Development of Progressive GAN based Image Rain Removal Technique

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

Master of Technology
in
Intelligent Automation and Robotics
(Electronics and Telecommunication Engineering Department)
Jadavpur University

By

SAIFUDDIN SK
Class Roll No: 002110704015
Examination Roll No: M6IAR24013
Registration No: 139724

Under the guidance of
Prof. Sheli Sinha Chaudhuri
Professor
Department of Electronics and Telecommunication Engineering
Jadavpur University
Kolkata-700032
West Bengal, India

DECLARATION OF ORIGINALITY AND COMPLIANCE OF ACADEMIC ETHICS

I hereby declare that this thesis contains Previous Work and original work by the under-signed candidate, as part of my Master of Technology in Intelligent Automation and Robotics under the Department of Electronics and Telecommunication Engineering. All information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that I have thoroughly cited and referenced all material and findings which are not original to this research, as provided by these rules and conduct.

Name: SAIFUDDIN SK
Exam Roll No: 002110704015
Thesis Title: Development of Progressive GAN based Image Rain Removal Technique

---

Signature of Candidate

FACULTY OF ENGINEERING & TECHNOLOGY
JADAVPUR UNIVERSITY

CERTIFICATE

This is to certify that the thesis entitled **— "Development of Progressive Gan based Image Rain Removal Technique"** has been carried out by **SAIFUDDIN SK** bearing **Class Roll No:02110704015, Examination Roll No**: **M6IAR24013** and **Registration No: 139724** of **2017-2018**, under my guidance and supervision and be accepted in partial fulfillment of the requirement for the degree of **Master of Technology** in **Intelligent Automation and Robotics** in the **Department of Electronics and Telecommunication Engineering, Jadavpur University.**

Prof. Sheli Sinha Chaudhuri

Supervisor
Department of Electronics and
Telecommunication Engineering
Jadavpur University

Prof. Manotosh Biswas

Head of the Department
Department of Electronics and
Telecommunication Engineering
Jadavpur University

Prof. Dipak Laha

Dean
Faculty Council of Engineering
Jadavpur University
Kolkata- 700032

FACULTY OF ENGINEERING & TECHNOLOGY
JADAVPUR UNIVERSITY

CERTIFICATE OF APPROVAL

The foregoing thesis titled **"Development of Progressive GAN based Image Rain Removal Technique"** is hereby approved as a creditworthy study of an engineering subject conducted and presented satisfactorily to warrant its acceptance as a precondition to the degree for which it was submitted. It is understood that the undersigned does not automatically support or accept any argument made, opinion expressed, or inference drawn in it by this approval, but only approves the thesis for the reason it was submitted.

Committee on Final Examination
for Evaluation of the Thesis

Signature of External Examiner

Signature of Supervisor

ACKNOWLEDGEMENT

This thesis, titled **"Development of Progressive GAN based Image Rain Removal Technique"**, is the culmination of my efforts, supported by many individuals, including my advisor and lab seniors. I am grateful for the opportunity to acknowledge their contributions.

First and foremost, I extend my heartfelt thanks to my thesis supervisor, **Prof. Sheli Sinha Chaudhuri**, Professor in the Department of Electronics and Telecommunication Engineering at Jadavpur University. Her invaluable guidance, insightful suggestions, and unwavering support were instrumental in both conducting this research and writing this thesis. I consider myself fortunate to have had such a positive and inspiring mentor. Her confidence and innovative ideas were key to the successful completion of my work, and she has been a true guardian throughout this journey. I am also deeply thankful to our Head of Department, **Prof. Manotosh Biswas**, for his constant assistance and for providing access to all necessary resources for my research.

I would like to extend my heartfelt gratitude to **Mr. Asfak Ali**, a research scholar in our department and a figure akin to an elder brother, for his invaluable assistance with ideation, implementation, and analysis.

Finally, I am profoundly grateful to my parents. Their sacrifices, love, encouragement, and support have been indispensable, shaping who I am today.

—----------

Saifuddin Sk

Intelligent Automation & Robotics
Department of Electronics and Telecommunication Engineering
Jadavpur University
Kolkata- 700032

Abstract

Deraining is a significant challenge in the field of computer vision, with the primary goal of removing rain streaks from images to improve visibility. This thesis introduces an innovative solution to deraining by employing a Progressive Generative Adversarial Network (GAN)-based deraining model. The proposed model comprises a generator and a discriminator that collaborate in an adversarial manner, continuously refining the deraining process. The generator is responsible for generating more realistic and visually appealing rain-free images, while the discriminator becomes increasingly proficient at distinguishing between real and generated images.

Initially, the U-Net architecture is utilized as the generator model. To further enhance the model's performance, a pair of U-Nets is intentionally used as the generator, resulting in a more effective deraining process.

The Rain13K dataset is employed for training and evaluation, offering a wide range of rainy scenarios to improve the model's ability to generalize to unseen rainy images. The model's performance is evaluated on benchmark datasets, including Rain100L, Rain100H, Test100, Rain1200, and Test2800, and compared against state-of-the-art deraining methods. Objective metrics, such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), are used to quantitatively assess image quality. This comprehensive evaluation allows for an analysis of the model's effectiveness and the identification of the optimal generator architecture for deraining tasks.

Keywords: Deraining, Computer Vision, Generative Adversarial Networks, Image Deraining, Progressive Deraining

Contents

1 Introduction **1**
- 1.1 Rain Imaging Model 2
- 1.2 Importance of Deraining 3
- 1.3 Potential of GAN based Deraining 5
- 1.4 Objective of the Thesis 7
- 1.5 Overview of the Thesis Work 8
- 1.6 Outline of the Thesis 9

2 Motivation **11**
- 2.1 Motivation 11

3 Literature Review **13**
- 3.1 Existing Deraining Techniques 13
  - 3.1.1 Prior-Based Deraining Techniques 13
  - 3.1.2 Learning-Based Deraining Techniques 13
  - 3.1.3 U-Net-Based Deraining Techniques 15
  - 3.1.4 GAN-based Deraining Techniques 16
- 3.2 Strengths and Limitations of Existing Techniques 17
  - 3.2.1 Strengths of Existing Techniques 17
  - 3.2.2 Limitations of Existing Techniques 19
- 3.3 Generative Adversarial Networks 21

4 Methodology **23**
- 4.1 Generator 26
  - 4.1.1 U-Net 27
  - 4.1.2 Progressive GAN 28
- 4.2 Discriminator 30
- 4.3 Loss Function 31

5 Experimental Setup **33**
5.1 Hyper-Parameter 33
5.2 Datasets 33
5.3 Training Details 34
5.4 Performance Metrics 35

6 Evaluation **37**
6.1 Quantitative Evaluation 37
6.1.1 PSNR 37
6.1.2 SSIM 38
6.2 Qualitative Evaluation 45
6.2.1 State-of-the-Art Comparison 45

7 Conclusion **49**
7.1 Future Scope 50

8 List of Publication

List of Tables

6.2.a : Detailed comparison of various test datasets using a single U-Net model …… 46

6.2.b : Detailed comparison of various test datasets using one pair of U-Net model…… 47

6.2.c : Detailed comparison of various test datasets using two pairs of U-Net model 48

6.2.d : Comprehensive Assessment of the Proposed Model in Comparison to Existing Models Across a Variety of Datasets …….. 48

List of Figures

1.1: The framework of the image deraining task …….. 2

1.2: An example of rain image and it's clear version …… 5

4.1: The Architecture of derain algorithm using GAN …….. 23

4.2: Illustration of Proposed Progressive Generator Architecture -(a) Single U-Net Architecture Block. (b) Pair of U-Net Architecture Block. (c) Two Pair of U-Net Architecture Block…… 25

4.3: Illustrating the detailed description of the Progressive GAN architecture ……. 29

4.4: Discriminator Architecture ……. 31

6.1.a: Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Test-2800 dataset …… 40

6.1.b: Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Test-1200 dataset …… 41

6.1.c: Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Test-100 dataset ……. 42

6.1.d: Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Rain-100L dataset ……. 43

6.1.e: Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Rain 100H dataset …….. 44

Chapter 1
Introduction

Rain is a natural weather condition that has a significant negative impact on not only human visual perception but also the performance of high-level vision tasks such as image classification, object detection, and video surveillance. That's why rain removal plays an important role in various computer vision applications, including image de-noising, decomposition, and restoration. Rain can make it challenging for computer vision algorithms to operate effectively, particularly outside, because it might hide crucial characteristics. Raindrops act as natural lenses, which cause images to become blurred and distorted. Distant rain appears foggy, whereas nearby rain appears as bright streaks, making it more difficult to see. Numerous techniques have been proposed to address rain removal, typically categorized into two groups: video-based and single-image-based methods. Generally, removing rain from a single image is more challenging due to the lack of temporal information. However, it is important since, in some cases, such as images downloaded from the Internet, only one single rainy image is available. On the other hand, the video rain removal problem can be solved by applying the single-image-based method frame by frame.

Deep learning techniques are increasingly gaining popularity in computer vision applications. They have proven to be very effective in tasks such as image recognition, de-noising of rain, classification, de-raining, and super-resolution. It's fascinating to see how rapidly this field is advancing and the exciting possibilities it holds for the future. However, it has been noted that deeper networks can face the challenge of degradation without a well-designed architecture. In certain applications such as image de-noising and de-raining, deep learning techniques can lead to over-smoothing. This is because the network can sometimes mistakenly remove important details from the images that resemble noise or rain. Basically, the method could accidentally make images blurry by removing some small details. To avoid this problem, it's important to choose the right network structure and settings when using deep learning for computer vision tasks. This helps prevent loss of quality and overly smooth results.

In this paper, we present a novel deep-learning method for removing rain from single images. Our approach is based on the U-Net architecture, which consists of an encoder-decoder structure. The encoder takes in an input image and captures its features while gradually reducing the spatial dimensions. Then, the decoder upsamples these features to reconstruct the image. To retain the spatial information necessary for effectively removing rain streaks, skip connections are established between the corresponding encoder and decoder layers. The network is trained on image pairs with and without rain to learn how to distinguish and remove rain patterns. This architecture allows for efficient rain removal by utilizing both low-level and high-level image features and preserves spatial details through skip connections.

The subsequent sections of this paper are structured as follows: Section 3 provides an extensive review of the existing literature on rain removal models. Following that, section 4 outlines the proposed methodology. In sections 5 and 6, the experimental setup and results are presented. Lastly, Section 6 offers the conclusion of this study.

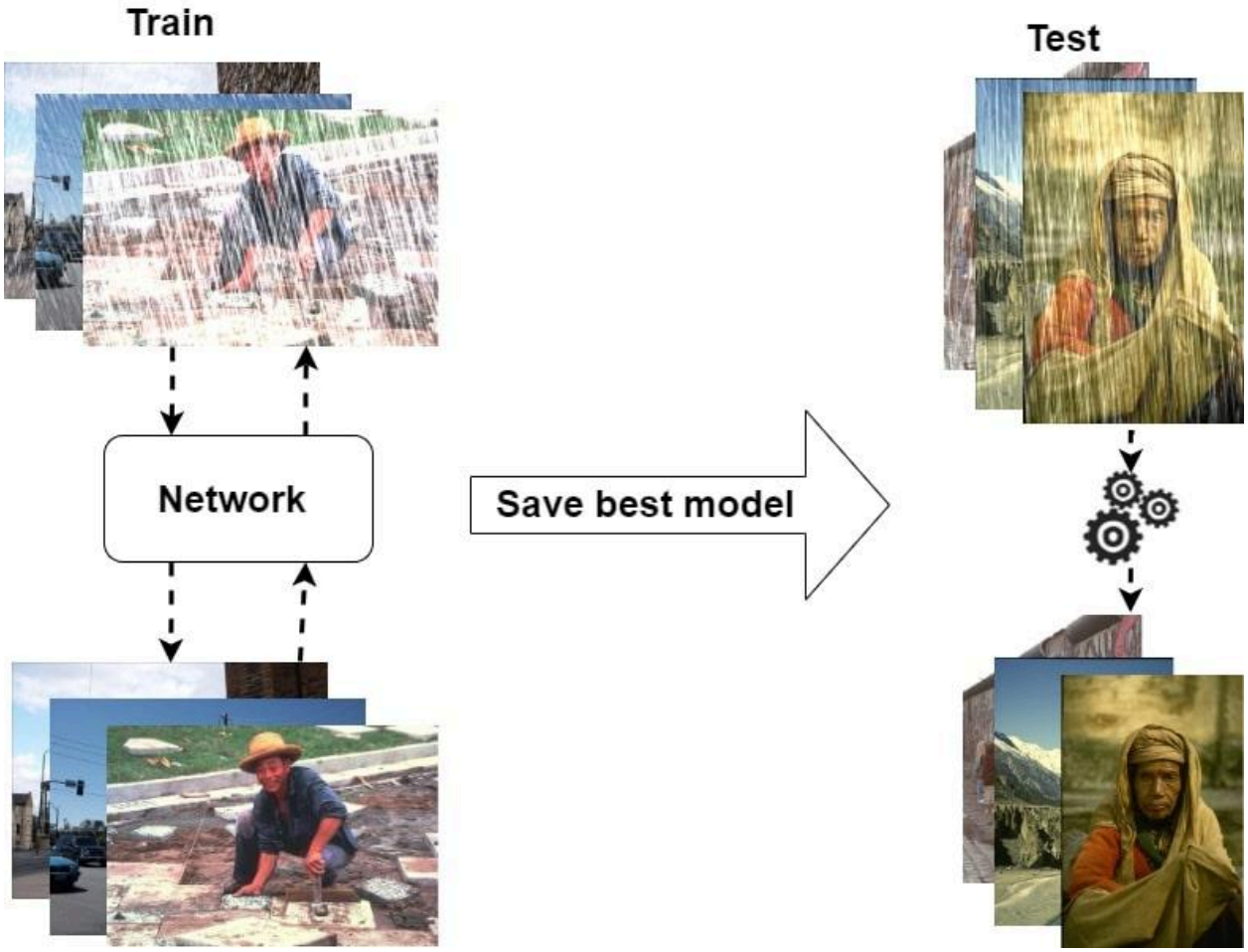


Fig 1.1 : The framework of the image-deraining task.

1.1 Rain Imaging Model

The rain imaging model operates on the premise that light experiences both scattering and absorption as it traverses a rain medium. Such engagements with atmospheric particles result in a decline in the quality of images, manifesting as diminished contrast, color accuracy, and overall perceptibility. The primary objective of the rain imaging model is to scientifically elucidate this deterioration phenomenon, facilitating proficient representation and subsequent refinement of images captured in rainy conditions.

