FUSION BASED GENERATIVE ADVERSARIAL NETWORK FOR SINGLE IMAGE DEHAZING

A DISSERTATION

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

MASTER OF TECHNOLOGY
IN
ARTIFICIAL INTELLIGENCE

Submitted by

DEEPANKER RAWAT

2K21/AFI/08

Under the supervision of

MR KAVINDER SINGH

COMPUTER SCIENCE & ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

MAY, 2023

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, **DEEPANKER RAWAT**, Roll No's – **2K21/AFI/08** students of M.Tech in Artificial Intelligence (Department of Computer Science & Engineering), hereby declare that the project Dissertation titled "**FUSION BASED GENERATIVE ADVERSARIAL NETWORK FOR SINGLE IMAGE DEHAZING**" which is submitted by me to the Department of Computer Science & Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology. This work has never previously served as the foundation for the award of a degree, associateship, diploma, fellowship, or other equivalent title or recognition.

Place: Delhi — Deepanker Rawat

Date: — 2K21/AFI/08

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

CERTIFICATE

I hereby confirm that the Project Dissertation titled **"FUSION BASED GENERATIVE ADVERSARIAL NETWORK FOR SINGLE IMAGE DEHAZING"** submitted by **DEEPANKER RAWAT**, Roll No's – **2K21/AFI/08**, Department of Computer Science & Engineering, Delhi Technological University, Delhi, is a testament to the project undertaken by the student under my guidance. This submission partially fulfils the requirements for the Master of Technology degree. To the best of my knowledge, this work has not been previously submitted, either in part or in its entirety, for any other degree or diploma at this university or elsewhere.

Place: Delhi

Date:

MR KAVINDER SINGH

SUPERVISOR

Assistant Professor

Department of Computer Science & Engineering

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I would like to extend my heartfelt appreciation to **Mr Kavinder Singh** for his unwavering guidance and mentorship throughout the project. His invaluable assistance paved the way for me to accomplish my objectives by thoroughly explaining the necessary tasks and emphasizing the significance of this project in an industrial context. He consistently offered his support and patiently addressed any challenges I encountered. I am truly grateful for his continuous encouragement and motivation, as this project would not have achieved success without him.

Place: Delhi — Deepanker Rawat

Date: — 2K21/AFI/08

Abstract

Hazy environments significantly reduce the quality of digital images since haze varies with the scene depth. Consequently, the removal of haze from original images assumes great importance. Deep learning approaches have greatly improved the efficiency of the dehazing process for images. This work aims to provide a detailed analysis of various deep learning-based approaches employed for image dehazing. The techniques employed are examined in detail, encompassing their respective methodologies and the utilization of specific loss functions. Furthermore, this work also delves into an exploration of diverse datasets that are available and commonly used for image dehazing tasks. By expanding our understanding of the methodologies, loss functions, and datasets associated with dehazing, we can advance the field and facilitate the development of more effective dehazing algorithms.

In this work, we proposed a model known as FAG-Net for image dehazing. We also proposed a new generator architecture which consists of a dense block, a transition block and a new feature attention (FA) block at each layer which enhance the realistic nature of the haze-free image. FA block consists of one channel attention (CA) and one pixel attention (PA) block. Channel attention aims to enhance the relevant information in different color channels of the input image, while pixel attention aims to selectively emphasize or suppress certain pixels in the image, which helps the models to concentrate on the most affected areas of an image by haze. Perceptual loss and reconstruction loss are used along with adversarial loss to give more attention to the pixel which contains more haze and to maintain the realistic nature of generated haze-free image. Our FAG-Net trained on RESIDE(ITS), O-HAZE and I-HAZE datasets to conduct the experimental analysis. Extensive experimental study demonstrates that our FAG-Net better than previous state-of-the-art methods.

Contents

List of Tables

List of Figures

List of Symbols

I	Intensity of Image
J	Scene Radiance
A	Transmission Map
β	Scattering Coefficient
$\sum$	Summation
l	Lightning of an image
c	Contrast
s	Structure of an image
G_1	Global sub-generator
G_2	Local sub-generator
D_1, D_2	Disciminator
L_{total}	Total loss
L_A	Adversarial loss
λ	Trade-off weights
L_{FM}	Feature matching loss
L_{vgg}	Perceptual loss
L_F	Fidelity loss
L^j	Joint-Discriminator loss
$L_{perceptual}$	Perceptual loss
$L_{cyclegan}$	Cycle-gan loss
L_{trans}	Translation loss
L_{rec}	Reconstruction loss
L_{idt}	Identity loss
ϕ	$VGG16$ feature extractor from 2nd and 5th pooling layer
$\otimes$	Element-wise multiplication
σ	Sigmoid
δ	ReLu
γ	Constant for cyclic perceptual-consistency loss

List of Abbreviations

DL	Deep Learning
FA	Feature Attention
DCPDN	Densely Connected Pyramid Dehazing Network
CA	Channel Attention
GAN	Generative Adversarial Network
FAG-Net	Feature Attention Generative Adversarial Network
ITS	Indoor Training Set
SSIM	Structural Similarity Index
PSNR	Peak Signal-To-Noise Ratio
ReLU	Rectified Linear Unit
DCP	Dark Channel Prior
CNN	Convolutional Neural Network
CAP	Color Attenuation Prior
PA	Pixel Attention

Chapter 1

INTRODUCTION

Dehazing is quite a significantly challenging thing to do in computer vision. The quality of a digital image degraded in a hazy environment. Whenever, we capture an image using a digital device like a smartphone, digital camera, or any other. Haze occurs due to the availability of tiny particles, mist, smog, fog, and moisture in the surrounding that can depreciate the quality of the image. There are numerous uses for dehazing, including, outdoor photography, object detection [33], outdoor surveillance [34], video data compression [35] etc.

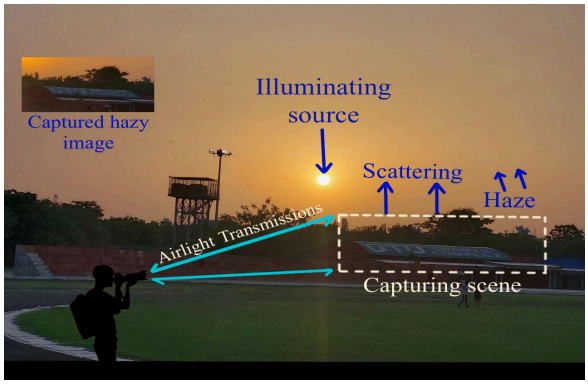

Figure 1.1: Image captured in the presence of illuminating source

Hazy environments pose a significant challenge to digital image quality, as the presence of haze varies across different depths within a scene. Consequently, the removal of haze from original images becomes crucial to restore their visual clarity and enhance their overall appeal. Deep learning approaches have recently revolutionised the field of image dehazing by making it possible for more effective and efficient solutions.

Deep learning-based approaches have shown remarkable promise when tackling the

complex task of dehazing. These techniques leverage the power of neural networks to learn intricate patterns and relationships within hazy images, allowing them to effectively remove unwanted haze and reveal hidden details. By training on huge datasets of hazy images and corresponding dehazed images, deep learning models can capture the underlying characteristics of haze and learn to generate visually pleasing dehazed outputs.

One of the key advantages of deep learning-based dehazing techniques is their ability to handle a wide range of haze conditions and scene complexities. Whether it's a light haze that mildly obscures the image or a dense fog that significantly degrades visibility, these models can adapt and produce satisfying results. This adaptability is particularly useful in real-world scenarios where atmospheric conditions can vary greatly.

Network architectures and loss functions must be carefully designed in order for deep learning-based dehazing techniques to be successful. Researchers have proposed various architectures, such as single-image-based models and multi-scale models, each with their unique strengths and limitations. Additionally, the choice of loss functions plays a crucial role in guiding the training process and ensuring the preservation of important image features during dehazing.

To facilitate the evaluation and benchmarking of dehazing algorithms, several datasets specifically designed for image dehazing have been made available. These datasets consist of pairs of hazy images and their corresponding ground truth haze-free images, allowing researchers to assess the performance of their dehazing models quantitatively. Some datasets also include images captured in different weather conditions and diverse scenes, enabling the evaluation of algorithm robustness.

As the field of image dehazing continues to evolve, researchers are continuously exploring novel techniques and methodologies to further improve the quality and efficiency of dehazing algorithms. By leveraging the power of deep learning and utilizing comprehensive datasets, we can expect significant advancements in dehazing technology, leading to visually stunning images with enhanced clarity and detail, even in the challenging atmospheric conditions.

The atmospheric scattering model [25], defines a relationship between the captured

scene, atmospheric light and as:

$$I(y) = J(y)t(y) + A(1 - t(y)) \tag{1}$$

wherein I represent intensity of an image at y, J represents scene radiance. T represents transmission map and A represents atmospheric light. When homogeneous atmosphere is there, transmission map of an image can be defined by:

$$t(z) = e^{-\beta \cdot d(z)} \tag{2}$$

here β signifies the scattering coefficient. The value β is constant when there is a homogeneous haze. $d(z)$ indicates the scene depth of the captured image with haze.

