resolution images. These features are then used to calculate the content loss, which measures the difference between the feature representations of the generated and real images. In this implementation, the EfficientNetB7 model, which is one of the larger variants in the EfficientNet family, is used as the feature extractor. The EfficientNetB7 model has been pre-trained on the ImageNet dataset, which allows it to learn rich and meaningful feature representations from natural images.

By incorporating the EfficientNet feature extractor into the SRGAN model, the generator is encouraged to produce high-resolution images that not only fool the discriminator but also have feature representations that are similar to those of real high-resolution images. This guidance from the content loss helps the generator to generate sharper and more realistic high-resolution images, as the EfficientNet model has been trained to capture rich and discriminative features from natural images. The Efficientnet feature extractor's architecture diagram is shown in Figure 5.3.

5.3 User Interface Design

A wireframe for the project webpage was created and is shown from Figure 5.4 to 5.7

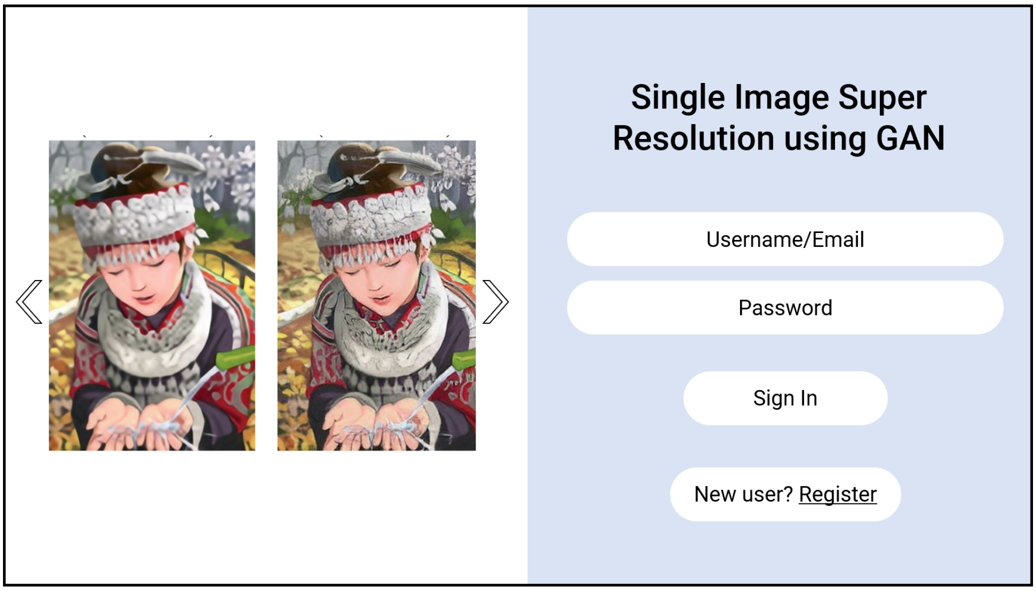


Figure 5.4: Website landing page

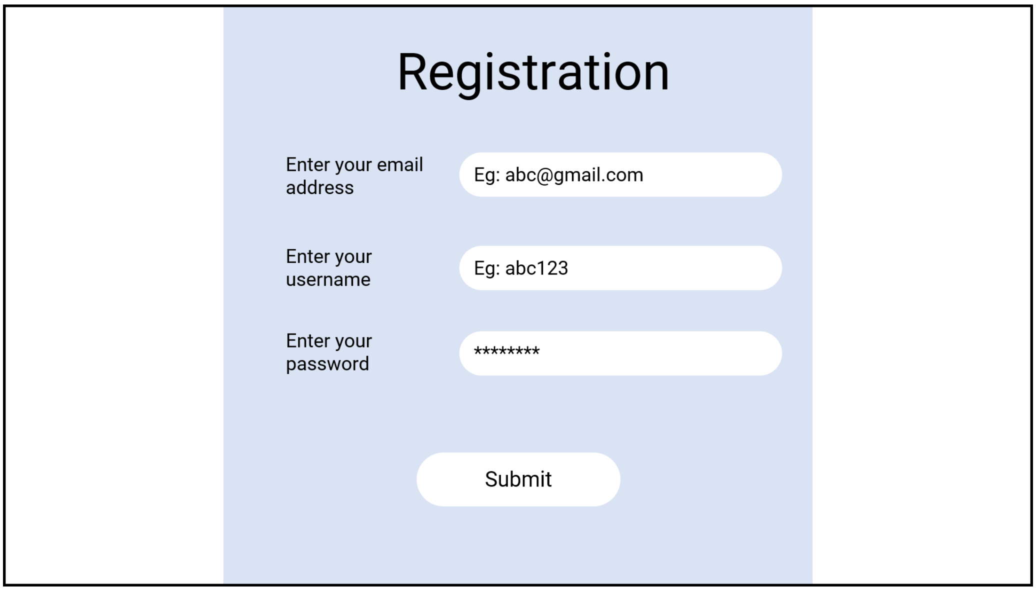


Figure 5.5: Website registration page

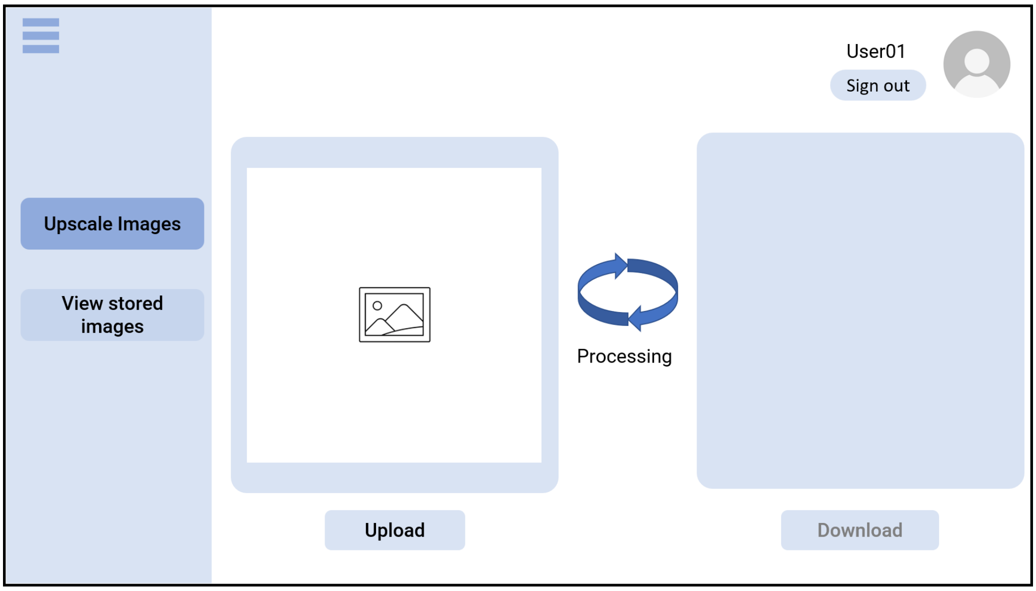


Figure 5.6: Website home page

Figure 5.7: Saved Images can be viewed in the website

5.4 Database Design

The work makes use of an SQL-based database that stores both the details of users registered into the service as well as various images uploaded by each user from their account for upscaling. The SQL database makes use of two tables to do so:-

- Users table : Which stores the username and password of each user. A unique id value is assigned to each user which is auto-incremented every time a new account is created.
- Images table : Which stores the details of images uploaded by each user in blob format. In this table, a foreign key user_id is established taking the primary key id from the users table. This enables to administrator to associate each image with a specific user so that they can be retrieved easily.

A detailed schema diagram depicting the database structure is shown in Fig.5.8.

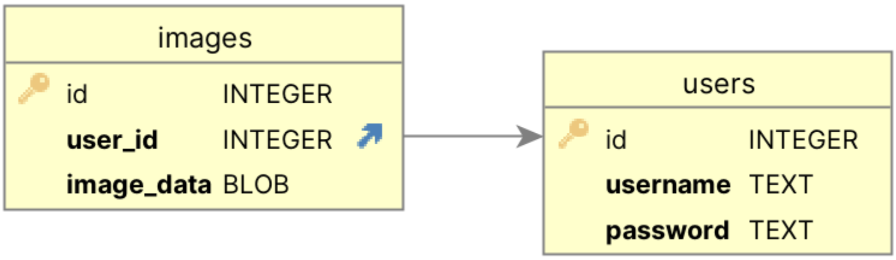


Figure 5.8: Database schema diagram

5.5 Description of Implementation Strategies

The super resolution GAN model is trained with the help of different modules such as the Degradation module, Generator, Discriminator, and Efficientnet feature extractor.

5.5.1 Implementation of the Degradation module

The different degradations are applied on a dataset of high-resolution images using python in the following order:

- Blur addition: Gaussian blur is applied using cv2.GaussianBlur() function using a standard deviation (sigma) value of 15 in both x and y directions.
- Downscaling: Downscaling is the process of reducing the spatial resolution of an image. Images here are downscaled using Pillow Image.resize() function, utilizing bicubic interpolation. This is used for making the hr and lr images used for training.
- Noise additon: Two different variants of noise called Gaussian and Poisson are added. Gaussian noise models the sensor noise present in digital images, following a normal distribution with a mean of zero and a specified standard deviation, while Poisson noise accounts for the shot noise inherent in the image acquisition process, following a Poisson distribution where the variance is equal to the intensity value of the pixel. The noise is added using using numpy and pillow functions. This gives more realistic noisy images.
- Image Compression: In real world scenarios, images are compressed for efficient storage. hence adding compression would increase the effectiveness of the dataset. Images are compressed using JPEG compression by using Pillow compression function.

5.5.2 Implementation of the Generator

The generator architecture follows a sub-pixel convolution model where the image is only up-scaled as the last step in it's pipeline. The generator first extracts high level features using 64 convolutional filters which are 9x9 in dimension using the Conv2D() function in Keras, Python followed by PReLU activation function to introduce non-linearlity in the output.

```
layers = Conv2D(64, (9,9), padding="same")(gen_ip)
layers = PReLU(shared_axes=[1,2])(layers)
```
(5.1)

This is followed by n-Residual blocks each consisting of 2D convolutions, PReLU activations and Batch Normalizations, which extract low-level features from the input image as shown in Fig 5.10.

```
res_model = Conv2D(64, (3,3), padding = "same")(ip)
res_model = BatchNormalization(momentum = 0.5)(res_model)
res_model = PReLU(shared_axes = [1,2])(res_model)

res_model = Conv2D(64, (3,3), padding = "same")(res_model)
res_model = BatchNormalization(momentum = 0.5)(res_model)
```
(5.2)

Image features are then only upscaled towards the end of the pipeline using upscale blocks that consist of 2D Convolutions with 256 filters followed by Upsampling2D and PReLU blocks as shown in Fig 5.11.

```
up_model = Conv2D(256, (3,3), padding="same")(ip)
up_model = UpSampling2D( size = 2 )(up_model)
up_model = PReLU(shared_axes=[1,2])(up_model)
```
(5.3)

5.5.3 Implementation of the Discriminator

The discriminator is a binary classifier that predicts whether the images given to it as fake (produced by the generator) or real (from the actual HR dataset). It attains a single scalar value for binary classification through sigmoid activation function after a series of convolution layers for feature extraction, flattening to convert multi-dimensional feature maps produced by the convolutional layers into a one-dimensional vector and Dense layers followed by LeakyReLU activation to calculate weighted sum of inputs.

5.5.4 Efficientnet feature extractor

The efficientnet feature extractor is a model pre-trained on the Image-net dataset that creates feature maps which are used to generate perceptual loss function. Here, EfficientnetB7 model is built first using tf.keras.applications.EfficientNetB7() function. Then an intermediate output from the 69th layer of efficientnet is extracted to generate the feature maps.

5.5.5 GAN training process

The GAN training process involves training the generator and discriminator in an adversarial manner. The training occurs by enumerating over batches of high resolution and low resolution images in every epoch. The generator first predicts upscaled images for a batch of low resolution images using the generator.predict_on_batch() method. These images are assigned a label 0 to indicate fake whereas the batch of high resolution images are assigned label 1 to indicate real. Here, the discriminator is trained first before training the generator using these batches of real and fake images and the discriminator's loss in successfully classifying them are compiled as d_loss_real and d_loss_gen respectively. The discriminator's cumulative loss is then found by averaging these two losses. The following is shown in Fig 5.12.

```
fake_imgs = generator.predict_on_batch(lr_imgs)

discriminator.trainable = True
d_loss_gen = discriminator.train_on_batch(fake_imgs, fake_label)
d_loss_real = discriminator.train_on_batch(hr_imgs, real_label)

discriminator.trainable = False

d_loss = 0.5 * np.add(d_loss_gen, d_loss_real)
```

(5.4)

The next step in training is predicting feature maps from high-resolution images using a pre-trained efficientnet feature extractor. These feature maps are then used to train the generator whose overall loss function consists of two different parts - adversarial loss and content loss, together called perceptual loss. Here adversarial loss is calculated each time using the validity score obtained from the discriminator model trained earlier and content loss is calculated as the mean squared error difference in the feature maps from the actual high-resolution images and generated high-resolution images. These two losses are compiled in the ratio 0.001:1 to obtain the total perceptual loss as shown in Fig 5.13.

```
image_features = efficientnet.predict(hr_imgs)

g_loss, _, _ = gan_model.train_on_batch([lr_imgs, hr_imgs], [real_label, image_features])
```

(5.5)

The aim in GAN training is to minimise these two losses so as to get the most perceptually accurate images.