Rainy imagery is fundamentally comprised of two distinct categories of elements: the precipitation-excluded resultant image and the rain streaks. The physical representation of rainfall can be elucidated through mathematical formulation[33].

$$I(x)=O(x)+R(x)$$

Where I(x) is the input rainy image, O(x) is the rain-free image, R(x) is the rain-streaks, and x represents the pixel intensity of the image.

In this model:

Input Rainy Image $I(x)$: This image captured under rainy conditions encompasses the visual manifestations of precipitation.

Rain-free Image $O(x)$: Serving as the hypothetical image devoid of rain, this exemplifies the pristine visual representation.

Rain-streaks $R(x)$: These streaks or patterns are formed by raindrops in descent, overlaying on the rain-free visual depiction.

The rain imaging model functions as a foundational framework for a variety of rain removal and image enhancement methodologies. These techniques strive to effectively depict the impact of rain by separating the rain-streaks R(x) from the initial rainy image I(x), consequently reinstating the authentic scene radiance O(x). The significance of this separation lies in the enhancement of the standard and precision of images taken under rainy circumstances, thereby enabling more effective visual examination and comprehension.

1.2 Importance of Deraining

De-raining plays a pivotal role in computer vision applications through the elimination of rain streaks from images, thereby enhancing visibility and preserving background details[24][27]. Several sophisticated de-raining techniques have been introduced to tackle this demanding task, including the progressive dilation dense residual fusion network (PDDRFN) and effective de-raining algorithms utilizing dual adjacent indexing and multi-deraining layers[3][30]. These methodologies exploit novel approaches such as feature fusion blocks, detail compensation memory mechanisms, and multi-loss functions to attain exceptional outcomes in the removal of rain streaks and image denoising. Furthermore, recent strategies concentrate on predictive filtering devoid of intricate rain model assumptions, integrating uncertainty-aware cascaded predictive filtering and weight-sharing multi-scale dilated filtering to efficiently manage diverse rain patterns. In general, de-raining is indispensable for enhancing image quality and facilitating enhanced performance across a variety of computer vision applications.

De-raining methodologies play a crucial role in improving the quality of images and visibility through the efficient elimination of the deteriorating influence of rain. The occurrence of rain, which arises from the dispersion and assimilation of light by particles in the atmosphere, notably diminishes the sharpness, differentiation, and accuracy of colors in images. Such degradation in image quality is capable of exerting adverse repercussions on a multitude of applications spanning diverse domains.

Deraining techniques play a pivotal role in enhancing visibility by addressing the impact of rain on visual clarity. The presence of rain can obscure crucial details and diminish the visibility of various objects, scenes, and structures. Conversely, deraining techniques are designed to alleviate the effects of rain, thereby uncovering concealed details and enriching the visibility of objects and features. Such advancements hold significant importance in fields like surveillance, remote sensing, and aerial photography, where the acquisition of clear and detailed images is imperative for precise analysis and interpretation.

Another notable facet pertains to the improvement of image fidelity. Rainy images frequently exhibit diminished contrast, faded coloration, and decreased clarity. De-raining methodologies reinstate the initial visual fidelity through the retrieval of contrast, reinstatement of colors, and refinement of particulars. Consequently, this results in aesthetically pleasing and increasingly captivating images, rendering them appropriate for a multitude of uses including photography, film production, and promotional campaigns.
Deraining methodologies are essential for the precise functioning of computer vision systems. Inclement weather, particularly rain, presents obstacles for various computer vision applications like object identification, classification, and monitoring. The presence of rain diminishes the clarity and definition of objects, thereby complicating the task of algorithms in precisely detecting and localizing them. Utilizing deraining techniques enhances the efficacy of computer vision algorithms through augmenting object visibility and refining their visual attributes, consequently yielding outcomes that are more dependable and accurate.

In the realm of remote sensing applications, including satellite imagery and aerial surveys, rainfall has the potential to impede the derivation of significant insights from the acquired data. Deraining methodologies facilitate the discernment and evaluation of notable characteristics on the Earth's terrain, such as terrain coverage, flora, aquatic expanses, and built structures. Such processes are imperative for endeavors like environmental surveillance, calamity appraisal, and city development.

Moreover, the implementation of deraining methods plays a crucial role in augmenting visual discourse. Precipitation, such as rain, can exert a substantial influence on the visual representation of various forms of media, such as images, videos, and multimedia materials. Methods for removing rain effects aim to reinstate the lucidity and liveliness of the content, thereby enhancing its visual efficacy and communicative potential. This factor holds significance across diverse sectors, encompassing media production, advertising, and virtual reality, particularly in the pursuit of visually stimulating and immersive encounters.

Overall, deraining techniques play a crucial role in enhancing image quality and visibility at a holistic level. They facilitate the retrieval of significant data, enhance algorithmic precision and effectiveness, and elevate the overall perceptual encounter across diverse application contexts. The evolution of efficient deraining methodologies remains a pivotal focus of investigation, fostering progressions within fields such as computer vision, remote sensing, image analysis, and related domains that heavily rely on visual information.



Fig 1.2 : An example of a rain image and its clear version

1.3 Potential of GAN based Deraining

Generative Adversarial Networks (GANs) have demonstrated considerable potential in the field of deraining applications, as indicated by numerous research investigations. Various GAN-based models, such as DerainGAN, have been created to proficiently eliminate rain droplets from individual images, resulting in noteworthy enhancements in Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) metrics when compared to conventional approaches[29]. Furthermore, the integration of wavelet thresholding with GANs has exhibited promise in improving rain removal processes, particularly by leveraging feature fusion and attention mechanisms to enhance image clarity after deraining[31]. Moreover, innovative frameworks like GRASPP-GAN have been introduced to tackle the issue of generalization observed in current methodologies, demonstrating superior performance in eliminating diverse types of rain streaks from images through sophisticated network architectures and gradient-based supervision[32]. So, GAN-based deraining strategies present efficient and effective solutions for augmenting image quality under rainy conditions.

Traditional methods for deraining typically depend on assumptions and prior knowledge regarding the degradation process induced by rainfall. Image enhancement algorithms may lead to the degradation of image details, while algorithms based on physical models necessitate manual feature extraction and lack universal applicability. The Dark Channel Prior (DCP) and

Color Attenuation Prior (CAP) are examples of such methods, leveraging statistical properties and priors associated with rain and scene attributes. Although these techniques have demonstrated promise in specific contexts, they encounter challenges with intricate scenes, diverse rain conditions, and non-uniform rain distributions, highlighting the necessity for more sophisticated and adaptable deraining approaches. Despite the favorable results achieved by deep learning-based deraining techniques, many of them still estimate ambient light and transmittance values utilizing a physical scattering model within a convolutional neural network (CNN)[23]. Furthermore, these approaches commonly employ mean square loss on a per-pixel basis, resulting in image smoothing, loss of detail, and diminished image sharpness.

In order to overcome these constraints, there is a potential opportunity for the enhancement of learning algorithms and loss functions. Several promising strategies involve the utilization of Generative Adversarial Networks (GANs) and conditional GANs, which have exhibited effectiveness in a variety of image conversion and reconstruction assignments. Furthermore, the implementation of perceptual loss functions holds the potential to enhance the visual fidelity of images. Drawing upon these factors, a patch GAN based network that is trainable has been devised to remove rain from a singular image of a particular scene.

GAN-based models have become prominent in the deraining domain, showing great promise in enhancing the quality and perceptibility of rain-afflicted images. The employment of Generative Adversarial Networks (GANs) in the context of deraining has garnered considerable interest owing to their capacity to capture intricate visual patterns and generate outputs that are remarkably true-to-life. GANs capitalize on the concept of adversarial training, wherein a generator network engages in a competitive process with a discriminator network to yield derain images that closely mirror authentic clear images.

One of the principal benefits of employing GAN-based models for deraining revolves around their capacity to acquire knowledge from data and produce authentic and aesthetically satisfying outcomes. Through the utilization of extensive datasets containing paired rain and clear images, GANs are able to grasp the intricate correlations between rain and environmental characteristics. The generator network is trained to convert rain inputs into derain outputs, effectively eliminating the rain and amplifying visibility. This learning-oriented strategy empowers GANs to address diverse forms and intensities of rain, adapt to varying scenes and lighting circumstances, and generate derain images that manifest enhanced contrast, color accuracy, and sharpness.

Comparative analysis plays a critical role in the evaluation of the efficacy of GAN-based models in the context of deraining. Given the novelty and fast-paced progression of GAN-based deraining, it becomes imperative to evaluate and juxtapose various architectural selections, loss functions, and training methodologies. The utilization of comparative analysis aids in the comprehension of the capabilities and constraints inherent in diverse GAN models and their constituents. This process empowers scholars to pinpoint the most efficient configurations and methodologies for deraining, consequently propelling advancements within the field.

Through the implementation of comparative analysis, scholars are able to assess the effectiveness of various GAN architectures, including U-Net, in relation to their capacity for producing derain images of superior quality. An exploration can be conducted into the influence of employing individual U-Net blocks as well as pairs of such blocks, essentially forming a progressive GAN, on the performance of deraining. Comparative analysis facilitates the recognition of the most effective models, thereby offering guidance for forthcoming research endeavors and the advancement of applications.

Moreover, conducting a comparative analysis aids in the evaluation of the effectiveness of GAN-based deraining models across different datasets. By subjecting the models to a variety of datasets, including widely-used benchmarks like Rain100H, Rain100L, Test100, Test1200, and Test2800, valuable insights can be gained regarding their ability to generalize and maintain robustness. This approach enables the assessment of their performance in various rain conditions, scene categories, and imaging scenarios. Additionally, comparative analysis simplifies the process of comparing deraining results based on metrics such as PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index), offering objective measurements of the deraining quality.

In conclusion, the utilization of GAN-based models exhibits considerable promise in the domain of deraining tasks owing to their capacity to produce authentic and visually attractive derain images. Nevertheless, in order to fully exploit this potential, it is imperative to conduct a thorough comparative analysis. Through a systematic examination of U-Net architectural selections, pairs of U-Net blocks, loss functions, and datasets, scholars can pinpoint the most efficient configurations and methodologies for GAN-based deraining. Comparative analysis plays a pivotal role in driving progress in the field and laying the foundation for the creation of resilient and high-performing deraining models.

1.4 Objective of the thesis

The aims of this research study are outlined below:

1. A GAN-based model is suggested for de-raining with the aim of efficiently eliminating rain in images. U-Net architecture as a generators, will be utilized by the model. The aim is to investigate the efficacy of these architectures in capturing and eliminating rain, as well as to assess their capacity to improve image quality and visibility.

2.By first passing the rain different dataset through a single U-net and then passing that result through a prior of U-Net which have been used as generator, and differences in De-Rain results or output result are shown.

3. In order to train and assess the suggested GAN model utilizing the Rain13kdataset dataset, encompassing a wide array of rain images obtained in diverse real-life situations. Through the

utilization of this dataset, the model can be effectively trained on an exhaustive range of rain conditions, leading to enhanced generalization across various environmental settings.

4. In order to evaluate the efficacy of the GAN model put forward, it is imperative to conduct testing on various benchmark datasets like Rain100H, Rain100L, Test100, Test1200, and Test2800. These datasets are indicative of diverse degrees of rain and complex scenarios. This assessment is aimed at gauging the model's effectiveness and resilience when exposed to different datasets, thus offering a thorough analysis of its performance.

5. In order to juxtapose the results of various models utilizing objective metrics such as PSNR and SSIM values, an evaluation will be conducted. The quantitative assessment will be utilized to appraise the efficiency of the suggested GAN model in comparison to other cutting-edge deraining methodologies. This examination is aimed at pinpointing the merits and demerits of each model, ultimately establishing the optimal strategy for rain mitigation.

6. In order to perform a qualitative examination through the visual juxtaposition of the de-rain images produced by the GAN model proposed with those derived from alternative methodologies, visual appraisals will be employed to assess the lucidity, disparity, and comprehensive visual excellence of the de-rain results. This examination will offer valuable perspectives on the authenticity and perceptual excellence of the de-rain images.

7. Assessing the generalization capacity of the GAN model under consideration through its evaluation on unseen or test images is imperative. This evaluation aims to gauge the model's efficacy in eliminating rain from images that were not included in the training set. Such an analysis will showcase the model's viability for practical scenarios, especially when faced with novel and unfamiliar rain patterns.

By achieving these objectives, this research paper aims to contribute to the field of deraining by proposing a Progressive GAN-based approach. The findings will provide insights regarding the impact of using a series of U-Net architectures in the order of 2^n along with the ground truth into the discriminator.

1.5 Overview of the thesis work

In this dissertation, a generative adversarial network (GAN) -based deraining framework is suggested for the purpose of improving image quality and visibility under rainy conditions. The framework employs GANs to acquire the relationship between images with rain and images without rain, facilitating the production of superior derain results.

The utilization of generator architectures such as U-Net is incorporated in this study owing to their efficacy in capturing spatial and contextual information. The primary objective entails examining the impact of these generator architectures on the deraining efficacy and determining the most suitable architecture for the proposed model. Furthermore, an investigation is

conducted into the effect of using a series of U-Net architectures in the order of 2^n for the deraining process.

To conduct training and assessment of the proposed model, the Rain13K dataset (Realistic Single Image Deraining) is employed. This dataset offers a wide array of rain and rain-free images, facilitating model training on diverse rain scenarios and improving its ability to generalize to unseen rain images. Throughout the training phase, a blend of loss functions is utilized to steer the model, including Mean Squared Error (MSE) loss, perceptual loss, and adversarial loss. Through the incorporation of these loss functions, the generator is motivated to produce derain outputs that are not only visually precise but also realistic. Furthermore, apart from training on the Rain13K dataset, the model's performance is assessed on various benchmark datasets such as Rain100H, Rain100L, Test100, Test1200, and Test2800. This assessment allows for the evaluation of the model's generalization capacity across different rain scenarios and facilitates a comparison with other cutting-edge deraining techniques. The comparison is grounded on quantitative metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), which furnish objective assessments of image quality and support a thorough examination of the model's efficacy.

Outline of the Thesis

This dissertation is structured into multiple sections, each dedicated to examining a particular facet of the investigation on deraining methods based on GANs. The framework of the dissertation is outlined as follows:

1. Introduction

- Overview of rain removal and its significance in various applications.
- Explanation of the raining model and its impact on image quality and visibility.
- Importance of deraining techniques in enhancing image quality and restoring visibility.
- Potential of GAN-based approaches in addressing the challenges of rain removal.
- Clear statement of objectives for the thesis.