Generally, atmospheric light along with the transmission map is unknown. Many attempts are made to calculate the transmission map. Some successful attempts which researchers came up with are Color attenuation prior [8], Non-local prior [1], Haze line prior [2], dark channel prior [6] and others [37, 38]. However, the prior-based solutions show inconsistent results in real-world situations due to the restriction of the priors. The priors sometimes lead to incorrect values for atmospheric light or transmission maps. Some approaches [1, 2, 6, 8, 15] fail to account for similarities to actual lightning, which leads to poor dehazing performance for bright items.

Using deep learning techniques [3–5, 10, 11], the evaluation of transmission map and atmospheric light becomes much easier. Some of the Deep learning techniques like cycle-dehaze [5], and AECR-Net [16] didn't even need the estimations map to yield results. We begin by discussing the fundamentals of haze and its various forms before performing a comparative examination of various deep learning-based dehazing methods. The history of the haze is then provided. Finally, we discuss the related work with respect to dehazing like datasets, work done in image dehazing till now, performance metrics etc.

Constituent	Mole percent
Oxygen	20.9
Nitrogen	78
CO_2	0.035
Argon	0.934
Methane	0.00017
Hydrogen	0.00053
Neon	0.001818

Table 1.1: The constituents of the dry air [47].

Weather condition	Particle type	Radius (μm)	Concentration (cm^{-3})
Haze	Aerosol	10^{-2} - 1	10^{3} - 10
Rain	Water drop	10^{2} - 10^{4}	10^{-2} - 10^{-5}
Cloud	Water droplet	1 - 10	300 - 10
Fog	Water droplet	1 - 10	100 - 10
Air	Molecule	10^{-4}	10^{19}

Table 1.2: Various atmospheric conditions and the corresponding sizes, types and concentrations of particles [48].

Haze occurs as a result of many particles present in the air like fog, smoke, mist, snow etc as shown in Table 1.1 and Table 1.2. In general, the air is composed of around 78% of NO_2, 20% O_2 gas, 0.04% of CO_2 gas, and small quantities of other gases. Fig. 1.1 depicts the photos taken during adverse weather.

Haze deteriorates the image quality, however, haze is not homogenous in nature. In nature [35] haze can be found in three forms as given below:

(i) **Uniform Haze**: In this type, the haze in an image is uniformly distributed. The hazy images look homogeneous distributed with haze from every single pixel.

(ii) **Plumes**: These are created when air flows steadily for local reasons.

(iii) **Layered Haze**: It occurs when a temperature inversion happens very close to the ground.

1.1 RELATED WORK

This section provides an introduction to different techniques used for single-image dehazing. We also provide information regarding the different image datasets, both real-world and synthetic, available for dehazing.

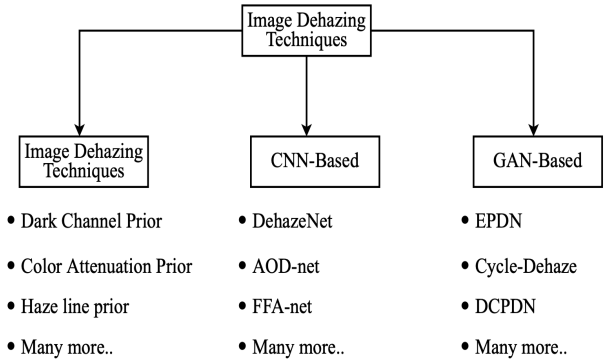

Figure 1.2: Various single Image dehazing techniques

1.1.1 Image Dehazing Techniques

Obtaining a dehazed image from the hazy one poses a significant challenge. Researchers have introduced many methods to tackle this task (Fig 1.2). Deep learning-based methods and statistical-based methods are the two main categories into which image dehazing techniques may be divided. Deep learning-based can further divided into two subcategories: (i) CNN-based and (ii) GANs based.

Statistical based methods

Statistical-based methods use assumptions to estimate statistical models and mitigate the effects of haze in images. These approaches involve analyzing the statistical properties of haze and employing algorithms to calculate parameters such as transmission and ambient light. By utilizing these estimations, the haze can be effectively reduced or removed, resulting in a clearer image. However, statistical-based methods may struggle when dealing with complex scenes and varying levels of haze.

Tan [15] proposed a model for the contrast maximization of an image since it is observed that the dehazed images show stronger contrast. A depth estimation method for transmission maps was introduced by He *et al.* [6] by using a DCP. [6] When the image is haze-free, DCP assumes that each pixel in one colour channel is near to zero. A technique for calculating information from scene depth was proposed by Zhu *et al.* [8] using a color attenuation prior. A technique which assumes that an image with no haze can be roughly represented using a haze line prior [2]. The prior-based methods produce positive out-

comes, however, the priors are dependent on the target scene and relative assumptions, and they are less robust in the real complicated scenario. For instance, dark channel prior [6] cannot effectively dehaze the sky area.

CNN-based methods

Deep learning-based techniques have gained significant attention and success in recent years. These techniques use neural networks' direct learning capabilities to identify and extract valuable elements from hazy images. Deep-learning-based techniques didn't use priors to calculate atmospheric light as well as the transmission map. Deep learning-based approaches can be further categorized into two subcategories: CNN-based methods and GANs-based methods. To calculate the transmission map for image dehazing, a CNN-based model is used by Cai *et al.* [3]. A multi-scale model was proposed by Ren *et al.* [26] that estimates transmission map by learning a mapping between transmission map and hazy inputs using a coarse-to-fine learning technique. AOD-Net [7] reformulated the physical scattering model. A fusion-based technique known as "Gated Fusion Network" was introduced by Ren *et al.* [13] that recovered the dehazed image from the hazy one.

GAN-based methods

Recently, Adversarial-based techniques, i.e., GAN has made significant development [27]. The discriminator and the generator are the two components of GAN. Both are trained concurrently to create a realistic image that confounds the discriminator. Many computer vision applications make use of GAN. In particular, GAN has shown promise in image synthesis [28–30]. The Success of generative adversarial networks gives us the motivation so that, we can use them for dehazing images. Zhang *et al.* [11] introduced a new model named DCPDN, i.e., "Densely connected pyramid dehazing network" for transmission map estimation. The disentangled dehazing network, which makes use of unpaired supervision, was proposed by Yang *et al.* [9]. A total of Three generators are included in the given adversarial model suggested by Yang *et al.* [9] one for haze-free image, one for the atmospheric light and one for the transmission map. DehazeGAN [31] uses a generative

adversarial network for the simultaneous estimation of atmospheric light and the transmission map. Single-image dehazing using different GANs is still in the initial stages. For all of the present dehazing techniques using GAN, the physical scattering model [25] is still a prerequisite. Little has been said about handling image dehazing without including the physical scattering model [25].

Both CNN-based and GAN-based methods possess their own strengths and limitations. CNN-based methods excel in accurately estimating the haze removal process but may encounter challenges in preserving fine details. GAN-based methods, on the other hand, excel in generating visually pleasing outputs but may introduce certain artifacts. Ongoing research and development efforts are focused on refining and improving both approaches to achieve better performance and address the specific challenges associated with each method.

In conclusion, the task of obtaining a dehazed image from the hazy one is a complex problem and researchers have introduced various techniques to address this challenge. These techniques can be broadly categorized into statistical-based techniques and deep learning-based techniques, which are later further divided into CNN-based and GAN-based approaches. Continued progress in these fields will help to improve dehazing techniques, enabling the creation of clearer and more aesthetically pleasing images even when haze is present.

1.2 Dataset available for image dehazing task

Deep learning methods often improve accuracy since the volume of training data increases drastically. In the case of single-image dehazing, there is various dataset available which can be used for testing as well as training the model.

(i) **HazeRD dataset** [19]: A collection of 14 clear outdoor images which are captured in five different conditions. All the dehazed images are captured along with the actual depth.

(ii) **I-Haze Dataset** [20]: Ancuti *et al.* introduced a collection of 35 indoor hazy images. For validation, the ground truth images are given with exactly five correspond-

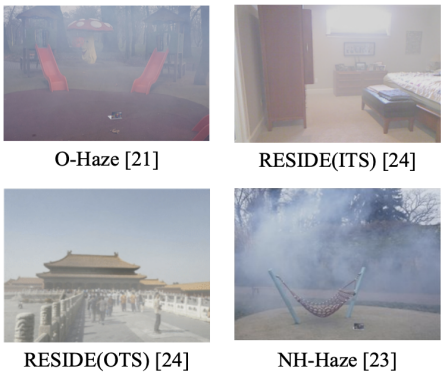

Figure 1.3: Sample images from different dehazing datasets

ing hazy images with different levels of haze depth. Resolution for both the pair of images(hazy and non-hazy) is 2833 x 4657 pixels.

