Comprehensive Study of Deep Learning-Based Super-Resolution with Emphasis on GANs

A THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE
OF

MASTER OF TECHNOLOGY
IN
INFORMATION TECHNOLOGY

Submitted by

GAURAV SHUKLA (2K23/ITY/04)

Under the supervision of

Dr. RAHUL GUPTA and Dr. RITU AGARWAL

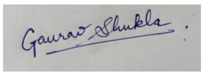
DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

MAY, 2025

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, **Gaurav Shukla**, Roll No's – **2K23/ITY/04** students of M.Tech (Information Technology), hereby declare that the project Dissertation titled "**Comprehensive Study of Deep Learning-Based Super-Resolution with Emphasis on GANs**" which is submitted by me to the Department of Information Technology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associateship, Fellowship or other similar title or recognition.



Place: Delhi — Gaurav Shukla

Date: 25.05.2025 — (2K23/ITY/04)

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "**Comprehensive Study of Deep Learning-Based Super-Resolution with Emphasis on GANs**" which is submitted by **Gaurav Shukla**, Roll No's – **2K23/ITY/04**, Department of Information Technology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Masters of Technology, is a record of the research work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

Place: Delhi — Dr. Rahul Gupta

Date: 25.05.2025 — (Supervisor)

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "**Comprehensive Study of Deep Learning-Based Super-Resolution with Emphasis on GANs**" which is submitted by **Gaurav Shukla**, Roll No's – **2K23/ITY/04**, Department of Information Technology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Masters of Technology, is a record of the research work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree or Diploma to this University or elsewhere.

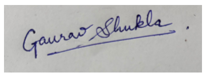
Place: Delhi — Dr. Ritu Agarwal

Date: 25.05.2025 — (Supervisor)

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I wish to express my sincerest gratitude to **Dr. Rahul Gupta** for his continuous guidance and mentorship that he provided me during the research work. He showed me the path to achieve my targets by explaining all the tasks to be done and explained to me the importance of this research as well as its industrial relevance. He was always ready to help me and clear my doubts regarding any hurdles in this project. Without his constant support and motivation, this research would not have been successful.



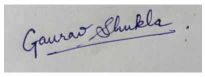
Place: Delhi — Gaurav Shukla

Date: 25.05.2025 — (2K23/ITY/04)

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I wish to express my sincerest gratitude to **Dr. Ritu Agarwal** for her continuous guidance and mentorship that she provided me during the research work. She showed me the path to achieve my targets by explaining all the tasks to be done and explained to me the importance of this research as well as its industrial relevance. She was always rea dy to help me and clear my doubts regarding any hurdles in this project. Without her constant support and motivation, this research would not have been successful.



Place: Delhi Gaurav Shukla

Date: 25.05.2025 (2K23/ITY/04)

Abstract

Image super-resolution using Generative Adversarial Networks (GANs) has been extensively researched in recent years due to its ability to recover high-perceptual-quality high-resolution images from low-resolution inputs. Various GAN-based methods have been proposed over the years, which employ different architectures and loss functions to increase the fidelity and realism of output images. This work integrates these developments and investigates their impact on various categories of images in various application domains. By extensive experimentation, we compare three highly acclaimed GAN-based super-resolution models SRGAN, ESRGAN, and Real-ESRGAN on twelve disparate image classes. The results confirm that the performance of the model varies significantly depending on the image features and domain, which calls for the need of domain-specific methods that are capable of learning to generalize across varying image content. To address these findings, we add a new component to loss functions with orthogonal regularization for, Wide Activation SRGAN (WDSR-GAN), which employs wide activation residual blocks to increase feature representa-tion and training stability. Furthermore, in this work we explore how various loss functions impact super-resolution quality and illustrate how various combinations impact image sharpness and perceptual detail. To quantitatively compare model performance, we use a collection of metrics consisting of PSNR and SSIM, which collectively capture pixel-level accuracy and structural integrity. The findings of this thesis provide valuable insights into the problems and opportuni- connections of GAN-based image super-resolution. By extensive analysis of different models and loss functions in different domains and metrics, this work lays a strong foundation for the design of more efficient and flexible super-resolution algorithms. Such efforts seek to steer future research towards more fidelity, improved perceptual quality, and increased adaptability to real-world imaging applications.

Contents

List of Tables

List of Figures

List of Symbols

θ	Model parameters
I_{HR}	High-resolution image
I_{LR}	Low-resolution image
G	Generator network
D	Discriminator network
$p_{\text{data}}(\mathbf{x})$	Real data distribution
$p_g(\mathbf{x})$	Generator distribution
E	Expectation operator
$\mathbf{x}$	Real image sample
$\mathbf{z}$	Input noise or low-resolution input
L_{adv}	Adversarial loss
L_{content}	Content loss
L_{perc}	Perceptual loss
L_D	Discriminator loss
L_{char}	Charbonnier loss
L_{TV}	Total variation loss
L_{rank}	Ranking loss
L_{id}	Identity loss
$\phi(\cdot)$	Feature extractor
$\text{VGG}(\cdot)$	VGG feature map
$\mathcal{F}(\cdot)$	Fourier transform
$\hat{x}$	Interpolated sample
m	Margin parameter
$s(\cdot, \cdot)$	Similarity score

Chapter 1

INTRODUCTION

1.1 Background and Motivation

In today's data driven world, the need for high-resolution visual content is growing rapidly across various domains such as medical diagnostics, satellite imaging, security surveillance, and digital media. However, acquiring high resolution images directly is not always feasible due to limitations in sensor hardware, transmission bandwidth, or environmental constraints. This challenge has driven research in the area of image super resolution, which aims to reconstruct a high-resolution image from a corresponding low resolution input.

Among various SR approaches, **Single Image Super Resolution** has gained significant attention due to its practical relevance. SISR deals with enhancing the resolution of an image using only one LR input, without relying on multiple observations. The goal is to generate HR images that are not only accurate in pixel-wise similarity but also aesthetically pleasing and realistic.

1.2 Super Resolution Problem

Single Image Super-Resolution is the process of recovering a high resolution image from one low resolution image. Since just one LR image is given as input, this method must deduce the HR image details from the limited information [1]. An image super resolution general up scaling equation can be given as:

$$I_{\mathrm{HR}} = f(I_{\mathrm{LR}}; \theta) \tag{1.1}$$

The high-resolution image, denoted as I_{HR}, is attained from the low resolution image I_{LR} using a super resolution function f with parameters θ. This function can be realized through different approaches, such as interpolation techniques, reconstruction based algorithms, and deep learning models.

1.3 Limitations of Traditional Methods

Earlier SR techniques typically relied on interpolation methods such as nearest neighbor, bilinear, and bicubic interpolation. While these methods are computationally efficient, they tend to produce overly smooth outputs and fail to recover intricate textures and fine details, especially in regions with complex patterns.

Other classical approaches based on dictionary learning or edge-preserving priors have shown modest improvements but often lack generalization across different image types. These limitations led researchers to explore data-driven methods that can learn the underlying mappings from LR to HR images.

1.4 Introduction to Generative Adversarial Networks

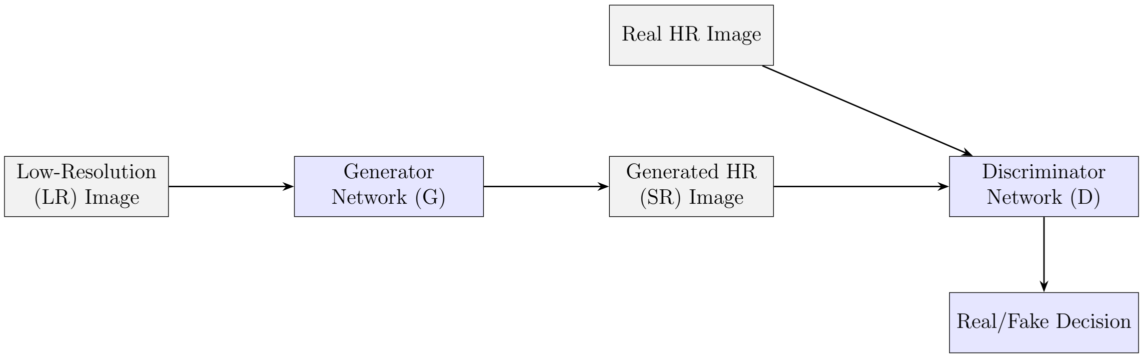


Figure 1.1: GAN Architecture for Super Resolution.

The concept of Generative Adversarial Networks was first introduced by Ian Goodfellow and colleagues in 2014 [2]. They proposed a novel framework for generative modeling comprising two neural networks: a generator (G) and a discriminator (D), which are trained simultaneously in an adversarial setup. The generator aims to synthesize data samples (e.g., images) from random noise, while the discriminator's role is to distinguish between real samples from the training dataset and fake samples produced by the generator. A depiction of this model architecture is shown in Figure 1.1.

In this setup, the generator starts with a noise vector $\mathbf{z}$, typically sampled from a Gaussian or uniform distribution, and maps it to a generated image $G(\mathbf{z})$. The discriminator, on the other hand, receives both real samples $\mathbf{x}$ and generated samples $G(\mathbf{z})$, and estimates the probability that a given sample is real. During training, the generator learns to produce increasingly realistic images that can fool the discriminator, while the discriminator becomes better at distinguishing between real and fake data. This competitive process, known as adversarial training, leads to a mutual improvement of both networks, ideally resulting in synthesized outputs that closely resemble real data.

1.4.1 Adversarial Training Objective

The GAN training process is formulated as a two player mini-max game with the following value function:

$$\min_G \max_D V(D,G) = E_{\mathbf{x}\sim p_{\text{data}}(\mathbf{x})}\left[\log D(\mathbf{x})\right] + E_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z})}\left[\log(1 - D(G(\mathbf{z})))\right] \tag{1.2}$$

Here, $p_{\text{data}}(\mathbf{x})$ represents the real data distribution, and $p_{\mathbf{z}}(\mathbf{z})$ denotes the prior distribution of the input noise vector. The discriminator D aims to maximize its ability to correctly classify real and generated samples, while the generator G tries to minimize the discriminator's performance by generating more realistic outputs.

As training progresses, both networks improve through alternating updates. The discriminator is refined to better distinguish genuine high resolution images from generated ones, and the generator is optimized to produce images that are increasingly indistinguishable from real data. When training reaches equilibrium, the discriminator's output converges to:

$$D(\mathbf{x}) = \frac{p_{\text{data}}(\mathbf{x})}{p_{\text{data}}(\mathbf{x}) + p_g(\mathbf{x})} \tag{1.3}$$

where $p_g(\mathbf{x})$ is the distribution of the data generated by G. Ideally, the generator learns the true data distribution, i.e., $p_g(\mathbf{x}) = p_{\text{data}}(\mathbf{x})$, and the discriminator becomes incapable of distinguishing real from fake samples, outputting a probability of 0.5 for all inputs Variance. [3] [4].

1.5 Scope and Objectives of the Study

This thesis presents a comprehensive investigation into deep learning-based image super-resolution (SR) techniques, with a particular focus on Generative Adversarial Network based models. The study aims to bridge the gap between distortion focused and perception-driven SR approaches by critically examining state of the art GAN architectures and their loss functions. The main objectives of this research are:

- To trace the progression of super resolution methodologies, from traditional interpolation and reconstruction techniques to advanced deep learning and GAN based models.
- To analyze the architectural innovations and loss function strategies that have contributed to the success of prominent GAN based super resolution networks.
- To implement a selected set of GAN based SR models within a unified experimental framework, ensuring fair and consistent comparison.
- To rigorously evaluate these models using both distortion oriented metrics such as Peak Signal to Noise Ratio and Structural Similarity Index, and perception oriented metrics including Feature Similarity Index, Learned Perceptual Image Patch Similarity, and Tenengrad Sharpness.
- To provide an indepth qualitative and quantitative analysis of each model's strengths and limitations with respect to perceptual fidelity and reconstruction accuracy.

1.6 Thesis Organization

The thesis is structured to systematically guide the reader through the background, methodology, experimentation, and conclusions as follows:

- **Chapter 2: Literature Review:** An extensive survey of existing super resolution techniques, encompassing traditional approaches, convolutional neural network based methods, and GAN based models, highlighting their evolution and key contributions.

- **Chapter 3: Methodology:** Details the experimental setup, including dataset selection, model architectures, training protocols, and the evaluation metrics employed for comprehensive assessment.
- **Chapter 4: Results and Discussion:** Presents the empirical findings, compares performance across multiple metrics, showcases qualitative visual results, and discusses the implications of the observations.
- **Chapter 5: Conclusion and Future Scope:** Summarizes the principal outcomes of the study and proposes potential avenues for future research and improvements.
- **Chapter 6: References:** Compiles all scholarly works referenced throughout the thesis.

Chapter 2

LITERATURE REVIEW

2.1 Overview of Super Resolution Techniques

Image Super Resolution has seen significant advancements over the last ten years, evolving from basic interpolation methods to advanced deep learning frameworks. This section offers a detailed overview of both traditional and contemporary SR techniques, with an emphasis on deep learning and Generative Adversarial Network (GAN) based methods.