2. Motivation and Contribution

- Discussion of the motivation behind the research on rain removal techniques.
- Explanation of the specific contributions of the thesis.

3. Related Work

- Review of prior-based, learning-based, deep learning-based, and GAN-based approach.
- Explanation of the principles, techniques, and contributions of each approach.
- Comparison of strengths and weaknesses of different methods.
- Identification of limitations in existing approaches.
- Addressing the research gap in the field.

4. **Proposed Method**

- Description of the GAN-based method for deraining.
- Explanation of the architecture of the generator and discriminator networks.
- Elaboration on the architecture employed for the generator (U-Net).
- Discussion of the progressive GAN used (U-NetR).
- Detailed explanation of the loss function employed.

5. **Experimental Setup**

- Overview of the datasets used for training and evaluation (Rain 13K).
- Description of preprocessing steps applied to the datasets.
- Discussion of training details (optimizer, learning rate schedule, number of epochs).
- Specification of performance metrics used for evaluation (PSNR, SSIM).

6. **Evaluation**

- Presentation of quantitative results using performance metrics.
- Comparison of outcomes among various methods.
- Qualitative evaluation through visual comparisons with original rain images and state-of-the-art techniques.

7. **Conclusion**

- Summary of contributions and key findings of the research.
- Suggestions for future research directions.

Chapter 2
Motivation

This chapter presents the motivation behind the research conducted in this thesis and outlines the specific contributions made in the field of deraining using GAN-based models.

2.1 Motivation

Rain is a common atmospheric phenomenon that significantly impacts the quality and visibility of images.The phenomenon arises from the dispersion and assimilation of electromagnetic radiation by particulates and contaminants within the atmospheric medium, leading to diminished sharpness, chromatic aberrations, and degradation of intricate features. Images affected by such phenomena present notable hurdles across multiple fields, such as machine vision, aerial surveillance, and self-governing navigation technologies.

In the field of computer vision and image processing, precipitation hinders the ability to detect, recognize, and track objects, resulting in a decline in performance and accuracy. Within remote sensing and aerial imaging, rain diminishes the clarity of significant elements like topography, structures, and points of interest, impeding functions such as environmental observation, crisis response, and cartography. Rainfall in autonomous vehicle systems obstructs the identification of barriers, traffic signs, and pedestrians, thereby creating safety concerns. Additionally, precipitation in photography can detrimentally impact the visual allure and artistic impression of photographs, influencing both novice and expert photographers.

The significance of deraining methods resides in their capacity to enhance image fidelity, increase clarity, and augment the interpretability of visual data acquired under rainy circumstances. Through the efficient elimination of rain elements, these methods facilitate a more thorough examination and comprehension of visual data. They play a crucial role in unveiling obscured intricacies, reinstating authentic color representation, and enhancing the holistic perceptual attributes of images, rendering them indispensable across various applications.

The rationale for this study arises from the crucial necessity to develop efficient deraining methodologies that can recover the authentic visual representation of rain-afflicted images. Conventional deraining techniques have played a significant role in alleviating the impacts of rain; nevertheless, they often depend on manually crafted priors or assumptions regarding the environment and rain characteristics, thereby restricting their applicability to diverse and intricate rain scenarios. Moreover, they might encounter difficulties in preserving intricate details, leading to color distortions or displaying limited resilience to fluctuating rain conditions, consequently yielding substandard deraining outcomes and constraining their pragmatic utility. Deep learning strategies, especially Generative Adversarial Network (GAN)-based models, have exhibited

considerable potential in various image synthesis tasks, such as deraining. GANs provide a robust framework for acquiring intricate mappings between rainy and clear images through adversarial training, enabling them to capture the inherent distribution of rain-free images and produce visually authentic derained results. The rationale behind this investigation is to assess the capabilities of GAN-based models in confronting the challenges associated with rain elimination, while also enhancing image quality and visibility.

The rationale behind the investigation of deraining techniques based on GANs arises from the potential benefits they provide. GANs exhibit the capacity to acquire intricate relationships between rainy and clear images without directly representing the deterioration process. They possess the capability to encompass detailed image features, patterns, and configurations, resulting in visually authentic and top-notch derained results. Moreover, GANs demonstrate proficiency in generalizing effectively to novel or evaluation images, rendering them appropriate for practical real-world scenarios.

Change–By proposing a GAN-based deraining model in this thesis, the goal is to contribute to the development of advanced techniques that overcome the limitations of traditional methods and provide superior deraining results. The use of generator architectures, such as U-Net along with a series of U-Net(s) in the order of 2^n is used to observe the deraining performance.

The incentive expands to the assessment and juxtaposition of the proposed framework and U-Net-based progressive GAN using datasets. Through evaluating the training results on datasets such as Rain13K datasets, the dissertation aims to offer a comprehensive comprehension of the model's performance in diverse rain scenarios. Objective performance metrics, such as peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM), will be employed to quantitatively assess the quality and faithfulness of the drained images.

Chapter 3
Literature Review

3.1 Existing Derainig Techniques

Various approaches are covered by existing deraining techniques, such as prior-based derain techniques, learning-based derain techniques, U-Net-based derain techniques, and GAN-based derain techniques. The prior-based methods, exemplified in the work of Cheng et al., exploit prior knowledge of rain patterns to effectively eliminate rain textural features and restore high-quality rain-free images[33]. Learning-based techniques, as discussed by Zhang et al., employ deep neural networks and extensive training data to significantly enhance deraining performance compared to conventional methods, emphasizing aspects like data, rain models, and network architecture to improve generalization ability and real-world usability[34][35]. Furthermore, GAN-based image restoration, as investigated by Poirier-Ginter and Lalonde, exhibits resilience in addressing various degradations without requiring extensive fine-tuning, surpassing other StyleGAN-based inversion techniques and producing realistic outcomes in inpainting, upsampling, denoising, and deartifacting tasks[36].

3.1.1 Prior-Based Derain Techniques

Rain removal from images has been a difficult task due to the negative impact on image quality. In recent years, several techniques have been proposed to address this problem, with a focus on leveraging prior knowledge and advanced network architectures.

One notable approach is to incorporate multi-sub-networks, inner-scale fusion, and dense blocks into the de-raining process. By incorporating these elements, these methods aim to significantly improve rain removal performance. For example, using dense blocks with skip connections improves feature extraction while preserving spatial information during rain removal.

The Cross-Scale Fusion Network represents a significant advancement in contemporary research, aimed at overcoming the constraints of earlier methodologies through the enhancement of rain depiction capabilities. This architectural design surpasses established methods in both artificial and authentic data collections. Moreover, the integration of Gate Recurrent Units enhances the network's capacity to apprehend chronological interdependencies, consequently enhancing the overall rain elimination performance.

Notwithstanding these achievements, obstacles persist, such as the potential occurrence of data loss in the course of individual network training and the inadequate acknowledgment of cross-scale correlations in precipitation eradication. Prospective investigations in this domain

could explore approaches to tackle these challenges and refine the efficacy of preceding de-rain methodologies.[1]

3.1.2 Learning-Based Derain Techniques

In recent years, learning-based derain techniques have made significant advances, revolutionizing the field of single image derain. These approaches combine data-driven and model-based methodologies to address the difficult task of removing rain streaks, outperforming traditional methods.

One notable approach uses an enhanced weighted guided image filter (iWGIF) in conjunction with adaptive feature learning and restoration networks that include attention mechanisms [2]. This method outperforms existing techniques in both qualitative and quantitative terms by efficiently extracting high-frequency information and learning rain streak patterns. The combination of data-driven and model-based strategies allows the algorithm to handle spatially varying rain streaks in single rainy images, overcoming a critical limitation of CNN-based methods.

A high-order collaborative network with multi-scale compact constraints and a bidirectional scale-content similarity mining module [3] also make important contributions. This innovative framework uses deep network features both internally and externally, imposing constraints on intermediate layers and exploring features at different scales. Furthermore, an online-update learning approach improves the model's adaptability to real-world images, resulting in better performance metrics like PSNR and SSIM on synthetic datasets. The proposed method outperforms eleven cutting-edge techniques on various datasets, demonstrating its effectiveness in rain streak removal.

By combining and refining these approaches, a comprehensive literature survey emerges, providing a holistic understanding of learning-based derain techniques. These methodologies represent a significant step toward more sophisticated and effective solutions for single image deraining, addressing issues such as data dependency, insufficient interpretation in CNN-based methods, and the presence of spatially varying rain streaks. The use of advanced filtering techniques, attention mechanisms, and collaborative networks emphasizes the importance of combining data-driven and model-based strategies to achieve optimal results.

Its learning-based derain techniques represent a paradigm shift in the field of image restoration, opening up promising avenues for future research and development. By combining data-driven and model-based approaches, these methods enable improved visual quality and accuracy in single image deraining applications.

3.1.3 U-Net-based Derain Techniques

U-Net-based derain techniques have emerged as effective tools in the field of image deraining, demonstrating significant progress in rain removal tasks. These methodologies use the U-Net architecture or its extensions to outperform conventional methods in terms of PSNR, SSIM metrics, and visual quality. Several U-Net variants have been developed, including U-Net++, AU-Net, and Asymmetric Dual-decoder U-Net (ADU-Net), each with its own set of advantages for rain streak detection and removal

The U-Net++ model expands on the original U-Net architecture and has shown high efficacy in rain removal tasks. By incorporating a kernel prediction network and pixel-wise multi-dilated filtering, this model outperforms traditional methods in terms of PSNR and SSIM metrics. The use of a loss function that includes SSIM, edge loss, and L1 loss increases rain removal efficiency while maintaining image structure. Furthermore, the U-Net++ model outperforms existing methods in both synthetic and real-world data benchmarks, demonstrating its robustness and versatility in different scenarios. [13][16][18].

The dual-decoder U-Net architecture, which includes one decoder for contamination residual and another for scene residual, addresses rain and rain removal in scene images. This method introduces the Asymmetric Dual-decoder U-Net (ADU-Net), which outperforms prior-based and deep learning approaches to image derain. ADU-Net improves PSNR values significantly on benchmark datasets such as RainCityscapes and SPA-data by combining a Contamination Residual Net with a cCFF module for scene image localization [12][17].

.Furthermore, the AU-Net model, which combines U-Net and dual attention mechanisms, has demonstrated promise in capturing contextual information for rain streak detection. AU-Net detects and removes cyclic rain streaks with high accuracy by combining dual-attention mechanisms and multi-convolution networks. Detailed experimental evaluations and dataset assessments validate AU-Net's effectiveness in rain removal tasks, emphasizing its potential for real-world applications [15].

Overall, U-Net-based derain techniques use advanced network architectures, attention mechanisms, and loss functions to achieve superior rain removal performance. These methodologies show significant improvements in both quantitative metrics and visual quality, outperforming traditional methods and providing promising solutions for a variety of deraining applications.

3.1.4 GAN-based Derain Techniques

Recent advancements in Generative Adversarial Networks (GANs) have resulted in significant progress in the field of single-image detection. This literature review provides a comprehensive overview of notable contributions from various methodologies, such as DerainAttentionGAN, RR-GAN, Cycle-attention-derain, AI-GAN, CycleGAN-based CNN, and GAN with ViT discriminator, in addressing critical challenges in rain streak removal.

DerainAttentionGAN is an unsupervised approach to single-image deraining that uses attention-guided GANs to outperform traditional unsupervised methods [4]. DerainAttentionGAN achieves significant quantitative and qualitative improvements by incorporating attention-guided generators, multiscale discriminators, cycle-consistency, and perceptual-consistency losses, effectively reducing residual rain lines and enhancing background color [4][11][19]. However, it faces challenges such as the slow development of unsupervised Single Image Deraining (SID) due to irregular rain lines and the insufficient use of PatchGAN to improve image quality [4].

RR-GAN proposes an end-to-end adversarial model for single-image rain removal based on unpaired supervision, multiscale attention memory generators, and deeply supervised discriminators [5]. This approach outperforms nine state-of-the-art methods [5][14]. However, it faces challenges such as hand-designed priors that are easily violated and the difficulty of collecting paired training data [5].

Cycle-attention-derain focuses on rain removal in multimedia applications by using an asynchronous interactive two-branch structure with mutual adversarial optimization [6]. In this framework, AI-GAN outperforms traditional deraining methods in multimedia applications [6][20]. However, existing methods overlook the shape, scale, and density of rain streaks, and CNN-based methods overlook the interdependence between rain streaks and background [6].

CycleGAN-based CNN addresses the single-image deraining problem with a CycleGAN-based approach that includes dual learning, residual image learning, and the Squeeze-and-Excitation (SE) block to prevent overfitting on synthetic training data [7][21][22]. Despite its effectiveness, this method faces difficulties due to defects in brightness loss, overfitting on synthetic data, and the complexity of rain streak removal and background preservation [21][22].

The RCA-cGAN model focuses on image de-raining, capturing spatial relationships through capsule structures in the generator and discriminator [9]. This model outperforms other methods for de-raining synthetic images [9][26]. Rain feature variability, as well as the depth limitation of traditional Convolutional Neural Networks (CNNs), pose challenges for global information extraction [9].

GAN with ViT discriminator is a method for single-image deraining that employs attention mechanisms and a small-data training method [10]. To improve the rain removal effect, this approach uses a progressive recurrent loss function and a data reselection algorithm [10][28]. While promising, challenges remain, such as the need for more explicit mention of limitations [10].

Thus, GAN-based derain techniques demonstrate promising advances in single-image deraining, providing unsupervised solutions that outperform traditional methods. While these methodologies are highly effective, they also face challenges such as irregular rain lines, hand-designed priors, and errors in rain streak characteristics. Addressing these limitations is critical to increasing the robustness and applicability of GAN-based derain techniques.

3.2 Strengths and limitations of Existing techniques

Removal of rain streaks from individual images poses a foundational challenge within the field of computer vision, carrying considerable real-world implications. Progress in the realm of single-image deraining techniques has been notably propelled by recent developments in Generative Adversarial Networks (GANs). This review of literature delivers a thorough examination of the current landscape of GAN-based derain methodologies, emphasizing their advantages, constraints, and prospects for enhancement.

3.2.1 Strengths of Existing Techniques

Recent advancements in both GAN and U-Net-based derain techniques have brought significant improvements to rain removal tasks, showcasing several strengths that contribute to their effectiveness and widespread adoption.

One primary strength of GAN-based derain techniques lies in their ability to perform unsupervised rain removal. Unlike traditional methods reliant on paired datasets containing both rainy and rain-free images, GAN-based approaches can learn from unpaired datasets. This flexibility in data requirements extends the applicability of deraining models to scenarios where obtaining paired data may be challenging or expensive.