(iii) **O-Haze dataset** [21]: Ancuti *et al.* introduced a collection of 35 outdoor hazy images. For validation, the ground truth images are given with exactly five corresponding hazy images with different levels of haze depth. A professional haze machine is used to generate hazy environment. The resolution for both the pair of images(hazy and non-hazy) is 2833 x 4657 pixels.

(iv) **Dense-Haze dataset** [22]: Ancuti *et al.* introduced a new dataset with 33 real hazy images. Both images contain the same visual information recorded under the same lighting conditions.

(v) **NH-Haze dataset** [23]: Ancuti *et al.* introduced a collection of 55 outdoor hazy images. The given hazy images are non-homogenous in nature which means the haze is non-uniformly distributed. The non-homogenous nature provides a greater depth for image dehazing tasks.

(vi) **RESIDE dataset** [24]: The dataset becomes the latest benchmark for image dehazing tasks which contains five categories of images for testing and training.

1.3 Performance Metrics

The following are the most often used Validation Metrics for quantitatively comparing models:

(i) **PSNR Score:** PSNR is a metric which compares the haze-free image created by a dehazing technique to the actual image. When PSNR value becomes high, the quality of the reconstructed image gets increases. Mathematically it can be represented as,

$$\text{PSNR} = 10 \log_{10} \frac{(2^d - 1)}{MSE} \tag{3}$$

Where d represents the maximum bit value in an image. MSE is the Mean Square Error. It evaluates the level of statistical model accuracy. Calculated is the average squared difference between observed and predicted values. Mathematically it can be represented as:

$$MSE = \frac{1}{nm} \sum_{i=1}^{n} \sum_{j=1}^{m} (I(z) - J(z))^2 \tag{4}$$

Where n and m represent height and width of an image respectively. I represent the ground truth and J represent the output. The quality of the reconstructed image will improve with greater PSNR values.

(ii) **SSIM Score:** The Structural Similarity Index (SSIM) is a metric used to assess the degradation of an image caused by processes like data compression or image reconstruction. It evaluates the similarity between two images based on their contrast, structure, and lighting. he SSIM score goes from -1 to 1, with a score of 1 denoting complete identity between the images. Mathematically, it can be represented as follows:

$$SSIM = f(l(z), c(z), s(z)) \tag{5}$$

Where $l(z)$ represents lightning in an image, $c(z)$ represents the contrast and $s(z)$ represents the structure of an image.

Chapter 2

LITERATURE REVIEW

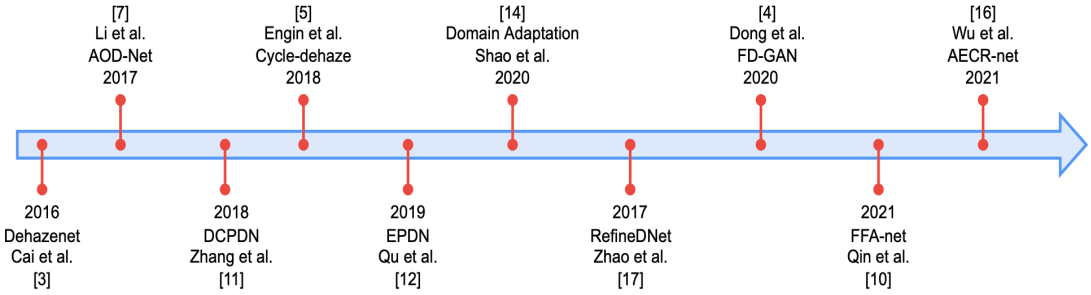

Figure 2.1: Deep learning-based image dehazing models' timeline

In this chapter, various learning-based approaches proposed so far for single-image dehazing are discussed (Fig. 2.1). The following approaches are discussed in terms of intuition, methodology, and loss functions. The popular dehazing methods discussed below are Dehaze-Net [3], AOD-Net [7] , FFA-Net [10], EPDN(i.e. Enhanced Pix2pix Dehazing Network) [12], DCPDN (i.e. Densely Connected Pyramid Dehazing Network) [11], Cycle-Dehaze [5], Domain Adaptation for Image Dehazing [14], RefineDNet [17], FD-GAN [4], AECR-Net [16].

2.1 DehazeNet method

An end-to-end network was proposed by Cai *et al.* [3] for calculating the transmission map which is used to construct dehazed image. It used a CNN based architecture which generates all the features which are relevant to hazy images. To imropve the reconstructed image quality, a bilateral rectified linear unit is also introduced.

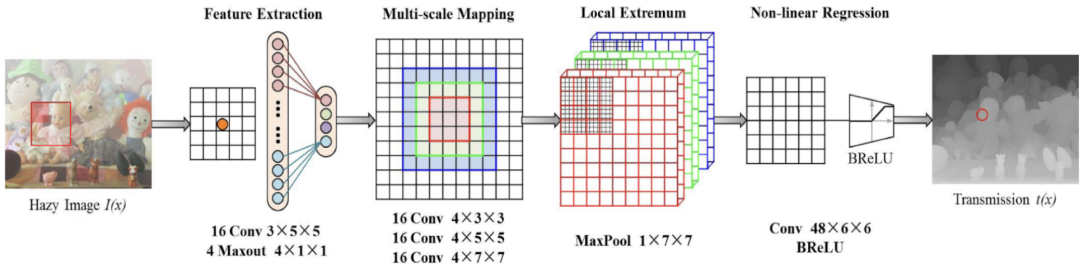

Figure 2.2: Dehaze-Net Architecture [3]

DehazeNet uses four main components to calculate transmission map: a feature extraction component, local extremum, a multi-scale mapping component and a nonlinear regression component. A maxout layer [18] is used by the feature extraction component which reduces the dimensionality. Multiple-scale mapping is used which extracts dense features in an input image at multiple levels. Local Extremum is used to reduce the noise in the network. For the non-linear regression component, Cai *et al.* [3] also introduced a BreLu(Bilateral Rectified Linear Unit) to increase the convergence rate.

The Dehaze-Net architecture (Fig. 2.2) is made up of several pooling and convolutional layers. SGD (Stochastic Gradient Descent) is used to train DehazeNet. MSE is utilised to calculate the loss.

2.2 AOD-Net method

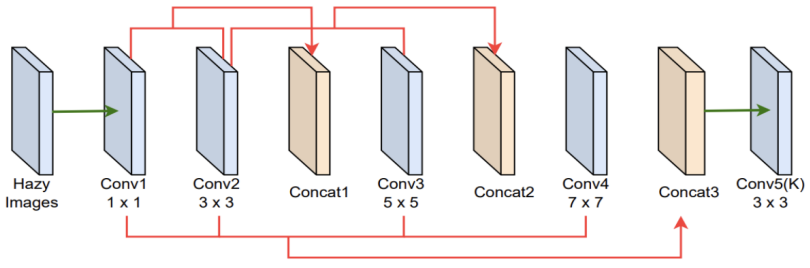

Figure 2.3: K-estimation components of AOD-Net [7].

A CNN-based dehazing technique is introduced by Li *et al.* [7]. AOD-Net generates the dehazed image without calculating transmission map or atmospheric light. [7] propose a k-estimation component which directly estimates the errors occurs while reconstructing

a dehazed image. The introduction of k-estimation will update the (1) as:

$$J(z) = K(z)I(z) - K(z) + b \tag{6}$$

$$K(z) = \frac{\frac{1}{t(z)}(I(z) - A) + (A - b)}{I(z) - 1} \tag{7}$$

Here $k(z)$ is the k-estimation module. $I(z)$ is the given dehazed image. b is the bias with 1 as the default value. A is the air light. $t(z)$ is the transmission map. The architecture of AOD-net consists five convolution layers with different filter sizes and concat these layers as shown in Fig. 2.3. Since this model can generate image in a single run, it can be used in other image enhancement-related tasks.

2.3 Feature Fusion Attention Network(FFA-Net) method

A feature fusion attention network was introduced by Qin *et al.* [10] for directly generating dehazed images. This feature attention module treats pixels and features of an image differently. This feature attention module contains two components: channel attention and pixel attention (Fig. 2.4).

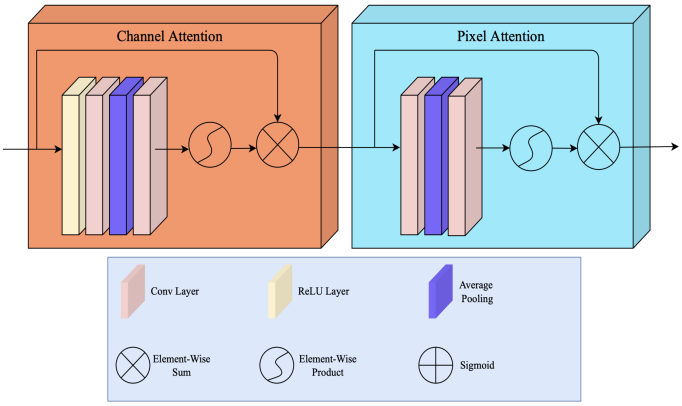

Figure 2.4: Channel and Pixel attention architecture of FFA-Net [10].