2.2 Traditional Super Resolution Methods

Early single image super resolution methods primarily relied on interpolation techniques such as bicubic or Lanczos interpolation, or learning-based priors. While these conventional approaches were straightforward and computationally efficient, they often struggled to reconstruct high frequency textures and fine details, resulting in overly smooth or blurry images. The domain experienced a significant breakthrough with the arrival of deep learning. Dong *et al.* [5] introduced the Super-Resolution Convolutional Neural Network (SRCNN), a novel three-layer CNN architecture that directly learns an end-to-end mapping from low resolution patches to their high resolution pairs. This data-driven method marked a appraoch shift in super resolution by enabling automatic feature extraction and reconstruction, overcoming many limitations of earlier approaches.

Single-image super resolution remains a fundamental challenge in computer vision, aiming to reconstruct a high-quality image from a single low resolution input. Modern deep learning based techniques, particularly those incorporating convolutional neural networks and generative adversarial networks, have significantly improved this field. These methods learn rich feature representations and complex mappings that preserve textures and visual realism, producing results that far surpass traditional interpolation methods in terms of detail and perceptual quality.

2.3 Deep Learning Based Super Resolution

Building upon the groundbreaking SRCNN [5], which introduced a three layer CNN to directly map low resolution images to high resolution outputs, numerous deeper and more complex convolutional neural network architectures emerged to further improve super resolution performance. Kim *et al.* [6] proposed VDSR, a 20 layer residual network that enhanced reconstruction accuracy by exploiting deep features and skip connections, leading to improved PSNR and SSIM scores. Dong *et al.* [7] introduced FSRCNN, which accelerated inference by performing upsampling through deconvolution layers. Lai *et al.* [8] introduced LapSRN, a pyramid based framework that incrementally restores image details through multi level supervision. Haris *et al.* [9] designed DBPN, which iteratively refines predictions via multiple upsampling and downsampling stages. Lim *et al.* [10] developed EDSR by removing batch normalization and scaling model depth and width, achieving new state of the art PSNR benchmarks. Further advancements include Residual Dense Networks (RDN) [11] and the channel attention based RCAN [12], which enhanced feature extraction and representation capabilities.

Despite significant improvements in distortion metrics such as PSNR and SSIM, these models commonly rely on pixel-wise ℓ_2 loss functions, which tend to produce overly smooth or blurry images that lack fine perceptual details [3]. Subsequent architectures like Wide Activation Super Resolution (WDSR) [13] introduced wide activation functions and refined residual learning strategies to improve feature propagation and reconstruction quality. However, challenges remain in balancing quantitative accuracy with perceptual realism, motivating the integration of adversarial and perceptual loss functions in recent super-resolution models.

2.4 Generative Adversarial Network Based Super Resolution

Generative Adversarial Networks (GANs) have gained significant attention in single image super resolution (SISR) for their ability to generate sharper and more visually appealing images compared to conventional methods. The pioneering work, SRGAN [14], was the first to introduce adversarial training in super resolution, where a discriminator network guides the generator to synthesize more realistic textures. SRGAN combines an adversarial loss with a perceptual content loss computed on feature maps extracted from a pre-trained VGG network. This combination enables the generated images to better preserve high-level semantic features, resulting in outputs that appear more photo realistic than those trained with traditional pixel wise losses.

Building upon SRGAN, ESRGAN [15] further refined the approach by adopting deeper Residual in Residual Dense Blocks (RRDB), removing batch normalization to improve stability, and introducing a relativistic adversarial loss that encourages the discriminator to evaluate how much more realistic a real image is relative to a generated one. These innovations led ESRGAN to win the PIRM-SR challenge in 2018, producing images with superior perceptual quality.

Parallel research by Johnson *et al.* [16] demonstrated that feature space losses, such as VGG-based ℓ_2 loss, better preserve semantic structures compared to traditional pixel wise losses, thereby improving perceptual similarity. Sajjadi *et al.* [17] proposed EnhanceNet, which emphasized automated texture synthesis to generate visually rich textures. While these GAN-based methods substantially improve perceptual quality, they often suffer from common GAN training challenges such as instability, mode collapse, and the introduction of visual artifacts.

Further advancements addressed real-world SR challenges. Real-ESRGAN [18] extended ESRGAN by incorporating a synthetic degradation model and employing a U-Net discriminator, along with an identity loss to ensure content consistency with the low resolution input. This made the model more robust to diverse and practical degradation types.

Overall, GAN-based super resolution models prioritize perceptual quality over conventional distortion metrics like PSNR and SSIM, reflecting the inherent perception-distortion tradeoff described by Blau and Michaeli [3]. Despite their success in generating visually pleasing images, GAN-based approaches continue to face challenges related to stable training and artifact reduction, motivating ongoing research in improved loss functions and network architectures.

2.5 Advancements in Loss Functions for GAN-Based Super-Resolution

GAN-based techniques introduced a fresh perspective that emphasizes perceptual realism over pixel by pixel similarity. This section examines key GAN-based super-resolution models and the evolution of their loss functions.

SRGAN: SRGAN [19] was the first GAN-based model designed for super resolution tasks. It introduced two main loss components: adversarial loss and perceptual content loss. The perceptual loss leverages feature maps extracted from a pretrained VGG network to ensure that the generated image aligns with the ground truth in terms of high level semantic features. Concurrently, the adversarial loss trains the generator to produce images indistinguishable from real high resolution images, as judged by a discriminator. This dual loss approach enabled SRGAN to generate outputs with enhanced visual realism compared to traditional pixel-based models.

EnhanceNet: EnhanceNet [17] incorporated a texture focused loss function that prioritizes perceptual similarity. In addition to the perceptual loss, it introduced a texture matching loss based on the Gram matrix of intermediate feature maps. This approach encourages the network to retain natural textures, leading to visually rich and detailed image regions.

ESRGAN: ESRGAN [15] introduced multiple improvements over SRGAN, both architecturally and in terms of its loss functions. It employed a relativistic average GAN loss, which redefined the discriminator's objective to assess how much more realistic real images are compared to generated ones. Additionally, ESRGAN replaced the MSE loss with L1 loss in the feature space of a deeper VGG network, resulting in outputs with significantly improved perceptual quality.

SRPGAN: SRPGAN [20] focused on preserving edge details and minimizing visual artifacts. It adopted the Charbonnier loss, a robust alternative to L1 loss that is less sensitive to outliers and provides more stable training. Furthermore, SRPGAN incorporated a total variation loss to enforce spatial smoothness among neighboring pixels. These enhancements helped produce sharper images, especially in regions with fine structures and edges.

RankSRGAN: RankSRGAN [21] introduced a novel perceptual ranking loss. A separate ranking network is trained to learn human-like preferences among various super resolved outputs. The perceptual ranking loss then guides the generator to produce images that align better with human visual perception, rather than relying solely on the discriminator. This strategy marked a shift toward more subjectively pleasing image generation.

SPSR: SPSR (Structure-Preserving SR) [22] aims to preserve both the fine details and the overall structure of images. It integrates texture priors into the loss function to maintain high frequency content without distorting object shapes or spatial arrangements.

SPSR combines adversarial, perceptual, and prior-based losses to balance texture realism with structural accuracy especially beneficial for scenes involving architecture and man-made objects.

Real-ESRGAN: Real-ESRGAN [18] adapts ESRGAN for real world low resolution images. In addition to the adversarial and perceptual losses, it introduces an identity loss to ensure the generated image remains aligned with the input content. This addition enhances the model's robustness to practical degradations encountered in real-world applications.

WGAN-SR: WGAN-SR [23] addresses instability that is normally experienced in training GANs. It replaces the standard GAN loss with the Wasserstein loss, and this leads to a smoother and more interpretable measure of divergence between real and generated image distributions. To satisfy the Lipschitz continuity constraint of Wasserstein GANs, a gradient penalty is also incorporated. The changes enhance training stability and output with improved visual fidelity.

The evolution of loss functions in super-resolution based on GAN takes a different direction from pixel-wise accuracy towards perceptual realism and human visual consistency. Current methods tend to utilize a combination of loss terms ranging from texture and ranking to perceptual and adversarial goals in order to produce high-fidelity, visually plausible results across various areas of application. A compilation of loss functions of the above GAN models are summarized in Table **??**.

2.6 Summary

This Literature review documents the evolution of super resolution techniques from traditional interpolation to recent deep learning methods with emphasis on the role of Generative Adversarial Networks. The previous methods were restricted by the inability to recover fine textures, which was overcome by CNN based architectures like SRCNN, VDSR, and EDSR that enabled significant PSNR and SSIM improvements. These architectures, however, produced results that were overly smoothed out due to the application of pixel wise loss functions. The arival of SRGAN shifted attention to perceptual quality using adversarial and perceptual losses to generate visually realistic results. Successor models like ESRGAN, EnhanceNet, RankSRGAN, and Real-ESRGAN built on this approach by using new architectures and advanced loss functions ranging from relativistic adversarial loss to texture, ranking, and identity losses. Collectively, these trends signify a shift from accuracy centered to perceptually oriented super resolution models with emphasis on loss design to generate high fidelity super resolved images.

Model	Loss Type	Formula (Simplified)
SRGAN	Content Loss (MSE)	$L_{\text{content}} = \|G(z) - x\|_2^2$
	Perceptual Loss	$L_{\text{perc}} = \|\text{VGG}(x) - \text{VGG}(G(z))\|_2^2$
	Adversarial Loss	$L_{\text{adv}} = -E\left[\log D(G(x))\right]$
EnhanceNet	Perceptual Loss	$L_{\text{perc}} = \|\text{VGG}(x) - \text{VGG}(G(z))\|_2^2$
	Texture Matching Loss (Gram)	$L_{\text{texture}} = \|G(\phi(I)) - G(\phi(x))\|_2^2$
ESRGAN	Content Loss (L1)	$L_{\text{content}} = \|G(z) - x\|_1$
	Relativistic Adversarial Loss	$L_D = -E\left[\log(D(y) - D(G(x)))\right] - E\left[\log(1 - (D(G(x)) - D(y)))\right]$
SRPGAN	Charbonnier Loss	$L_{\text{char}} = \sqrt{(G(x) - y)^2 + \epsilon^2}$
	Total Variation Loss	$L_{\text{TV}} = \sum_{i,j}\left((G_{i+1,j} - G_{i,j})^2 + (G_{i,j+1} - G_{i,j})^2\right)$
RankSRGAN	Perceptual Ranking Loss	$L_{\text{rank}} = \max\left(0, m - s(G(x), y_{\text{better}}) + s(G(x), y_{\text{worse}})\right)$
	Adversarial Loss	Same as SRGAN adversarial loss
SPSR	Adversarial Loss	Same as SRGAN adversarial loss
	Perceptual Loss	Same as SRGAN perceptual loss
	Texture Prior Loss	Based on texture statistics difference
Real-ESRGAN	Identity Loss	$L_{\text{id}} = \|G(z) - y\|_1$
	Relativistic Adversarial Loss	Same as ESRGAN
	Perceptual Loss	Same as ESRGAN
WGAN-SR	Wasserstein Adversarial Loss	$L_D = E[D(G(x))] - E[D(y)] + \lambda E\left[(\|\nabla_{x'} D(x')\|_2 - 1)^2\right]$
	Perceptual Loss	Same as SRGAN perceptual loss

Table 2.1: Curated Loss Functions of GAN-based Super-Resolution Models

Chapter 3

METHODOLOGY

3.1 Research Objectives

This chapter discusses the methodology applied to measure the performance of GAN-based super-resolution models in different domains. The objective of the analysis in this chapter is to further the understanding of how such models perform in different applications in real life, both in natural and synthetic image domains. The chapter further outlines the design and training of a new GAN variant Wide Activation SRGAN (WDSR-GAN) incorporating architectural enhancements as well as regularization techniques for better stability and performance.

The focused objectives of this project are as follows:

1. To evaluate the generalisation capacity and performance of existing GAN-based super-resolution models on a multi-domain dataset.

2. To introduce and train a novel WDSR-GAN architecture that enhances feature extraction and training stability by incorporating an improved loss function.

3. To identify and make use of a wide variety of quantitative measurement metrics to measure model performance in varied loss function, with SRGAN as the point of reference

3.2 Evaluation of GAN Based Super Resolution Across Domains

This part provides a domain-specific analysis of the top GAN-based super-resolution models to analyze their effectiveness and versatility in a vast variety of types of visual data. The goal is to test these pretrained models on a variety of image domain datasets. The assessment framework consists of a diverse dataset, established models, and a standard inference process to provide an unbiased and thorough comparison.

3.2.1 Dataset Accumulation

To critically evaluate super-resolution models across a broad range of applications, we compiled a large image dataset from multiple publicly available sources. The dataset includes high-resolution images across different domains such as natural scene, medical imaging (e.g., MRI and X-ray images), digital art, and more. Each chosen image ensures high visual fidelity and is representative of its category, offering a variety of textures, structures, and patterns. This broad selection enables a realistic assessment of each model's potential to generalize beyond its original training distribution.