Furthermore, GAN-based derain techniques often demonstrate superior quantitative and qualitative performance compared to traditional methods. Metrics such as PSNR and SSIM, along with perceptual metrics and visual inspection, consistently show the effectiveness of GAN-based techniques in producing visually pleasing results with reduced rain artifacts and enhanced image quality.

The incorporation of attention mechanisms is another notable strength of many GAN-based derain techniques. By selectively focusing on regions affected by rain streaks, attention-guided generators can effectively suppress rain streaks while preserving essential image features like edges and textures. This attention mechanism significantly improves the efficiency and effectiveness of rain removal, resulting in more accurate and visually appealing results.

Techniques such as RR-GAN and AI-GAN leverage multi-scale discriminators and dual learning strategies to enhance the deraining process further. Multi-scale discriminators enable the model to capture both local and global features, enhancing its ability to distinguish between rain streaks and background details. Dual learning strategies allow the model to simultaneously learn from both rain-affected and rain-free images, leading to more robust and generalizable deraining performance.

Moreover, methodologies such as DerainAttentionGAN, RR-GAN, and AI-GAN consistently demonstrate superior performance metrics compared to state-of-the-art methods. Achieving higher PSNR and SSIM scores indicates better preservation of image quality and structural similarity between rain-free and derained images. Additionally, perceptual metrics and user studies often confirm the visual superiority of GAN-based derain techniques, further validating their effectiveness.

In the realm of U-Net-based derain techniques, several strengths stand out. The U-Net++ model, for instance, exhibits remarkable effectiveness in rain removal tasks, surpassing conventional methods in terms of PSNR and SSIM metrics. Its proposed algorithm enhances the Peak Signal-to-Noise ratio and incorporates a loss function that includes SSIM, edge loss, and L1 loss, leading to efficient rain removal and better preservation of image structure.

The dual-decoder U-Net architecture and the AU-Net model further contribute to the strengths of U-Net-based derain techniques. The dual-decoder U-Net architecture addresses rain and rain removal in scene images, outperforming existing methods in various benchmarks. The AU-Net model effectively captures contextual information for rain streak detection, enhancing the accuracy of rain removal processes with its dual attention mechanisms and multi-convolutional channel.

the strengths of both GAN and U-Net-based derain techniques lie in their capacity to execute unsupervised rain removal, exhibit superior quantitative and qualitative results, integrate attention mechanisms, employ multi-scale discriminators and dual learning strategies, and consistently attain superior performance metrics when contrasted with conventional approaches. Together, these strengths significantly enhance the efficiency and suitability of derain techniques in diverse real-world contexts.

3.2.2 Limitations of Existing Techniques

Existing GAN-based derain techniques face several limitations that impact their effectiveness and practical utility in real-world scenarios.

One significant limitation is the difficulty in handling irregular rainlines and variations in rain streak characteristics. Traditional CNN-based methods may struggle to effectively capture and remove complex rain patterns, leading to residual artifacts and suboptimal deraining results. Additionally, hand-designed priors and assumptions about rain streak characteristics may limit the generalization capabilities of deraining models, particularly in real-world scenarios where rain patterns can be highly variable.

Furthermore, numerous derain techniques based on Generative Adversarial Networks (GANs) heavily depend on synthetic data during the training phase. This heavy reliance on synthetic data can potentially lead to overfitting and a lack of generalization when applied in real-world settings. The synthetic data used may not sufficiently capture the intricate variations and diversity present in actual rain streaks, hence resulting in subpar performance when these techniques are applied to real images. Consequently, models that overfit on synthetic data may exhibit sensitivity towards variations within synthetic datasets, yet struggle to generalize effectively when faced with unseen real-world data, consequently limiting their practical applicability in deployment scenarios.

Another constraint of derain techniques based on Generative Adversarial Networks (GANs) is their intricate computational nature, which may pose a hindrance for applications that require real-time processing or operate within resource-limited settings. The duration for training and making predictions with GAN models can be substantial, necessitating substantial computational power and storage capacity. This computational burden has the potential to restrict the scalability and feasibility of GAN-based derain techniques when applied in systems that are integrated, portable devices, or contexts with stringent latency demands.

Furthermore, existing GAN-based derain techniques may overlook certain aspects of rain streak characteristics, such as shape, scale, and density. CNN-based methods, in particular, may struggle to capture the intricate details of rain streaks and their interactions with the background scene. Failure to accurately model these characteristics can result in incomplete or inaccurate rain removal, leading to residual artifacts and reduced image quality.

In addition to these technical limitations, GAN-based derain techniques face challenges in real-world deployment. Factors such as varying lighting conditions, camera settings, and scene complexity can impact the performance of deraining models, requiring robustness and adaptability to diverse environmental conditions. Additionally, integrating deraining models into

existing image processing pipelines or applications may require additional engineering efforts and considerations to ensure seamless operation and compatibility.

Similarly, U-Net-based derain techniques also exhibit several limitations that hinder their effectiveness and practical applicability.

One limitation is the computational complexity, especially with advanced architectures such as U-Net++. The incorporation of multiple attention mechanisms and convolutional layers can increase computational overhead, potentially limiting real-time application in certain scenarios.

Moreover, while U-Net-based models often demonstrate superior performance on benchmark datasets, their generalization to diverse environmental conditions and image types remains a challenge. Models trained on specific datasets may struggle to effectively remove rain streaks in unseen or complex scenarios, necessitating further research on data augmentation and domain adaptation techniques.

Another concern is the potential generation of artifacts or distortions in derained images, particularly when removing heavy rain streaks or complex patterns. This issue may arise due to limitations in the network's capacity to accurately reconstruct image details while removing rain, necessitating ongoing refinement of loss functions and network architectures.

Additionally, U-Net-based derain techniques often rely on large-scale labeled datasets for training, posing challenges in data collection and annotation. The availability of diverse and representative datasets is crucial for ensuring model robustness and generalization across different environmental conditions and image domains.

So, both GAN and U-Net-based derain techniques offer promising solutions for rain removal tasks, but they also face several limitations that need to be addressed to improve their effectiveness and practical applicability in real-world scenarios. Addressing these limitations through continued research and innovation will be essential for advancing the field of single-image deraining.

3.3 Generative Adversarial Networks

Generative Adversarial Networks (GANs) represent a category of sophisticated deep learning models comprising two distinct neural networks, namely a generator and a discriminator. The inception of GANs dates back to 2014 when Ian Goodfellow and his research associates unveiled this innovative concept, subsequently leading to widespread recognition and adoption within the realm of machine learning.

The generator network utilizes random noise or a latent vector as an input and converts it into a sophisticated output sample, often an image. Its primary aim is to comprehend the fundamental data distribution and produce artificial samples that mimic genuine data from a specified distribution. Comprising various layers like fully connected layers, convolutional layers, or a blend of both, the generator gradually changes the initial noise input into a higher-dimensional output. The design of the generator may differ based on the particular task and dataset at hand. Its objective is to create samples that are compelling enough to deceive the discriminator into categorizing them as authentic. Throughout the training phase, the generator acquires the ability to generate samples that encapsulate the patterns, configurations, and statistical features of the training data. By refining its parameters, the generator strives to decrease the distinguishability between the generated samples and real data. Conversely, the discriminator network, also known as the critic in certain contexts, is tasked with discriminating between authentic and artificial samples. It accepts either a genuine sample from the target distribution or a synthetic sample produced by the generator as input. The discriminator's function is to determine whether the input sample is genuine or fake. Typically, the discriminator acts as a binary classifier that issues a probability score denoting the probability of the input being authentic or artificial.

The discriminator is trained on a mixture of authentic and synthesized samples. Its primary goal is to effectively differentiate between genuine and fabricated samples, thereby acquiring the ability to capture the distinctive features and statistical characteristics of the authentic data. As the training advances, the discriminator's capacity to discern between real and synthetic samples enhances. The training procedure entails a confrontational interaction between the generator and discriminator. The generator's objective is to produce samples that are indistinguishable from authentic data by the discriminator, whereas the discriminator's aim is to precisely categorize the samples as either authentic or fabricated. Both the generator and discriminator are recurrently adjusted, with each modification bringing them closer to their respective goals. This adversarial learning mechanism stimulates the generator to create more authentic samples, while the discriminator refines its skill in distinguishing between authentic and fabricated samples. The dynamic interaction between the generator and discriminator throughout the training process results in a competitive dynamic, where the generator strives to generate samples that closely resemble authentic data, and the discriminator enhances its ability to detect synthesized samples. This cyclic process persists until a desired level of authenticity is attained, or until convergence is achieved.

The collaborative interaction between the generator and discriminator establishes a robust framework for generative modeling. The generator is designed to understand the inherent data distribution and produce samples, whereas the discriminator serves as a crucial assessor, offering evaluations to the generator in order to enhance the quality of the generated samples. This antagonistic dynamic propels the Generative Adversarial Network (GAN) towards generating a wide array of high-quality samples that closely mirror the authentic data distribution. The training procedure of GANs comprises a competitive scenario involving the generator and discriminator. The primary objective of the generator is to create artificial samples that are virtually indistinguishable from real ones, while the discriminator endeavors to enhance its capacity to discern between authentic and synthetic samples. Through an iterative training process, the generator and discriminator are incrementally refined, bringing them closer to their individual objectives. This process of adversarial learning motivates the generator to enhance its capability to produce more authentic samples, whereas the discriminator evolves to better differentiate between real and fabricated samples.

One of the principal benefits of Generative Adversarial Networks (GANs) lies in their capacity to produce highly authentic and varied samples. The generator model effectively grasps the fundamental data distribution, enabling the creation of samples that mirror the characteristics of the training data. GANs have found successful applications across diverse domains, encompassing image synthesis, text generation, music creation, and beyond. Moreover, the realm of GANs has seen expansions into Conditional GANs (CGANs), wherein supplementary information or constraints are furnished to both the generator and discriminator. This facilitates a more precise and focused generation process, allowing for tasks like producing images based on specific attributes or generating visuals in accordance with textual descriptions. The impact of GANs on various research fields has been profound, leading to notable strides in image synthesis, data enrichment, and related endeavors. Furthermore, their integration with other methodologies is evident in applications like GAN-based deraining, where GANs are utilized to eliminate rain and enhance visibility in visual content. Nonetheless, the training of GANs poses challenges that demand meticulous hyperparameter optimization, substantial training datasets, and extensive computational resources. The training phase of GANs can be turbulent, giving rise to issues like mode collapse, a scenario in which the generator falters in producing a diverse range of outputs. To combat these hurdles and enhance the reliability and efficacy of GANs, researchers are actively exploring innovative strategies.

Overall, Generative Adversarial Networks (GANs) represent a potent framework for the task of generative modeling, facilitating the production of authentic and varied samples. The capacity of GANs to comprehend intricate data distributions and produce top-notch samples has rendered them a favored option across multiple domains, such as image synthesis, style conversion, enhanced image resolution, and removal of rain effects, among various others.

Chapter 4

Methodology

This paper introduces an innovative strategy for addressing the problem of removing rain streaks in computer vision through the utilization of a Progressive Generative Adversarial Network (GAN)-based approach. The primary objective is the enhancement of visibility and image quality by eliminating rain streaks from images. The suggested framework comprises a generator and a discriminator operating within an adversarial context to progressively refine the deraining procedure. This section delineates the methodology, encompassing the structure, training regimen, datasets, and performance metrics employed for the development and evaluation of the model.

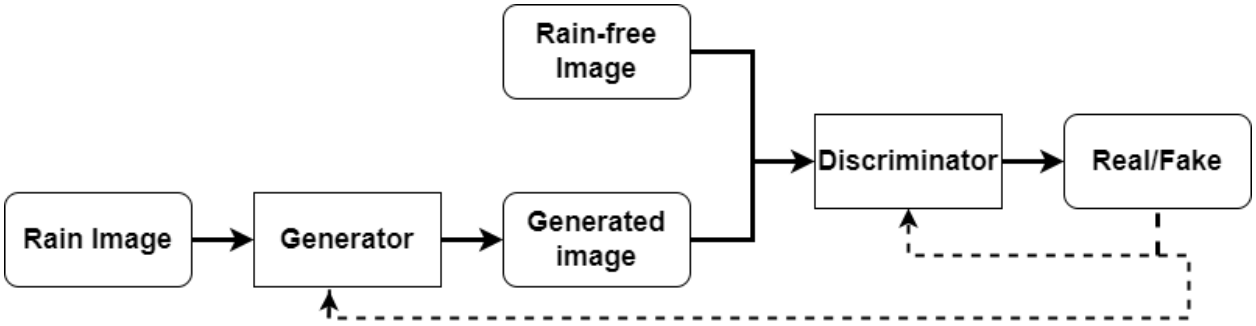


Fig 4.1 : The Architecture of derain algorithm using GAN

The instruction of the deraining model based on GAN involves the simultaneous optimization of the generator and discriminator using an adversarial training approach. Throughout the training of the generator, the objective is to reduce the gap between the produced images and the actual images, while also tricking the discriminator into categorizing the produced images as authentic. On the other hand, the training of the discriminator focuses on developing the capability to accurately differentiate between real and produced images, thus improving its ability to distinguish between them. The training procedure utilizes a variety of loss functions.

Adversarial Loss: The loss function encourages the generator to generate images without rain that closely resemble real images, thereby tricking the discriminator. Content Loss: Typically determined through Mean Squared Error (MSE), this loss guarantees visual similarity between the produced images and the actual images, concentrating on accuracy at the pixel level. Perceptual Loss: Computed with a pretrained network like VGG, this loss ensures that the high-level characteristics of the generated images match those of the actual images, highlighting semantic coherence.

The Rain13K dataset functions as the principal dataset utilized for both training and assessment purposes, providing a wide array of rainy scenarios to improve the model's ability to generalize to unseen rainy images. Furthermore, the model's efficacy is measured across multiple standard datasets, such as Rain100L, Rain100H, Test100, Rain1200, and Test2800. These datasets offer varied and complex rainy conditions, allowing for a thorough evaluation of the model's performance across different scenarios.

The execution of the proposed Generative Adversarial Network (GAN)-based model for deraining involves a series of fundamental procedures. Initially, there is a preprocessing stage carried out on both images affected by rain and their corresponding ground truth rain-free images. This preprocessing phase encompasses activities like resizing, normalization, and data augmentation, all of which are designed to improve the resilience and efficacy of the model. Subsequent to this, the generator and discriminator undergo training utilizing specified loss functions, with meticulous adjustment of hyperparameters to optimize the training procedure and attain optimal performance. Post training, the effectiveness of the model is assessed on standard datasets using objective metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). These metrics offer quantitative evaluations of the accuracy and authenticity of the derained images, enabling comparisons with cutting-edge deraining methodologies.