The concept behind the channel attention module is that different channels in an image contain various types of information. In order to get this spatial and feature information,

global average pooling is used. Mathematically, channel attention can be represented as:

$$g_{\mathrm{c}} = H_{\mathrm{p}}(F_{\mathrm{c}}) = \frac{1}{H * W} \sum_{m=1}^{W} \sum_{n=1}^{H} I_{\mathrm{c}}(m, n) \tag{8}$$

$$CA_{\mathrm{c}} = \sigma(Conv(\delta(Conv(g_{\mathrm{c}})))) \tag{9}$$

$$F^*{}_{\mathrm{c}} = F_{\mathrm{c}} \otimes CA_{\mathrm{c}} \tag{10}$$

Here, I_c represents a pixel of an Image at (i, j) in c^{th} channel. F_c represents the input. H_p represents global average pooling. σ signifies sigmoid functions. δ signifies ReLu function. CA_c represents the channel weights.

Pixel attention module is introduced to deal with thick hazy pixels in an image. It takes the output generate by channel attention as input and performs two convolution operation followed by the activation functions.

$$PA = \sigma(Conv(\delta(Conv(F^*)))) \tag{11}$$

$$\tilde{F} = F^* \otimes PA \tag{12}$$

Here, F^* represent input which is coming from chanel attention's output. PA denotes the pixel attention's output and $\tilde{F}$ is the output of FA module.

2.4 Enhanced Pix2pix Dehazing Network method

EPDN [12] is a GAN-based technique which uses pix2pixHD GAN [32] architecture that converts an image dehazing problem into an image-to-image translation problem. However, the dehazed image is not directly provided to pix2pixHD [32] GAN. EPDN first provides some modification to the original code by introducing an enhancer block to the generator part. This method doesn't rely on the estimation of transmission maps or atmospheric light for generating haze-free images.

Three main components comprise the architecture of EPDN: a multi-scale discriminator module, an enhanced module, and a multi-resolution generator module.

Multi-resolution generator: The multi-resolution generator of EPDN is a two-generator architecture. The first one is a global sub-generator, G_1, and the other one is a local sub-generator, G_2. G_1 is incorporated in G_2. A coarse-scale image is generated G_1 while G_2 generates an image on a fine-scale. A coarse-to-fine image is generated when the two generators combine.

Enhanced block: Due to the overcolored and lacked details dehazed images produced by the pix2pixHD, an enhancing block was introduced by Que et al. [12]. The enhancing block can extract information on a fine scale by using a receptive field model.

Multi-scale Discriminator: The discriminator of the networks also follows a two-discriminator architecture D_1 and D_2. Both follow similar architecture. However, D_1 guides the generator to generate the result on coarse scale whereas D_2 directs the generator to produce the fine-scale images. The overall loss of the EPDN architecture is shown in equation (3).

$$L_{total} = L_A + \lambda L_{FM} + \lambda L_{VGG} + L_F \tag{13}$$

Here, L_A refers to the adversarial loss of generative adversarial networks. L_{FM} refers to feature matching loss which helps the generator to create realistic multi-scale statistical data. L_{VGG} is the perceptual loss. L_F is the fidelity loss which provides euclidean distance between actual and generated haze-free images. EPDN didn't rely on the scattering model [25] for generating the haze-free image. When the method runs on heavily hazed images, the performance of the model falls.

2.5 Densely Connected Pyramid Dehazing Network

By predicting ambient light and transmission map, Zhang *et al.* [11] presented a GAN-based network that creates haze-free images. The architecture of DCPDN majorly divides into these categories: A joint discriminator, a densely connected pyramid network and an atmospheric light estimation network. Fig. 2.5 shows the architecture of DCPDN. This architecture is used to calculate the transmission map and a U-net architecture [39] is used to estimates the atmospheric light. After estimating both ambient light and transmission

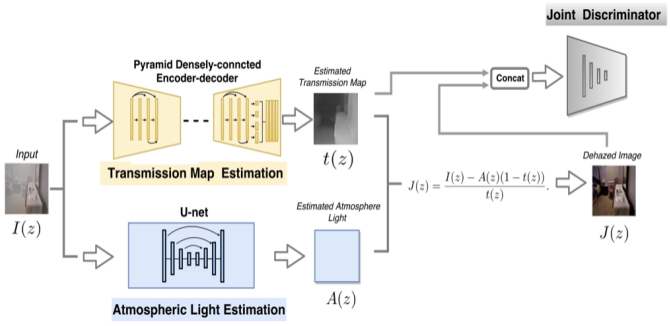

Figure 2.5: Architecture of DCPDN [11]

map, we generate the dehazed image using the atmospheric scattering equation (1). The joint discriminator determines if the created image is fake or real by computing the loss. [11] first combined both the transmission map and the dehazed image before feeding it to a joint discriminator for training. The total loss of the DCPDN architecture is given as.

$$L_{total} = \lambda_j L^j + L^t + L^d + L^a \tag{14}$$

Here, L^j refers to the joint-discriminator loss from generative adversarial networks, where the generator produces haze-free image and discriminator is trained to verify if an image is real or false. L^a refers to the traditional MSE loss, which is used to calculate transmission map. L^t refers edge-preserving loss. I is the newly introduced loss by the author which inside uses three different losses: L_2 loss, feature loss and two-directional gradient loss. L^d represents the dehazing loss.

2.6 Cycle-Dehaze method

Engin *et al.* proposed a Cycle-Dehaze technique which is based on the concept of Cycle GAN [30]. An entire architecture that doesn't require combining the corresponding ground truth image with the real ground truth image for model training. The transmission map for estimation is not necessary since the method is based on GANs.

The cycle-dehaze method train on an unpaired set of hazy and clean images. As we know, the deep learning method generally works on low-resolution images. So, this model

downsamples the high-quality images first as cycle-GAN works better on low-resolution images. For down-sampling bicubic downsampling is used and for up-sampling Laplacian pyramid is used. Cycle-dehaze consists of two generators along with two discriminators. The architecture of cycle-dehaze is shown in Fig 2.6.

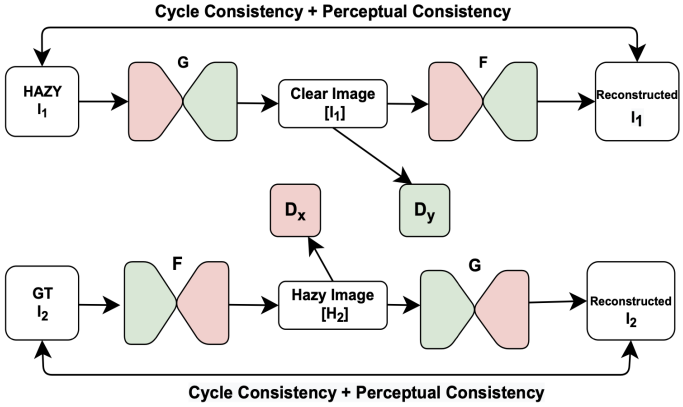

Figure 2.6: Architecture of cycle-dehaze [5]

To improve the quality of an image cycle-consistency and perceptual loss are used. The overall loss for the cycle-dehaze architecture can be written as:

$$L(F,\ G,\ D_y,\ D_x) = \ \gamma * \mathcal{L}_{Perceptual}(F,\ G) + \ L_{CycleGAN}(F,\ G,\ D_y,\ D_x) \qquad (15)$$

Here, F and G represent the generator-one and generator-two respectively. D_x, D_y represents discriminators one and two. $L_{Perceptual}$ represents perceptual loss and $\mathcal{L}_{CycleGAN}$ represents cycle-gan loss. The L1-regularization is computed by perceptual-consistency loss. The following model compares the reconstructed image to the original image. To preserve the sharpening of an image, perceptual loss is used and to ensure a higher PSNR value, consistency loss is used.

2.7 Domain Adaptation for Image Dehazing

A domain adaptation technique was proposed by Shao *et al.* [14]. Domain adaptation method for image dehazing [14] tries to lessen the discrepancy across various domains because the majority of approaches [**?**, 3, 12] typically employed synthetic hazy images for

training. The aim of domain adaptation in this case is to lessen disparity across various domains.

This method is composed of two components: an image translation component and two dehazing components (one for the real hazy inputs and the other for the synthetic hazy inputs). At first, the image is translated from one domain to another and then uses the resultant images with the actual image in order to conduct dehazing on both the synthetic domain as well as the real domain.

The overall loss function for the domain adaptation [14] architecture can be written as:

$$L = \lambda_d(L_{rd} + L_{sd}) + \lambda_m(L_{rm} + L_{sm}) + \ \lambda_t(L_{rt} + L_{st}) + \lambda_c L_c + L_{tran} \tag{16}$$

Here, L_{trans} represents the translation loss. L_{rm} and L_{sm} represent the mean squared error loss for dehazed images, both real and synthetic respectively. L_{rd}, L_{sd} represents the dark channel loss for both real as well as synthetic dehazed images respectively. L_rt, and L_st represents the total variation loss for synthetic hazy images along with the real ones respectively. L_c represents a cyclic-consistency loss.