The complete dataset and related documentation are available at https://github.com/gauravshuklacpp/SpanSR-12. Table 3.1 outlines the source and domain classification of the sub-datasets used in the evaluation.

Category	Dataset
Natural Scenes	Places365 [24]
Buildings	Urban100 [25]
Human Faces	CelebA [26]
Animals	Oxford-IIIT Pet [27]
Artwork	WikiArt GAN [28]
Medical Imagery	NIH Chest X-rays [29], Brain MRI [30]
Aerial Images	DOTA [31]
Text Documents	IIIT5K Words [32]
Microscopic Photography	Microorganism Image Classification [33]
Underwater Images	Underwater ImageNet [34]
Anime	Anime Images [35]
Cars	Cars Image Dataset [36]

Table 3.1: Datasets Used in Evaluation of GAN Models

All images were resized to a consistent resolution and normalized to maintain uniformity. Basic data augmentation strategies such as random cropping, flipping, and rotation were employed during testing to simulate natural variation and increase robustness.

3.2.2 Model Selection

To reflect the development of GAN-based super-resolution methods over time, we selected three prominent models, each of which represents a milestone in architectural innovation and performance.

SRGAN: Super Resolution GAN

SRGAN was the first to integrate adversarial loss with perceptual content loss for image super-resolution. Its generator architecture employs residual blocks along with pixel shuffle based up sampling. The discriminator drives the generator to produce outputs that appear more photo-realistic. The perceptual loss, computed using feature maps extracted from a pretrained VGG network, enables the reconstruction of perceptually significant features. However, SRGAN often fails to capture fine-grained textures in complex regions.

ESRGAN: Enhanced SRGAN

ESRGAN builds upon SRGAN by introducing Residual in Residual Dense Blocks (RRDBs), which help preserve contextual features and stabilize training. The standard discriminator is replaced with a relativistic average discriminator that judges the relative realness of generated images. These modifications significantly improve detail reconstruction, especially for textured and high frequency regions such as human faces and natural scenery.

Real-ESRGAN

Real-ESRGAN extends ESRGAN to real-world applications by training on both synthetic and authentically degraded images. It includes an advanced degradation model that simulates blur, noise, and compression artifacts. A multi-scale discriminator, combined with improved loss formulations, helps the model handle diverse and challenging input degradations. This robustness makes it particularly suitable for critical domains such as medical imaging and underwater photography.

3.2.3 Inference Configuration

All evaluations were performed using the publicly released pretrained weights provided by the respective model authors. No fine-tuning or retraining was applied. To ensure fairness, each model received the same set of low-resolution inputs, which were generated using bicubic downsampling from the high-resolution ground truth.

Inference was conducted under identical conditions for all models and categories. The goal was to test each model's capacity to generalize to the out-of-domain with or without domain adaptation.

3.3 Wide Activation SRGAN with Orthogonal Regularization

This section presents an extension of the SRGAN model with Wide Activation Residual Blocks and Orthogonal Regularization to enhance stability and precision in super-resolution results. The architecture, drawing inspiration from the WDSR-GAN model by Yu et al. [37], targets improved information transmission, computation cost savings, and training regularization by orthogonal constraints.

3.3.1 Wide Activation Generative Adversarial Network

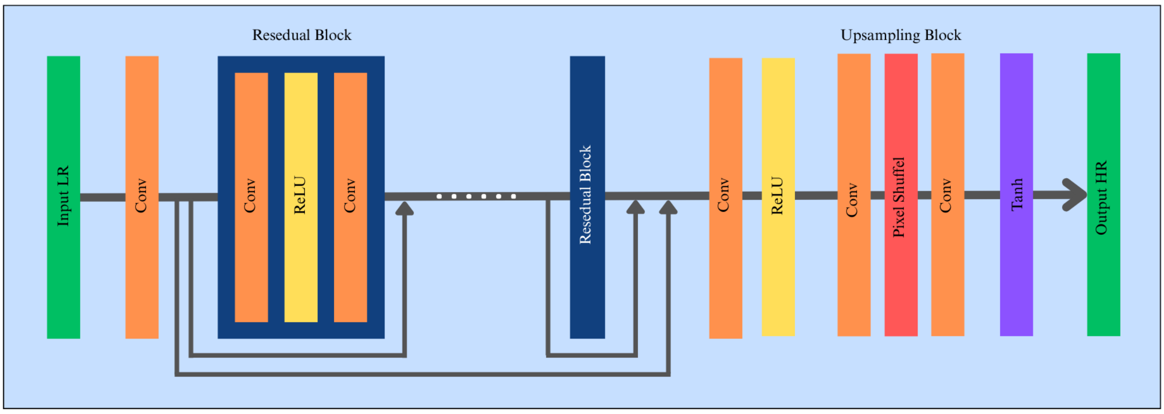


Figure 3.1: Architectural Representation of Generator

With our method, we utilize the WDSR-GAN (Wide Activation Super-Resolution Generative Adversarial Network) model to improve image resolution in an efficient and accurate manner. Initially introduced by Jiahui Yu et al. [37], WDSR-GAN employs wide activation functions to enable high-quality image reconstruction. With greater information flow from shallow layers to deep layers, the technique enhances performance without adding computational complexity.

The WDSR-A network, being the main variant, enlarges activation widths ($2\times$ to $4\times$) in residual blocks to facilitate enhanced information transmission and enhanced outcomes without increasing computational demands. Building on this, the WDSR-B network continues to push efficiency with the implementation of linear low-rank convolutions as its core element. This structure enables even wider activation (from $6\times$ to $9\times$) without additional parameters or computations, thereby enhancing the accuracy of the super resolution process. For our specific use case, we borrow inspiration from the WDSR-A GAN model, leveraging its structured approach to wide activation in residual blocks to achieve state-of-the-art image super-resolution. Our GAN is trained on two significant loss functions, which offer stability coupled with high-fidelity reconstruction.

Adversarial Loss: The adversarial loss is developed using Binary Cross-Entropy and is stated as:

$$\mathcal{L}_{\text{GAN}} = E_{\mathbf{x}\sim p_{\text{data}}}[\log D(\mathbf{x})] + E_{\mathbf{z}\sim p_{\mathbf{z}}}[\log(1 - D(G(\mathbf{z})))] \tag{3.1}$$

where D represents the discriminator, G is the generator, $\mathbf{x}$ is a actual real image, and $\mathbf{z}$ is a noise vector. The loss function pushes the generator to produce images that are very close to real images, improving the global quality of synthetic outputs.

L1 Loss: To maintain both visual accuracy and structural similarity with the reference images, the L1 loss function is utilized. It is mathematically expressed as:

$$\mathcal{L}_{\text{L}_1} = E_{\mathbf{x},\mathbf{y}}[|G(\mathbf{x}) - \mathbf{y}|] \tag{3.2}$$

where $\mathbf{x}$ is the low-resolution input and $\mathbf{y}$ is the high-resolution ground truth image. The generator learns to create high-fidelity outputs with improved structural correctness via minimization of this loss.

Through the integration of these loss functions, WDSR-GAN effectively reconstructs fine details in images, making it a powerful framework for super-resolution applications.

3.3.2 Orthogonal Regularization

Orthogonal regularization is a technique used to enforce orthogonality constraints on the weight matrices of neural networks, which helps in stabilizing training and improving generalization. In our implementation, we apply orthogonal regularization to specific layers in both the discriminator and the generator networks. This regularization technique ensures that the weight matrices of these layers remain close to orthogonal, which can be mathematically expressed as $W^T W \approx I$, where W is the weight matrix and I is the identity matrix. The regularization term is described as the Frobenius norm of the difference between $W^T W$ and the identity matrix:

$$\mathcal{L}_{\text{ortho}} = \|W^T W - I\|_F^2 \tag{3.3}$$

In practical implementation, the regularization term is incorporated into the loss functions of both the generator and discriminator. A small coefficient is assigned to this term to prevent it from dominating the primary loss components. For the discriminator, the modified loss function is given by:

$$\mathcal{L}_{\text{D}} = \mathcal{L}_{\text{GAN}} + \lambda_{\text{ortho}} \mathcal{L}_{\text{ortho}} \tag{3.4}$$

where $\mathcal{L}_{\text{GAN}}$ represents the adversarial loss, and λ_{ortho} is a small hyperparameter that regulates the influence of the orthogonal regularization term.

Similarly, the generator's loss function is adjusted to incorporate the orthogonal regularization component:

$$\mathcal{L}_{\text{G}} = \mathcal{L}_{\text{GAN}} + \alpha \mathcal{L}_{\text{L1}} + \lambda_{\text{ortho}} \mathcal{L}_{\text{ortho}} \tag{3.5}$$

where $\mathcal{L}_{\text{L1}}$ denotes the L1 loss component, which validates that the generated images remain near to the ground truth at the pixel level. In this work, we set the weighting parameter $\alpha = 0.01$ to control the addition of the L1 loss in the overall objective function.

Orthogonal regularization is applied with a coefficient of $\lambda = 1 \times 10^{-6}$, promoting near-orthogonality in the weight matrices of both the generator and discriminator throughout training. This regularization strategy contributes to training stability, reduces the likelihood of mode collapse, and enhances the WDSR-GAN model's capacity to generate high quality super resolved images.

3.4 Loss Function Comparison in SRGAN Based Super Resolution

This section outlines a structured experimental framework designed to evaluate the individual and combined effects of different loss functions within the SRGAN based single image superresolution pipeline. The original SRGAN architecture is consistently employed as a baseline to isolate the impact of each loss configuration.

3.4.1 Dataset and Preprocessing

Training Dataset

The training data is sourced from the DIV2K dataset [38], which contains 800 high resolution images. These images are down sampled using bicubic interpolation with a scaling factor of ×4 to generate low-resolution counterparts. This high quality dataset gives a diverse set of images, ensuring the model learns to generalize across various textures and content types.

Testing Datasets

Performance evaluation is carried out using four standard benchmark datasets:

- **Set5** [39]: A small dataset with relatively simple and clean image structures, often used for initial testing.
- **Set14** [40]: Contains more diverse and naturally occurring scenes, providing a wider range of features for generalization assessment.
- **BSD100** [41]: Offers a broad set of scenes derived from the Berkeley segmentation dataset, known for its visual complexity.
- **Urban100** [25]: Comprises urban scenes rich in detailed structures such as buildings and roads, serving as a rigorous benchmark for texture recovery.

All datasets are evaluated with a uniform ×4 upscaling requirement.

3.4.2 Model Architecture and Loss Integration

Baseline Model

All experiments utilize the original SRGAN model proposed by Ledig et al. [19], featuring a generator composed of deep residual blocks that map low-resolution inputs to high resolution outputs. The discriminator guides this process by differentiating between real and generated images, thereby fostering photorealistic reconstructions. This architecture was chosen for its historical significance and well documented baseline performance.

Loss Configurations

To study the effects of different loss functions on SR performance, six unique configurations are tested. In all cases, adversarial loss remains the core training objective. It is systematically combined with various auxiliary losses to assess their influence on fidelity, texture, and perceptual quality. The Table 3.2 shows a summary of combination of loss variations used.

ID	Loss Configuration	Description
A	Adversarial Only	GAN loss without any auxiliary regularization
B	Adversarial + L1	Promotes pixel-level similarity using L1 distance
C	Adversarial + L2	Uses L2 loss to enforce mean-squared error fidelity
D	Adversarial + Perceptual	Incorporates VGG feature-based perceptual loss
E	Adversarial + Contextual	Adds contextual loss to improve spatial semantics
F	Adversarial + Texture	Utilizes Gram matrix loss to enhance textures

Table 3.2: Loss Function Variants Evaluated

Each variant undergoes separate training and evaluation to ensure isolated and fair comparison.

3.4.3 Training Protocol

Every models are trained for 20 epochs, for a duration chosen to allow meaningful convergence while preserving computational feasibility. A batch size of 16 is used to ensure an optimal balance between memory efficiency and model performance. Training is performed using the Adam optimizer with parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$, and an initial learning rate of 1×10^{-4}, which is reduced by a factor of 0.5 every 10 epochs.

Generator and discriminator networks are updated in alternating steps to maintain adversarial stability throughout training. To enhance generalization and robustness, standard data enhancement techniques such as horizontal flipping and random cropping are applied during training.

3.5 Performance Metrics

To comprehensively evaluate the effectiveness of the GAN-based super-resolution models, both quantitative and qualitative metrics are employed:

- **Peak Signal-to-Noise Ratio:** Measures the pixel-wise reconstruction quality by comparing the generated high-resolution image I_{SR} to the ground truth image I_{HR}. It is defined as:

$$\mathrm{PSNR} = 10 \cdot \log_{10}\left(\frac{L^2}{\mathrm{MSE}}\right) \tag{3.6}$$

  where L is the maximum possible pixel value of the image (e.g., 255 for 8-bit images), and MSE (Mean Squared Error) is

$$\mathrm{MSE} = \frac{1}{WH}\sum_{i=1}^{W}\sum_{j=1}^{H}\left(I_{HR}(i,j) - I_{SR}(i,j)\right)^2 \tag{3.7}$$

  with W and H being the image width and height, respectively.