In essence, this methodology elucidates a systematic strategy for the creation, instruction, and assessment of the envisioned Progressive GAN-based deraining model. Through the utilization of sophisticated GAN methodologies and thorough evaluation criteria, this strategy strives to significantly augment clarity in images influenced by rain, consequently driving progress in the realm of image deraining within computer vision.

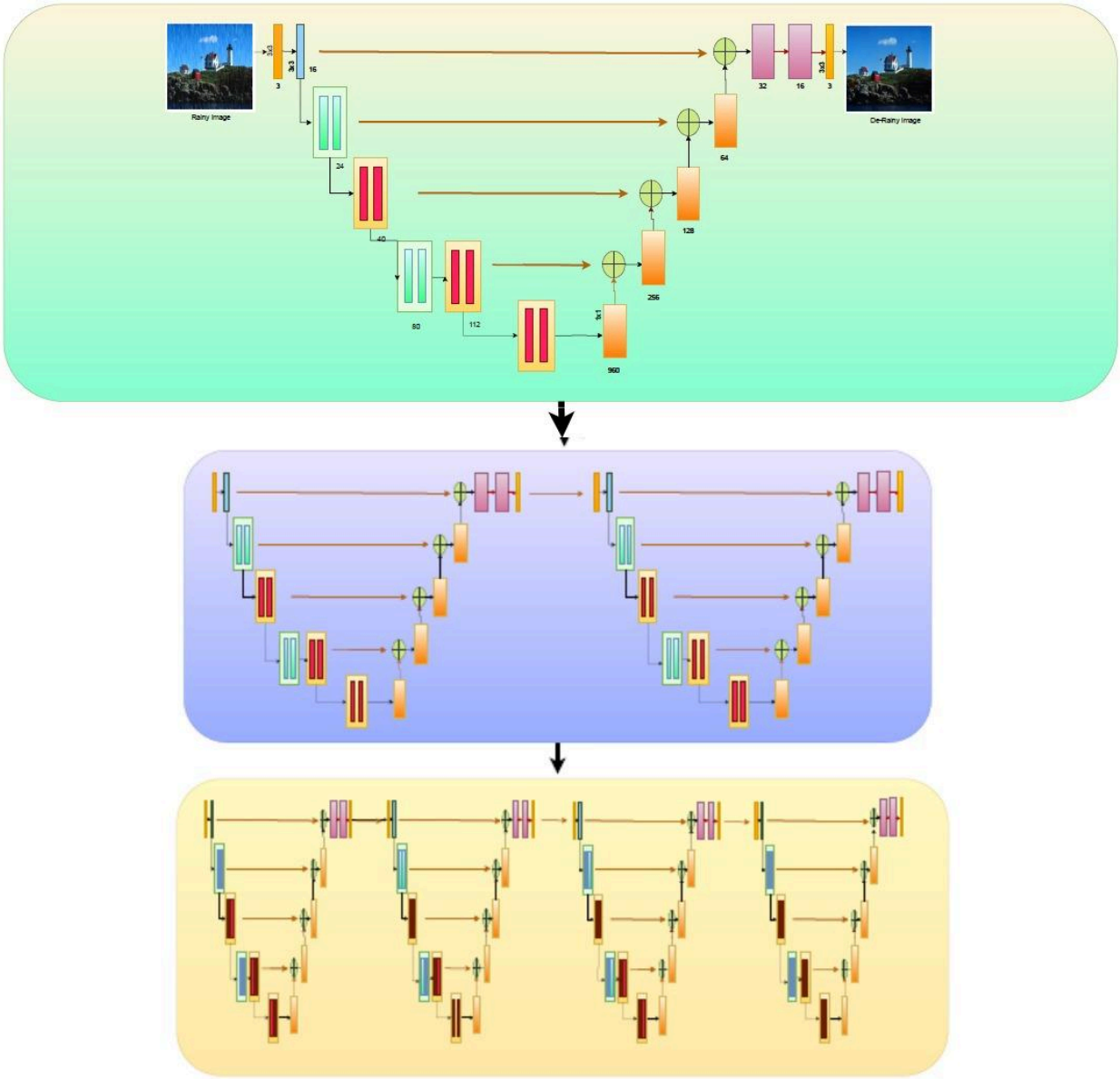


Figure 4.2: Illustration of Proposed Progressive Generator Architecture - (a) Single U-Net Architecture Block. (b) Pair of U-Net Architecture Block. (c) Two Pair of U-Net Architecture Block.

4.1 Generator

The generator within the suggested Generative Adversarial Network (GAN) framework aimed at de-raining is accountable for converting input images containing rain into distinct and aesthetically pleasing de-rain results. It adheres to an architecture comprising an encoder-decoder structure featuring various generator blocks and encoder blocks.

The commencement of the generator involves an initial encoder block tasked with encoding the rainy input image to extract high-level features that encapsulate the fundamental information pertaining to the presence of rain. At this stage, a variety of encoder blocks like CNN, VGG, MobileNet, EfficientNet, ResNet, and Vision Transformer can be utilized [23][25]. These distinct encoder blocks exhibit diverse architectural configurations tailored to effectively capture and represent image features. Subsequently, the encoded features derived from the encoder block are transmitted to the decoder block, which is accountable for the reconstruction of the de-rain output. The decoder block adopts architectures such as [23] CNN, Residual, U-Net, PSPNet, MANet, or FPN, that are explicitly crafted to grasp and merge contextual information, facilitating precise and thorough reconstruction of the de-rain image. Through a gradual process, the decoder block performs upsampling on the encoded features by employing a sequence of upsampling layers or transposed convolutions. This procedure results in the expansion of the spatial dimensions of the encoded features, enabling the generator to produce a de-rain output that aligns with the dimensions of the original input image. Throughout the upsampling phase, the incorporation of skip connections is commonplace. These connections serve to establish direct links between corresponding encoder and decoder layers. By making use of skip connections, the generator is able to exploit both low-level and high-level features concurrently, thereby capturing intricate details and contextual information to enhance the quality of the de-rain outcomes.

Various loss functions can be utilized to enhance the performance of the generator and steer its training process. Loss functions commonly employed encompass Mean Squared Error (MSE) loss, Structural Similarity Index (SSIM) loss, and adversarial loss. These functions gauge the disparity between the generated de-rain image and the authentic clear image, motivating the generator to yield outputs that exhibit perceptual resemblance to genuine clear images. The primary aim of the generator is to comprehend the mapping between rain and clear images, capturing the intricate relationships and patterns within the training dataset. By means of an iterative training regimen, the generator's parameters are fine-tuned to diminish the distinction between the produced de-rain images and the authentic clear images. Consequently, this process empowers the generator to generate high-quality de-rain results when deployed on unfamiliar rain images during the inference stage.

Through the integration of various generator blocks, encoder blocks, and loss functions, the suggested GAN model demonstrates proficiency in removing rain from images and improving their clarity. The structure of the generator and the methods employed during training enable the

model to understand and depict the fundamental attributes of rain, ultimately resulting in the production of aesthetically pleasing de-rained outputs.

The suggested de-rain model based on GAN employs a variety of architectures, specifically U-Net, as the generator component. The selection of these architectures is done thoughtfully, considering their unique traits and capabilities in capturing and restoring image details throughout the de-rain process. Each architecture presents a distinct methodology for feature extraction and image reconstruction, enabling a thorough investigation of their effectiveness in rain elimination as elaborated below:

4.1.1 U-Net

The U-Net architecture, a prominent deep learning framework, has been extensively applied in diverse image-to-image conversion assignments such as deraining. While initially developed for segmenting biomedical images, its efficacy in deraining tasks has been remarkable.

The U-Net architecture is distinguished by its symmetrical encoder-decoder configuration featuring skip connections. The encoder segment of U-Net comprises multiple convolutional layers succeeded by downsampling procedures like max pooling or strided convolutions. This downsizing operation diminishes the spatial resolution of the feature maps while augmenting their depth or channel count, enabling the network to capture hierarchical and abstract representations of the input image. The decoder component of U-Net, symmetric to the encoder, is designed to restore the spatial resolution of the feature maps through upsampling techniques such as transposed convolutions or bilinear interpolation to enhance the spatial dimensions. These upsampling methods are commonly succeeded by convolutional layers for feature refinement. The standout feature of U-Net is the incorporation of skip connections that link corresponding encoder and decoder layers, providing a direct passage for low-level features from the encoder to reach the decoder layers. By integrating detailed low-level information with high-level contextual data, U-Net can maintain fine-grained details throughout the deraining process. The presence of skip connections empowers U-Net to effectively capture and utilize both local and global features. The early layers' capture of low-level specifics assists in recovering subtle structures, while the later layers' comprehension of high-level context contributes to understanding the entire scene and rain distribution. The amalgamation of local and global insights plays a pivotal role in U-Net's precise deraining capabilities.

The U-Net architecture, characterized by its encoder-decoder design and skip connections, has demonstrated efficacy in capturing intricate details and maintaining a holistic perspective, rendering it a widely favored option for tasks involving the removal of rain.

4.1.2 Progressive GAN :

Generative Adversarial Networks (GANs) represent a category of neural networks widely employed in the realm of unsupervised learning. These networks consist of a generator and a discriminator neural network, where the generator model is tasked with producing novel images through the assimilation of knowledge from an established image dataset. The term "adversarial" in GAN originates from the fact that it involves the training of two distinct networks engaged in a competitive process. One network manipulates a given data sample to generate new data to the fullest extent feasible, while the other network endeavors to ascertain the authenticity of the generated data. The Progressive Growing GAN stands out as a notable progression within the domain of Generative Adversarial Networks (GANs), facilitating the creation of high-fidelity, large-scale images. This technique implements an innovative strategy that involves incrementally enhancing the dimensions of the images generated by the generator, commencing from 4×4 pixels and progressively scaling up to 8×8, 16×16, and so forth, until attaining the desired output resolution.The progression is facilitated through a comprehensive training schedule comprising phases aimed at enhancing the model with a specific output resolution, followed by phases focused on gradually integrating a new model with a higher resolution. Throughout this iterative process, all pre-existing and newly incorporated layers will retain their trainability. The utilization of the Progressive Growing GAN framework has gained popularity for the generation of high-quality images employing convolutional neural networks. This methodology encompasses the employment of two models, namely a generator and a discriminator, both characterized by similar architectures. The commencement of the training procedure involves small images and subsequently incorporates additional blocks of convolutional layers to enhance both the generator and discriminator architectures as training advances. This step-by-step augmentation of layers allows for the efficient capture of coarse-level details by the models before honing their capacity to capture finer details. Through the application of this method, the models are capable of producing high-quality images that exhibit realism and aesthetic appeal.

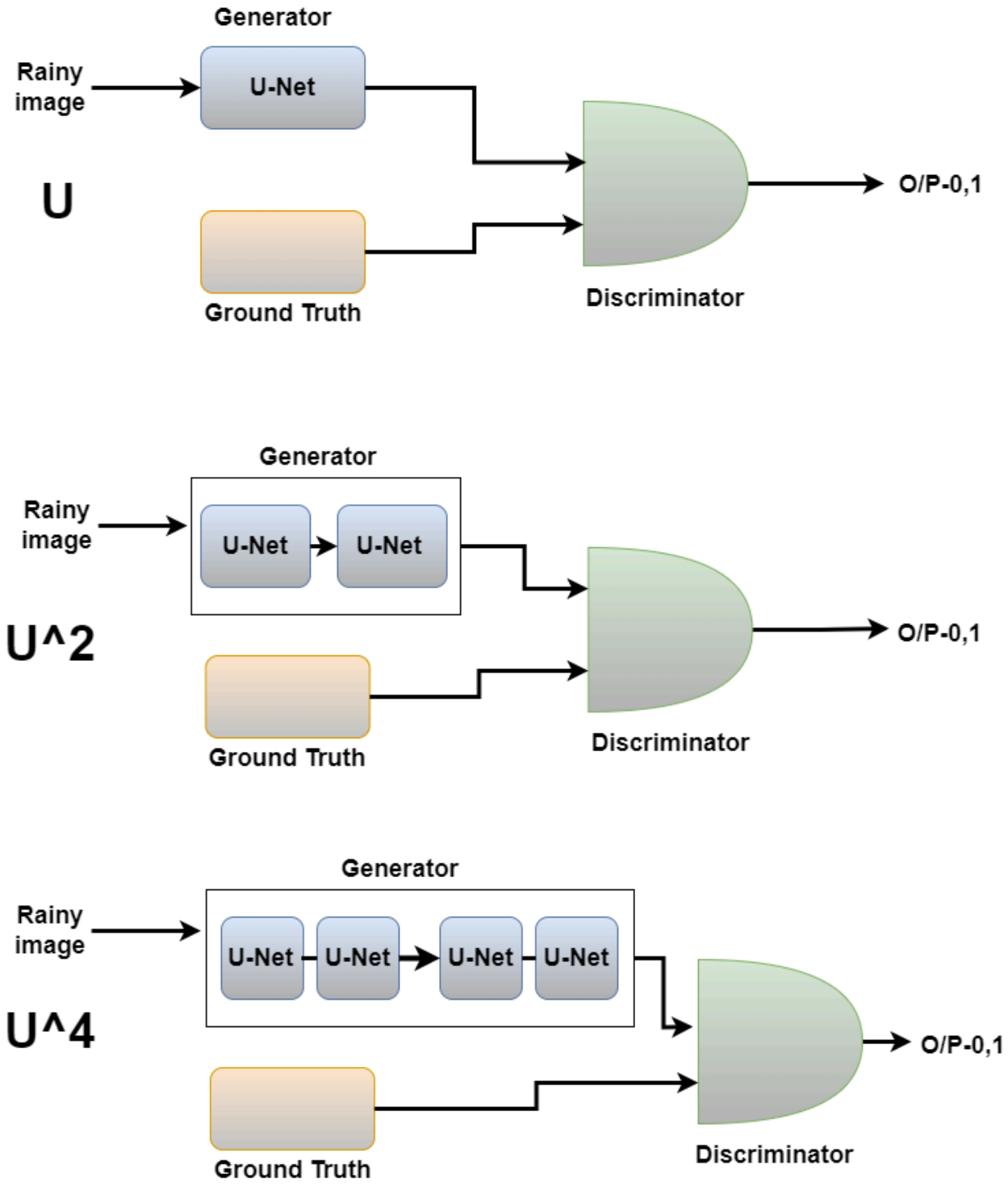


Figure 4.3: Illustrating the detailed description of the Progressive GAN architecture.