2.8 RefineDNet

Zhao *et al.* [17] proposed a refinement architecture which is weakly supervised in nature. As we know GAN methods produce more realistic results. However, the capabilities of earlier haze removal systems are constrained owing to the absence of real-world paired datasets (both actual and hazy). The weakly supervision help to generate the good results and also outperformed many deep learning approaches based on supervision.

The method combines the advantages of both the learning-based as well as statistical-based techniques by categorizing the dehazing task into two sub-problem: realness improvement and visibility restoration. The weakly supervised dehazing method has two stages. In stage I author uses DCP [6] to restore image visibility. In stage II GANs-based learning with an unpaired dataset is used to improve the realness of an image. The method also proposes a strategy, known as perceptual fusion, to blend the refined and

reproduced dehazed image, to get more qualified results. The strategy assigns weights which are close to the natural images. Fig. 2.7 illustrates the working of the perceptual fusion strategy.

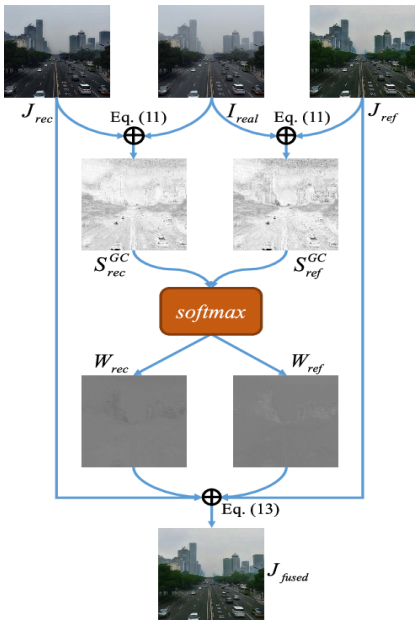

Figure 2.7: Working of perceptual fusion strategy in RefineDnet [17]

The overall loss for the refineDnet architecture can be written as:

$$L_{total} = \arg \min_{R_T, R_J} \max_{D} \lambda L_G + L_{rec} + L_{idt} \tag{17}$$

Here, L_{rec} represents the reconstruction loss which is used to normalize the reconstructed hazy images. L_G represents the adversarial loss. Mean absolute error is used to calculate the error between real and reconstructed hazy images. L_{itd} represents the identity loss which is to decrease the number of artefacts in the dehazed image which got introduced after the refinement part. λ represents the weights as a hyperparameter.

2.9 AECR-Net

A contrastive learning-based approach was proposed by Wu *et al.* [16] in 2021. For positive samples, previous deep learning-based techniques primarily used clear images to

train dehazing network, leaving negative information unexploited. Wu *et al.* uses an auto-encoder and contrastive regularization, in short AECR-net, to deal with both negative and positive samples of an image.

The ground truth image is portrayed by the authors as positive, the fuzzy image as negative, and the reconstructed images are used as an anchor. The suggested contrastive regularisation approach makes sure that the reconstructed picture is pushed further away from the original hazy image while being moved closer to the ground truth image. The autoencoder (AE) architecture also contains an adaptive mixing operation and a dynamic feature enhancement module, both of which are used to improve network transformability. Contrastive regularization is frequently employed in self-supervised learning and can enhance the quality of translation from unpaired images to images. The overall loss for the AECR-net architecture can be written as,

$$L_{total} = \ \beta \cdot \rho(G(\phi(I, w)), G(I), G(J)) + \ \min \|J - \phi(I, w)\| \tag{18}$$

Here, J represents dehazed image, I represents a hazy image and G represents the actual ground truth. $\min ||J - \phi(I, w)||$ represents reconstruction loss that is used to find displacement between the original image and the restored image. The $\phi(I, w)$, J and I represents the contrastive regularization term.

Chapter 3

METHODOLOGY

In this work, we proposed a new FAG-Net for single-image dehazing. When the dense haze is present, FAG-Net outperforms existing state-of-the-art dehazing techniques. The recovered image not only surpasses the previous models quantitatively but also generates visually appealing and highly detailed images. In FAG-Net, a new generator architecture is introduced to dehaze the images.

Figure 3.1: Hazy image (left) and its corresponding clean image (right) generated from our proposed model on RESIDE [24] ITS and I-Haze [20] datasets respectively.

The new generator consists of transition block, dense block (Fig. 3) and a new block known as feature attention (FA) block at each layer to enhance the realistic nature of haze-free image. FA block consists one channel attention (CA) block and one pixel attention (PA) block [10]. Channel attention enhance the relevant information in different color channels of the input image. While pixel attention aims to selectively emphasize or suppress certain pixels in the image which helps the models to concentrate on the most affected areas of an image by haze. A traditional CNN-based discriminator [28] is used,

which checks whether the output image generated by our FAG-Net is near to the actual dehazed image or not.

The following section discusses the proposed FAG-Net, which can compete with significant existing research approaches and perform considerable dehazing. The components of the network, in particular generator and discriminator components, are discussed in the subsections along with the loss functions. A dense block, transition block, and feature attention (FA) block are the three different sorts of blocks in our model.

3.1 Dense Block

The utilization of a dense block proves beneficial in extracting intricate features from hazy images, thereby enhancing the accuracy of transmission map estimation. Within the encoder part, a dense block encompasses a sequence of convolution layers, ReLU activation, Batch Normalization, and a dropout layer. Conversely, the decoder section of the network consists of a series of ReLU activation and Convolution Layers, forming the dense block. The incorporation of skip connections addresses the challenge of vanishing gradients, ensuring that crucial features are not disregarded during the learning process. By employing skip connections, the network can effectively propagate gradients and preserve vital information, leading to improved performance in haze removal tasks.

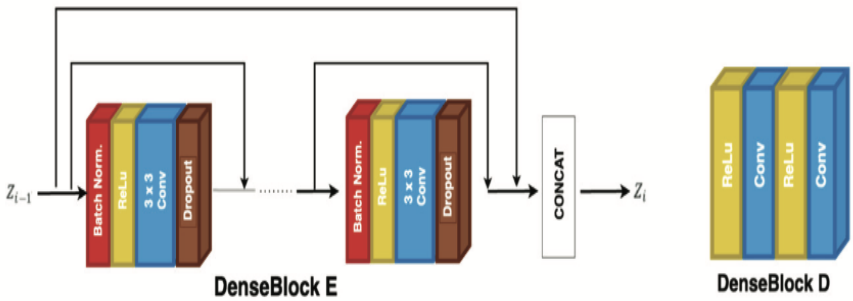

Figure 3.2: DenseBlock Architecture [50].

3.2 Transition Block

It is used to lower the spatial dimension of output feature maps from the convolution layers. It consists of convolution layers, ReLu, average pooling and Batch normalization. The components of the transition Block are the convolution layer, Batch normalization, ReLu and average pooling. The main advantage of using a transition block in image dehazing is that the number of parameters gets reduced drastically in a network, which can improve the speed of training and reduce the risk of overfitting. A transition block in the decoder part is used for upscaling an image which helps in gradually restoring the feature map's size.

3.3 Feature Attention (FA) Module

The FAG-Net uses a feature attention (FA) block which uses channel attention (CA) and pixel attention (PA) [10]. Combining both can provide us with a mechanism which can deal with the images which have non-uniform haze distribution. The main aim of channel attention is to enhance the relevant information in different color channels of the input image. Channel Attention performs a global average pooling at each channel and at each pixel. The Pixel attention aims to selectively emphasize or suppress certain pixels in the image which helps in aiming the most affected areas by haze. It targets the dense haze regions in a hazy image and tries to pass the information to the next layer of our model. Mathematically, the output of our feature-fusion block can be represented as:

$$\tilde{O}_{ff} = (CA_o \otimes I) \otimes PA_o \tag{19}$$

$$CA_o = \sigma(Conv(\delta(Conv(I_g)))) \tag{20}$$

$$PA_o = \sigma(Conv(\delta(CA_o \otimes I))) \tag{21}$$

where $\tilde{O_{ff}}$ represents the output of the feature attention block which is of shape $1 \times H \times W$. I represent input hazy image. CA_o represents channel attention's output. CA_o is the combination of sigmoid, ReLu and two convolution layers applied on I_g. Here,

I_g represents the input image after the global average pooling of shape $C \times 1 \times 1$. PA_o is the pixel attention's output. It is an element-wise multiplication of the actual input image and channel attention's output followed by the combination of the sigmoid, ReLU and convolution layer.