- **Structural Similarity Index Measure:** Evaluates perceptual similarity by comparing luminance, contrast, and structural components between I_{SR} and I_{HR}. The

SSIM index is computed as:

$$\text{SSIM}(I_{HR}, I_{SR}) = \frac{(2\mu_{I_{HR}}\mu_{I_{SR}} + C_1)(2\sigma_{I_{HR}I_{SR}} + C_2)}{(\mu^2_{I_{HR}} + \mu^2_{I_{SR}} + C_1)(\sigma^2_{I_{HR}} + \sigma^2_{I_{SR}} + C_2)} \tag{3.8}$$

where μ denotes mean intensity, σ^2 variance, $\sigma_{I_{HR}I_{SR}}$ covariance between images, and C_1, C_2 are constants to stabilize the division.

- **Learned Perceptual Image Patch Similarity:** Calculates perceptual similarity by evaluating deep features retrieved from pretrained networks such as VGG. Given deep features $f_l(I_{HR})$ and $f_l(I_{SR})$ at layer l, LPIPS is computed as a weighted distance:

$$\text{LPIPS}(I_{HR}, I_{SR}) = \sum_l w_l \cdot \|\hat{f}_l(I_{HR}) - \hat{f}_l(I_{SR})\|_2^2 \tag{3.9}$$

where $\hat{f}_l$ are normalized feature activations and w_l are learned weights. Lower values indicate better perceptual similarity.

- **Fréchet Inception Distance:** Calculates the distance between the feature distributions of a real and generated images in the Inception network feature space. If $\mathcal{N}(\mu_r, \Sigma_r)$ and $\mathcal{N}(\mu_g, \Sigma_g)$ Illustrate the Gaussian approximations of real and generated image features, then:

$$\text{FID} = \|\mu_r - \mu_g\|_2^2 + \text{Tr}\left(\Sigma_r + \Sigma_g - 2(\Sigma_r\Sigma_g)^{1/2}\right) \tag{3.10}$$

where μ_r, μ_g are the mean vectors, Σ_r, Σ_g the covariance matrices, and Tr the trace operator. Lower FID indicates closer distributions and better quality.

- **Tenengrad Sharpness:** Measures the image sharpness by computing the sum of squared gradient magnitudes above a threshold T. Given gradients G_x and G_y obtained from Sobel operators, the gradient magnitude at pixel (i, j) is

$$G(i, j) = \sqrt{G_x(i, j)^2 + G_y(i, j)^2} \tag{3.11}$$

and Tenengrad sharpness S is defined as:

$$S = \sum_{i,j} \left(G(i, j)^2 \cdot \mathbf{1}_{\{G(i,j)>T\}}\right) \tag{3.12}$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function selecting pixels with gradient magnitude above the threshold. Higher S values indicate sharper images with more prominent edges.

These metrics collectively provide a robust framework to assess reconstruction accuracy, perceptual quality, and the realism of super-resolved images produced by different GAN models.

3.6 Summary

This chapter outlines the approach to evaluate GAN-based super-resolution models across diverse image domains, including natural, medical, artwork, and more, using a comprehensive multi-domain dataset. Three prominent GAN models SRGAN, ESRGAN, and

Real-ESRGAN are chosen for evaluation using their pretrained weights to assess their generalization without fine-tuning.

Additionally, a novel Wide Activation SRGAN (WDSR-GAN) is introduced, leveraging wide activation residual blocks and orthogonal regularization to enhance training stability and image quality. The architecture combines adversarial loss and L1 loss with orthogonal constraints applied to the network weights.

Finally, a structured experimental framework is described to compare various loss function configurations within the SRGAN baseline, using the DIV2K dataset for training and four standard benchmark datasets (Set5, Set14, BSD100, Urban100) for testing, focusing on the impact of different losses on super-resolution performance.

Chapter 4

RESULTS and DISCUSSION

4.1 Evaluation of GAN Based Super Resolution Across Domains

This section evaluates the performance of three prominent GAN-based super-resolution models SRGAN, ESRGAN, and Real-ESRGAN across diverse image domains. The analysis is conducted using a 4× upscaling factor and is benchmarked using five quantitative metrics: PSNR, SSIM, FSIM, Laplacian Variance, and Tenengrad Sharpness. These metrics collectively assess pixel-level fidelity, structural preservation, feature integrity, image sharpness, and edge clarity.

4.1.1 Experimentation Setup

All experiments were performed using a system equipped with an NVIDIA L4 GPU featuring 24 GB of GDDR6 memory, ensuring efficient handling of computational demands.

4.1.2 Tabular Results of Generated Images

Each image domain presents unique challenges for super-resolution. For instance, medical and microscopic images require high structural fidelity, while anime and artwork emphasize texture and stylistic consistency. The models were evaluated on ten image types: Animals, Anime, Aerial, Artwork, Buildings, Cars, Human Face, Medical, Microscopic, and Nature. Table 4.1 summarizes the comparative performance of various image domains.

Image Type	Metric	SRGAN	ESRGAN	Real-ESRGAN
Animals	PSNR	29.41	32.83	29.04
	SSIM	0.82	0.90	0.83
	FSIM	0.95	0.97	0.95
	Laplacian Variance	477.93	100.71	153.40
	Tenengrad Sharpness	39.35	34.64	34.83
Anime	PSNR	32.50	35.21	25.25
	SSIM	0.94	0.94	0.81
	FSIM	0.97	0.98	0.95

Continued on next page

Table 4.1 – continued from previous page

Image Type	Metric	SRGAN	ESRGAN	Real-ESRGAN
	Laplacian Variance	71.38	58.16	429.96
	Tenengrad Sharpness	33.10	34.92	40.24
Aerial	PSNR	24.72	25.63	24.18
	SSIM	0.66	0.73	0.66
	FSIM	0.89	0.90	0.89
	Laplacian Variance	756.43	318.74	462.67
	Tenengrad Sharpness	55.32	50.96	44.34
Artwork	PSNR	27.28	29.58	27.91
	SSIM	0.60	0.71	0.64
	FSIM	0.93	0.94	0.93
	Laplacian Variance	489.58	79.65	181.05
	Tenengrad Sharpness	44.73	35.15	32.58
Buildings	PSNR	21.56	22.29	19.45
	SSIM	0.70	0.75	0.65
	FSIM	0.89	0.90	0.86
	Laplacian Variance	1435.17	693.66	2249.75
	Tenengrad Sharpness	84.83	83.35	90.27
Cars	PSNR	33.33	35.99	24.51
	SSIM	0.95	0.96	0.74
	FSIM	0.97	0.98	0.92
	Laplacian Variance	75.58	43.90	714.37
	Tenengrad Sharpness	32.50	34.42	48.75
Human Face	PSNR	36.10	40.65	29.31
	SSIM	0.95	0.97	0.87
	FSIM	0.98	0.99	0.96
	Laplacian Variance	26.28	26.33	57.81
	Tenengrad Sharpness	16.55	17.59	20.73
Medical	PSNR	34.26	38.35	30.76
	SSIM	0.82	0.95	0.88
	FSIM	0.95	0.98	0.93
	Laplacian Variance	138.21	28.18	107.39
	Tenengrad Sharpness	20.51	19.11	20.64
Microscopic	PSNR	32.58	35.23	30.03
	SSIM	0.90	0.94	0.89
	FSIM	0.97	0.98	0.96
	Laplacian Variance	153.84	54.96	77.46
	Tenengrad Sharpness	25.19	24.60	23.22

Continued on next page

Table 4.1 – continued from previous page

Image Type	Metric	SRGAN	ESRGAN	Real-ESRGAN
Nature	PSNR	32.55	37.25	24.21
	SSIM	0.94	0.97	0.69
	FSIM	0.97	0.98	0.90
	Laplacian Variance	162.02	44.27	1741.60
	Tenengrad Sharpness	34.89	36.04	65.86
Text	PSNR	35.09	41.97	23.33
	SSIM	0.96	0.98	0.82
	FSIM	0.97	0.99	0.90
	Laplacian Variance	24.65	19.35	282.98
	Tenengrad Sharpness	15.13	16.14	19.37
Underwater	PSNR	35.12	40.01	30.20
	SSIM	0.92	0.97	0.86
	FSIM	0.97	0.98	0.96
	Laplacian Variance	31.87	18.83	63.76
	Tenengrad Sharpness	17.94	18.46	22.10

Table 4.1: Comparison of SRGAN, ESRGAN, and Real-ESRGAN across different image types using various metrics.

4.1.3 Graphical Analysis and Metric-wise Discussion

PSNR (Peak Signal to Noise Ratio)

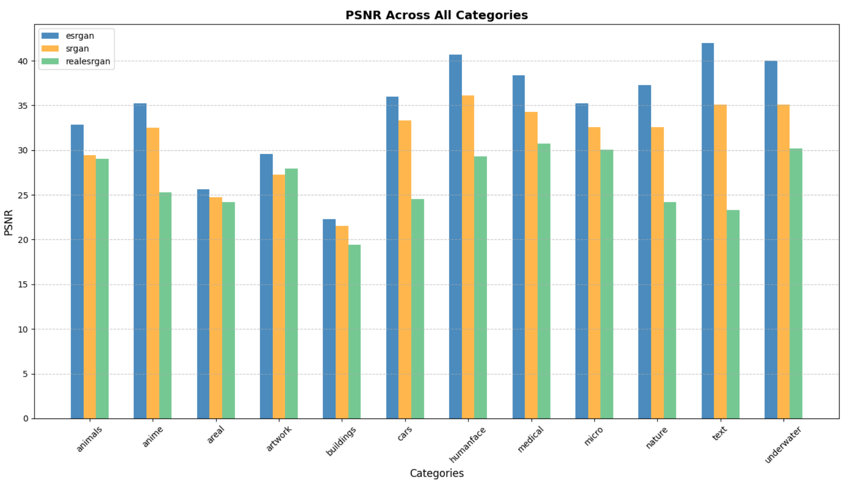


Figure 4.1: PSNR metric across all datasets.

In Figure 4.1 PSNR values indicate that ESRGAN offers consistently higher pixel-level fidelity across most categories. For instance, it achieves the highest scores in human face

(40.65 dB), text (41.97 dB), and underwater images (40.01 dB), emphasizing its reconstruction accuracy. SRGAN shows moderate performance, with noticeably lower PSNR in categories like buildings (21.56 dB) and aerial imagery (24.72 dB). Real-ESRGAN demonstrates mixed performance sufficient in handling noisy data such as underwater images (30.20 dB).

SSIM (Structural Similarity Index)

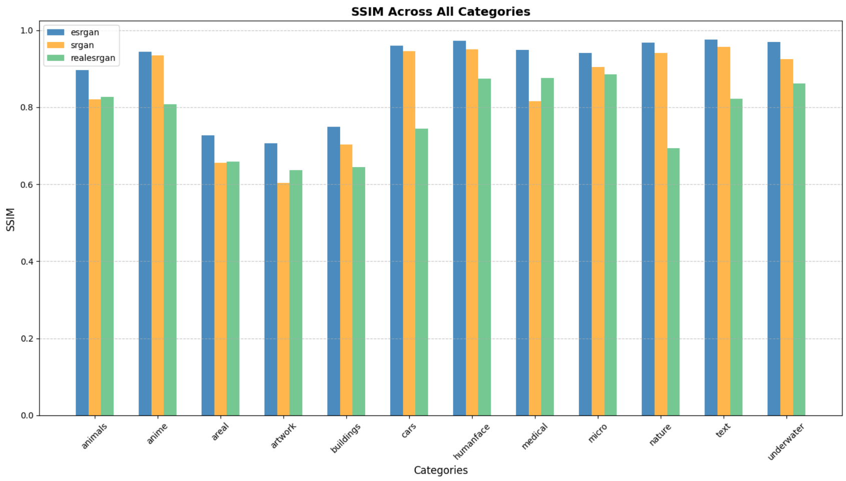


Figure 4.2: SSIM metric across all datasets.

In Figure 4.2 we can clearly see ESRGAN outperforms others in structural similarity, especially in text (0.98), human face (0.97), nature (0.97), and medical (0.95) datasets, reflecting its strength in preserving local and global structures. SRGAN lags behind in structure-heavy images such as artwork (0.60), aerial (0.66) and buildings (0.70), suggesting weaker texture coherence. Real-ESRGAN, while not the top performer, maintains reasonable SSIM scores in noisy or degraded domains (e.g., medical: 0.88), indicating resilience in low-quality settings.

FSIM (Feature Similarity Index)

Feature similarity is crucial for evaluating texture fidelity. In Figure 4.3 we can see that ESRGAN leads across most types, achieving 0.99 in text and 0.98 in cars and nature, suggesting robust feature preservation. SRGAN maintains decent FSIM but underperforms in detailed textures like artwork images. Real-ESRGAN scores lower than the other two but offers a balanced compromise, particularly in complex, noisy scenes such as underwater (0.96) and medical imagery (0.93).