5.6 Conclusion

In conclusion the system implementation chapter provides detailed explanation on the datasets used and explains the proposed methodology with the help of diagrams and code snippets. The chapter also highlights the wireframes created during user-interface prototyping and also describes the database design involved.

Chapter 6

Results and Discussions

6.1 Overview

After training the Super-Resolution GAN model for many epochs and saving the generator at regular intervals, the saved model was utilized in a web application enabling registered users to upload low-resolution images, upscale them using the trained generator, view the high-resolution output alongside the original, and download the upscaled image.

6.2 Testing

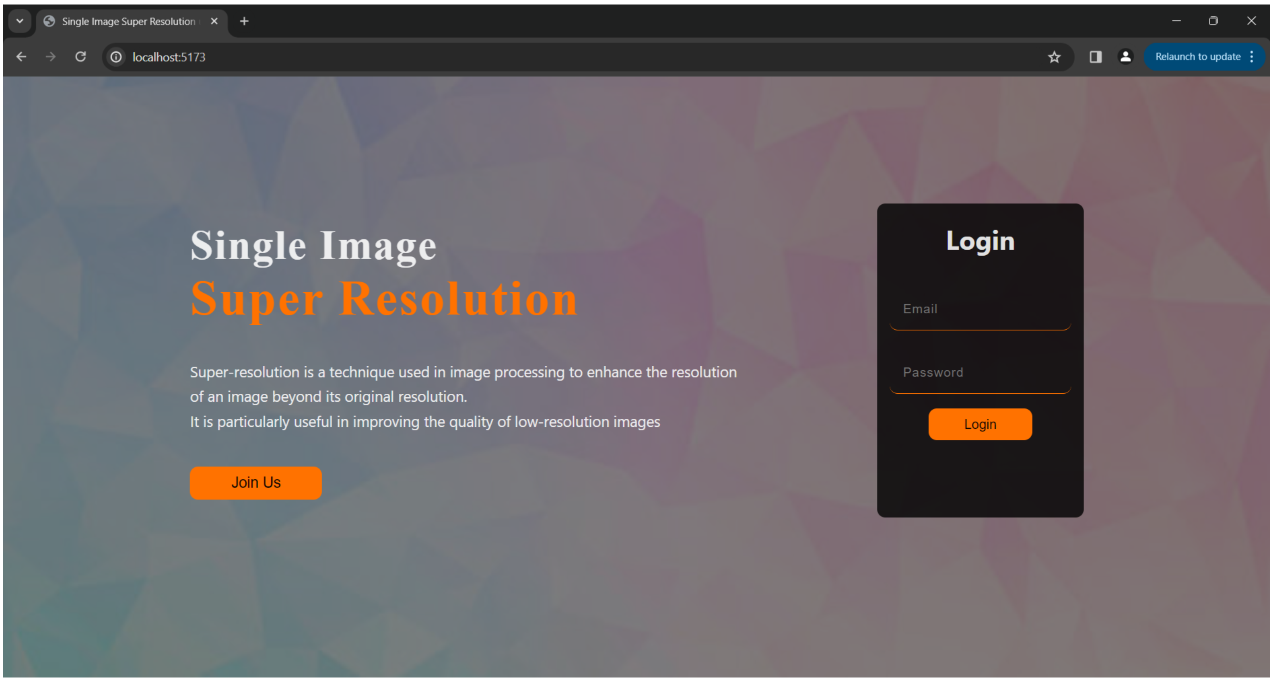


Figure 6.1: Login Page

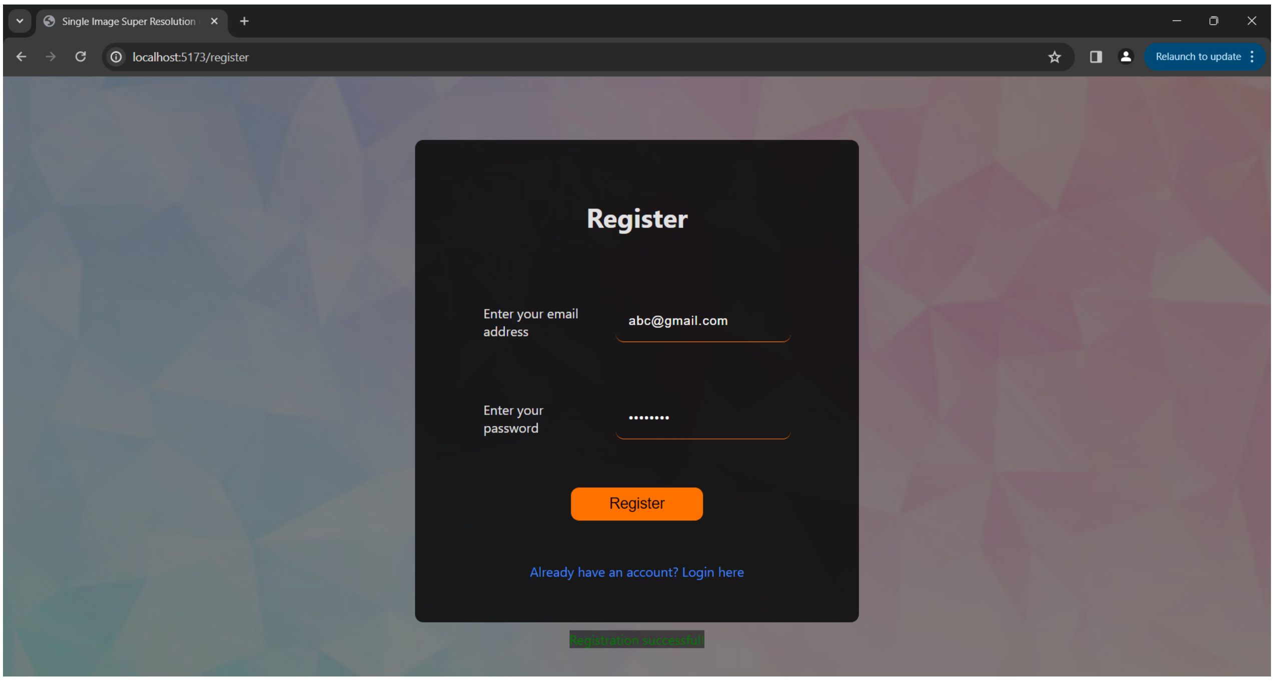


Figure 6.2: Registration Page

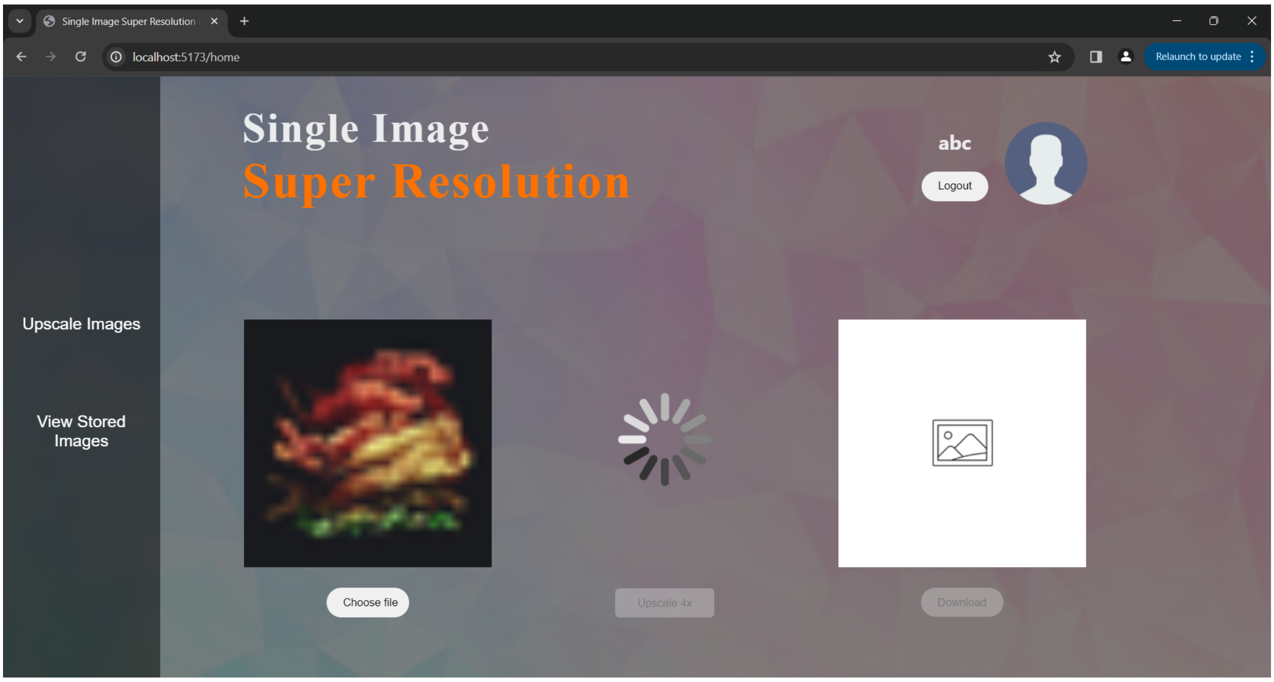


Figure 6.3: Home Page - Upscale Images

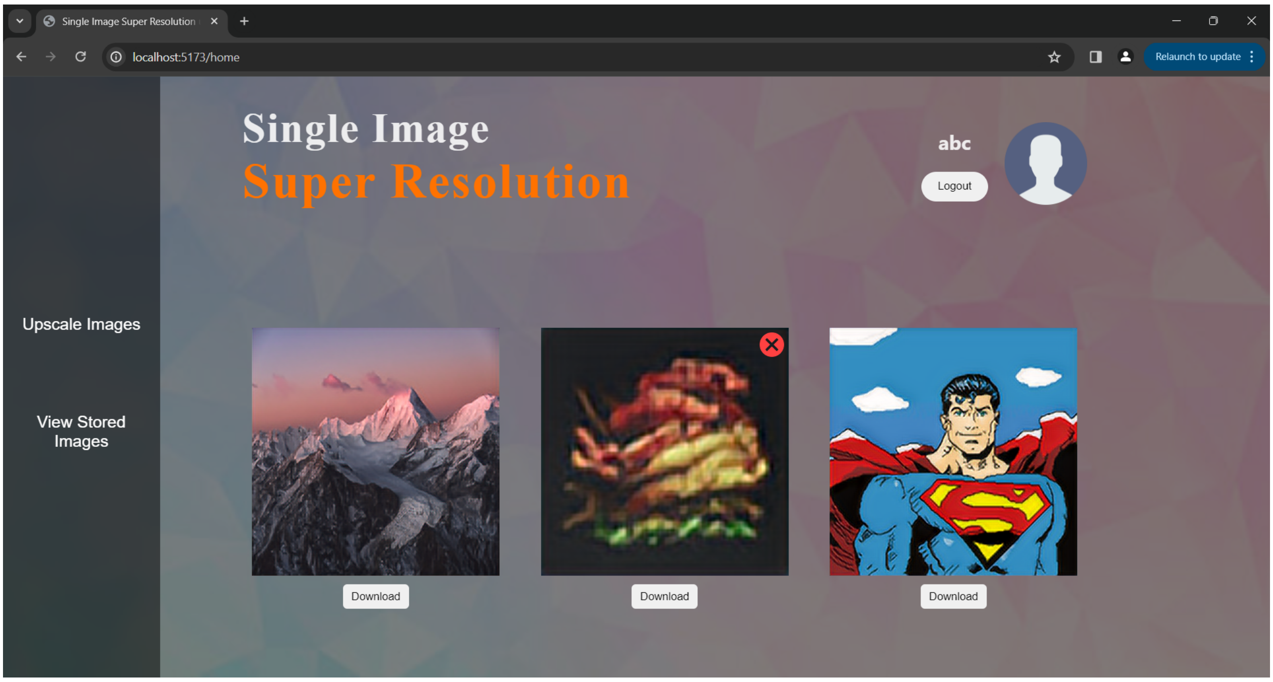


Figure 6.4: Home Page - View Stored images

6.3 Quantitative Results

The super resolved image was assessed in comparison to standard image upscaling methods like Nearest neigbhor, Bilinear Interpolation and Bicubic Interpolation. The images were assessed with standard methods like PSNR and SSIM as shown in Figure 6.1.