3.4 Generator

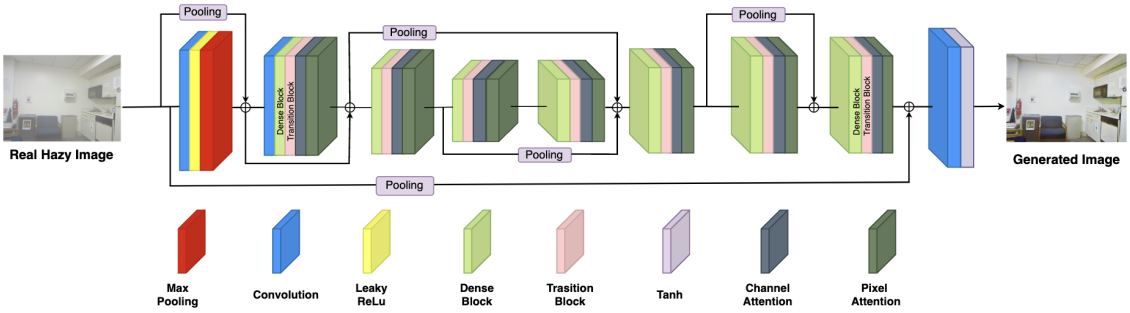

Figure 3.3: Generator architecture for FAG-Net.

Our new generator uses an encoder-decoder architecture [4] that input's a hazy image and output's a dehazed image. Both the encoder and the decoder consist of dense Blocks, transition Blocks and feature attention (FA) blocks. The encoder consists of three dense blocks and three transition blocks followed by a feature attention (FA) block. Dense block and transition block use DenseNet-121 CNN architecture's pre-trained weights [51]. The encoder blocks downscale input image size while extracting the important information. The feature-fusion (FA) block is used to enhance realistic nature of dehazed image after each transition block. The decoder architecture converts the encoded image to its original shape and size. The transition block in the decoder also contains a upsample block which progressively upscales the image to its original size. Tanh is used as an activation function at the end of the decoder to provide non-linearity and produce more accurate results.

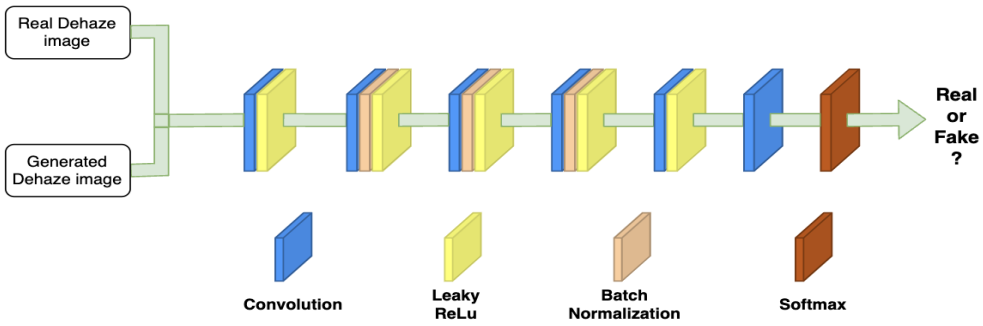

Figure 3.4: Discriminator Architecture

3.5 Discriminator Architecture

The discriminator (Fig. 4) takes the real dehaze images from the dataset and the generated dehaze images from the proposed generator as inputs and checks whether the generated image is real or fake. Then give this feedback to the generator which again produces a fake dehaze image. And this process will go on and on until we get some pleasant eye-catching results. The discriminator uses a general CNN patched-based discriminator which is used in conditional-gan [6]. The discriminator consists of six convolution Layers, 5 leaky ReLu, 3 batch normalization and softmax. In each input image, the discriminator uses 14×14 patches to determine whether it is real or fake.

To determine the authenticity of the images, the discriminator employs a combination of different loss functions, including pixel-wise loss, adversarial loss, and perceptual loss. These diverse loss functions enable the discriminator to comprehensively evaluate the similarity between the generated images and the real images within the dataset. By integrating multiple loss functions, the discriminator provides a holistic assessment of the realism and quality of the generated dehazed images.

3.5.1 Loss Functions

a) **Adversarial Loss:** The generative networks employ adversarial loss to enhance the realism of the dehazed images they produce. The objective of the generator is to generate dehazed images that closely resemble real ones, while the discriminator's role is to differentiate between genuine dehazed images and those generated by the generator.

Mathematically, the adversarial loss can be expressed as follows:

$$L_{adv} = \nabla_{\theta adv}(1/n) \sum_{j=1}^{n} [log(1 - D(G(z_i))) + logD(x_i)] \tag{22}$$

Here $G(z_i)$ signifies the generator's output when at i^{th} pixel, noise z is given, *log(D(x))* signifies the probability estimate of discriminator that an actual data instance is real and *D(G(z))* represents discriminator's probability estimate that a generated instance is actual or not.

b) **Reconstruction Loss:** The given loss measures how well the reconstructed dehazed images match the actual dehazed image by examining the pixel-wise difference. To determine how well an image has been reconstructed from the input, it compares the pixels between the clear and dehazed images. Since it reduces the impact of large errors, we select mean absolute error loss (MSE) as reconstruction loss. Mathematically, we can represent reconstruction loss as:

$$L_{recon} = \sum_{j=1}^{n} \|G(I_i) - J_i\| \tag{23}$$

where *G(I_i)* is the generator's output, J_i the actual ground truth, I_i represents the given input hazy image.

c) Perceptual Loss: This loss is used when we are comparing two different images which look identical and only shifted by just one pixel. Perceptual loss examines semantic and perceptual differences between images at the highest level. It determines the difference between generated dehazed images and the original ones using features derived from a pre-trained network.

Chapter 4

RESULTS and DISCUSSION

4.1 Experimental Analysis

4.1.1 Datasets

We utilised the following datasets to train our FAG-Net:

1) **I-Haze Dataset** [20]: It is a collection of 35 indoor hazy images. For validation, the ground truth images are given with exactly five corresponding hazy images with different levels of haze depth. Resolution for both the pair of images(hazy and non-hazy) is 2833 x 4657 pixels.

2) **O-Haze dataset** [21]: Ancuti *et al.* introduced a collection of 35 outdoor hazy images. For validation, the ground truth images are given with exactly five corresponding hazy images with different levels of haze depth. A professional haze machine is used to generate hazy environment. The resolution for both the pair of images(hazy and non-hazy) is 2833 x 4657 pixels.

2) **RESIDE-ITS dataset(Indoor Training Set)** [24]: It becomes a benchmark dataset which contains indoor images (Indoor Training Dataset). We select 1400 images which covered the majority of hazy cases.

4.1.2 Training Details

We used google colab pro to train our proposed FAG-Net. The GPU used is a Tesla T4 with 32GB of RAM. We trained our model for 10^4 epochs for each dataset. Each dataset set takes 2 days to train. We set the learning rate to $2 \times 10^{-4}$and gets decayed by 0.5

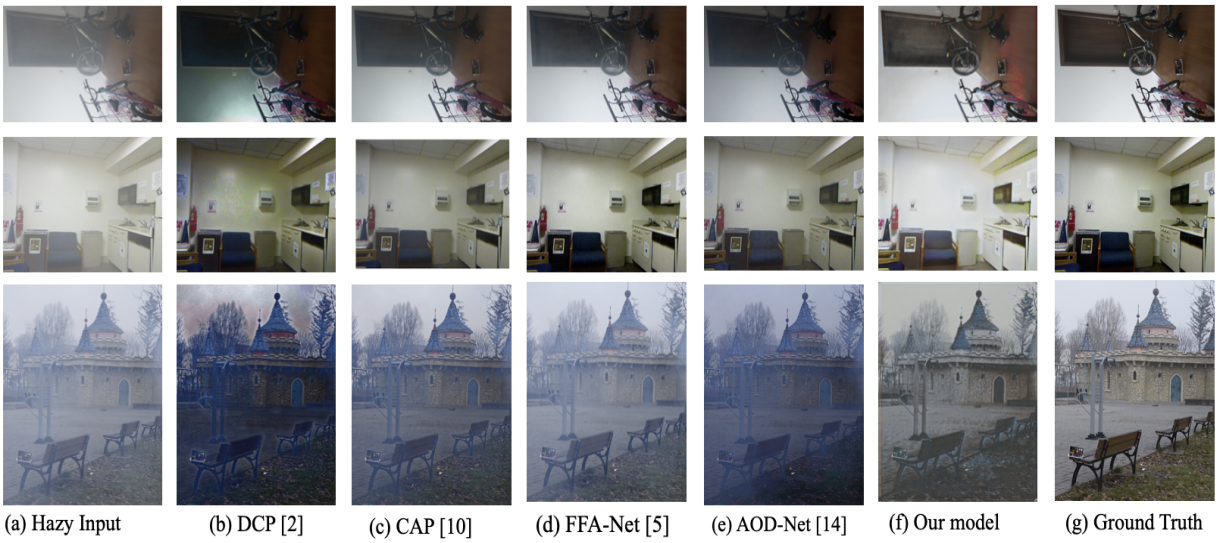

Figure 4.1: Results generated on different datasets: I-HAZE [20], RESIDE(ITS) [24],O-HAZE [21](Top to bottom).

after every 2000 epochs. After thorough analysis, the adversarial loss and perceptual loss weight values are set to 0.5. Adam optimizer [49] is used to train FAG-Net with a batch size of 1.