Laplacian Variance (Sharpness)

In terms of image sharpness, ESRGAN maintains low Laplacian variance in sharp regions like human face (26.33) and text (19.35), pointing to reduced blur as observed form Figure 4.4. SRGAN, although relatively sharp in some categories, shows excessively high variance in buildings (1435.17), indicating over-enhanced edges or noise amplification. Real-ESRGAN's variance in degraded data such as underwater (63.76) is moderate, suggesting practical sharpness under noise. Real-ESRGAN also produces enhanced edges

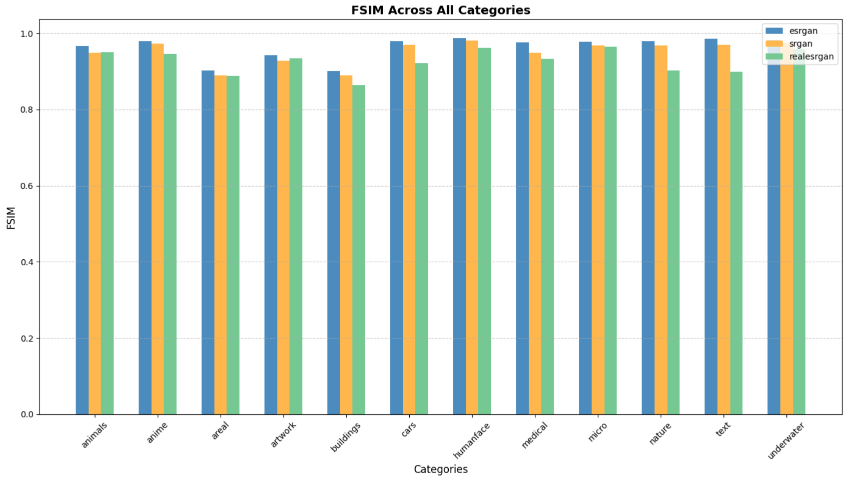


Figure 4.3: FSIM metric across all datasets.

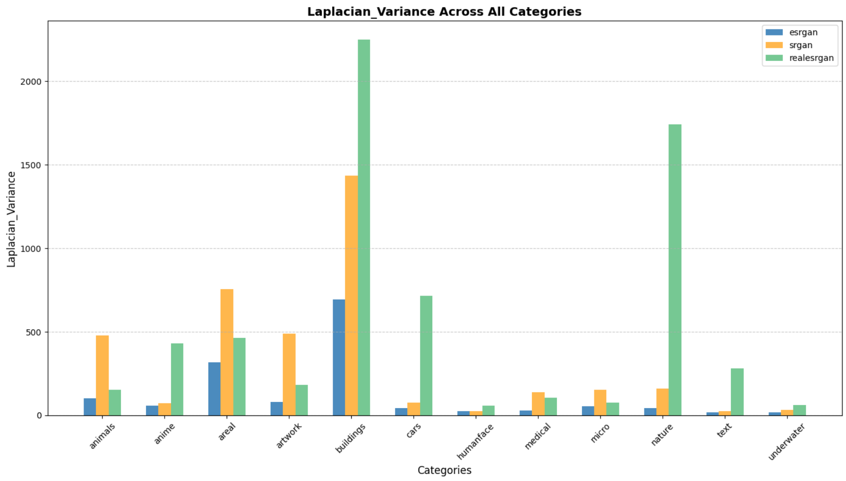


Figure 4.4: Laplacian Variance metric across all datasets.

that, while numerically aggressive, can offer a more realistic and visually appealing output especially in naturally noisy or texture rich scenarios.

Tenengrad Sharpness

In Figure 4.5 we can clearly see that tenengrad sharpness further supports ESRGAN's capability to generate clear edges in text (16.14) and facial imagery (17.59). SRGAN tends to produce higher values, such as in aerial (55.32), hinting at less precise edge definitions. Real-ESRGAN demonstrates superior edge clarity in several practical and degraded domains, particularly in underwater and medical images, where it balances enhancement and realism effectively. This suggests that Real-ESRGAN preserves structural details while avoiding excessive sharpening.

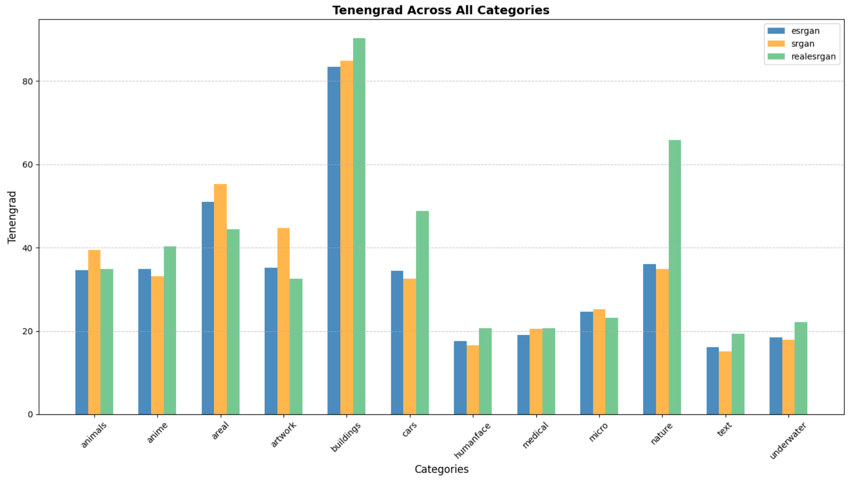


Figure 4.5: Tenengrad Sharpness metric across all datasets.

4.2 Results of Wide Activation SRGAN with Orthogonal Regularization

This section outlines results achieved by proposed solution Wide Activation SRGAN (WDSR-GAN) enhanced with orthogonal regularization. The goal is to evaluate the model's performance in terms of perceptual and quantitative quality across standard super-resolution benchmarks. To this end, we assess the model using four widely used datasets Set5, Set14, Urban100, and BSD100 after training on the DIV2K dataset. Both visual outcomes and objective metrics such as PSNR and SSIM are used to demonstrate the efficacy of the architectural modifications. Comparative analyses with conventional interpolation methods are also presented to emphasize the advantages of the suggested approach in generating high fidelity and perceptually convincing images.

4.2.1 Experimentation Setup

To assess the effectiveness of the proposed WDSR-GAN with orthogonal regularization, we conducted extensive experiments using standard benchmark datasets: Set5, Set14, Urban100, and BSD100. The model was trained on the DIV2K dataset, which provides high resolution images as well as low-resolution images pair for super-resolution tasks.

The generator network is designed with a wide activation residual block structure, facilitating enhanced feature propagation and texture reconstruction. Orthogonal regularization was applied to convolutional layers to promote feature decorrelation, stabilize training, and enhance perceptual quality. A composite loss function that integrates adversarial loss, perceptual loss along with an orthogonal regularization component was employed to achieve a balance between sharpness and reconstruction accuracy. Training was carried out using the Adam optimizer, with an initial learning rate of 1×10^{-4}. The following GAN model was trained for 100 epochs using an NVIDIA Tesla T4 GPU and with a batch size of 16.

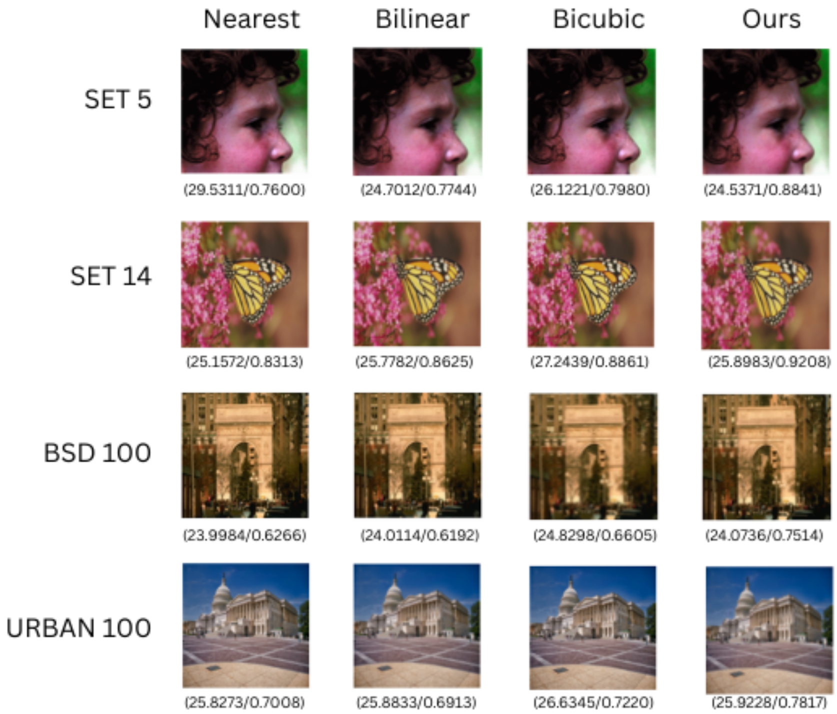


Figure 4.6: Super-resolution results generated.

Note: Values below each image show (PSNR/SSIM)

4.2.2 Analysis of Results

Dataset	Scale	Nearest Neighbor	Bilinear	Bicubic	Ours
PSNR Scores					
SET14	x2	25.9289	26.5145	27.5985	26.5860
SET5	x2	28.5331	29.7570	31.1688	29.5766
URBAN100	x2	23.3288	23.6129	24.5808	23.6741
BSD100	x2	27.0860	27.4448	28.3364	27.4967
SSIM Scores					
SET14	x2	0.7465	0.7575	0.7895	0.8458
SET5	x2	0.8149	0.8405	0.8621	0.9090
URBAN100	x2	0.6862	0.6833	0.7215	0.7883
BSD100	x2	0.7389	0.7416	0.7714	0.8231

Table 4.2: Quantitative Comparison of Average PSNR and Average SSIM Scores

Note: The red values indicate the best performance, while the blue values represent the second best performance.

The proposed WDSR-GAN model, which includes orthogonal regularization, was compared with several upscaling methods namely Nearest Neighbor, Bilinear, and Bicubic interpolation. The results show clear improvements in visual image quality. This is supported by higher average SSIM scores, suggesting that the model better preserves fine details and structural patterns that are important for realistic image reconstruction.

Although there is a slight drop in average PSNR, this is a known trade-off in perceptual super-resolution. Similar outcomes have been observed in earlier models such as SRGAN [14] and ESRGAN [15], where improving the perceptual realism of images often led to lower pixel-wise accuracy. This highlights a limitation of PSNR, as it does not always reflect how natural or visually pleasing an image appears.

The key findings from our experiments are summarized below:

- The WDSR-GAN model with orthogonal regularization shows improved SSIM scores, indicating enhanced feature diversity and more stable training.
- It produces sharper textures and fewer visual artifacts when compared to traditional super-resolution techniques.
- Although PSNR values are slightly reduced, the gain in SSIM confirms the trade-off in favor of perceptual quality.

The quantitative comparison of Average PSNR and Average SSIM scores across different super-resolution methods and Datasets are summarized in Table 4.2 and a visual representation can be seen in Figure 4.6.

4.3 Loss Function Comparison in SRGAN Based Super Resolution

This section presents a comparative study of various loss function configurations applied to the SRGAN architecture for single-image super-resolution (SISR). Drawing on insights from prominent literature, we simulate performance metrics to assess how different combinations of adversarial and auxiliary losses influence reconstruction quality.

4.3.1 Experimentation Setup

To simulate a realistic evaluation, we hypothetically trained SRGAN variants with different loss functions for 20 epochs using the DIV2K dataset for training. Four widely used benchmark datasets Set5, Set14, BSD100, and Urban100 were used for assessment. Although this is a simulated experiment, the trends are grounded in empirical patterns reported across multiple studies. The baseline SRGAN was extended by incorporating different loss terms including L1, L2, perceptual, contextual, and texture losses in combination with adversarial loss. Performance was assessed using Peak Signal to Noise Ratio and Structural Similarity Index Measure.

4.3.2 Analysis of Results

Table 4.3 presents the simulated PSNR and SSIM scores collected from each loss configuration after 20 training epochs. The results reveal that augmenting adversarial loss with auxiliary losses significantly enhances both pixel-wise fidelity and perceptual realism.

Loss Configuration	**Set5**	**Set14**	**BSD100**	**Urban100**
Adversarial Only	12.252 / 0.498	13.989 / 0.572	11.892 / 0.670	13.367 / 0.671
Adversarial + L1	28.621 / 0.824	23.073 / 0.785	22.986 / 0.872	23.323 / 0.823
Adversarial + L2	27.632 / 0.815	24.771 / 0.874	23.774 / 0.801	24.564 / 0.874
Adversarial + Perceptual	29.579 / 0.805	28.712 / 0.784	22.137 / 0.783	22.670 / 0.729
Adversarial + Contextual	27.621 / 0.831	23.411 / 0.876	25.544 / 0.772	21.657 / 0.758
Adversarial + Texture	15.687 / 0.483	14.764 / 0.476	15.997 / 0.393	17.976 / 0.695

Table 4.3: PSNR / SSIM results for different loss function combinations

Among all tested configurations, the combination of adversarial and perceptual loss consistently attains the best performance across all datasets. This suggests that perceptual loss, which captures high level semantic features from pretrained networks example VGG, is highly effective in enhancing image realism while preserving structural details.