Metric	SR Value	NN Value	Bilinear Value	Bicubic Value
PSNR index ($-\infty$ to ∞)	23.6789	23.3524	24.1277	24.635
SSIM index (-1 to 1)	0.5945	0.5514	0.5847	0.6143

Table 6.1: Comparison of the proposed SR technique with other spatial upscaling techniques using standard metrics.

The scores for the SR model performs worse than the standard method, this is because it is shown that PSNR and SSIM aren't good at appropriately grading perceptual quality.

Using Perceptually appropriate image assessment methods which correctly analyses human perception in images. Using this we can see that the SR images are better as shown in Figure 6.2

Metric	SR Value	NN Value	Bilinear Value	Bicubic Value
DSS index (0 to 1)	0.8361	0.6813	0.7051	0.7587
FSIM index (0 to 1)	0.8202	0.6956	0.7581	0.785
HaarPSI index (0 to 1)	0.7143	0.5384	0.6894	0.7118
IW-SSIM index (0 to 1)	0.9089	0.8724	0.8569	0.8871
LPIPS (0 to 1)	0.3933	0.5303	0.4062	0.3956
MS-SSIM index (0 to 1)	0.924	0.9075	0.9051	0.922
PieAPP loss (0 to ∞)	1.4256	3.1009	3.4729	3.0959
VIFp index (0 to 1)	0.3016	0.2528	0.2728	0.2962
VSI index (0 to 1)	0.9459	0.9034	0.9325	0.9394

Table 6.2: Comparison of the proposed SR technique with other spatial upscaling techniques using metrics that consider human perceptual accuracy.

6.4 Graphical Analysis

VGG19 is the standard feature extractor used for Super Resolution tasks but it is more computationally heavy when compared to EfficientNet. The computational difference between VGG19 and EfficientNet can be seen by looking at the GMACs, GLOPs and Number of Parameter of each model as shown in Figure 6.5.

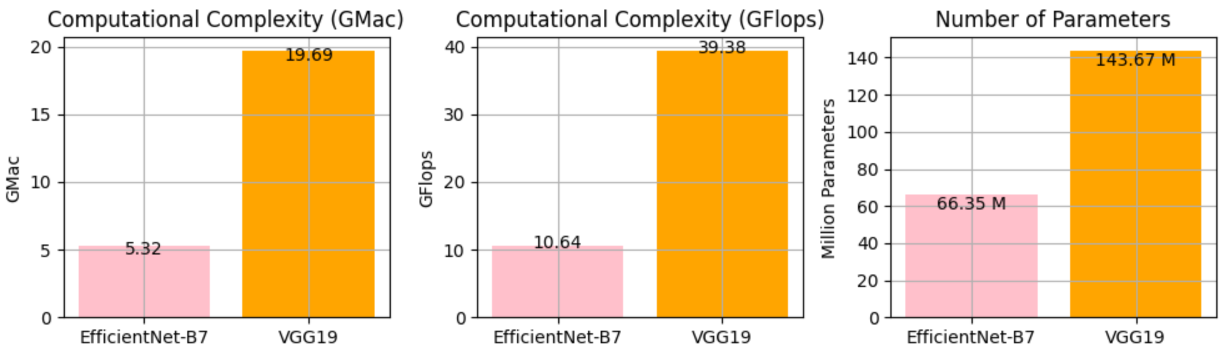


Figure 6.5: Comparison between the GMACs, GFlops and Number of Parameters between VGG19 and EfficientNet

The GPU usage can also be used to check the computational requirements of each and can be used to infer details from the GPU utilisation, that VGG19 has a higher peak and more inconsistencies when compared to EfficientNet which has a more consistent and low utilisation as shown in Figure 6.6 and Figure 6.7

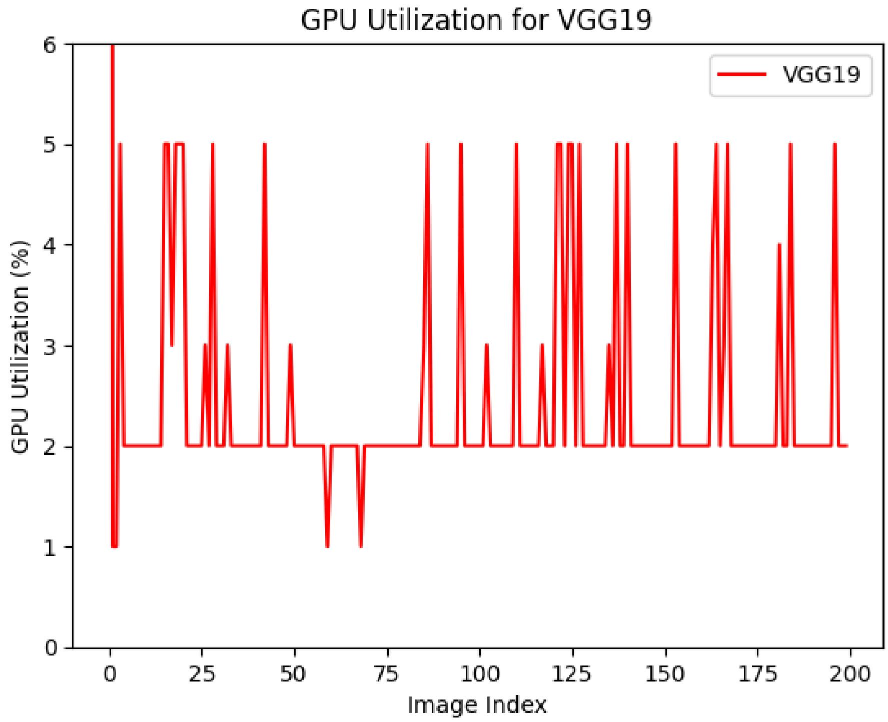


Figure 6.6: GPU utilisation of VGG19

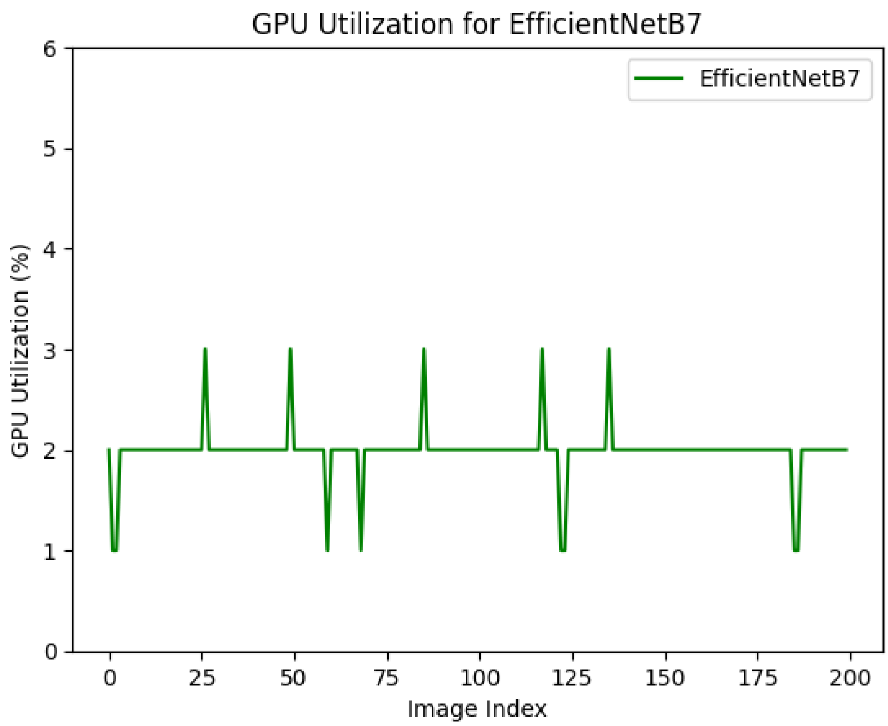


Figure 6.7: GPU utilisation of EfficientNet

6.5 Discussion

EfficientNet was shown to be more compuatationally lighter than VGG19 and also shown to be more consistent with less peaks in GPU utilisation. We were able to train a model which gives better results when compared to standard upscaling methods like Nearest Neigbhor, Bilinear Interpolation and Bicubic Interpolation. While using standard methods like PSNR and SSIM, the SR model didn't perform that well because Super resolved models are shown to be perceptually better but get a lesser score in standard methods like PSNR and SSIM. Hence we have utilised various perception based image assessment metrics to show how the Super Resolves version is perceptually better.

6.6 Risks & Challenges

GAN-based Super Resolution involves several challenges and considerations:

1. **Complexity of Real-World Degradations:** Modeling diverse real-world degradations accurately poses a significant challenge due to the complex and varied nature of degradation types.

2. **Adversarial Attacks:** Adversarial attacks pose a risk, with real-world degradation models potentially vulnerable to subtle manipulations leading to unexpected or incorrect high-resolution outputs.

3. **Trade-off Between Realism and Generalization:** Striking the right balance between realism in degradation modeling and the model's generalization across different scenarios presents a challenging trade-off.

4. **Dataset Bias Risk:** Dataset bias is a risk, as models trained on datasets biased towards specific degradation types may not perform well in scenarios with different characteristics.

Chapter 7

Conclusions

In conclusion we were able to make a Super Resolution model using GAN which can upscale any image to a higher resolution while not losing the important details and information in the image. We were able to integrate it into a user friendly web application which can be used by anyone for their upscaling needs with various features. This trained model was assessed with various image assessment metrics to see it performance compared to normal image upscaling methods like Nearest Neighbour, Bilinear Interpolation and Bicubic Interpolation.

References

[1] J. Zhang, T. Xu, J. Li, S. Jiang, and Y. Zhang, "Single-image super resolution of remote sensing images with real-world degradation modeling," *Remote Sensing*, vol. 14, no. 12, p. 2895, 2022.

[2] Y. Anagun, "Smart brain tumor diagnosis system utilizing deep convolutional neural networks," *Multimedia Tools and Applications*, pp. 1–27, 2023.

[3] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295–307, 2015.

[4] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 1646–1654.

[5] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4681–4690.

[6] M.-Y. Liu, X. Huang, J. Yu, T.-C. Wang, and A. Mallya, "Generative adversarial networks for image and video synthesis: Algorithms and applications," in *Proceedings of the IEEE*, vol. 109, no. 5, May 2021, pp. 839–858.

[7] P. Isola and J.-Y. Zhu, "Image-to-image translation with conditional adversarial networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 109, no. 5, May 2021, pp. 1–9.

[8] H. Emami, M. Moradi Aliabadi, M. Dong, and R. B. Chinnam, "Spa-gan: Spatial attention gan for image-to-image translation," in *IEEE Transactions on Multimedia*, 2019, pp. 1–10.

[9] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," *arXiv preprint arXiv:1604.07379*, pp. 2536–2544, 2016.

[10] M.-C. Sagong, Y.-G. Shin, S.-W. Kim, S. Park, and S.-J. Ko, "Pepsi : Fast image inpainting with parallel decoding network," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, June 2019, pp. 10 265–10 273.

[11] T. Nguyen-Phuoc, C. Li, L. Theis, C. Richardt, and Y.-l. Yang, "Hologan: Unsupervised learning of 3d representations from natural images," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 1459–1468.

[12] E. R. Chan, C. Z. Lin, M. A. Chan, K. Nagano, B. Pan, S. De Mello, O. Gallo, L. Guibas, J. Tremblay, S. Khamis, T. Karras, and G. Wetzstein, "Efficient geometry-aware 3d generative adversarial networks," in *Proceedings of the IEEE*, 2021, pp. 16 123–16 129.

[13] C. Tian, M. Zheng, W. Zuo, B. Zhang, Y. Zhang, and D. Zhang, "Multi-stage image denoising with the wavelet transform," *Pattern Recognition*, vol. 134, p. 109050, 2023.

[14] E. Agustsson and R. Timofte, "Ntire 2017 challenge on single image super-resolution: Dataset and study," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 2016–2025.

[15] R. Timofte, E. Agustsson, L. Van Gool, M.-H. Yang, L. Zhang, B. Lim, S. Son, H. Kim, S. Nah, K. M. Lee *et al.*, "Ntire 2017 challenge on single image super-resolution: Methods and results," in *2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2017, pp. 2016–2025.