4.1.3 Qualitative and Quantitative Evaluation

We evaluate the FAG-Net both quantitatively and qualitatively against following state-of-the-art methods: DCP [6], CAP [8], FFA-Net [10], AOD-Net [7]. For measuring our FAG-Net quantitatively, PSNR and SSIM scores are used. PSNR compares the haze-free image created by a dehazing technique to the actual image. SSIM evaluates image degradation induced by processing such as data compression or image reconstruction.

Table 4.1: Quantitative comparisons with O-HAZE [21], I-HAZE [20] and RESIDE(ITS) [24] Datasets.

Methods	**O-HAZE**		**I-HAZE**		**RESIDE(ITS)**	
	SSIM	**PSNR (dB)**	**SSIM**	**PSNR (dB)**	**SSIM**	**PSNR (dB)**
DCP [2]	0.65	16.78	0.75	14.42	15.20	0.82
CAP [11]	0.59	16.08	0.62	12.23	13.36	0.75
FFA-Net [5]	0.86	17.1	1.73	12.36	14.45	0.78
AOD-Net [14]	0.82	15.20	0.67	11.72	13.98	0.73
Ours	**0.86**	**15.20**	**0.92**	**24.3**	**0.85**	**19.2**

Table I shows the quantitative performance of the FAG-Net with other state-of-the-art models. From the table, we can clearly observe that our model performs quantitatively on par with other dehazing methods. Fig. 5 shows the qualitative comparison of FAG-Net with other dehazing models [2, 5, 10, 14]on different datasets [20, 21, 24]. For indoor datasets [20, 24], our model outperforms all the datasets by producing visually great results. However, for the outdoor dataset [21] further improvement is still required as the output result has some low contrast.

Chapter 5

CONCLUSION AND FUTURE SCOPE

The research focused on exploring multiple deep-learning-based approaches for image dehazing. These approaches were discussed briefly, shedding light on their respective methodologies and mathematical equations from various perspectives. Examining different techniques offers valuable insights into the strengths and limitations of image dehazing methods. Deep learning has brought about a revolutionary impact on image dehazing tasks in recent years, and the introduction of generative adversarial networks (GANs) has further spurred researchers to generate new ideas for tackling image dehazing challenges.

Deep learning has demonstrated remarkable results in image dehazing; however, there is still room for improvement in terms of space and time complexity. Further efforts are required to combine the advantages of different proposed statistical-based methods with deep learning approaches. This is because certain models perform well in specific environments, and researchers aim to develop weakly supervised methods capable of producing excellent results in both real hazy and synthetic environments.

The development of effective weakly supervised methods is an area of research. The weakly supervised methods aim to leverage limited or imperfect annotations to train deep learning models for image dehazing. Researchers might lessen their dependency on fully annotated datasets, which are often time-consuming and costly to develop, by investigating weak supervision strategies. This enables more realistic and scalable solutions in real-world circumstances.

Furthermore, combining statistical-based methods with deep learning approaches has the potential to improve the performance and robustness of dehazing algorithms. Researchers might potentially transcend the limitations of individual methodologies and

generate more complete and effective solutions by using the strengths of both paradigms.

To summarise, while deep learning has greatly advanced image dehazing jobs, there is still much work to be done. Improving the space and time complexity, exploring weakly supervised methods, and combining statistical-based and deep learning approaches are key areas of focus for researchers. By addressing these challenges, the goal is to develop innovative techniques that can generate high-quality dehazed images in a variety of real-world environments.

In this work, we also introduced a new feature attention generative network (FAG-Net) for image dehazing and demonstrated its visual capabilities. To extract more information and to give the results much more depth, a newly proposed feature-fusion (FA) block is introduced in the generator. Perceptual Loss and reconstruction Loss along with adversarial loss are used to provide more visually attractive natural outcomes including superior structure and edge identification while also considerably reducing the appearance of artefacts. Our proposed FAG-Net generated aesthetically appealing dehazed images with less color distortion surpassing existing state-of-the-art dehazing techniques. Extensive experimental analysis demonstrates that our FAG-Net surpasses existing state-of-the-art models both quantitatively and qualitatively. In the future, this work can be extended for other image enhancement tasks like low-light image enhancement, image denoising, image deraining etc.

Bibliography

[1] D. Berman, T. Treibitz and S. Avidan, "Non-local Image Dehazing," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 1674-1682.

[2] D. Berman, T. Treibitz and S. Avidan, "Single Image Dehazing Using Haze-Lines," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 3, pp. 720-734, 1 March 2020.

[3] B. Cai, X. Xu, K. Jia, C. Qing and D. Tao, "DehazeNet: An End-to-End System for Single Image Haze Removal," in IEEE Transactions on Image Processing, vol. 25, no. 11, pp. 5187-5198, Nov. 2016.

[4] Y. Dong, Y. Liu, H. Zhang, S. Chen, and Y. Qiao, "FD-GAN: Generative Adversarial Networks with Fusion-Discriminator for Single Image Dehazing", AAAI, vol. 34, no. 07, pp. 10729-10736, Apr. 2020.

[5] D. Engin, A. Genc and H. K. Ekenel, "Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 2018, pp. 938-9388, doi: 10.1109/CVPRW.2018.00127.

[6] K. He, J. Sun and X. Tang, "Single Image Haze Removal Using Dark Channel Prior," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 12, pp. 2341-2353, Dec. 2011, doi:10.1109/TPAMI.2010.168.

[7] B. Li, X. Peng, Z. Wang, J. Xu and D. Feng, "AOD-Net: All-in-One Dehazing Network," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 4780-4788, doi: 10.1109/ICCV.2017.511.

[8] L. Liu, G. Cheng and J. Zhu, "Improved Single Haze Removal Algorithm Based on Color Attenuation Prior," 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), Chongqing, China, 2021, pp. 1166- 1170, doi: 10.1109/ICIBA52610.2021.9687882.

[9] X. Yang, Z. Xu, and J. Luo, "Towards Perceptual Image Dehazing by Physics-Based Disentanglement and Adversarial Training", AAAI, vol. 32, no. 1, Apr. 2018.

[10] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, "FFA-Net: Feature Fusion Attention Network for Single Image Dehazing", AAAI, vol. 34, no. 07, pp. 11908-11915, Apr. 2020.

[11] H. Zhang and V. M. Patel, "Densely Connected Pyramid Dehazing Network," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 3194-3203, doi: 10.1109/CVPR.2018.00337.

[12] Y. Qu, Y. Chen, J. Huang and Y. Xie, "Enhanced Pix2pix Dehazing Network," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 2019, pp. 8152-8160, doi: 10.1109/CVPR.2019.00835.

[13] W. Ren et al., "Gated Fusion Network for Single Image Dehazing," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 3253-3261, doi: 10.1109/CVPR.2018.00343.

[14] Y. Shao, L. Li, W. Ren, C. Gao and N. Sang, "Domain Adaptation for Image Dehazing," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020, pp. 2805-2814, doi: 10.1109/CVPR42600.2020.00288.

[15] R. T. Tan, "Visibility in bad weather from a single image," 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, 2008, pp. 1-8, doi: 10.1109/CVPR.2008.4587643.

[16] H. Wu et al., "Contrastive Learning for Compact Single Image Dehazing," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 2021, pp. 10546-10555, doi: 10.1109/CVPR46437.2021.01041.

[17] S. Zhao, L. Zhang, Y. Shen and Y. Zhou, "RefineDNet: A Weakly Supervised Refinement Framework for Single Image Dehazing," in IEEE Transactions on Image Processing, vol. 30, pp. 3391-3404, 2021, doi: 10.1109/TIP.2021.3060873.

[18] I. Goodfellow, D. Warde-Farley, M. Mirza, A. Courville, and Y. Bengio, "Maxout Networks," 30th International Conference on Machine Learning, ICML 2013, vol. 1302, Feb. 2013.

[19] Y. Zhang, L. Ding and G. Sharma, "HazeRD: An outdoor scene dataset and benchmark for single image dehazing," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 2017, pp. 3205-3209, doi: 10.1109/ICIP.2017.8296874.

[20] A. Codruta, C. Ancuti, R. Timofte, and C. Vleeschouwer, "I-HAZE: a dehazing benchmark with real hazy and haze-free indoor images," Apr. 2018.

[21] A. Codruta, C. Ancuti, R. Timofte, and C. Vleeschouwer, "O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images," Apr. 2018.

[22] C. O. Ancuti, C. Ancuti, M. Sbert and R. Timofte, "Dense-Haze: A Benchmark for Image Dehazing with Dense-Haze and Haze-Free Images," 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 1014-1018, doi: 10.1109/ICIP.2019.8803046.

[23] C. O. Ancuti, C. Ancuti and R. Timofte, "NH-HAZE: An Image Dehazing Benchmark with Non-Homogeneous Hazy and Haze-Free Images," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 1798-1805, doi: 10.1109/CVPRW50498.2020.00230.

[24] B. Li et al., "Benchmarking Single-Image Dehazing and Beyond," in IEEE Transactions on Image Processing, vol. 28, no. 1, pp. 492-505, Jan. 2019, doi: 10.1109/TIP.2018.2867951.