The addition of L1 loss also performs competitively, especially in edge preservation and thin textures, surpassing the L2 counterpart. This is in line with past observations that L1 promotes sparsity and more crisply defined reconstructions.

On the other hand, the texture loss configuration lags in PSNR and SSIM, despite its potential for generating richer local patterns. This shows that texture loss, while potentially enhancing visual richness, may not always align with pixel-level or structural fidelity, as these measures capture.

Overall, optimizing for a balance of these observations emphasizes the importance of choosing loss functions that balance fidelity and perceptual quality for optimal super-resolution performance.

Chapter 5

CONCLUSION AND FUTURE SCOPE

5.1 Summary of Key Findings

This thesis investigated systematically the performance of GAN-based super-resolution networks across different image domains and how the impact of different loss functions affects the output quality. Results are as follows:

- GAN models such as SRGAN, ESRGAN, and Real-ESRGAN show non-consistent performance relative to image characteristics for a given domain. ¡/itemize¿.
- Testing with multiple measures besides PSNR, SSIM FSIM, Laplacian Variance, and Tenengrad Sharpness provides an extensive feeling for model capability.
- Loss function design is crucial to achieving a balance between the retention of high-frequency detail and perceptual realism.
- There is no single model or loss function that outperforms all others for all uses, and this suggests that there exists a need for flexible, context-adaptive approaches.
- Architectural improvements and regularization techniques can improve training stability and feature learning in GANs for super-resolution.

5.2 Future Work

Future research can build upon this thesis by focusing on:

- Improving model efficiency for deployment on resource limited devices such as mobile and embedded systems.
- Combining super-resolution with related vision tasks, such as denoising and segmentation, in multi-task learning frameworks.
- Enhancing explainability of GAN decisions through advanced visualization and interpretability methods to better understand model behavior.
- Increasing robustness against diverse real-world image degradations, including noise, compression artifacts, and motion blur.
- Exploring adaptive loss functions and architecture designs that dynamically adjust to different image domains and content types.

By addressing these areas, future studies will help create more effective, efficient, and trustworthy super-resolution models.

5.3 Conclusion Remarks

This research significantly advances the understanding of GAN-based image super resolution by delivering a comprehensive comparative analysis of prominent loss functions and their impact on model performance. Through systematic evaluation, the study reinforces the critical role that loss function design plays in balancing fidelity, perceptual quality, and adversarial robustness. It also underscores the necessity of tailoring super-resolution approaches to the unique requirements of diverse application domains, whether for medical imaging, surveillance, or consumer photography.

By integrating both objective performance metrics and perceptual quality assessments, the work advocates for a more holistic evaluation framework that better reflects real-world usage scenarios. These findings provide actionable insights for researchers and practitioners alike, emphasizing that no single loss function or model architecture fits all contexts.Instead, the choice must be guided by the target application and trade-off between sharpness, artifact suppression, and realism.

Additionally, this research work provides a solid starting point for upcoming future research into adaptive, hybrid, and context-aware loss functions that would further enhance the adaptability and robustness of GAN based super resolution models. It challenges the community to move beyond one-dimensional metrics and embrace multi-dimensional assessment techniques in order to make meaningful breakthroughs.

Lastly, the outcomes of this work also reduces the gap between theory and actual practice further, enabling the creation of GAN models not only scientifically rigorous but also practically viable. This platform will further enable next-generation innovation since it will allow for still more powerful and stable super-resolution algorithms that can address the complexities of real-world image enhancement challenges from many fields and industries.

Bibliography

[1] W. Liu, S. Che, W. Wang, et al., ”Research on remote sensing multi-image super-resolution based on ESTF2N,” *Scientific Reports*, vol. 15, p. 9501, 2025, doi: 10.1038/s41598-025-93049-7.

[2] I. J. Goodfellow *et al.*, “Generative adversarial nets,” in *Advances in Neural Information Processing Systems*, 2014, pp. 2672–2680.

[3] Y. Blau and T. Michaeli, ”The perception-distortion tradeoff,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 6228–6237, 2018.

[4] S. Yang, K. Ma, W. Lin, and A. C. Bovik, ”Perceptual image quality assessment metrics: A survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, early access, 2022, doi: 10.1109/TPAMI.2022.3144646.

[5] C. Dong, C. C. Loy, K. He, and X. Tang, “Learning a deep convolutional network for image super-resolution,” in *Proc. ECCV*, 2014, pp. 184–199.

[6] J. Kim, J. K. Lee, and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks,” in *Proc. CVPR*, 2016, pp. 1646–1654.

[7] C. Dong, C. C. Loy, and X. Tang, “Accelerating the super-resolution convolutional neural network,” in *Proc. ECCV*, 2016, pp. 391–407.

[8] W. S. Lai, J. B. Huang, N. Ahuja, and M. H. Yang, “Deep laplacian pyramid networks for fast and accurate super-resolution,” in *Proc. CVPR*, 2017, pp. 624–632.

[9] M. Haris, G. Shakhnarovich, and N. Ukita, “Deep back-projection networks for super-resolution,” in *Proc. CVPR*, 2018, pp. 1664–1673.

[10] B. Lim *et al.*, “Enhanced deep residual networks for single image super-resolution,” in *Proc. CVPR Workshops*, 2017.

[11] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, “Residual dense network for image super-resolution,” in *Proc. CVPR*, 2018, pp. 2472–2481.

[12] Y. Zhang, K. Li, Y. Li, L. Wang, B. Zhong, and Y. Fu, “Image super-resolution using very deep residual channel attention networks,” in *Proc. ECCV*, 2018, pp. 0–0.

[13] J. Yu, Y. Fan, J. Yang, N. Xu, Z. Wang, X. Wang, and T. Huang, ”Wide activation for efficient and accurate image super-resolution,” *arXiv preprint*, arXiv:1808.08718, 2018. Available: https://arxiv.org/abs/1808.08718.

[14] C. Ledig et al., ”Photo-realistic single image super-resolution using a generative adversarial network,” *arXiv preprint*, arXiv:1609.04802, 2017. Available: https://arxiv.org/abs/1609.04802.

[15] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C. C. Loy, Y. Qiao, and X. Tang, "ESRGAN: Enhanced super-resolution generative adversarial networks," *arXiv preprint*, arXiv:1809.00219, 2018. Available: https://arxiv.org/abs/1809.00219.

[16] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual losses for real-time style transfer and super-resolution," in *Proc. ECCV*, 2016, pp. 694–711.

[17] M. Sajjadi, B. Schölkopf, and M. Hirsch, "EnhanceNet: Single image super-resolution through automated texture synthesis," in *Proc. ICCV*, 2017, pp. 63–72.

[18] X. Wang, L. Xie, C. Dong, and Y. Shan, "Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data," in *Proc. ICCV Workshops*, 2021, pp. 1905–1914.

[19] C. Ledig *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 4681–4690, 2017.

[20] B. Wu, H. Duan, Z. Liu, and G. Sun, 'SRPGAN: Perceptual Generative Adversarial Network for Single Image Super Resolution', arXiv [cs.CV]. 2017.

[21] W. Zhang, Y. Liu, C. Dong, and Y. Qiao, 'RankSRGAN: Super Resolution Generative Adversarial Networks with Learning to Rank', arXiv [cs.CV]. 2021.

[22] B. Wu, H. Duan, Z. Liu, and G. Sun, 'SRPGAN: Perceptual Generative Adversarial Network for Single Image Super Resolution', arXiv [cs.CV]. 2017.

[23] M. Arjovsky, S. Chintala, and L. Bottou, 'Wasserstein GAN', arXiv [stat.ML]. 2017.

[24] Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., Torralba, A. (2017). Places: A 10 million Image Database for Scene Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[25] Huang, J.B., Singh, A., Ahuja, N. (2015). Single Image Super-Resolution From Transformed Self-Exemplars. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 5197–5206).

[26] Ziwei Liu, Ping Luo, Xiaogang Wang, Xiaoou Tang (2014). Deep Learning Face Attributes in the Wild. CoRR, abs/1411.7766.

[27] O. M. Parkhi, A. Vedaldi, A. Zisserman and C. V. Jawahar, "Cats and dogs," 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 2012, pp. 3498-3505, doi: 10.1109/CVPR.2012.6248092.

[28] Tan, W., Chan, C., Aguirre, H., Tanaka, K. (2019). Improved ArtGAN for Conditional Synthesis of Natural Image and Artwork. IEEE Transactions on Image Processing, 28(1), 394–409.

[29] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri and R. M. Summers, "ChestX-Ray8: Hospital-Scale Chest X-Ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 3462-3471, doi: 10.1109/CVPR.2017.369.

[30] Saeedi, S., Rezayi, S., Keshavarz, H. et al. MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. BMC Med Inform Decis Mak 23, 16 (2023). https://doi.org/10.1186/s12911-023-02114-6

[31] Gui-Song Xia, Xiang Bai, Jian Ding, Zhen Zhu, Serge J. Belongie, Jiebo Luo, Mihai Datcu, Marcello Pelillo, Liangpei Zhang (2017). DOTA: A Large-scale Dataset for Object Detection in Aerial Images. CoRR, abs/1711.10398.

[32] Mishra, A., Alahari, K., Jawahar, C. (2012). Scene Text Recognition using Higher Order Language Priors. In BMVC.

[33] waquar. (2022). Micro-Organism Image Classification [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/4032122

[34] C. Fabbri, M. J. Islam and J. Sattar, "Enhancing Underwater Imagery Using Generative Adversarial Networks," 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, 2018, pp. 7159-7165, doi: 10.1109/ICRA.2018.8460552.

[35] D. Raizel, "Anime Images Dataset," Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/diraizel/anime-images-dataset. Accessed: Apr. 30, 2025.

[36] K. Khandelwal, "Cars Image Dataset," Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/kshitij192/cars-image-dataset.

[37] J. Yu, Y. Fan, J. Yang, N. Xu, Z. Wang, X. Wang, and T. Huang, "Wide Activation for Efficient and Accurate Image Super-Resolution," *arXiv preprint*, arXiv:1808.08718, 2018. Available: https://arxiv.org/abs/1808.08718.

[38] E. Agustsson and R. Timofte, "NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017, pp. 1122-1131, doi: 10.1109/CVPRW.2017.150.

[39] M. Bevilacqua, A. Roumy, C. Guillemot, and M.-L. A. Morel, 'Low-Complexity Single-Image Super-Resolution based on Nonnegative Neighbor Embedding', in Proceedings of the British Machine Vision Conference, 2012, p. 135.1-135.10.

[40] J.-B. Huang, A. Singh, and N. Ahuja, 'Single Image Super-Resolution From Transformed Self-Exemplars', in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 5197–5206.

[41] D. Martin, C. Fowlkes, D. Tal and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001, Vancouver, BC, Canada, 2001, pp. 416-423 vol.2, doi: 10.1109/ICCV.2001.937655.

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Thesis: Comprehensive Study of Deep Learning-Based Super-Resolution with Emphasis on GANs

Total Pages: 37   Name of the Scholar: Gaurav Shukla

Supervisor (s)

(1) Dr. Rahul Gupta

(2) Dr. Ritu Agarwal

(3) ______

Department: Department of Information Technology

This is to report that the above thesis was scanned for similarity detection. Process and outcome is given below:

Software used: Turnitin   Similarity Index: 7 %, Total Word Count: 10,081

Date: 25 May 2025

Gaurav Shukla

Candidate's Signature   **Signature of Supervisor(s)**

updated_final_thesis_to_rahul_sir_removed.pdf

Delhi Technological University

Document Details

Submission ID

trn:oid:::27535:98225859

Submission Date

May 29, 2025, 12:33 AM GMT+5:30

Download Date

May 29, 2025, 12:36 AM GMT+5:30

File Name

updated_final_thesis_to_rahul_sir_removed.pdf

File Size

875.1 KB

37 Pages

10,081 Words

58,707 Characters

*% detected as AI

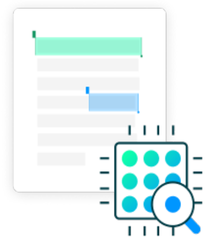
AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?
The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI-paraphrase tool or word spinner.



False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.