[16] X. Wang, K. Yu, C. Dong, and C. C. Loy, "Recovering realistic texture in image super-resolution by deep spatial feature transform," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 606–615.

[17] S. Bell-Kligler, A. Shocher, and M. Irani, "Blind super-resolution kernel estimation using an internal-gan," in *Advances in Neural Information Processing Systems*, vol. 32, 2019, pp. 1–10.

Appendix A: Presentation

Single Image SUPER RESOLUTION Using GAN

Final presentation

Abel John Mathew

Ashwin Saji 220

Athif Ahamed P V

Didin Shibu

Guide : Dr. Jincy J. Fernandez

Contents

Problem Definition

- Enhancing Image Quality in Low-Resolution Scenarios through Advanced High-Resolution Reconstruction Techniques.
- Transforms low resolution image into a high resolution one.
- Information extraction from images becomes easier and faster.
- Can be used in many sensitive applications like medical field, satellite imagery etc

Project Objectives

- Develop an ML-based single image super-resolution application that would enhance the clarity of low resolution images.
- Ensure that the important details and features in the image are preserved and not distorted after the upscaling process.
- Train the novel network to produce accurate higher resolution image from the input that is a lower resolution image.
- Conduct extensive experiments on benchmark datasets to evaluate and analyze the performance of models.

Novelty of Idea

- Utilisation of a Degraded dataset to capture realism.
- Utilisation of KernelGAN for more realistic degraded dataset.
- Implementation of EfficientNet for Perceptual loss calculation.

Scope of Implementation

- **GAN model:** Utilisation of Generative Adversarial Networks for best Generative results.
- **Enhanced image clarity**: Improved quality of upscaled images as a result of using a machine learning model trained in an adversarial manner.
- **Arbitrary sizes**: The system would work with images of any shape and size, provided enough computational power is available.
- **User-Friendly Interface**: The implementation involves creating a user-friendly interface, making the system accessible to users with varying levels of expertise.

Literature Survey

Sl.no	Title	Methodology	Advantages	Disadvantages
1	**Single-image super resolution of remote sensing images with real-world degradation modeling**, Remote Sensing, vol. 14, no. 12, p. 2895, 2022, J. Zhang et al	real-world degradation modeling framework and a residual balanced attention network with modified UNet discriminator (RBAN-UNet) .	-residual learning enables the network to learn the residual mapping between the LR and HR images, which improves the reconstruction quality -balanced attention modules (BAMs) to selectively enhance important features and suppress irrelevant ones	-requires a large amount of training data to achieve good performance. -This method may be computationally expensive.
2	**Srdiff: Single image super-resolution with diffusion probabilistic models**, Neurocomputing, vol. 479, pp. 47-59, 2022, H. Li et al	Diffusion based model for single image super-resolution using markov chain and Gaussian noise	-SRDiff does not rely on adversarial losses, which can sometimes lead to over-smoothening.	-not as effective on images with large-scale structures or complex textures -requires more training time and computational resources than some other methods.

Literature Survey

Sl. no	Title	Methodology	Advantages	Disadvantages
3	**Self-supervised cycle-consistent learning for scale-arbitrary real-world single image super-resolution,** Expert Systems with Applications, vol. 212, p. 118657, 2023, H. Chen et al	scale-arbitrary super-resolution network (SASRN) and a scale-arbitrary resolution-degradation network (SARDN)	- does not require paired LR-HR training data, which can be difficult to obtain for real-world images - using a cycle-consistency loss, which helps to prevent overfitting	- May not work well for images with severe degradation
4	**Real-world single image super-resolution: A brief review,** Information Fusion, vol. 79, pp. 124-145, 2022, H. Chen et al	Deep learning-based super-resolution,(degradation modeling, image pairs, domain translation, and self learning-based algorithms)	- Compared with conventional interpolation, SR is a more effective way to obtain high-resolution images.	-Require large amounts of training data and computational resources. –Degradation methods: computationally expensive and require accurate modeling of the degradation process.
5	**A review of image super-resolution approaches based on deep learning and applications in remote sensing**, Remote Sensing, vol. 14, no. 21, p. 5423, 2022, X. Wang et al	Comparison based on Interpolation-based methods, Reconstruction-based methods and Learning-based methods	- All methods can improve resolution of a lower resolution image	- low model inference efficiency, the unsatisfactory reconstruction of real-world images, and a single approach to measuring the quality of images

Literature Survey

Sl. no	Title	Methodology	Advantages	Disadvantages
6	**Smart brain tumor diagnosis system utilizing deep convolutional neural networks**, Multimedia Tools and Applications,pp.1-27,2023, Springer,Y. Anagun	Convolutional neural network based on EfficientNetv2 model.	The model was trained using the cutting-edge optimizer Ranger, which improved the stability and convergence of the network and reduced the variance in the process.	Training deep learning models can be computationally intensive and require specialized hardware, such as GPUs, to achieve reasonable training times.
7	**A lightweight CNN-based network on COVID-19 detection using X-ray and CT images**, Computers in Biology and Medicine,vol.146,p.105604,2022,M.-L. Huang and Y.-C. Liao	Lightweight CNN-based network that is trained on a new dataset of chest X-ray and CT images using transfer learning.	Computationally efficient and saves computation and running time	The proposed technology may require further validation and testing on larger and more diverse datasets to confirm its effectiveness and generalizability.
8	**Smart brain tumor diagnosis system utilizing deep CNN**," Multimedia Tools and Applications,pp.1-27,2023,Springer,Y. Anagun	CNN model with supervised classifier algorithms and Transfer learning	The ensemble method is able to combine the strengths of different CNN models and reduce the impact of their weaknesses, resulting in improved performance	The proposed method requires a large amount of training data and computational resources to train the ensemble model

Literature Survey

Sl. no	Title	Methodology	Advantages	Disadvantages
9	**Detecting brain tumors using deep learning convolutional neural network with transfer learning approach**, *International Journal of Imaging Systems and Technology*, vol. 32, no. 1, pp. 307–323, 2022,S. Anjum *et al*	CNN with transfer learning usings BraTS dataset.	Automatically extracts key features making it more robust for further image analysis.	Large amount of data for training, which can be time consuming and computationally expensive.
10	**COVID-19 detection in X-ray images using convolutional neural networks**', *Machine Learning with Applications*, vol. 6, p. 100138, 2021,D. Arias-Garzón *et al*	VGG16-19(type of CNN) which uses transfer learning.	High accuracy in image classification,transfer learning capabilities.	Large size which can make it computationally expensive to train and use, tendency to overfit on smaller data sets.

Literature Survey

Sl. no	Title	Methodology	Advantages	Disadvantages
11	**Multi-stage image denoising with the wavelet transform.** Pattern Recognition, vol. 134, p. 109050, 2023.	3-stage CNN -Dynamic Convolutional Block (DCB),wavelet transform and enhancement blocks (WEBs) and a residual block (RB)	Adaptability, Tradeoff between denoising and computational cost, Improved denoising	Lack Specific Wavelet transform, Complexity and Computational Requirement
12	**Recorrupted-to-recorrupted:Unsupervised deep learning for image denoising**,IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2043-2052, 2021.	Unsupervised deep denoiser,data augmentation technique, called recorrupted-to-recorrupted (R2R) to address the overfitting caused by the absence of truth images	No need of clean/noisy image pair, address overfitting	Performance than supervised counterpart
13	**Repaint: Inpainting using denoising diffusion probabilistic models,**IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11461-11471, 2022.	Unconditional Denoising Diffusion Probabilistic Model (DDPM) as the generative prior	Generalization, Semantically meaningful generation, high quality	Lack of fine-grained control, Computational cost

Literature Survey

Sl. no	Title	Methodology	Advantages	Disadvantages
14	**Comprehensive overview of backpropagation algorithm for digital image denoising,** Electronics, vol. 11, no. 10, p. 1590, 2022.	Using ANN to denoise through backpropagation algorithm	Once trained,no prior knowledge of noise is required, Generalization	Overfitting
15	**Srdiff: Single image super-resolution with diffusion probabilistic models**, Neurocomputing, vol. 479, pp. 47-59, 2022, H. Li et al	Diffusion based model for single image super-resolution using markov chain and Gausian noise	SRDiff does not rely on adversarial losses, over-smoothening.	not as effective on images with large-scale structures or complex textures More training time and computational resources than some other methods.
16	**Generative Adversarial Networks for Visual Synthesis: Algorithms and Applications**, PROCEEDINGS OF THE IEEE,Vol. 109, No. 5, May 2021|,MING-YU LIU et al.	Comparative Analysis of different methods based on GAN used in Visual synthesis tasks.	GAN based methods show exceptional accuracy in generating content	Difficulty in training

Literature Survey

Sl. no	Title	Methodology	Advantages	Disadvantages
17	**Image-to-Image Translation with Conditional Adversarial Networks**, 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-9,Phillip Isola et al.	pix2pix GAN for image translation	Patch-wise discriminator	Requires paired training data.
18	**Context Encoders: Feature Learning by Inpainting**,IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016,pp. 1-12,Deepak Pathak et al.	Context Encoder(CE) For image inpainting	Realistic generation, handle irregularly shaped inpainting masks	Over smoothening, Artifacting
19	**HoloGAN: Unsupervised Learning of 3D Representations From Natural Images**, arXiv preprint arXiv:1904.01326, 2019,Thu Nguyen-Phuoc, Chuan Li et al	Holographic GAN for neural rendering	Disentangled representation	Artifact Generation

Literature Survey

Sl. no	Title	Methodology	Advantages	Disadvantages
20	**SPA-GAN: Spatial Attention GAN for Image-to-Image Translation**, IEEE Transactions on Multimedia Vol 23,pp. 1-10 Hajar Emami et al.	SPA-GAN for image translation	Attribute Consistency	Lesser Versatility
21	**PEPSI : Fast Image Inpainting with Parallel Decoding Network**, IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, June 2019, pp. 10265-10273,M.-C. Sagong, Y.-G. Shin, S.-W. Kim, S. Park, and S.-J. Ko	PEPSI (parallel extended-decoder path for semantic inpainting network)	Reduced Overfitting	Increased Complexity
22	**Efficient Geometry-aware 3D Generative Adversarial Networks**, Proceedings of the IEEE, pp. 16123-16129, Eric R. Chan et al.	EG3D-tri-plane-based 3D GAN framework for neural rendering	Decouples feature generation from neural rendering	computationally expensive

Methodology

1) Dataset Degradation Module

- Utilisation of a Degradation Pipeline to create a degraded dataset for more realistic super resolution.

- The high resolution images from the dataset are degraded and a new realistic dataset is formed.

Methodology

1.1. **Downscaling to 256 x 256 (HR image)**

- The blurred image is downscaled to 256 x 256 and this is considered as the High Resolution image.

- The downscaling of the image is used by the Pillow Library, particularly the resize function which utilises bicubic downscaling to downscale the image.

Methodology

1.2: **Blur Estimation:**

- The Blur Kernel Estimation is done using the KernelGAN framework. It is used for finding realistic blur kernels with the help of a single input image.

- The KernelGAN consists of a 7 layer generator which is trained on an input image and it is trained to successfully produce a blur kernel.

- The Discriminator consists of convolution layers which output a 32 x 32 D-map of the range 0-1, which shows how real each image is compared to its surrounding pixels.

Methodology

1.3: **Noise Addition**

- The noise addition is done using gaussian noise and poisson noise.
- Gaussian noise is a type of statistical noise that follows a Gaussian (normal) distribution.

$$P(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$$

- 'μ' is the mean and 'σ' is the standard deviation.
-

Methodology

- Poisson noise follows a poisson distribution. Commonly found as camera sensor noise in low light conditions.

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

- 'λ' is the average rate of events in a given interval of time and '*x*' is a non negative integer indicating number of events occurred.

Methodology

1.4: **JPEG Compression**

- Image size is reduced using JPEG compression

1.5: **Second order downscaling**

- The final image is further downscaled to 64x64 resolution and this is considered as the Low Resolution image.