4.3 Discriminator

The discriminator plays a vital role in the proposed GAN framework designed for de-raining applications. Its primary function lies in distinguishing authentic clear images from the synthetic de-rain images generated by the generator. Comprising a set of convolutional layers, the discriminator is responsible for feature extraction from input images and rendering judgments on their genuineness.

The structural design of the discriminator commonly consists of multiple convolutional layers succeeded by activation functions like ReLU (Rectified Linear Unit) or LeakyReLU (Leaky Rectified Linear Unit). These convolutional layers play a crucial role in acquiring hierarchical representations of the input images, encompassing both low-level and high-level characteristics. Each convolutional layer employs a collection of filters to convolve with the input feature maps, extracting distinct patterns and information. The quantity and dimensions of the filters may differ based on the specific discriminator architecture. During the transition through these convolutional layers, the feature maps progress towards a more abstract form, representing progressively intricate features associated with the genuineness of the input images. In order to diminish the spatial dimensions of the feature maps while expanding the receptive field, pooling layers like max pooling or average pooling can be utilized. These layers reduce the size of the feature maps, consolidating the acquired information and empowering the discriminator to capture broader patterns. Following the convolutional layers and downsampling procedures, the discriminator may opt for fully connected layers or global pooling to further combine the extracted features. The purpose of these layers is to grasp high-level representations and succinctly summarize the acquired information from the convolutional layers. The ultimate layer of the discriminator incorporates a sigmoid activation function, which transforms the aggregated features into a probability score. This score signifies the confidence level of the discriminator regarding whether the input image is real or generated. A score nearing 1 signifies a high probability of the input being an authentic clear image, while a score nearing 0 signifies a high probability of it being a generated de-rain image.

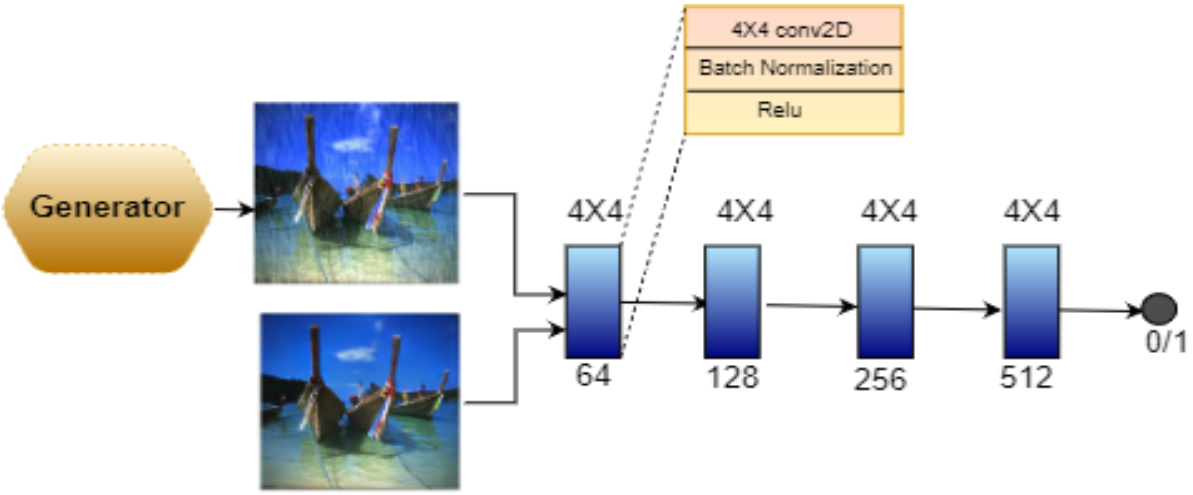


Fig 4.4 : Discriminator Architecture

During the training phase, the discriminator undergoes simultaneous training with the generator in an adversarial fashion. The primary objective of the generator is to produce de-rain images that are able to deceive the discriminator into categorizing them as authentic, whereas the discriminator's goal is to effectively differentiate between genuine clear images and artificially created de-rain images. Through the utilization of backpropagation and gradient descent for optimizing the discriminator's parameters, it gradually enhances its ability to discern between real and synthetic images. Consequently, this iterative process motivates the generator to enhance its de-raining capabilities and generate results that closely resemble real clear images.

The discriminator assumes a pivotal function within the GAN framework, delivering evaluations to the generator and enabling the adversarial training procedure. The capacity to differentiate authentic, unaltered images from synthetic de-rain images is instrumental in enhancing the training regimen and refining the GAN model tailored for de-raining purposes.

4.4 Loss Function

The suggested approach employs several loss functions in order to enhance the training of the networks. The formulation of the total loss for the generator is as follows:

The loss function known as binary cross-entropy with logits (BCEWithLogits) serves to measure the discrepancy between the predicted probability from the discriminator network and the actual ground truth label. This particular loss function is mathematically defined as follows:

$$L_{BCE} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(D(J_i)) + (1 - y_i)\log(1 - D(G(I_i)))]$$

Here, D represents the discriminator network, yi denotes the ground truth label (where 1 signifies clear images and 0 denotes rain images), and N stands for the size of the training dataset.

The loss function utilized to calculate the discriminator loss is exclusively the BCEWithLogits loss function. Its primary objective is to facilitate the discriminator in distinguishing between the derain images produced by the generator and the original ground truth images. The loss can be formulated as follows:

$$L_D = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(D(J_i)) + (1 - y_i)\log(1 - D(G(I_i)))]$$

where D represents the discriminator network, yi denotes the ground truth label (1 for clear images and 0 for rain images), and N stands for the training size.

Chapter 5

Experimental Setup

The experimental result of the proposed method is shown in Figs. () and Tables (). To evaluate the proposed method, extensive tests have been been trained and tested on a Google Collab virtual machine equipped with Python 3.8.1 and ... library and a system configuration of an Intel Core i5 1.60 GHz CPU with 8GB of RAM.

5.1 Hyper-Parameter :

The rain removal model undergoes three training iterations, with each composed of 75 epochs, in order to incrementally acquire knowledge from the dataset. Within every cycle, the model handles 16 samples in grouped batches of 16. The utilization of a learning rate set at 0.0001 is crucial in the training process to control the magnitude of weight updates, thus promoting stable convergence and mitigating the potential for overshooting.

The Adam optimizer is employed to optimize by dynamically adjusting learning rates for individual parameters, amalgamating the advantages of Adagrad and RMSprop. The hyper-parameters play a crucial role in shaping the training dynamics of the rain removal model, affecting its ability to generalize patterns from data and efficiently mitigate rain distortion.

5.2 Datasets

The Rain13k dataset, developed by Quang Trn Minh and accessible on Kaggle, represents a substantial collection of 13,712 image pairs designed to tackle the challenge of eliminating rain interference in the field of computer vision. Every pair comprises an unaltered clear image alongside its corresponding rainy rendition, sourced from a range of datasets that encompass different environments and meteorological scenarios.

This dataset assumes a pivotal role in the advancement and validation of algorithms aimed at mitigating the influence of precipitation on images. Specialized datasets such as Rain100H, Rain100L, Test100, Test2800, and Test1200 offer a diverse array of circumstances for comprehensive assessment. Apart from practical implementations, the Rain13k dataset holds significance for state-of-the-art investigations in image processing and computer vision, particularly in tackling obstacles introduced by inclement weather conditions. Its contribution to enhancing algorithms for rain elimination heightens the perceptual aspect in self-driving technology and enhances surveillance systems under varying weather conditions. Scholars can

utilize the dataset's 13,712 pairs of images for thorough examination and enhancement, thereby acquiring valuable insights into practical hurdles. Essentially, the Rain13k dataset emerges as an indispensable asset propelling scholarly exploration and creativity in the dynamic realm of computer vision.

Here are descriptions of the Test datasets Rain100H, Rain100L, Test100, Test2800, and Test1200:

1. **Rain100H**: This dataset consists of 1800 training image pairs and 100 testing image pairs, each containing a rainy image and its corresponding rain-free ground truth. It features heavy rain streaks in five different directions, making it suitable for evaluating training algorithms designed to handle complex rain patterns.
2. **Rain100L**: The Rain100L dataset includes 1800 training image pairs and 100 testing image pairs with light rain streaks. This dataset focuses on simpler rain patterns, providing a contrast to the more complex Rain100H dataset. It is useful for testing algorithms that may perform better with less dense rain conditions.
3. **Test100**: This dataset contains 100 rainy images and their corresponding ground truth images. It is used for evaluating the performance of deraining algorithms, providing a standard benchmark for comparing different methods under consistent conditions.
4. **Test2800**: Comprising 2800 rainy images paired with their rain-free ground truth, the Test2800 dataset offers a large-scale benchmark for assessing deraining algorithms. The variety of scenes and rain patterns helps in evaluating the generalization capabilities of different models.
5. **Test1200**: This dataset includes 1200 image pairs, featuring a range of rain conditions. It serves as another benchmark for single image deraining, allowing for comprehensive evaluation across different models and techniques.

5.3 Training Details

Rain13K, Rain100H, Rain100L, Test100, Test2800, and Test1200 are the traditional datasets that are commonly used for the purpose of both training and testing. These datasets, specifically Rain100H, Rain100L, Test100, Test2800, and Test1200, function as the testing datasets while Rain13K functions as the training dataset for applications related to image deraining. Together, they consist of a total of 13,712 pairs of clean and rain images gathered from a variety of datasets such as Rain14000, Rain1800, Rain800, and Rain12. These datasets provide a wide range of precipitation scenarios, which include rain streaks and their accumulation.

The assessment of image de-raining models relies on PSNR and SSIM metrics calculated on the Y-channel within the YCbCr color space. The suggested model is comprised of 113,593

parameters for training. Throughout the training process, this model undergoes three training cycles, each encompassing 75 epochs, to progressively learn from the dataset. Within each iteration, the model handles 16 samples in batches of 16. In order to regulate the step size during weight updates, a learning rate of 0.0001 is utilized during training, thereby ensuring consistent convergence and diminishing the likelihood of overshooting.

The model underwent training utilizing the Adam optimization technique featuring an initial learning rate set at 0.001. This approach involves the dynamic adjustment of learning rates based on pertinent parameters, merging attributes of Adagrad and RMSprop. These hyper-parameters play a crucial role in the training of the rain removal model, impacting its ability to effectively reduce rain interference and generalize patterns from the dataset. To provide a comparison, contemporary single image deraining methodologies such as DDN, JORDER, JORDERE, SIRR, PRN, PReNet, GraNet, BRN, and SSDRNet were evaluated in terms of quality and quantity against our proposed model.

5.4 Performance Metrics

In order to assess the effectiveness of the de-raining technique we have introduced, a variety of established metrics were utilized to offer quantitative evaluations of the visual quality. The subsequent metrics for performance evaluation were employed:

1. Peak Signal-to-Noise Ratio (PSNR) serves as a commonly employed metric for evaluating the fidelity of image reconstructions. This metric quantifies the relationship between the maximum achievable signal power, determined by the range of pixel intensities, and the degradation resulting from the deraining procedure. Elevated PSNR values correspond to superior image quality, reflecting diminished levels of distortion.

2. The Structural Similarity Index (SSIM) is a metric used to quantify the level of structural similarity present between two images. This index considers factors such as luminance, contrast, and overall image structure. Ranging from 0 to 1, SSIM values of 1 signify a complete match in structural likeness. Enhanced SSIM values are indicative of superior image quality, as they indicate a stronger similarity to the original clear images.

By assessing our deraining outcomes through the utilization of PSNR and SSIM metrics, our aim was to quantitatively gauge the efficacy of our proposed approach in enhancing image quality and visibility. These metrics afford an objective evaluation of image fidelity and perceptual resemblance, facilitating a comparison of diverse deraining methodologies' efficiency and an examination of the influence of different architectural decisions and training approaches

on the derain images' quality. It is important to highlight that despite the common use of PSNR and SSIM metrics for image quality evaluation, they do possess certain constraints. PSNR concentrates on pixel-level distinctions and may not completely encapsulate an image's perceptual quality. On the other hand, SSIM, while being more inclined towards perceptual aspects, might not consistently correlate with human perception. Hence, complementary to these metrics, we also undertook visual inspections and qualitative evaluations to acquire a comprehensive comprehension of our deraining technique's performance.

Chapter 6

Evaluation

6.1 Qualitative Evaluation:

A thorough qualitative assessment of the proposed model is carried out through the scrutiny of significant metrics including PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index). The aim of this examination is to offer a holistic understanding of the model's efficacy while maintaining authenticity and rigor during the evaluation process.

6.1.1 PSNR :

The Peak Signal-to-Noise Ratio (PSNR) serves as a significant measure in the assessment of image quality, particularly within the realm of digital image and video compression. This parameter delineates the relationship between the highest signal power and the perturbing impact induced by noise, expressed in a logarithmic manner in decibels. PSNR stands as a conventional instrument for assessing the extent of quality disparity between an initial image and its compressed iteration.

In practical application, a greater PSNR value signifies enhanced quality within the compressed image. It is commonly employed alongside Mean Squared Error (MSE) for the evaluation of the comprehensive quality of image compression. MSE calculates the total squared error between the compressed and initial images, with reduced values indicating a lesser degree of error.

In the realm of image compression, the attainment of a higher Peak Signal-to-Noise Ratio (PSNR) and a lower Mean Squared Error (MSE) concurrently signifies enhanced quality within the compressed image in relation to the original one, thereby establishing a dependable yardstick for assessing the efficacy of compression algorithms.

The Peak Signal-to-Noise Ratio (PSNR) is a widely used metric to evaluate the quality of an image restoration process, specifically for assessing the effectiveness of rain removal. PSNR compares the original rain-free image $O(x)$ with the restored image $\hat{O}(x)$ after rain removal. It is defined as:

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

Where:

L is the maximum possible pixel value of the image. For 8-bit images,L=255

MSE (Mean Squared Error) is calculated as:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} \left[O(x_i) - \hat{O}(x_i) \right]^2$$

Here,

N is the total number of pixels in the image, $O(x_i)$ is the pixel value of the original image at position i, and $\hat{O}(x_i)$ is the pixel value of the restored image at position i.