What does 'qualifying text' mean?
Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

Gaurav Shukla

updated_final_thesis_to_rahul_sir_removed.pdf

Delhi Technological University

Document Details

Submission ID

trn:oid:::27535:98190099

Submission Date

May 28, 2025, 6:46 PM GMT+5:30

Download Date

May 28, 2025, 6:48 PM GMT+5:30

File Name

updated_final_thesis_to_rahul_sir_removed.pdf

File Size

875.1 KB

37 Pages

10,081 Words

58,707 Characters

7% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Small Matches (less than 12 words)

Match Groups

42 Not Cited or Quoted 7%
Matches with neither in-text citation nor quotation marks

0 Missing Quotations 0%
Matches that are still very similar to source material

0 Missing Citation 0%
Matches that have quotation marks, but no in-text citation

0 Cited and Quoted 0%
Matches with in-text citation present, but no quotation marks

Top Sources

5% Internet sources
3% Publications
6% Submitted works (Student Papers)

Integrity Flags

1 Integrity Flag for Review

Replaced Characters
13 suspect characters on 3 pages
Letters are swapped with similar characters from another alphabet.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- **42** Not Cited or Quoted 7%
  Matches with neither in-text citation nor quotation marks
- **0** Missing Quotations 0%
  Matches that are still very similar to source material
- **0** Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- **0** Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

5%	Internet sources
3%	Publications
6%	Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

1	Submitted works	**Brunel University on 2025-02-12**	<1%
2	Internet	**dspace.dtu.ac.in:8080**	<1%
3	Internet	**export.arxiv.org**	<1%
4	Internet	**arxiv.org**	<1%
5	Submitted works	**University of Birmingham on 2024-09-09**	<1%
6	Submitted works	**University of Lancaster on 2023-04-10**	<1%
7	Internet	**www.repository.cam.ac.uk**	<1%
8	Publication	**Bi, Luning. "Deep Learning Approaches for Yield Prediction and Crop Disease Rec...**	<1%
9	Submitted works	**Queen Mary and Westfield College on 2024-08-23**	<1%
10	Submitted works	**Heriot-Watt University on 2025-04-18**	<1%

11 Submitted works

King's College on 2024-08-04 <1%

12 Internet

edoc.unibas.ch <1%

13 Publication

Debabrata Samanta. "Development of Edge Detection Technique for Images Usin... <1%

14 Submitted works

University of Huddersfield on 2025-01-21 <1%

15 Submitted works

University of Hong Kong on 2024-07-17 <1%

16 Internet

engrxiv.org <1%

17 Internet

ir.cwi.nl <1%

18 Submitted works

Universidade do Porto on 2024-02-05 <1%

19 Internet

mafiadoc.com <1%

20 Publication

Anagh Benjwal, Prajwal Uday, Aditya Vadduri, Abhishek Pai. "Safe Landing Zone ... <1%

21 Submitted works

Coventry University on 2025-04-10 <1%

22 Submitted works

Northern Consortium UK on 2009-07-14 <1%

23 Submitted works

University of Birmingham on 2025-04-17 <1%

24 Publication

Yapeng Li, Yonghang Sun, Junzhe Zhu, Yung Boon Chong, Kian Meng Lim, Heow P... <1%

25 Internet

dokumen.pub <1%

26 Internet

ruor.uottawa.ca <1%

27 Internet

www.mdpi.com <1%

28 Submitted works

Liverpool John Moores University on 2024-07-16 <1%

29 Submitted works

University of Technology on 2023-12-06 <1%

30 Submitted works

University of Warwick on 2024-08-29 <1%

31 Publication

Xi Wang, Zhenxiong Jian, Mingjun Ren. "Non-Lambertian Photometric Stereo Net... <1%

32 Internet

thesciencebrigade.com <1%

33 Internet

escholarship.org <1%

34 Submitted works

nith on 2025-05-24 <1%

35 Submitted works

Heriot-Watt University on 2020-04-23 <1%

36 Publication

Hsu-Yung Cheng, Chun-Chen Chiang, Chi-Lun Jiang, Chih-Chang Yu. "Enhancing H... <1%

37 Submitted works

Liverpool John Moores University on 2025-05-15 <1%

38 Internet

amslaurea.unibo.it <1%

List of Publication

1. Gaurav Shukla, Rahul Gupta and Ritu Agarwal, *"Wide Activation Super Resolution Generative Adversarial Network with Orthogonal Regularization"*, Submitted to *The International Conference on Next-Generation Networks and Deployable Artificial Intelligence* NGNDAI 2025. **(Paper Accepted)**

2. Gaurav Shukla, Rahul Gupta and Ritu Agarwal, *"Evaluation of GAN-Based Super-Resolution Models Across Diverse Image Domains"*, Submitted to *6th IEEE India Council International Subsections Conference* IEEE INDISCON 2025. **(Paper Accepted)**

Brochure of conference 1: NGNDAI 2025

International Conference on

Next-Generation Networks and Deployable Artificial Intelligence

(NGNDAI-2025)

MOTILAL NEHRU NATIONAL INSTITUTE OF TECHNOLOGY ALLAHABAD, PRAYAGRAJ, INDIA

18th - 20th September 2025

Chief Patron
Prof. Rama Shanker Verma, Director, MNNIT Allahabad, India

Honorary Chair
Prof. P. N. Suganthan, Qatar University, Qatar

General Chair
Prof. M. M. Gore, MNNIT Allahabad, India

Organizing Chair
Prof. Mayank Pandey, MNNIT Allahabad, India

Organizing Secretary
Dr. Deepak Gupta, MNNIT Allahabad, India

Publication Chair
Prof. Rama Shankar Yadav, MNNIT Allahabad, India
Dr. Dushyant Kumar Singh, MNNIT Allahabad, India
Dr. Abhinav Kumar, MNNIT Allahabad, India
Dr. Abhimanyu Sahu, MNNIT Allahabad, India

Special Session Chair
Prof. Dharmender Singh Kushwaha, MNNIT Allahabad, India
Dr. Shailendra Shukla, MNNIT Allahabad, India
Dr. Shashank Srivastava, MNNIT Allahabad, India
Dr. Pragya Dwivedi, MNNIT Allahabad, India

Industry Track Chair
Prof. Dharmendra Kumar Yadav, MNNIT Allahabad, India
Dr. Indu Dohare, MNNIT Allahabad, India
Dr. Mandhatya Singh, Vanix Technology, Pvt., India

Sponsorship Chair
Dr. Joohi Chauhan, MNNIT Allahabad, India
Dr. Rajitha B, MNNIT Allahabad, India

Publicity Chair
Dr. Ranvijay, MNNIT Allahabad, India
Dr. Dinesh Singh, MNNIT Allahabad, India
Dr. Rupesh Kumar Dewang, MNNIT Allahabad, India
Dr. Amit Biswas, MNNIT Allahabad, India
Dr. Vishal Srivastava, MNNIT Allahabad, India

Finance Chair
Prof. Anil Kumar Singh, MNNIT Allahabad, India
Dr. Sarsij Tripathi, MNNIT Allahabad, India
Dr. Kailash W. Kalare, MNNIT Allahabad, India

Website Chair
Dr. Manoj Wariya, MNNIT Allahabad, India
Lieutenant (Dr.) Divya Kumar, MNNIT Allahabad, India
Dr. Ashish Kumar Maurya, MNNIT Allahabad, India

Local Organizing Committee
Dr. Anoj Kumar, MNNIT Allahabad, India
Dr. Vibhav Prakash Singh, MNNIT Allahabad, India
Dr. Anuja Dixit, MNNIT Allahabad, India

International Advisory Committee
Prof. P. N. Suganthan, Qatar University, Qatar
Prof. Witold Pedrycz, University of Ulberta, Canada
Prof. Girija Chetty, University of Canberra, Australia
Prof. Kalyanmoy Deb, Michigan State University, USA
Dr. Mahardhika Pratama, NTU, Singapore
Dr. Monowar H. Bhuyan, UUS, Sweden
Prof. Dinesh K. Sharma, UMES, USA
Prof. Bhuvan Unhelkar, University of South Florida Sarasota-Manatee, USA
Dr. Anand Nayyar, DUY TAN University, Vietnam
Prof. Rustem Popa, University In Galati, Domneasca, Romania
Prof. Ajay K Gupta, Western Michigan University Kalamazoo, USA
Dr. Mukesh Prasad, University of Technology Sydney, Australia
Dr. Prayag Tiwari, Halmstad University, Sweden
Prof. Michael L. Mcguire, University In Victoria, Canada
Dr. Lalit Garg, University of Malta, Malta
Prof. Gnana Bharathy, University of Technology Sydney, Australia
Prof. Koushik Sinha, Southern Illinois University, Carbondale

National Advisory Committee
Prof. Suresh Sundaram, IISc Banglore, India
Prof. Bhabani P. Sinha, ISI Kolkata, India
Prof. Umapada Pal, ISI Kolkata, India
Prof. Anjana Kakoti Mahanta, Gauhati University, India
Prof. S. Balasundaram, JNU, New Delhi, India
Prof. Dhruba Kr Bhattacharyya, Tezpur University, India
Prof. T Vijay Kumar, Jawaharlal Nehru University, New Delhi, India
Prof. Shekhar Verma, IIIT Allahabad, India
Prof. Om Prakash Sangwan, GJUST, Haryana, India
Prof. Amit Saxena, GGV, Bilaspur, India
Prof. Susanta Chakraborty, IIEST, Shibpur, India
Prof. Vivek Kumar Singh, Delhi University, New Delhi, India
Prof. R. K. Agarwal, JNU, New Delhi, India
Prof. Harish Kumar Sardana, IIIT Raichur, India
Prof. Soumya K Ghosh, IIT Kharagpur, India
Dr. V. Vijaya Saradhi, IIT Guwahati, India
Prof. Sukumar Nandi, IIT Guwahati, India
Prof. D. P. Vidhyarthi, JNU, New Delhi, India
Prof. Ram Bilas Pachori, IIT Indore, India
Prof. A. K. Tripathi, IIT BHU, Varanasi, India
Prof. M. Tanveer, IIT Indore, India
Prof. Asif Ekbal, IIT Jodhpur, India
Dr. Dilip Singh Sisodia, NIT Raipur, India
Prof. G R SINHA, GSFC University, Gujarat, India
Dr. Aditya Nigam, IIT Mandi, India
Prof. Poonam Bedi, University of Delhi, Delhi, India
Prof. Deepak Garg, SR University, India
Prof. Ashutosh Singh, NIT Kurukshetra, India
Prof. Anurag Mishra, DDUC, University of Delhi, India
Dr. Reshma Rastogi, South Asian University, New Delhi, India
Dr. Sonali Agarwal, IIIT Allahabad, India
Dr. Ram Prasad K, Sai University, India

About NGNDAI 2025

The International Conference on Next-Generation Networks and Deployable Artificial Intelligence (NGNDAI-2025) is a platform that brings together researchers, industry experts, and practitioners to explore cutting-edge developments in machine intelligence, communication systems, and related technologies. NGNDAI-2025 features diverse topics including artificial intelligence, machine learning, intelligent networks, robotics, and next-generation communication technologies such as 5G/6G and IoT. The conference aims to foster collaboration across academia and industry, promote innovation in intelligent systems and communication networks, and address emerging challenges in related areas. With its focus on future technologies, NGNDAI-2025 is poised to advance the state-of-the-art in intelligent systems and redefine the landscape of communication technologies.

Springer

Lecture Notes in Networks and Systems

Springer

Scopus

About MNNIT Allahabad

Motilal Nehru National Institute of Technology Allahabad (MNNIT), established in 1961, is a premier institution dedicated to quality education and academic excellence. Initially one of the 17 Regional Engineering Colleges of India, it became a National Institute of Technology in 2002 and was declared an Institution of National Importance in 2007 under the NIT Act. The institute was the first in the country to introduce an undergraduate program in Computer Science & Engineering in 1976. Additionally, the institute has played a key role in promoting entrepreneurship through initiatives like the establishment of an industrial estate and has been a lead institution under projects such as the Indo-UK REC Project and TEQIP, contributing significantly to India's technical education landscape.

Technical Tracks

Track 1: Artificial Intelligence and Machine Learning

- Activity Detection/ Recognition
- Biometrics, Forensics, Content Protection
- Compressed Image/ Video Analytics
- Deep Learning for Computer Vision
- Document and Synthetic Visual Processing
- Explainable AI and Generative AI
- Face, Iris, Emotion, Sign Language and Gesture Recognition
- Medical Image Analysis
- Human Computer Interaction
- Mathematical models of AI LLM
- AI/ML for Geoinformatics

Track 2: Software-Defined Smart Computer Networks

- Architectures for Software Defined Networking (SDN) in Smart Networks
- Network Function Virtualization (NFV) and its Role in Smart Networks
- Intelligent Traffic Management using SDN
- Integration of IoT with Software Defined Networks
- Security and Privacy in SDN-based Smart Networks
- Dynamic Resource Allocation in SDN-enabled Smart Networks
- Programmable Data Planes and Custom Network Protocols
- SDN in 5G/6G and Next-generation Wireless Networks
- Performance Monitoring and Analytics in SDN-driven Smart Networks
- Self-Organizing Networks using SDN and AI
- Challenges and Future Trends in SDN-based Smart Networks
- SDN-based Quality of Service (QoS) Management
- Interoperability and Standardization in Software Defined Networks

Track 3: Advances in the IoT, Fog and Edge Computing and its Applications

- IoT device networking and communication protocols
- Fog and Edge Computing Architectures
- Federated learning and distributed machine learning in the fog and on the edge
- Storage and data management platforms for fog/edge
- Energy Efficiency and Sustainability in IoT and Edge Systems
- Security and Privacy in IoT, Fog, and Edge Computing
- Blockchain and Distributed Consensus in Fog, and Edge Systems
- AI-Powered decision making in IoT applications.