Methodology

Original high resolution image

RSET
RAJAGIRI SCHOOL OF ENGINEERING & TECHNOLOGY

Methodology

Image after realistic blur is applied through KernelGAN

Methodology

Image downscaled to 256x256 resolution

Methodology

Gaussian and Poisson noise applied to the downscaled image

Methodology

30 percent JPEG compression applied to the noisy image

Methodology

Second order degradation downscales image to 64x64 resolution

Methodology

2) Generative Adverserial Network Module:

The GAN module consists of :

1. Development of Generator
2. Development of Discriminator
3. Development of EfficientNet Feature Extractor

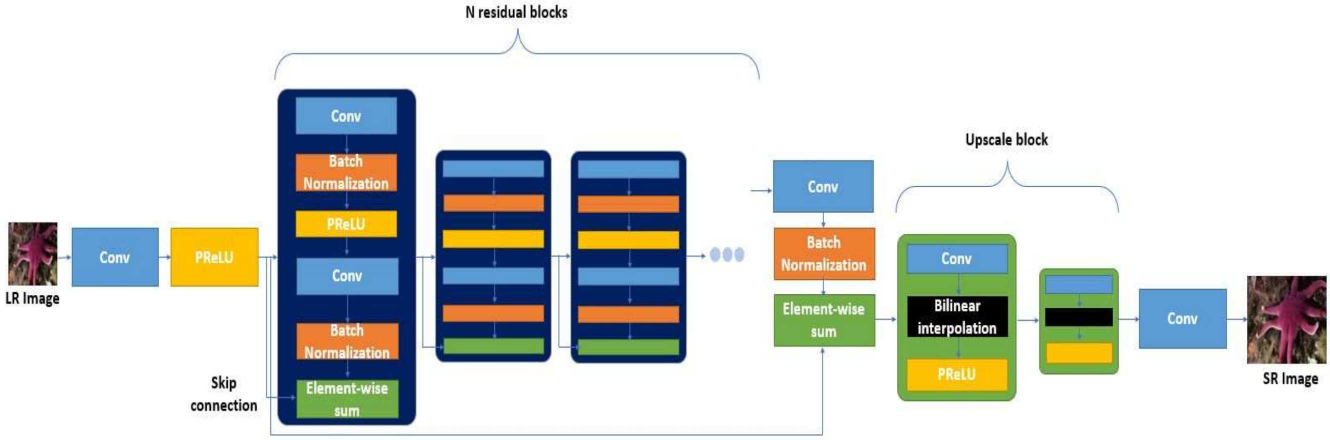
RSET RAJAGIRI SCHOOL OF ENGINEERING & TECHNOLOGY

Methodology

1. **Generator**

- The Generator is used for enhancing the features of the low res input image and upscaling it with respect to a Scaling factor.
- The Generator consists of Conv2D blocks for low level feature extraction, Residual blocks with skip connections for high level feature extraction and UpSampling2D layers for the final upscaling of the input image.

Methodology



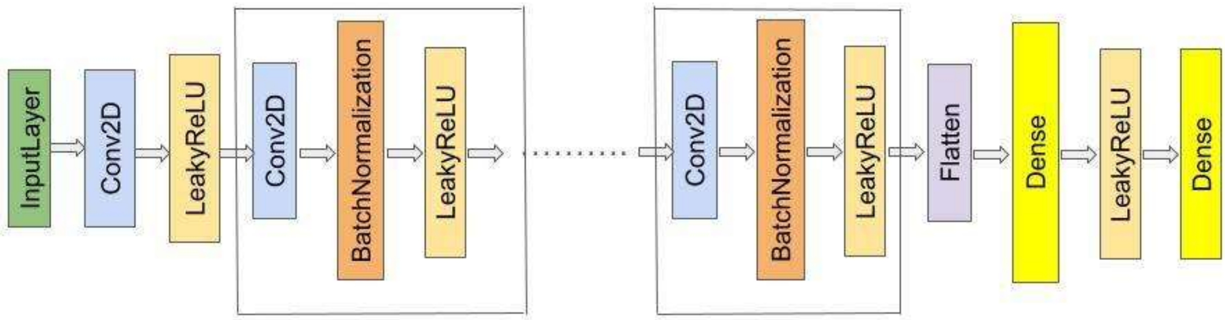
Methodology

2. Discriminator

- The Discriminator is used for differentiating whether an image given to it is real or not.

- The discriminator processes the input image through convolutional layers, downsampling operations, and fully connected layers to extract features and make a decision about the authenticity of the input.

Methodology

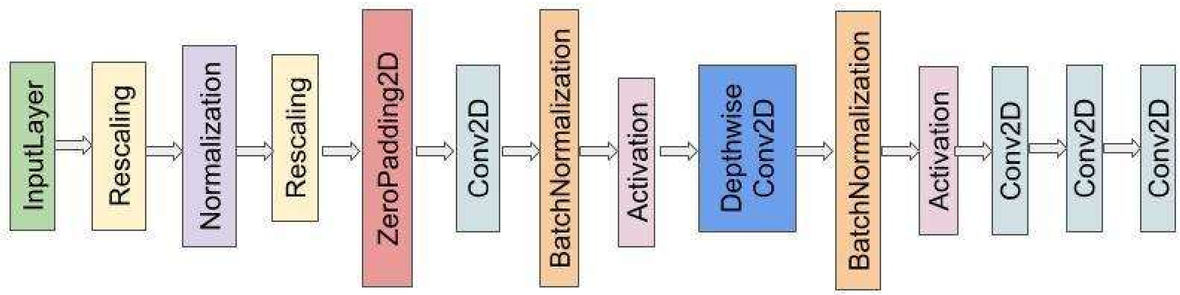


Methodology

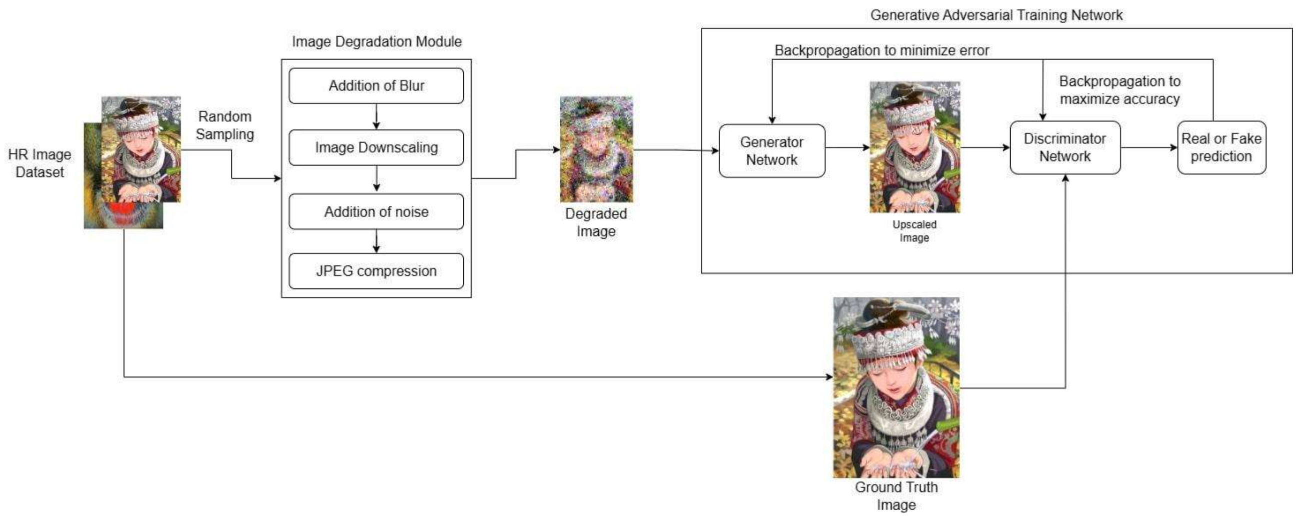
3. EfficientNet Feature Extractor for Perceptual Loss

- In this GAN model, we utilise perceptual loss while training the model.
- The EfficientNet which is pretrained on the imagenet dataset is used to extract the features of the generated image.
- This loss combined with adversarial loss is the main loss functions in this GAN.