6.1.2 SSIM :

The Structural Similarity Index (SSIM) is a significant metric driven by perception that quantifies visual similarity. It assesses the likeness between two images by taking into account crucial elements like brightness, contrast, and structural details. SSIM scrutinizes every pixel in a picture within an 11-by-11 vicinity and calculates a similarity index based on structure. The result is an image of identical dimensions to the one under assessment. SSIM is predominantly employed for the appraisal of deterioration in image quality resulting from processes such as data compression or transmission losses. It furnishes a numerical assessment. Typically, compression outcomes vary from -1 to 1, with superior values denoting higher quality. SSIM has gained acknowledgment for its straightforwardness, effectiveness, and reliability in comparison to alternative methodologies.

Nonetheless, it is imperative to acknowledge that SSIM could face challenges in assessing images that are heavily distorted or cross-distorted. Notwithstanding its advantages, there are situations in which SSIM may not provide precise evaluations when confronted with substantial modifications in the images.

The Structural Similarity Index (SSIM) measures the similarity between two images by comparing their luminance, contrast, and structure. SSIM is particularly useful for assessing the perceptual quality of the restored image. It is defined as:

$$\text{SSIM}(O, \hat{O}) = \frac{(2\mu_O\mu_{\hat{O}}+C_1)(2\sigma_{O\hat{O}}+C_2)}{(\mu_O^2+\mu_{\hat{O}}^2+C_1)(\sigma_O^2+\sigma_{\hat{O}}^2+C_2)}$$

Where:

- μ_O and $\mu_{\hat{O}}$ are the mean pixel values of the original and restored images, respectively.
- σ_O^2 and $\sigma_{\hat{O}}^2$ are the variances of the original and restored images, respectively.
- $\sigma_{O\hat{O}}$ is the covariance between the original and restored images.
- $C1$ and $C2$ are constants to stabilize the division with weak denominators, typically defined as

$$C_1 = (K_1L)^2 \text{ and } C_2 = (K_2L)^2, \text{with } K_1 \ll 1 \text{ and } K_2 \ll 1.$$

The metrics PSNR and SSIM play a vital role in quantitatively assessing the efficacy of rain removal algorithms. PSNR quantifies the fidelity of the recovered image to the source, whereas SSIM evaluates perceived quality through consideration of human visual perception aspects.

This figure(**6.1.a**, **6.1.b**, **6.1.c**, **6.1.d**, **6.1.e**) illustrates the output results of a single UNet, a pair of UNets, and a double-pair UNet on various datasets(**Rain100H, Rain100L, Test100, Test2800, and Test1200**). The conclusion drawn from the results indicates that the double-pair UNet model outperforms both the single UNet and the pair of UNets across different datasets.



Fig 6.1.a : Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Test-2800 dataset.



Fig 6.1.b : Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Test-1200 dataset.

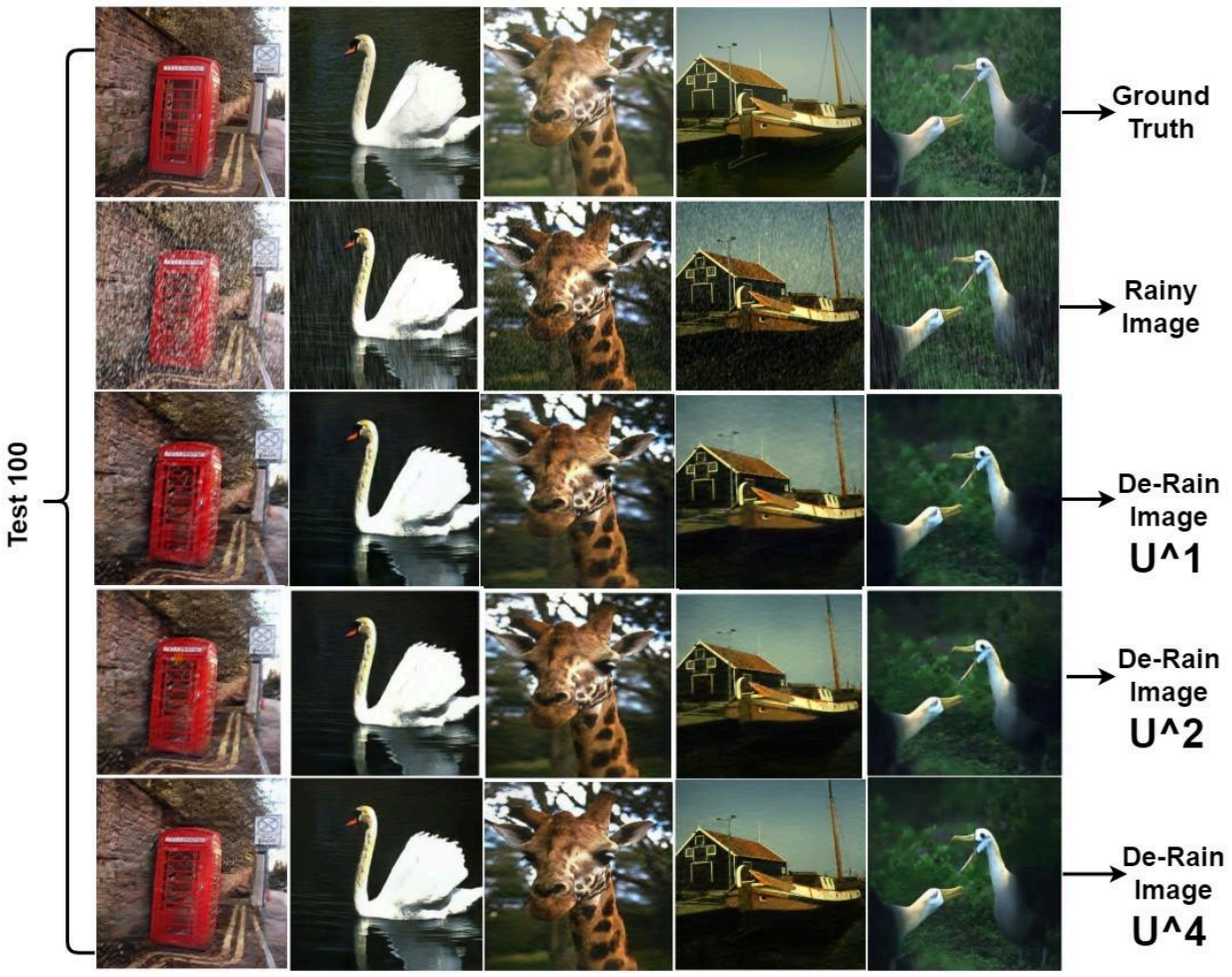


Fig 6.1.c : Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Test-100 dataset.



Fig 6.1.d : Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Rain 100L dataset.

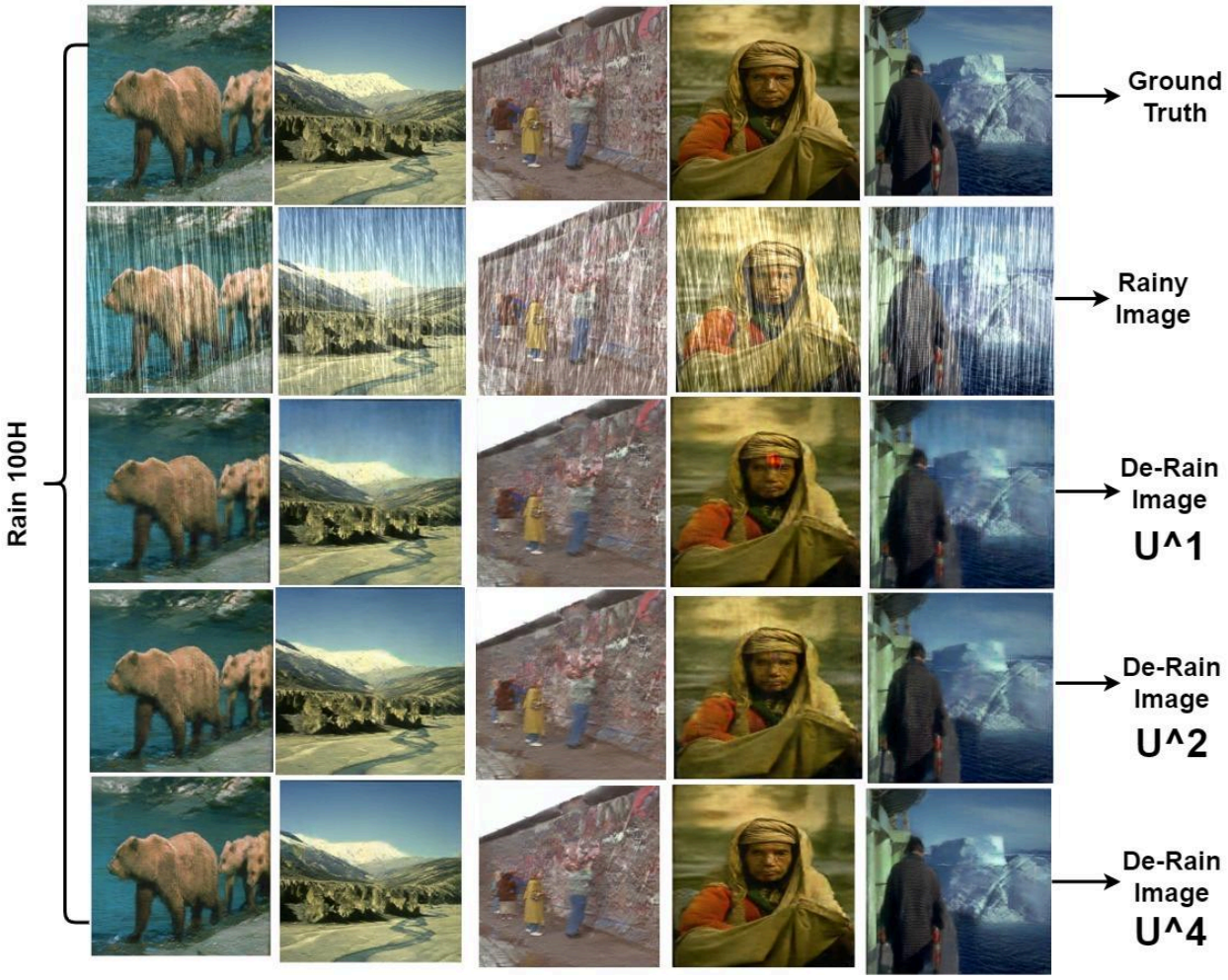


Fig 6.1.e : Detailed step-by-step (U^1,U^2,U^4) output comparison of the proposed De-rain model on the Rain 100 H dataset.

6.2 Quantitative comparison

6.2.1 State-of-the-art comparison :

Our study has conducted a comparison between the suggested approach and alternative single-image de-raining methods, for instance, (model name). Two evaluation metrics, namely PSNR and SSIM, were employed to assess the tracking algorithms listed in Table 2. The performance of these algorithms was evaluated across different datasets such as Rain100H, Rain100L, Test100, Test2800, and Test1200.

In the case of De-rain, the training set Rain30k is first run through a U-Net(U^1), and the results of different test datasets (Rain100H, Rain100L, Test100, Test2800, and Test1200) are shown figure (6.2.a)

In the second case, using two U-Net or a pair of U-Net(U^2) as generators, the results of various datasets(Rain100H, Rain100L, Test100, Test2800, and Test1200)are shown in figure(6.2.b).

In the third case, using 4 U-Net or 2 pair of U-Net as generators, the results of various datasets(Rain100H, Rain100L, Test100, Test2800, and Test1200) are shown in Fig(6.2.c).

Finally, if we summarize the three figures(fig-6.2.a, fig-6.2.b, fig-6.2.c), we can see that as the number of pairs of U-Net or U-Net increases, the results of D-Ren are better as expected for the different data sets used. And we show the difference between the results of our model and the results of models used by others in Figure(6.2.d)

No	U-Net	Data set	SSIM	PSNR
1	U^1	Test 2800	0.840401480389	31.8086
2	U^1	Test 1200	0.846831838190	32.5436
3	U^1	Test100	0.773436871718	27.5020
4	U^1	Rain100L	0.824970738291	30.9530
5	U^1	Rain100H	0.733482078909	28.8981

Table 6.2.a : Detailed comparison of various test datasets using a single U-Net model.

No	U-Net	Data set	SSIM	PSNR
1	U^2	Test 2800	0.87186974855405	32.8882
2	U^2	Test 1200	0.86771239101886	33.3415
3	U^2	Test100	0.80675518877652	28.2848
4	U^2	Rain100L	0.85854910612106	32.0567
5	U^2	Rain100H	0.78310902714729	30.4688

Table 6.2.b : Detailed comparison of various test datasets using one pair of U-Net mode

No	U-Net	Data set	SSIM	PSNR
1	U^4	Test 2800	0.87603460039411	33.0000
2	U^4	Test 1200	0.86666548624634	33.2074
3	U^4	Test100	0.81818560799773	28.5175
4	U^4	Rain100L	0.87042819678783	32.4100
5	U^4	Rain100H	0.79204773962497	30.7038

Table 6.2.c : Detailed comparison of various test datasets using two pair of U-Net model

Methods	Test 100 PSNR	Test 100 SSIM	Test 1200 PSNR	Test 1200 SSIM	Test 2800 PSNR	Test 2800 SSIM	Rain 100H PSNR	Rain 100H SSIM	Rain 100L PSNR	Rain 100L SSIM	Average PSNR	Average SSIM
DerainNet	22.77	0.81	23.38	0.835	24.31	0.861	14.92	0.592	27.03	0.884	22.482	0.796
SEMI	22.35	0.788	26.05	0.822	24.43	0.782	16.56	0.486	25.03	0.842	22.884	0.744
DIDMDN	22.5	0.818	29.29	0.901	28.13	0.867	17.35	0.524	25.23	0.741	24.584	0.77
UMRL	24.41	0.829	30.55	0.91	29.97	0.905	26.01	0.832	29.18	0.923	28.024	0.879
RESCAN	25	0.835	30.51	0.882	31.29	0.904	26.36	0.786	29.8	0.881	28.592	0.857
PreNet	24.81	0.851	31.36	0.911	31.75	0.916	26.77	858	32.4	0.95	29.418	0.897
MSPFN	27.5	0.876	32.39	0.916	32.82	0.930	28.66	0.86	32.4	0.933	30.754	0.903
MDARNet	—-	—-	33.08	0.919	—-	—-	29.7	—-	—-	—-	31.39	0.919
Proposed Model	28.51	0.818	33.2	0.866	33	0.876	30.7	0.792	32.4	0.87	31.562	0.844

Table 6.2.d : Comprehensive Assessment of the Proposed Best Model in Comparison to Existing Models Across a Variety of Datasets..