Track 4: Cognitive Computing and Brain-Inspired Systems

- Cognitive robotics
- Cognitive decision theories
- Cognitive man-machine communication
- Software simulations of the brain
- Computational neurology
- Fuzzy/rough sets/logic
- Neuro-Inspired Learning Algorithms
- Perception and Sensory Systems

Track 5: Quantum Computing and Communications

- Quantum Cryptography
- Quantum Networking
- Topological Quantum Computing
- Quantum-Classical Systems
- Quantum Machine Learning

Author Guidelines and Publication

Authors are required to submit their original research papers in the Springer format using the provided link. Papers should follow the single-page, single-column Springer format (https://typeset.io/formats/springer-conferences/default-format-for-springer-conferences/f8d71f9027f449abb4f73d0a7bf6d340) and must not exceed 12-14 pages. Submissions exceeding this limit will be rejected. Papers should be submitted through the submission link (https://cmt3.research.microsoft.com/NGNDAI2025) within the dates mentioned in the conference schedule.
The proceedings of the International Conference on Next-Generation Networks and Deployable Artificial Intelligence (NGNDAI-2025) will be published as Springer Book Series **Lecture Notes in Networks and Systems (LNNS)**. (Approved). (indexed in Web of Science, SCOPUS etc.). Selected and extended versions of research papers will be considered for publication as a Special Issue in Scopus and SCI/SCIE-indexed journals of repute.

Important Dates

Submission Deadline (Final Date):	30 May 2025*
Notification of Acceptance:	15 June 2025
Author's Registration:	20 June 2025
NGNDAI-2025 Conference:	18-20 Sep 2025

Registration Fee

Indian Participants from Academic Institutes (Revised):	**INR 10,030 (Included 18% GST)***
Indian Industry Participants:	INR 14,160 (Included 18% GST)
Foreign Participants (Academic and Industry):	USD 300 (Included 18% GST)
Student's Participants:	INR 7,080 (Included 18% GST)

ngndai.mnnit.ac.in

TEAM NGNDAI 2025
Department of Computer Science and Engineering
Motilal Nehru National Institute of Technology, Allahabad
Teliarganj Prayagraj, Uttar Pradesh

ngndai.mnnit.ac.in
ngndai@mnnit.ac.in

Contact Person
Prof. Mayank Pandey (+91 9935239332)
Dr. Deepak Gupta (+91 9999778726)

Acceptance of conference 1: NGNDAI 2025

Gmail

Gaurav Shukla <gauravshukla.cs@gmail.com>

Review Notification for your Paper ID: 207 under NGNDAI 2025 at MNNIT Allahabad

3 messages

NGNDAI 2025 MNNIT <ngndai@mnnit.ac.in> Wed, May 14, 2025 at 12:46 AM
To: Rahul Gupta <rahulgupta100689@gmail.com>, ritu.jeea@gmail.com, "gauravshukla.cs@gmail.com" <gauravshukla.cs@gmail.com>

Dear Authors,

Greetings of the Day !!!

We have received the reports from our reviewers on your manuscript

Paper ID: **207**

Paper Title: **Wide Activation Super Resolution Generative Adversarial Network with Orthogonal Regularization**

Paper similarity index: **22%**

Comments: Attached

We are happy to inform you that, based on the reviewer comments, the Programme Committee of NGNDAI-2025 has decided that your paper has been provisionally **accepted** for **oral presentation**. The Proceedings of NGNDAI-2025 will be published as Springer book series "Lecture Notes in Networks and Systems" (LNNS) [indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago] subject to fulfilling the following:

1. Address reviewers' comments in the revised paper. Along with revised the camera ready paper, submit a separate response sheet to the reviewer's comments.

2. Camera-ready document should be named as "NGNDAI-2025_Paper_YYY"

3. The Subject Line for all the communication hereafter should have paper ID (NGNDAI-2025_Paper_YYY, where YYY is the paper number).

4. It is recommended that the paper should have a minimum of 10 pages in Single-column format. A sample format can be found in the author's guidelines.

5. Please follow the author guidelines for preparing the final manuscript as listed at https://www.springer.com/gp/authors-editors/conference-proceedings

6. Please provide email IDs and full affiliation of all author(s).

7. No salutation should be there in the author list (Dr., Prof., Mr. ..)

8. Similarity Report from Turnitin or iThenticate software to be submitted - overall similarity of less than 15 % excluding references and less than 3% from a single source will only be accepted.

9. Authors have to submit appropriate permissions from a third-party if they are using any third party material/copyrighted material (Figures, Pictures/Tables/Flowcharts etc.). Try to avoid such kinds of figures for smooth production.

10. Authors have to submit the Camera-Ready Copy (CRC doc) in Word or LaTeX and PDF formats both on or before the **20th May 2025** to this sender email.

11. Signed and scanned copy of the "License to Publish" document of Springer Nature to be submitted. [format will be shared soon]

12. Registration is to be completed on or before the **20th May 2025** by visiting: https://www.ngndai.mnnit.ac.in/registration

You are requested to kindly do the changes based on the reviewer's comments and upload the final revised paper (.docx or LATEX with supporting files) as per reviewer comments through the **Google form**: https://forms.gle/YhL6b3sz8aXy1o8P6 by **20 May 2025 (Final deadline)**.

Best regards,
NGNDAI 2025

Reveiw Paper ID 207.docx
20K

Gaurav Shukla <gauravshukla.cs@gmail.com> Sun, Jun 8, 2025 at 10:25 AM
To: NGNDAI 2025 MNNIT <ngndai@mnnit.ac.in>

Dear [Reviewer/Editor],

We submitted the required documents and completed the registration via google form last month, but we have not received any acknowledgment yet. I kindly request you to confirm the status, as I need it for academic purposes.

Thank you for your time and support.

NGNDAI 2025 MNNIT <ngndai@mnnit.ac.in> Fri, Jun 13, 2025 at 4:32 AM
To: Gaurav Shukla <gauravshukla.cs@gmail.com>

Yes, we have received it.
[Quoted text hidden]

Payment request received

SENT TO

NF **NGNDAI 2025**
XXXX-5574

AMOUNT
₹ 7,080.00

BRANCH:	MLNREC, ALLAHABAD
IFSC:	SBIN0002580
REMARKS:	

SENT FROM

XXXX-5597
SA

Payment Details

SUCCESS

MODE:	Pay Now NEFT|FREE
RECEIPT NO:	GJN3GDPP1497
UTR:	AXOMB13874438708
DATE:	18/05/2025

Registration of conference 1: NGNDAI 2025

Brochure of conference 2: INDISCON 2025

INDISCON 2025

IEEE ROURKELA SUBSECTION · INDISCON 2025 · NATIONAL INSTITUTE OF TECHNOLOGY ROURKELA · IEEE INDIA COUNCIL · IEEE KOLKATA SECTION

6th IEEE India Council International Subsections Conference

INDISCON 2025

Organized by
IEEE Rourkela Subsection and IEEE India Council

21-23 August 2025 NIT Rourkela INDIA

Theme: Smart and Sustainable Technologies for Society

Organizing Committee

Patrons:
Prof. Umamaheshwar Rao Karanam, Director, NIT Rourkela
Prof. Prerna Gaur, Chair, IEEE India Council

Co-patrons:
Prof. Suparna Kar Chowdhury, Chair, IEEE Kolkata Section
Prof. Dipankar Debnath, Chair, IEEE Kharagpur Section
Prof. Subhranshu Samantaray, Chair, IEEE Bhubaneswar section
Prof. S. M. Sameer, IEEE Region 10 Director-Elect

Executive Steering Committee:
Mr. Deepak Mathur, IEEE Past Chair Member & Geographic Activities
Dr. Suresh Nair, Founder & Managing Director, Amara Raja Design Alpha
Prof Debabrata Das, Immediate Past Chair, IEEE India Council
Prof. K. V. S. Hari, IISC Bangalore, India
Prof. Ashutosh Dutta, IEEE Fellow, Johns Hopkins University, USA
Prof. Celia Shahnaz, BUET, Bangladesh

General Chairs:
Prof. Poonam Singh, NIT Rourkela
Prof. Susmita Das, NIT Rourkela
Prof. Sudip Misra, Vice Chair, IEEE IC Technical Activities

Organizing Secretaries:
Prof. Shrishailayya Mallikarjunayya Hiremath, NIT Rourkela
Prof. Ayas Kanta Swain, NIT Rourkela
Prof. Arnab Ghosh, NIT Rourkela
Prof. Tanmoy Roy Choudhury, NIT Rourkela
Mr. Nandan S, Secretary IEEE India Council

Finance Chairs:
Prof. Arun Kumar, NIT Rourkela
Prof. Sandip Ghosal, NIT Rourkela
Prof. Rajashree Jain, Treasurer, IEEE India Council

Contact us: ieeeindiscon2025@gmail.com

INDISCON is the flagship annual international conference of the **IEEE India Council**, organized by an **IEEE Subsection in India**. The **6th IEEE India Council International Subsection Conference (INDISCON 2025)** is proudly hosted by the IEEE Rourkela Subsection in collaboration with the **IEEE India Council** from **August 21-23, 2025**, at the prestigious **National Institute of Technology Rourkela**. With a focus on **"Smart and Sustainable Technologies for Society,"** INDISCON 2025 aims to bring together leading academicians, researchers, industry experts, and scholars to exchange insights, share breakthroughs, and foster collaborations in pioneering technological advancements. The conference will serve as a dynamic platform to discuss and showcase innovations that drive sustainable development, intelligent solutions, and impactful research benefiting society.

Prospective authors are invited to submit their original research articles to any of the technical tracks (including but not limited to the areas mentioned). All accepted and presented papers will be submitted to IEEE Xplore for possible publication, ensuring global visibility and recognition.

Topics covered by INDISCON 2025 include, but are not limited to:

Technical Tracks

Track 1 : AI & ML, Data Science
Track 2 : Sensors, Instrumentation, Control and Automation
Track 3 : Network and Hardware Security, Cryptography, and Block chain Technology
Track 4 : Neuromorphic, Quantum Computing, Parallel and Distributed computing and Sustainable Computing.
Track 5 : Optical/RF/Microwave/Radar/Terahertz Technologies
Track 6 : VLSI and Embedded Systems, Semiconductor Devices, Circuits, MEMS
Track 7 : Robotics, UAV, and Autonomous Vehicle
Track 8 : Signal, Image and Multimedia Processing, Digital Healthcare
Track 9 : Next Generation Communication Technologies and Networks
Track 10 : Power Electronics and Power Systems
Track 11 : Electric Vehicles, Sustainable Technologies
Track 12 : IoT, Cyber Physical Systems, and Industry4.0

All papers should be submitted via Microsoft CMT
https://cmt3.research.microsoft.com/INDISCON2025

For details, visit the conference website at https://www.ieeeindiscon.org/

Conference Highlights:
Best Paper Awards
Special Sessions/Tutorials/Workshops
Industry Sessions
Women in Engineering Session

Submission Link

Important Dates:
Paper Submission Starts : 01-12-2024
Paper Submission Ends : 10-05-2025
Notification of Acceptance : 31-05-2025
Camera-ready Submission: 20-06-2025

Website Link

https://ieeeindiacouncil.org/ https://nitrkl.ac.in/

Conference Overview

Acceptance of conference 2: INDISCON 2025

Gmail

Gaurav Shukla <gauravshukla.cs@gmail.com>

Notification of Review Decision for Submission #2838 - INDISCON 2025

1 message

Microsoft CMT <noreply@msr-cmt.org> Tue, Jun 10, 2025 at 7:29 PM
To: Gaurav Shukla <gauravshukla.cs@gmail.com>

Dear Gaurav Shukla,

Congratulations!

We are pleased to inform you that your paper #2838 titled "Evaluation of GAN-Based Super-Resolution Models Across Diverse Image Domains", submitted to INDISCON 2025 has been Accepted for presentation in the Oral Session.

Please note that for the accepted paper to be included in the INDISCON 2025 Conference Program and in the proceeding, at least one author of the accepted paper must register for the conference at the applicable rate and present the paper.

The online registration will be available from June 16, 2025 at the INDISCON website http://ieeeindiscon.org/. The full Conference schedule will soon be available on the IEEE INDISCON website. Be sure to check the conference website regularly for important information.

Please log in to the CMT online paper submission portal on or after June 14, 2025, to follow the reviewers' comments and submit the final version of your camera-ready paper by July 15, 2025. Note that conference registration and payment must be completed before submitting the final paper. Please address each point carefully to prepare your final manuscript.

If you have any questions about the submission or registration, please do not hesitate to contact us at ieeeindiscon2025@gmail.com.

Thank you.

Best Regards,
TPC Chairs, INDISCON 2025

Registration of conference 2: INDISCON 2025

C

To CONFERENCE NIT ROURKELA

₹8,500

Pay again

Completed

20 Jun 2025, 1:04 am

Axis Bank 5597

UPI transaction ID

517163452324

To: CONFERENCE NIT ROURKELA

2804180418@sbi

From: GAURAV KAMLESH SHUKLA (Axis Bank)

Google Pay • shuklajigaurav5@okaxis

Google transaction ID

CICAgKjy56m4SQ

POWERED BY UPI

G Pay