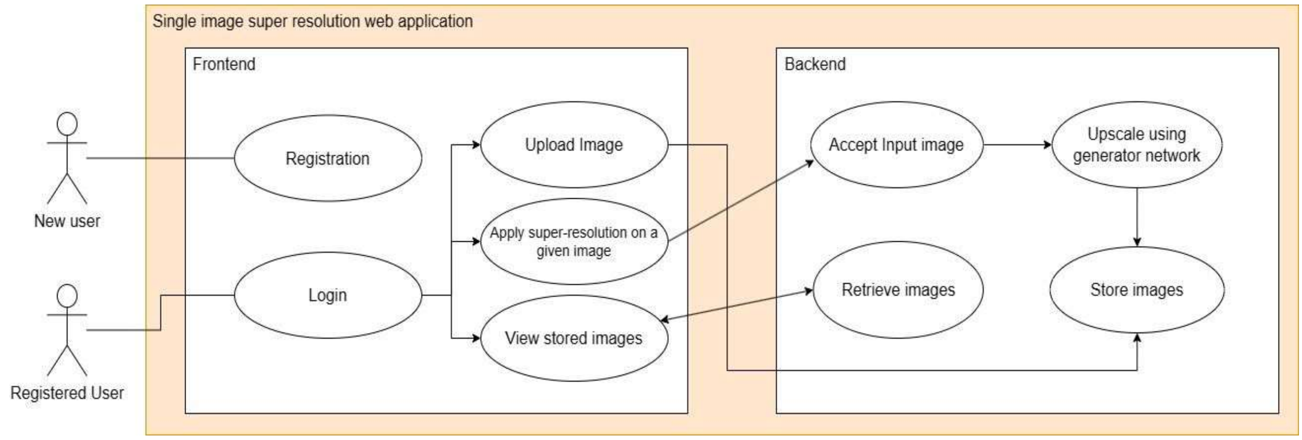
Methodology



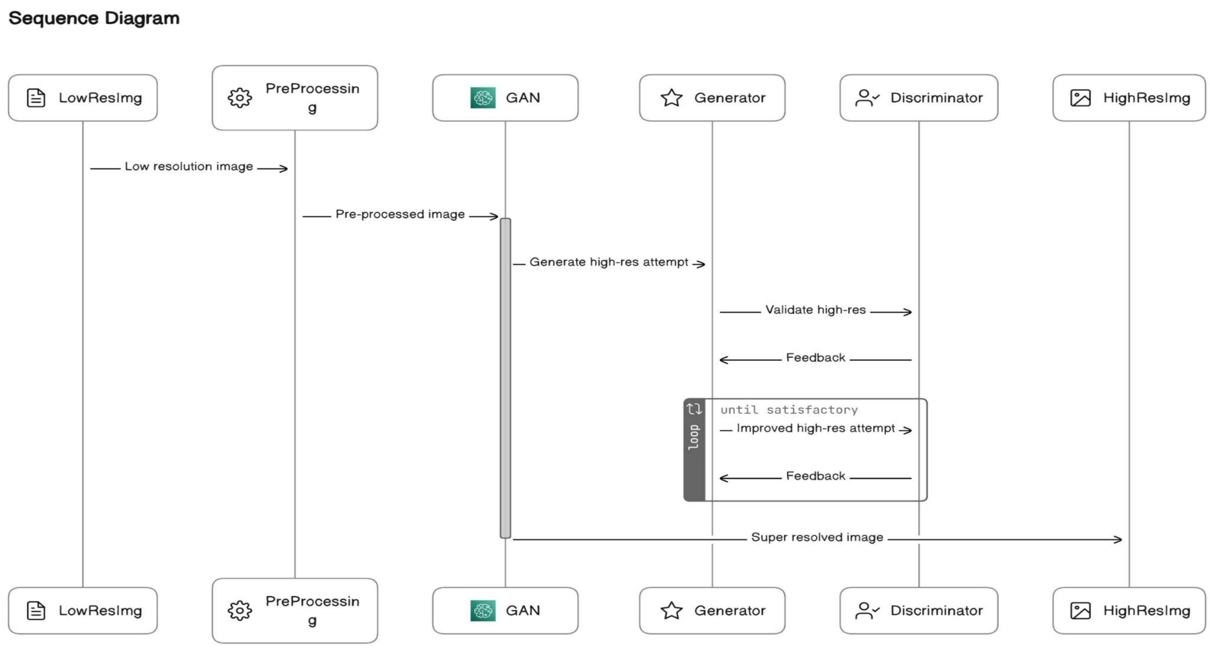
Architecture Diagram



UML Diagram

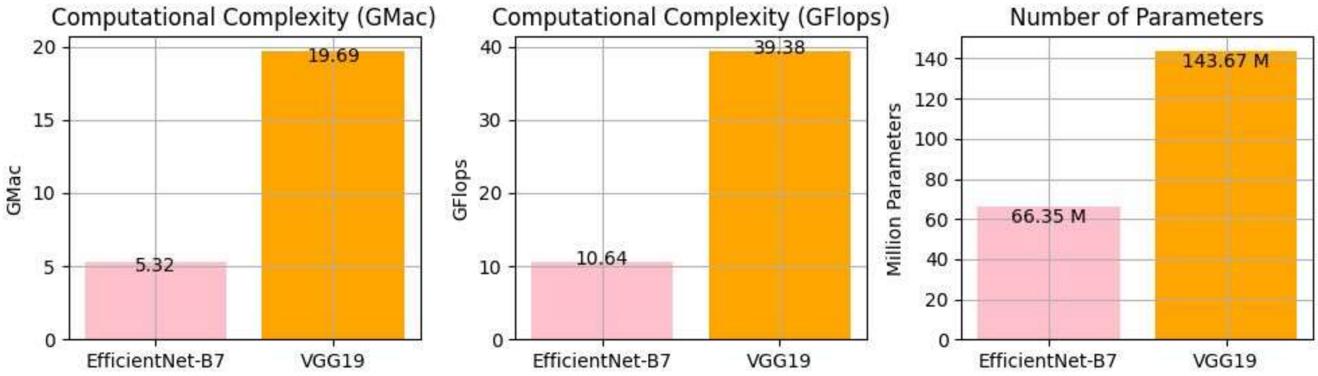


Sequence Diagram



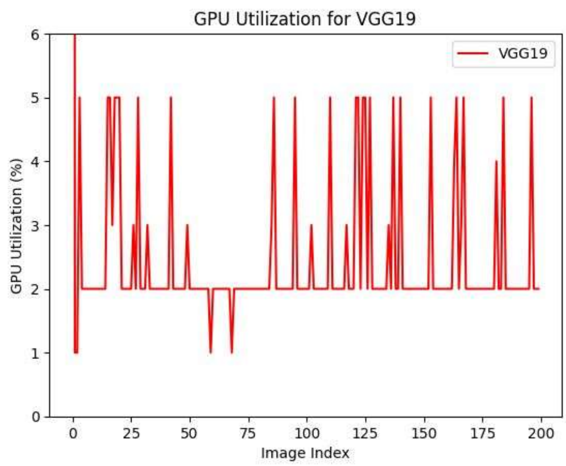
Results

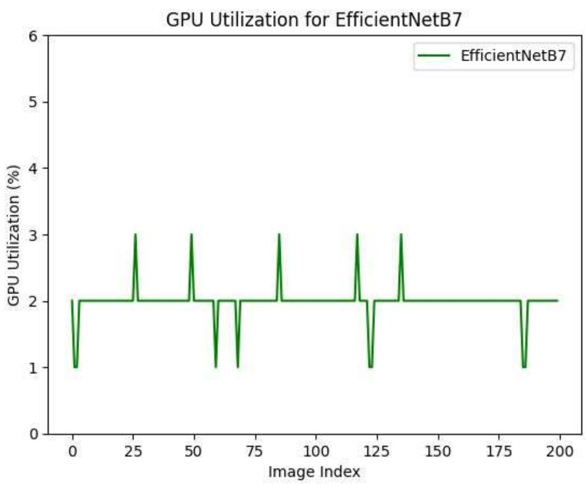
We can use GMac(Giga Multiply-Add Operations per Second) and GFlops (Giga Floating point operations per second) to find the number of computational operations performed.



Results

EfficientNet vs VGG19 computation comparison





Results

- We can assess the quality of the image with standard image quality metrics like PSNR and SSIM, but they don’t give the best results .

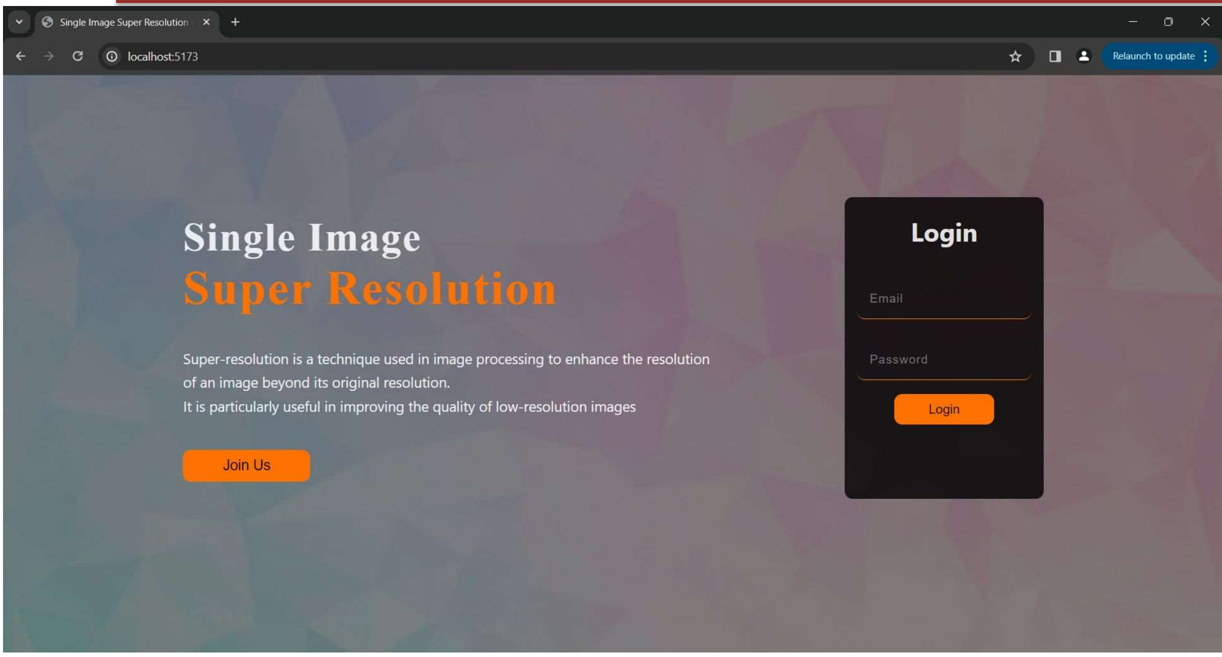
Metric	SR Value	NN Value	Bilinear Value	Bicubic Value
PSNR index (-∞ to ∞)	23.6789	23.3524	24.1277	24.635
SSIM index (-1 to 1)	0.5945	0.5514	0.5847	0.6143

Results

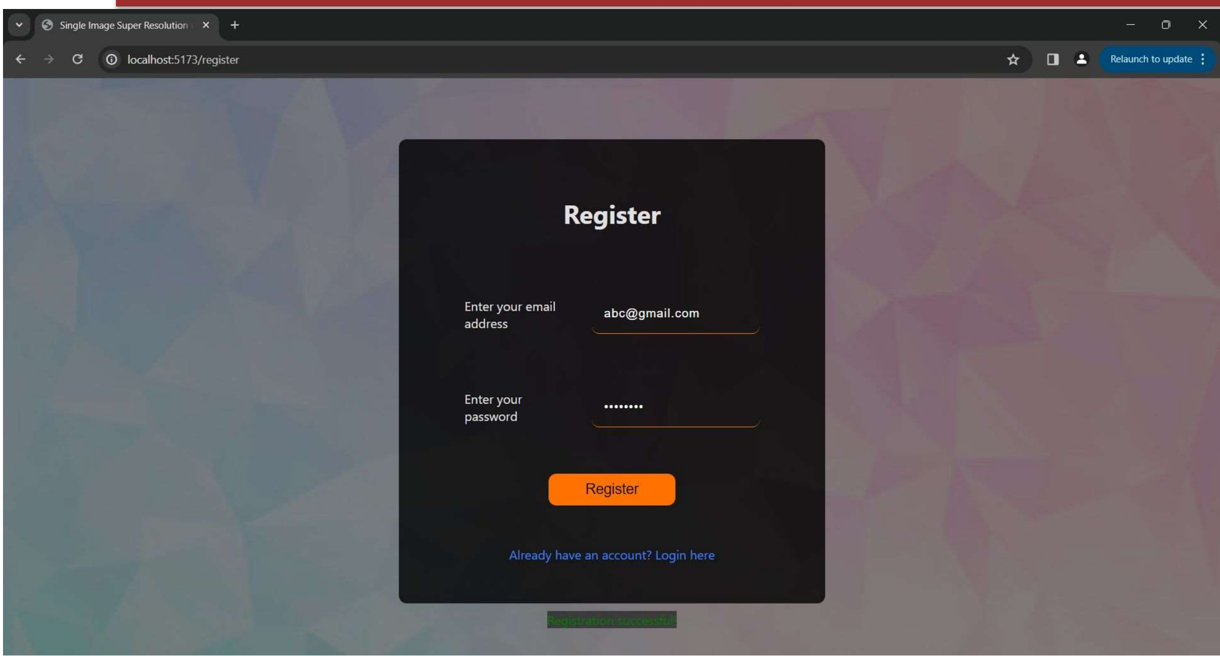
- Human Based Perception Metrics

Metric	SR Value	NN Value	Bilinear Value	Bicubic Value
DSS index (0 to 1)	0.8361	0.6813	0.7051	0.7587
FSIM index (0 to 1)	0.8202	0.6956	0.7581	0.785
HaarPSI index (0 to 1)	0.7143	0.5384	0.6894	0.7118
IW-SSIM index (0 to 1)	0.9089	0.8724	0.8569	0.8871
LPIPS (0 to 1)	0.3933	0.5303	0.4062	0.3956
MS-SSIM index (0 to 1)	0.924	0.9075	0.9051	0.922
PieAPP loss (0 to ∞)	1.4256	3.1009	3.4729	3.0959
VIFp index (0 to 1)	0.3016	0.2528	0.2728	0.2962
VSI index (0 to 1)	0.9459	0.9034	0.9325	0.9394