Chapter 7

Conclusion

In conclusion, this thesis introduces a novel methodology for enhancing image clarity and visibility in rainy conditions, utilizing a generative adversarial network (GAN)-based deraining framework. By leveraging the capabilities of GANs, the framework gains an understanding of the relationship between rain-affected images and their rain-free counterparts, ultimately leading to improved deraining results.

The implementation of generator structures, such as U-Net, has been proven effective in capturing both spatial and contextual details, which are essential for accurate deraining. An analysis is conducted to investigate the impact of these structures on deraining performance, with the aim of identifying the most suitable one for the proposed model. Additionally, a sequence of U-Net architectures following an exponential progression of 2^n is explored, further enhancing the deraining capabilities.

The training and evaluation processes employ the Rain13K dataset, which offers a diverse range of rain scenarios to ensure comprehensive training and model generalization. The generator is guided to produce visually accurate and realistic derain outputs through a combination of loss functions, including Mean Squared Error (MSE), perceptual loss, and adversarial loss.

Performance evaluation is carried out on benchmark datasets such as Rain100H, Rain100L, Test100, Test1200, and Test2800, allowing for the assessment of various rain scenarios. Quantitative image quality evaluations are performed using objective measures like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), enabling a thorough comparison with state-of-the-art deraining methods.

The proposed deraining framework demonstrates promising results in improving image quality and visibility under rainy conditions. The effectiveness of this framework, as validated by rigorous experimentation and evaluation, suggests its potential for real-world applications in the fields of computer vision and image processing.

7.1 Future Scope

The potential scope of this study on precipitation removal using Generative Adversarial Network (GAN)-based models is vast and encompasses various crucial domains for improvement and exploration.

Firstly, it is essential to expand the diversity of the dataset. Although the Reside dataset was utilized in the current study, obtaining a more extensive and diverse dataset with authentic rain images from real-world scenarios would strengthen the credibility and applicability of the research findings. Collaborations with academic institutions, industry organizations, or the scholarly community can facilitate data exchange and acquisition, providing access to a broader range of rain environments and scenarios for training and evaluation.

Furthermore, the optimization of loss function selection is pivotal in model training. The present research employed MS-SSIM, LD loss, and BCEWithLogits loss; however, future investigations could explore alternative or domain-specific loss functions. Perceptual loss measures, adversarial losses, or content-specific constraints could be considered to improve training procedures and enhance the visual quality of derain images.

Additionally, it is important to address the limitation of computational resources during model training. Strategies to improve the model's efficiency, such as simplifying its complexity or using methods like pruning or quantization for model compression, are key to overcoming this constraint.

Besides technical factors, the establishment of standardized assessment metrics and benchmark datasets is vital for ensuring fair comparisons and advancements in the field. The development of comprehensive evaluation protocols that consider perceptual fidelity, rain removal effectiveness, and resilience to varying rain scenarios will enhance the ability to objectively compare different de-raining techniques.

Lastly, investigating application-specific adaptations of deraining models is a significant avenue for future research. Customizing the de-raining algorithms to address distinct challenges and requirements in areas such as autonomous driving, surveillance systems, or underwater imaging could lead to substantial progress. Modifying the models to tackle issues like motion blur, depth estimation, or scene-specific attributes can increase their effectiveness and relevance in practical applications.

By exploring these potential areas, researchers have the opportunity to make remarkable advancements in the robustness, applicability, and efficiency of de-raining techniques. The continued investigation and improvement within these domains are set to result in enhanced visibility and better image quality under rainy conditions in various fields and contexts.

List Of Publication

PATENTS:

1. **S. Sk**, A. Ali, R. Das, P. M. Sk, and S. Chaudhuri, “Iot-powered water conditions monitoring system for fish farming,” Application No. : 202431007012, Journal No. : 08/2024, Feb. 2024.
2. A. Ali, **S. Sk**, B. Das, et al., “A real-time sign language translator device,” Application No. : 202431007372, Journal No. : 08/2024, Feb. 2024.

COPYRIGHT :

1. A. Ali, H. Shaw, **S. Sk**, B. Das, and S. Chaudhuri, Rainsegtrack dataset: Video object segmentation and tracking in rainy weather, Jan. 2024- 350/2024-CO/L
2. A. Ali, D. Hossain, **S. Sk**, and S. Chaudhuri, Video-rain99 dataset, Nov. 2023-8801/2023-CO/L
3. A. Ali, M. Sk, D. Pal, **S. Sk**, and S. Chaudhuri, Iurwdd: Indian urban and rural waste detection dataset and benchmark, May 2023.-L-121946/2023

Conference :

1. A. Ali, B. Das, **S. Sk**, S. Ganguli, and S. S. Chaudhuri, “Arabic sign language recognition and voice translation device using embedded machine learning,” in 2nd Doctoral Symposium on Human Centered Computing (HUMAN 2024), 2021.

Journal :

1. **S. Sk,** A. Ali, B. Das, S. Ganguli, and S. S. Chaudhuri, “Development of Progressive GAN based Image Rain Removal Technique”

Bibliography

[1] Wang, Cong, Xiaoying Xing, Yutong Wu, Zhixun Su, and Junyang Chen. "Dcsfn: Deep cross-scale fusion network for single image rain removal." In *Proceedings of the 28th ACM international conference on multimedia*, pp. 1643-1651. 2020.Wang, Cong, Xiaoying Xing, Yutong Wu, Zhixun Su, and Junyang Chen. "Dcsfn: Deep cross-scale fusion network for single image rain removal." In *Proceedings of the 28th ACM international conference on multimedia*, pp. 1643-1651. 2020.

[2] Zheng, Chaobing, Jun Jiang, Wenjian Ying, and Shiqian Wu. "Single Image Deraining via Feature-based Deep Convolutional Neural Network." In *2023 IEEE 18th Conference on Industrial Electronics and Applications (ICIEA)*, pp. 84-89. IEEE, 2023.

[3] Wang, Cong, Jinshan Pan, and Xiao-Ming Wu. "Online-updated high-order collaborative networks for single image deraining." In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 2, pp. 2406-2413. 2022.

[4] J. Pu, X. Chen, L. Zhang, Q. Zhou and Y. Zhao, "Removing rain based on a cycle generative adversarial network," 2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA), Wuhan, China, 2018, pp. 621-626, doi: 10.1109/ICIEA.2018.8397790. keywords: {Rain;Generators;Image color analysis;Training;Loss measurement;Visualization;TV;General adversarial network;single image de-raining;perceptual loss;dark},

[5] Guo, Z., Hou, M., Sima, M. *et al.* DerainAttentionGAN: unsupervised single-image deraining using attention-guided generative adversarial networks. *SIViP* **16**, 185–192 (2022). https://doi.org/10.1007/s11760-021-01972-9

[6] Fei Yang, Jianfeng Ren, Zheng Lu, Jialu Zhang, Qian Zhang, Rain-component-aware capsule-GAN for single image de-raining, Pattern Recognition,Volume 123,2022,108377,ISSN 0031-3203,https://doi.org/10.1016/j.patcog.2021.108377.
(https://www.sciencedirect.com/science/article/pii/S0031320321005574)

[7] Xin Jin, Zhibo Chen, Weiping Li, AI-GAN: Asynchronous interactive generative adversarial network for single image rain removal, Pattern Recognition, Volume 100, 2020, 107143, ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2019.107143.
(https://www.sciencedirect.com/science/article/pii/S0031320319304443)

[8] Chen, M., Wang, P., Shang, D. *et al.* Cycle-attention-derain: unsupervised rain removal with CycleGAN. *Vis Compute* **39**, 3727–3739 (2023). https://doi.org/10.1007/s00371-023-02947-2

[9] Zhu, H., Peng, X., Zhou, J. T., Yang, S., Chanderasekh, V., Li, L., & Lim, J.-H. (2019). Singe Image Rain Removal with Unpaired Information: A Differentiable Programming Perspective.

Proceedings of the AAAI Conference on Artificial Intelligence, *33*(01), 9332-9339. https://doi.org/10.1609/aaai.v33i01.33019332

[10] Wei B, Wang D, Wang Z, Zhang L. PRAGAN: Progressive Recurrent Attention GAN with Pretrained ViT Discriminator for Single-Image Deraining. Sensors (Basel). 2022 Dec 7;22(24):9587. doi: 10.3390/s22249587. PMID: 36559956; PMCID: PMC9785271.

[11] Koch, C, and S. Ullman. "Shifts in selective visual attention: towards the underlying neural circuitry. " Human Neurobiology 4.4(1985):219.

[12] Feng, Yuan, Yaojun Hu, Pengfei Fang, Sheng Liu, Yanhong Yang, and Shengyong Chen. "Asymmetric dual-decoder U-Net for joint rain and haze removal." *ACM Transactions on Multimedia Computing, Communications and Applications* 20, no. 3 (2023): 1-23.

[13] B. Hu, Y. Sun, Y. Yang, Z. Ouyang, and F. Zhang "Towards Robust Rain Removal with UNet++," *Comput. Mater. Contin.*, vol. 75, no. 1, pp. 879-890. 2023. https://doi.org/10.32604/cmc.2023.035858

[14] Engin, D., Genç, A., Kemal Ekenel, H.: Cycle-dehaze: Enhanced cyclegan for single image dehazing. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops,pp. 825–833 (2018)

[15] Zheng, Z., Chen, Z., Wang, S. *et al.* Dual-attention U-Net and multi-convolution network for single-image rain removal. *Vis Comput* (2023). https://doi.org/10.1007/s00371-023-03198-x

[16] P. C. Barnum, S. Narasimhan, and T. Kanade. Analysis of rain and snow in frequency space. International Journal of Computer Vision, 86(2-3):256–274, 2010.

[17] Yunhao Ba, Howard Zhang, Ethan Yang, Akira Suzuki, Arnold Pfahnl, Chethan Chinder Chandrappa, Celso M. de Melo, Suya You, Stefano Soatto, Alex Wong, and Achuta Kadambi. 2022. Not just streaks: Towards ground truth for single image deraining. In Proceedings of the 17th European Conference on Computer Vision (ECCV '22), Tel Aviv, Israel, Part VII. Springer-Verlag, Berlin, 723–740.

[18] W. Yang, R. T. Tan, J. Feng, Z. Guo and S. Yan, "Joint rain detection and removal from a single image with contextualized deep networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 6, pp. 1377–1393, 2019.

[19] V. Santhaseelan and V. K. Asari. Utilizing local phase information to remove rain from video. International Journal of Computer Vision, 112(1):71–89, 2015.

[20] X. Li , J. Wu , Z. Lin , H. Liu , H. Zha , Recurrent squeeze-and-excitation context aggregation net for single image deraining, European Conference on Computer Vision (2018) 262–277

[21]. Wu, Z., Li, S., Chen, C., Hao, A., Qin, H.: Recursive multi-model complementary deep fusion for robust salient object detection via parallel sub-networks. Pattern Recogn. 121, 108212 (2022)

[22]. Zhang, Z., Zhang, Y., Cheng, X., Lu, G.: Siamese network for object tracking with multi-granularity appearance representations.Pattern Recogn. 118, 108003 (2021)

[23]. Zheng C, Jiang J, Ying W, Wu S. Single Image Deraining via Feature-based Deep Convolutional Neural Network. In2023 IEEE 18th Conference on Industrial Electronics and Applications (ICIEA) 2023 Aug 18 (pp. 84-89). IEEE.

[24]. Zhao, Penghui, Haowen Zheng, Suigu Tang, Zongren Chen, and Yangyan Liang. 2023. "DAMNet: A Dual Adjacent Indexing and Multi-Deraining Network for Real-Time Image Deraining" *Fractal and Fractional* 7, no. 1: 24. https://doi.org/10.3390/fractalfract7010024

[25]. Y. Luo, Y. Xu and H. Ji, "Removing Rain from a Single Image via Discriminative Sparse Coding," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 3397-3405, doi: 10.1109/ICCV.2015.388. keywords: {Rain;Dictionaries;Visual effects;Image coding;Visualization;Machine vision},

[26] Chang, Y.; Yan, L.; and Zhong, S. 2017. Transformed low rank model for line pattern noise removal. In ICCV.

[27] @article{Kong2023ProgressiveDD, title={Progressive dilation dense residual fusion network for single-image deraining}, author={Xiaolin Kong and Tao Gao and Ting Chen and Jing Zhang}, journal={IET Image Process.}, year={2023}, volume={17}, pages={4102-4115}, url={https://api.semanticscholar.org/CorpusID:262203067}}

[28] J. Xiao , W. Zou , Y. Chen , W. Wang , J. Lei , Single image rain removal based on depth of field and sparse coding, Pattern Recognit. Lett. 116 (2018) 212–217

[29].Sahil, Yadav., Aryan, Mehra., Honnesh, Rohmetra., Rahul, Ratnakumar., Pratik, Narang. (2021). DerainGAN: Single image deraining using wasserstein GAN. Multimedia Tools and Applications, doi: 10.1007/S11042-021-11442-6

[30] Guo, Qing, Jingyang Sun, Felix Juefei-Xu, Lei Ma, Di Lin, Wei Feng, and Song Wang. "Uncertainty-aware cascaded dilation filtering for high-efficiency deraining." arXiv preprint arXiv:2201.02366 (2022).

[31] (2022). A comprehensive analysis on single image-based deraining using generative adversarial network. Global Journal of Engineering and Technology Advances, doi: 10.30574/gjeta.2022.11.3.0072

[32]Jiuyong, Li., Hao, Feng., Zhennan, Deng., Xintong, Cui., Hongxia, Deng., Haifang, Li. (2022). Image Derain Method for Generative Adversarial Network Based on Wavelet High Frequency Feature Fusion. Lecture Notes in Computer Science, doi: 10.1007/978-3-031-18916-6_14

[33]Yongqian, Jay, Cheng., Ze, Yang., Wenjun, Zhang., Ling, Yang., Jun, Wang., Tingzhao, Guan. (2023). A Single Image Derain Method Based on Residue Channel Decomposition in Edge Computing. Intelligent Automation and Soft Computing, doi: 10.32604/iasc.2023.038251

[34] (2023). Data-Driven Single Image Deraining: A Comprehensive Review and New Perspectives. doi: 10.36227/techrxiv.14944752

[35] (2023). Data-Driven Single Image Deraining: A Comprehensive Review and New Perspectives. doi: 10.36227/tech arxiv.14944752.v4

[36]Yohan, Poirier-Ginter., Jean-Francois, Lalonde. (2023). Robust Unsupervised StyleGAN Image Restoration. arXiv.org, doi: 10.48550/arXiv.2302.06733.