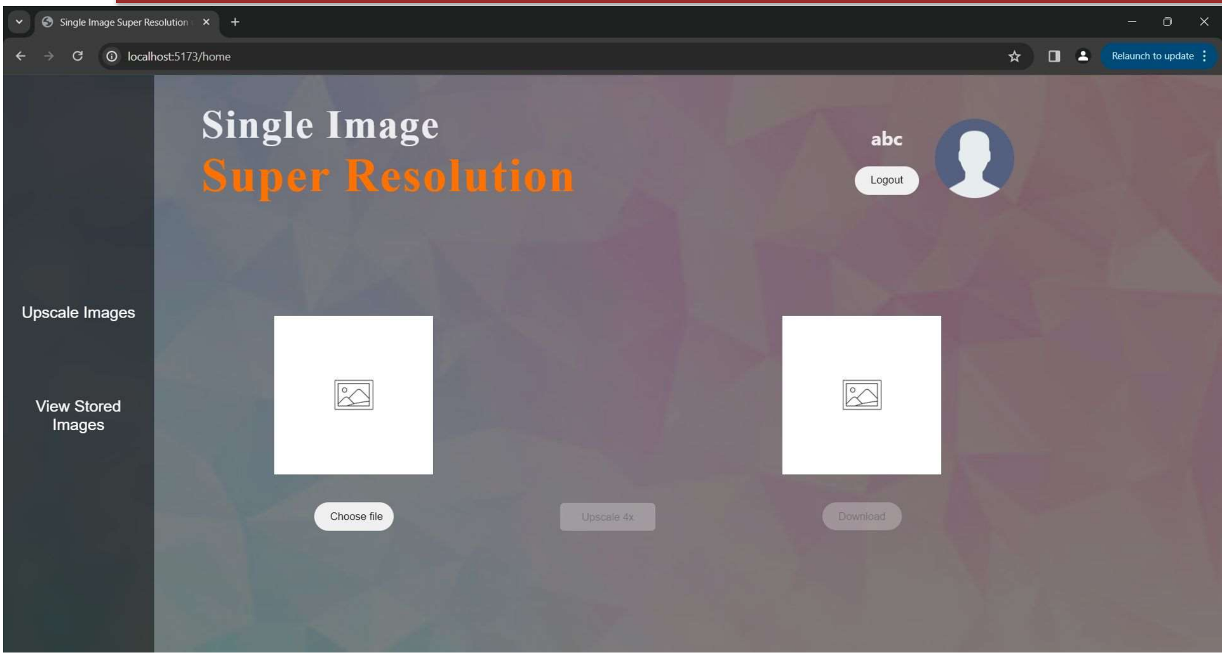
Results



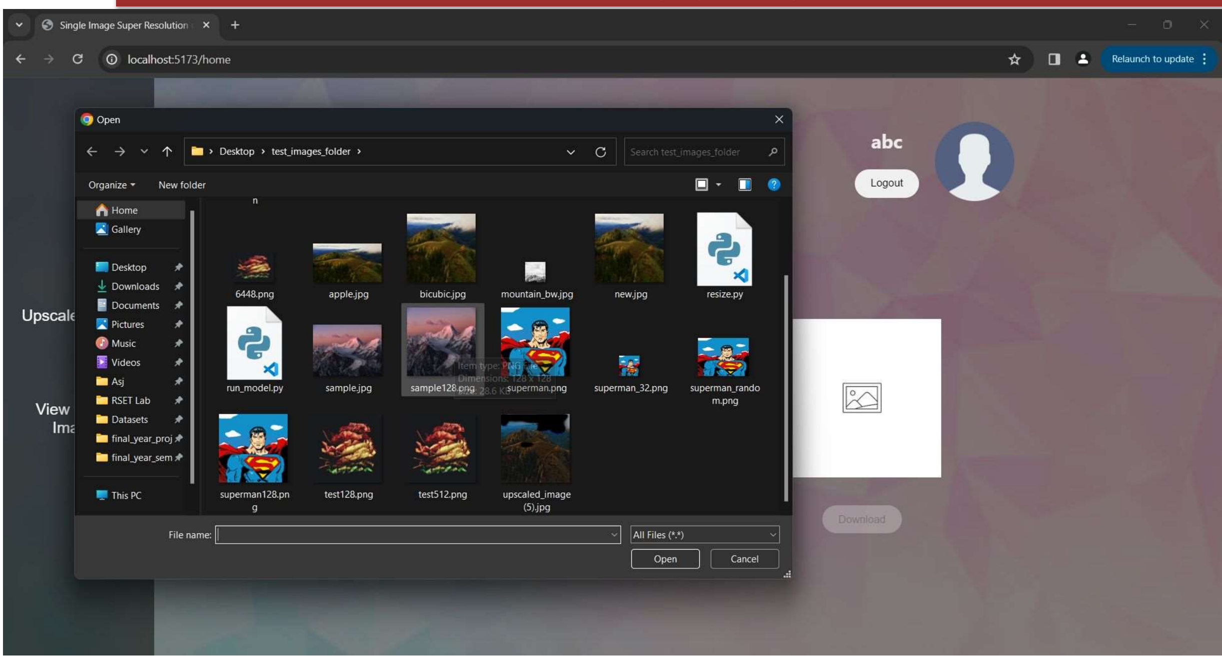
Results



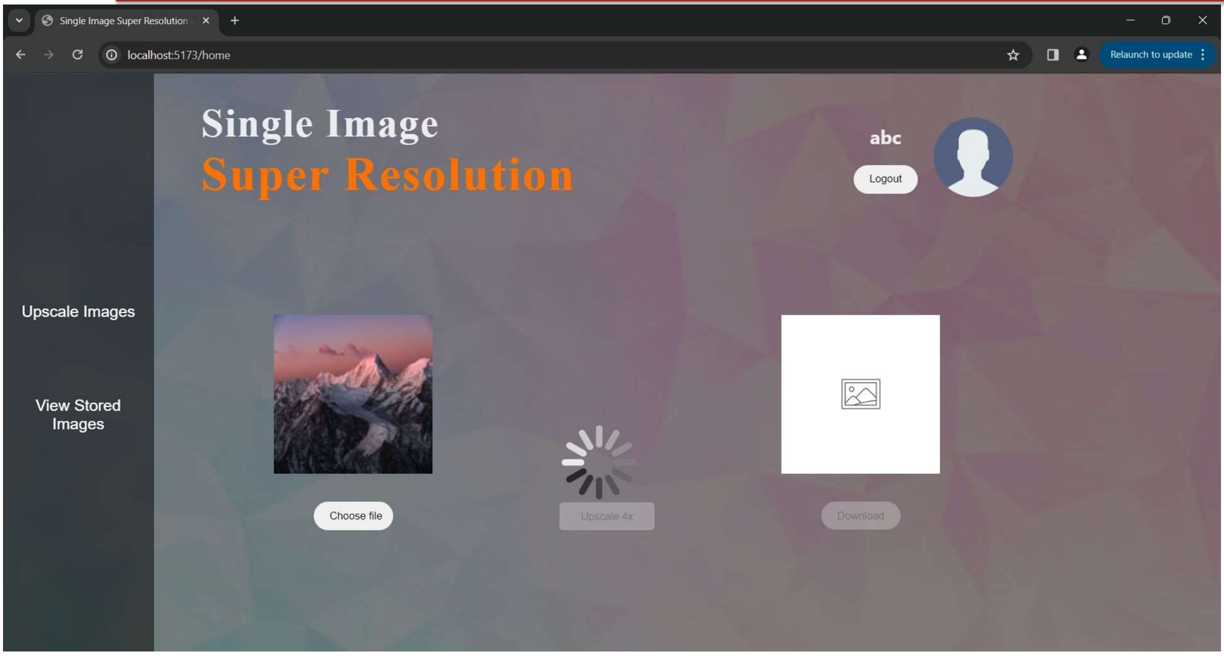
Results



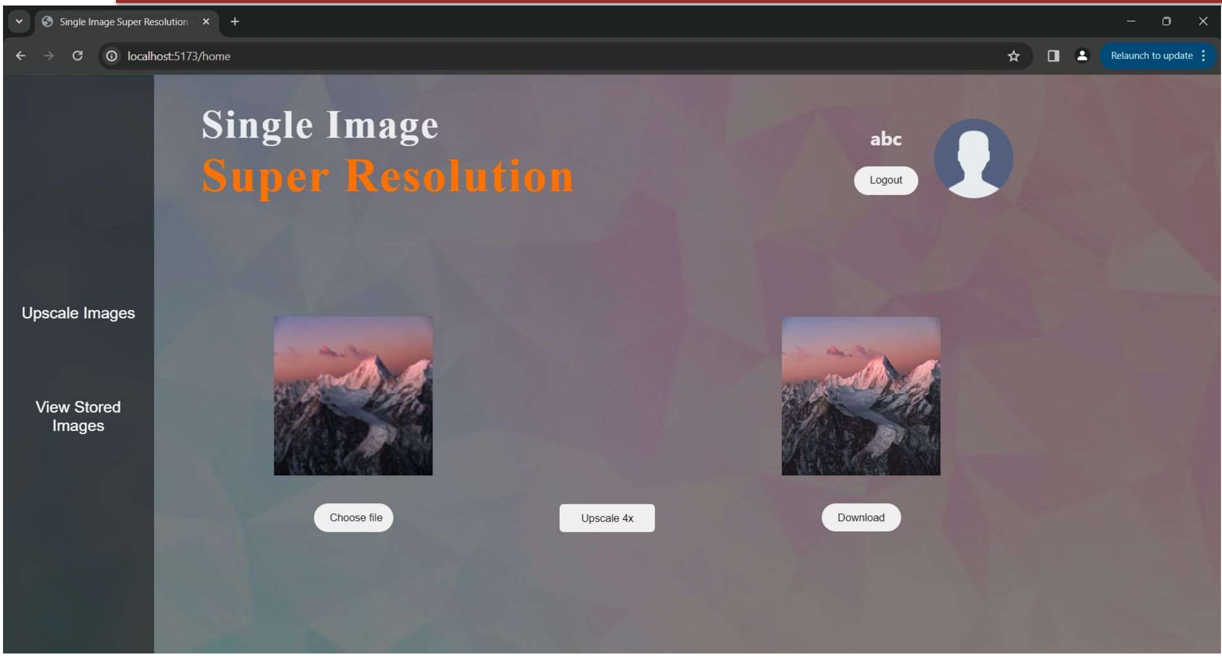
Results



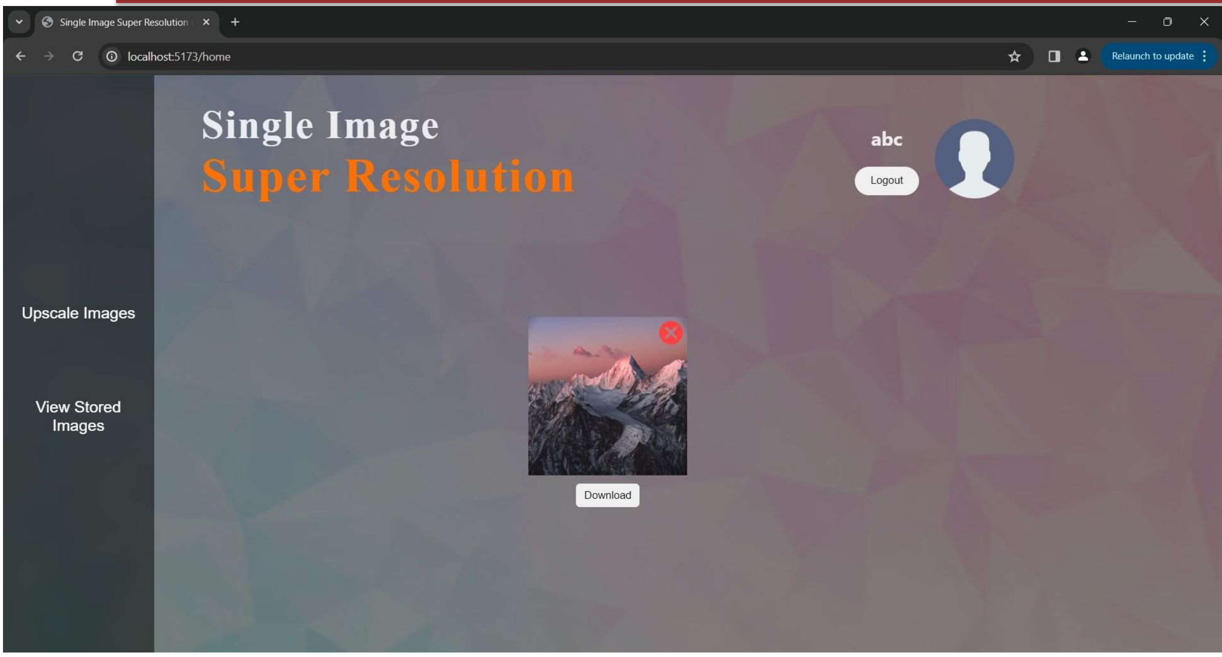
Results



Results



Results



Work Distribution

ABEL JOHN MATHEW	DIDIN SHIBU	ASHWIN SAJI	ATHIF AHAMED P V
Development of Generator and EfficientNet Feature Extractor	Development of Discriminator and EfficientNet Feature Extractor	Website backend, Model validation and integration.	Dataset Degradation , Website Frontend

Conclusion

In conclusion our project has attained the 100% completion goal by successfully developing the Super resolution GAN model and integrating it into a user friendly web application which can be used by anyone for their upscaling needs.

Future Scope

- **Further Refinement of Perceptual Loss**: Explore alternative perceptual loss functions beyond EfficientNet.
- **Multiple scaling factors**: Current models typically upscale by a fixed factor. Hence, investigate architectures that can handle variable upscale factors, allowing for more control over the final image resolution.
- **Privacy Controls:** Implement clear privacy policies and user controls regarding how uploaded images are stored and used.
- **Video Super Resolution**: Interpolate multiple frames of super resolved images for video upscaling.

References

[1] A. Lugmayr, M. Danelljan, F. Yu, R. Timofte, and L. Van Gool, "Repaint: Inpainting using denoising diffusion probabilistic models," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11461-11471, 2022.

[2] A. Mabrouk, R. P. Díaz Redondo,A. Dahou, M. Abd Elaziz, and M. Kayed, "Pneumonia detection on chest X-ray images using ensemble of deep convolutional neural networks," Applied Sciences, vol. 12, no. 13, p. 6448, 2022.

[3] A.Singh, S. Kushwaha, M.Alarfaj, and M. Singh, "Comprehensive overview of backpropagation algorithm for digital image denoising," Electronics, vol. 11, no. 10, p.1590, 2022.

[4] C. Tian, M. Zheng, W. Zuo, B.Zhang, Y. Zhang, and D. Zhang, "Multi-stage image denoising with the wavelet transform," Pattern Recognition, vol. 134, p. 109050, 2023.

References

[5] H. Chen, X. He, H. Yang, Y. Wu, L. Qing, and R. E. Sheriff, "Self-supervised cycle-consistent learning for scale-arbitrary real-world single image super-resolution," Expert Systems with Applications, vol. 212, p. 118657, 2023.

[6] H. Chen, X. He, L. Qing, Y. Wu,C. Ren, R. E. Sheriff, and C. Zhu, "Real-world single image super-resolution: A brief review," Information Fusion, vol. 79, pp.124-145, 2022.

[7] H. Li, Y. Yang, M. Chang, S. Chen, H. Feng, Z.Xu, Q. Li, and Y. Chen, "Srdiff: Single image super-resolution with diffusion probabilistic models," Neurocomputing, vol. 479, pp. 47-59, 2022.

[8] Hajar Emami, Majid Moradi Aliabadi, Ming Dong, Ratna Babu Chinnam. , “SPA-GAN: Spatial Attention GAN for Image-to-Image Translation”, IEEE TRANSACTIONS ON MULTIMEDIA,vol.23 pp. 391-401,February 2020

References

[9] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," Communications of the ACM, vol. 63, no. 11, pp. 139-144, 2020.

[10] J. Zhang, T. Xu, J. Li, S.Jiang, and Y. Zhang, "Single-image super resolution of remote sensing images with real-world degradation modeling," Remote Sensing, vol. 14, no. 12, p.2895, 2022.

[11] M.-C. Sagong et al, "PEPSI : Fast Image Inpainting with Parallel Decoding Network," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, June 2019, pp. 10265-10273.

[12] M.-L. Huang et al., "A lightweight CNN-based network on COVID-19 detection using X-ray and CT images," Computers in Biology and Medicine, vol. 146, p. 105604, 2022.

References

[13] M.-Y. Liu, X. Huang, J. Yu, T.-C. Wang, and A. Mallya, "Generative adversarial networks for image and video synthesis: Algorithms and applications," Proceedings of the IEEE, vol. 109, no. 5, pp. 839-862, 2021.

[14] P. Gupta, "Pneumonia detection using convolutional neural networks," Science and Technology, vol. 7, no. 01, pp. 77-80, 2021.

[15] Phillip Isola et al. , “Image-to-Image Translation with Conditional Adversarial Networks” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1-9

[16] S. Iqbal, M. U. Ghani, T. Saba, and A. Rehman, "Brain tumor segmentation in multi-spectral MRI using convolutional neural networks (CNN)," Microscopy Research and Technique, vol. 81, no. 4, pp. 419-427, 2018.

References

[17] S. Zhao, X. Zhu, D. Liu, F. Xu, Y. Wang, L. Lin, X. Chen, and Q. Yuan, "A hyperspectral image denoising method based on land cover spectral autocorrelation," International Journal of Applied Earth Observation and Geoinformation, vol. 123, p. 103481, 2023.

[18] T. Pang, H. Zheng, Y. Quan, and H. Ji, "Recorrupted-to-recorrupted: Unsupervised deep learning for image denoising," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2043-2052, 2021.

[19] X. Wang, J. Yi, J. Guo, Y. Song, J. Lyu, J. Xu, W. Yan, J. Zhao, Q. Cai, and H. Min, "A review of image super-resolution approaches based on deep learning and applications in remote sensing," Remote Sensing, vol. 14, no. 21, p. 5423, 2022.

[20] Y. Anagun, "Smart brain tumor diagnosis system utilizing deep convolutional neural networks," Multimedia Tools and Applications, pp. 1-27, 2023, Springer.

Status of Paper Publication

- Topic - EfficientNet Content Loss for Single Image Super resolution using GAN
- Working on Draft article
- Plan to submit article to upcoming journals like International Journal of Computational Vision and Robotics

Appendix B: Vision, Mission, Programme Outcomes and Course Outcomes

Vision, Mission, Programme Outcomes and Course Outcomes

Institute Vision

To evolve into a premier technological institution, moulding eminent professionals with creative minds, innovative ideas and sound practical skill, and to shape a future where technology works for the enrichment of mankind.

Institute Mission

To impart state-of-the-art knowledge to individuals in various technological disciplines and to inculcate in them a high degree of social consciousness and human values, thereby enabling them to face the challenges of life with courage and conviction.

Department Vision

To become a centre of excellence in Computer Science and Engineering, moulding professionals catering to the research and professional needs of national and international organizations.

Department Mission

To inspire and nurture students, with up-to-date knowledge in Computer Science and Engineering, ethics, team spirit, leadership abilities, innovation and creativity to come out with solutions meeting societal needs.

Programme Outcomes (PO)

Engineering Graduates will be able to:

1. Engineering Knowledge: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.

2. Problem analysis: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

3. Design/development of solutions: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.

4. Conduct investigations of complex problems: Use research-based knowledge including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.

5. Modern Tool Usage: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.

6. The engineer and society: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal, and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. Environment and sustainability: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. Ethics: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. Individual and Team work: Function effectively as an individual, and as a member or leader in teams, and in multidisciplinary settings.

10. Communication: Communicate effectively with the engineering community and with society at large. Be able to comprehend and write effective reports documentation. Make effective presentations, and give and receive clear instructions.

11. Project management and finance: Demonstrate knowledge and understanding of engineering and management principles and apply these to one's own work, as a member and leader in a team. Manage projects in multidisciplinary environments.

12. Life-long learning: Recognize the need for, and have the preparation and ability to engage in independent and lifelong learning in the broadest context of technological change.

Programme Specific Outcomes (PSO)

A graduate of the Computer Science and Engineering Program will demonstrate:

PSO1: Computer Science Specific Skills
The ability to identify, analyze and design solutions for complex engineering problems in multidisciplinary areas by understanding the core principles and concepts of computer science and thereby engage in national grand challenges.

PSO2: Programming and Software Development Skills
The ability to acquire programming efficiency by designing algorithms and applying standard practices in software project development to deliver quality software products meeting the demands of the industry.

PSO3: Professional Skills
The ability to apply the fundamentals of computer science in competitive research and to develop innovative products to meet the societal needs thereby evolving as an eminent researcher and entrepreneur.

Course Outcomes (CO)

Course Outcome 1: Model and solve real world problems by applying knowledge across domains (Cognitive knowledge level: Apply).

Course Outcome 2: Develop products, processes or technologies for sustainable and socially relevant applications (Cognitive knowledge level: Apply).

Course Outcome 3: Function effectively as an individual and as a leader in diverse teams and to comprehend and execute designated tasks (Cognitive knowledge level: Apply).

Course Outcome 4: Plan and execute tasks utilizing available resources within timelines, following ethical and professional norms (Cognitive knowledge level: Apply).

Course Outcome 5: Identify technology/research gaps and propose innovative/creative solutions (Cognitive knowledge level: Analyze).

Course Outcome 6: Organize and communicate technical and scientific findings effectively in written and oral forms (Cognitive knowledge level: Apply).

Appendix C: CO-PO-PSO Mapping

CO-PO AND CO-PSO MAPPING

	PO 1	PO 2	PO 3	PO 4	PO 5	PO 6	PO 7	PO 8	PO 9	PO 10	PO 11	PO 12	PSO 1	PSO 2	PSO 3
CO 1	2	2	2	1	2	2	2	1	1	1	1	2	3		
CO 2	2	2	2		1	3	3	1	1		1	1		2	
CO 3									3	2	2	1			3
CO 4					2			3	2	2	3	2			3
CO 5	2	3	3	1	2							1	3		
CO 6					2			2	2	3	1	1			3

3/2/1: high/medium/low

JUSTIFICATIONS FOR CO-PO MAPPING

MAPPING	LOW/ MEDIUM/HIGH	JUSTIFICATION
100003/ CS722U.1-PO1	M	Knowledge in the area of technology for project development using various tools results in better modeling.
100003/ CS722U.1-PO2	M	Knowledge acquired in the selected area of project development can be used to identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions.

100003/ CS722U.1-PO3	M	Can use the acquired knowledge in designing solutions to complex problems.
100003/ CS722U.1-PO4	M	Can use the acquired knowledge in designing solutions to complex problems.
100003/ CS722U.1-PO5	H	Students are able to interpret, improve and redefine technical aspects for design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
100003/ CS722U.1-PO6	M	Students are able to interpret, improve and redefine technical aspects by applying contextual knowledge to assess societal, health and consequential responsibilities relevant to professional engineering practices.
100003/ CS722U.1-PO7	M	Project development based on societal and environmental context solution identification is the need for sustainable development.
100003/ CS722U.1-PO8	L	Project development should be based on professional ethics and responsibilities.
100003/ CS722U.1-PO9	L	Project development using a systematic approach based on well defined principles will result in teamwork.
100003/ CS722U.1-PO10	M	Project brings technological changes in society.

100003/ CS722U.1-PO11	H	Acquiring knowledge for project development gathers skills in design, analysis, development and implementation of algorithms.
100003/ CS722U.1-PO12	H	Knowledge for project development contributes engineering skills in computing & information gatherings.
100003/ CS722U.2-PO1	H	Knowledge acquired for project development will also include systematic planning, developing, testing and implementation in computer science solutions in various domains.
100003/ CS722U.2-PO2	H	Project design and development using a systematic approach brings knowledge in mathematics and engineering fundamentals.
100003/ CS722U.2-PO3	H	Identifying, formulating and analyzing the project results in a systematic approach.
100003/ CS722U.2-PO5	H	Systematic approach is the tip for solving complex problems in various domains.
100003/ CS722U.2-PO6	H	Systematic approach in the technical and design aspects provide valid conclusions.
100003/ CS722U.2-PO7	H	Systematic approach in the technical and design aspects demonstrate the knowledge of sustainable development.

100003/ CS722U.2-PO8	M	Identification and justification of technical aspects of project development demonstrates the need for sustainable development.
100003/ CS722U.2-PO9	H	Apply professional ethics and responsibilities in engineering practice of development.
100003/ CS722U.2-PO11	H	Systematic approach also includes effective reporting and documentation which gives clear instructions.
100003/ CS722U.2-PO12	M	Project development using a systematic approach based on well defined principles will result in better teamwork.
100003/ CS722U.3-PO9	H	Project development as a team brings the ability to engage in independent and lifelong learning.
100003/ CS722U.3-PO10	H	Identification, formulation and justification in technical aspects will be based on acquiring skills in design and development of algorithms.
100003/ CS722U.3-PO11	H	Identification, formulation and justification in technical aspects provides the betterment of life in various domains.
100003/ CS722U.3-PO12	H	Students are able to interpret, improve and redefine technical aspects with mathematics, science and engineering fundamentals for the solutions of complex

		problems.
100003/ CS722U.4-PO5	H	Students are able to interpret, improve and redefine technical aspects with identification formulation and analysis of complex problems.
100003/ CS722U.4-PO8	H	Students are able to interpret, improve and redefine technical aspects to meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
100003/ CS722U.4-PO9	H	Students are able to interpret, improve and redefine technical aspects for design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
100003/ CS722U.4-PO10	H	Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools for better products.
100003/ CS722U.4-PO11	M	Students are able to interpret, improve and redefine technical aspects by applying contextual knowledge to assess societal, health and consequential responsibilities relevant to professional engineering practices.
100003/ CS722U.4-PO12	H	Students are able to interpret, improve and redefine technical aspects for demonstrating the knowledge of, and need for sustainable development.

100003/ CS722U.5-PO1	H	Students are able to interpret, improve and redefine technical aspects, apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
100003/ CS722U.5-PO2	M	Students are able to interpret, improve and redefine technical aspects, communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
100003/ CS722U.5-PO3	H	Students are able to interpret, improve and redefine technical aspects to demonstrate knowledge and understanding of the engineering and management principle in multidisciplinary environments.
100003/ CS722U.5-PO4	H	Students are able to interpret, improve and redefine technical aspects, recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.
100003/ CS722U.5-PO5	M	Students are able to interpret, improve and redefine technical aspects in acquiring skills to design, analyze and develop algorithms and implement those using high-level programming languages.
100003/ CS722U.5-PO12	M	Students are able to interpret, improve and redefine

		technical aspects and contribute their engineering skills in computing and information engineering domains like network design and administration, database design and knowledge engineering.
100003/ CS722U.6-PO5	M	Students are able to interpret, improve and redefine technical aspects and develop strong skills in systematic planning, developing, testing, implementing and providing IT solutions for different domains which helps in the betterment of life.
100003/ CS722U.6-PO8	H	Students will be able to associate with a team as an effective team player for the development of technical projects by applying the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
100003/ CS722U.6-PO9	H	Students will be able to associate with a team as an effective team player to Identify, formulate, review research literature, and analyze complex engineering problems
100003/ CS722U.6-PO10	M	Students will be able to associate with a team as an effective team player for designing solutions to complex engineering problems and design system components.
100003/ CS722U.6-PO11	M	Students will be able to associate with a team as an effective team player, use research-based knowledge and research methods including design of experiments, analysis and interpretation of data.

100003/ CS722U.6-PO12	H	Students will be able to associate with a team as an effective team player, applying ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
100003/ CS722U.1-PSO1	H	Students are able to develop Computer Science Specific Skills by modeling and solving problems.
100003/ CS722U.2-PSO2	M	Developing products, processes or technologies for sustainable and socially relevant applications can promote Programming and Software Development Skills.
100003/ CS722U.3-PSO3	H	Working in a team can result in the effective development of Professional Skills.
100003/ CS722U.4-PSO3	H	Planning and scheduling can result in the effective development of Professional Skills.
100003/ CS722U.5-PSO1	H	Students are able to develop Computer Science Specific Skills by creating innovative solutions to problems.
100003/ CS722U.6-PSO3	H	Organizing and communicating technical and scientific findings can help in the effective development of Professional Skills.