[25] S. Narasimhan and S. Nayar, “Vision and the Atmosphere,” International Journal of Computer Vision, vol. 48, pp. 233–254, Jul. 2002, doi: 10.1145/1508044.1508113.

[26] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, “Single Image Dehazing via Multi-scale Convolutional Neural Networks,” Oct. 2016, vol. 9906, pp. 154–169. doi: 10.1007/978-3-319-46475-6-10.

[27] I. Goodfellow et al., “Generative Adversarial Networks,” Commun. ACM, vol. 63, no. 11, pp. 139–144, Oct. 2020, doi: 10.1145/3422622.

[28] P. Isola, J. -Y. Zhu, T. Zhou and A. A. Efros, ”Image-to-Image Translation with Conditional Adversarial Networks,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 5967-5976, doi: 10.1109/CVPR.2017.632.

[29] A. Radford, L. Metz, and S. Chintala, “Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks,” in 4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings, 2016. [Online]. Available: http://arxiv.org/abs/1511.06434.

[30] J. -Y. Zhu, T. Park, P. Isola and A. A. Efros, ”Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks,” 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2242-2251, doi: 10.1109/ICCV.2017.244.

[31] H. Zhu, X. Peng, V. Chandrasekhar, L. Li, and J.-H. Lim, “DehazeGAN: When Image Dehazing Meets Differential Programming,” in Proceedings of the 27th International Joint Conference on Artificial Intelligence, 2018, pp. 1234–1240.

[32] T. -C. Wang, M. -Y. Liu, J. -Y. Zhu, A. Tao, J. Kautz and B. Catanzaro, ”High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs,” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 2018, pp. 8798-8807, doi: 10.1109/CVPR.2018.00917.

[33] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards RealTime Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031.

[34] Y. Miao, H. Hong and H. Kim, "Size and angle filter based rain removal in video for outdoor surveillance systems," 2011 8th Asian Control Conference (ASCC), Kaohsiung, Taiwan, 2011, pp. 1300-1304.

[35] P. Guruvareddiar and P. Prasad, "Artificial Intelligence Based Region of Interest Enhanced Video Compression," 2020 Data Compression Conference (DCC), Snowbird, UT, USA, 2020, pp. 373-373, doi: 10.1109/DCC47342.2020.00095.

[36] T. Sharma, T. Shah, N. K. Verma and S. Vasikarla, "A Review on Image Dehazing Algorithms for Vision based Applications in Outdoor Environment," 2020 IEEE Applied Imagery Pattern Recognition Workshop (AIPR), Washington DC, DC, USA, 2020, pp. 1-13, doi: 10.1109/AIPR50011.2020.9425261.

[37] A. Reza, "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement," VLSI Signal Processing, vol. 38, pp. 35–44, Aug. 2004, doi: 10.1023/B:VLSI.0000028532.53893.82.

[38] T. M. Bui and W. Kim, "Single Image Dehazing Using Color Ellipsoid Prior," in IEEE Transactions on Image Processing, vol. 27, no. 2, pp. 999-1009, Feb. 2018, doi: 10.1109/TIP.2017.2771158.

[39] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in LNCS, Oct. 2015, vol. 9351, pp. 234–241. doi: 10.1007/978-3-319-24574-4-28.

[40] A. Gupta, I. Khatri, A. Choudhry, and S. Kumar, "MCD: A modified community diversity approach for detecting influential nodes in social networks," in J Intell Inf Syst, pp. 1-23, 2023.

[41] S. Kumar, N. Kumar, A. Dev, and S. Naorem, "Movie genre classification using binary relevance, label powerset, and machine learning classifiers," in Multimedia Tools and Applications, vol. 82, no. 1, pp. 945-968, 2023.

[42] K. Singh, and A. S. Parihar, "Illumination Estimation for Nature Preserving low-light image enhancement," in The Visual Computer, pp. 1-16, 2023.

[43] A. S. Parihar, D. Varshney, K. Pandya, and A. Aggarwal "A comprehensive survey on video frame interpolation techniques," in The Visual Computer, vol. 38, pp. 295-391, 2022.

[44] K. Singh and A. S. Parihar, "Variational optimization based single image dehazing," in Journal of Visual Communication and Image Representation, vol. 79, pp. 103241, 2021.

[45] A. S. Parihar, and A. Java, "Densely connected convolutional transformer for single image dehazing," in Journal of Visual Communication and Image Representation, vol. 90, pp. 103722, 2023.

[46] A. S. Parihar, S. Kumar, and S. Khosla, "S-DCNN: stacked deep convolutional neural networks for malware classification," in Multimedia Tools and Applications, vol. 81, no. 21, pp. 30997-31015, Sept, 2022.

[47] N. Sissenwine, M. Dubin, and H. Wexler, 'The US standard atmosphere, 1962', Journal of Geophysical Research, vol. 67, no. 9, pp. 3627–3630, 1962.

[48] E. J. McCartney, 'Optics of the atmosphere: scattering by molecules and particles', New York, 1976.

[49] D.P.KingmaandJ.Ba, 'Adam: A method for stochastic optimization', arXiv preprint arXiv:1412. 6980, 2014.

[50] A. S. Parihar, K. Singh, A. Ganotra, A. Yadav and Devashish, "Contrast Aware Image Dehazing using Generative Adversarial Network," 2022 2nd International

Conference on Intelligent Technologies (CONIT), Hubli, India, 2022, pp. 1-6, doi: 10.1109/CONIT55038.2022.9847710.

[51] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243.

List of Publications

[1] D. Rawat, K. Singh, "FAG-Net: Feature Attention Generative Network for Single Image Dehazing", communicated and accepted at 5th International Conference on Advances in Computing, Communication Control and Networking(ICAC3N-23).

[2] D. Rawat, K. Singh, "A Comparative Study on Single Image Dehazing using Deep Learning-Based Techniques", communicated and accepted at 5th International Conference on Advances in Computing, Communication Control and Networking(ICAC3N-23).

Gmail Deepanker Rawat <deepanker.cse.rawat@gmail.com>

Acceptance Notification 5th IEEE ICAC3N-23 & Registration: Paper ID 605

2 messages

Microsoft CMT <email@msr-cmt.org> Sat, May 20, 2023 at 1:26 PM
Reply-To: Vishnu Sharma <vishnu.sharma@galgotiacollege.edu>
To: Deepanker Rawat <Deepanker.cse.rawat@gmail.com>

Dear Deepanker Rawat,
Delhi Technological University

Greetings from ICAC3N-23 ...!!!

Congratulations....!!!!!

On behalf of the 5th ICAC3N-23 Program Committee, we are delighted to inform you that the submission of "Paper ID- 605 " titled " FAG-Net: Feature Attention Generative Network for Single Image Dehazing " has been accepted for presentation and further publication with IEEE at the ICAC3N- 23 subject to incorporate the reviewers and editors comments in your final paper. All accepted papers will be submitted to IEEE for inclusion into conference proceedings to be published on IEEE Xplore Digital Library.

For early registration benefit please complete your registration by clicking on the following Link:
https://forms.gle/8e6RzNbho7CphnYN7 on or before 31 May 2023.

Registration fee details are available @ https://icac3n.in/register.
https://drive.google.com/file/d/1RGu6i4eGUI5B07zOfgRmDRfyjOhOhiC6/view?usp=share_link

You can also pay the registration fee by the UPI. (UPI id - icac3n@ybl) or follow the link below for QR code:
https://drive.google.com/file/d/1Ry-sF0apvy_0zUjM8lNW02lZgWAsE0YD/view?usp=sharing

You must incorporate following comments in your final paper submitted at the time of registration for consideration of publication with IEEE:

Reviewers Comments:
The theme and title of the article "FAG-Net: Feature Attention Generative Network for Single Image Dehazing" is suitable and appropriate for publication.
Paper is well written and technically sound.
Paper formatting is very good.
Overall flow of paper is very good.
English language is satisfactory. Though, spellings and grammar need to be checked throughout manuscript.

If you have any query regarding registration process or face any problem in making online payment, you can Contact @ 8168268768 (Call) / 9467482983 (Whatsapp/UPI) or write us at icac3n23@gmail.com.

Regards:
Organizing committee
ICAC3N – 2023

Deepanker Rawat <deepanker.cse.rawat@gmail.com> Sat, May 20, 2023 at 1:38 PM
To: kavinder85@gmail.com

---------- Forwarded message ---------
From: **Microsoft CMT** <email@msr-cmt.org>
Date: Sat, May 20, 2023 at 1:27 PM
Subject: Acceptance Notification 5th IEEE ICAC3N-23 & Registration: Paper ID 605
To: Deepanker Rawat <Deepanker.cse.rawat@gmail.com>

Gmail Deepanker Rawat <deepanker.cse.rawat@gmail.com>

Acceptance Notification 5th IEEE ICAC3N-23 & Registration: Paper ID 773

1 message

Microsoft CMT <email@msr-cmt.org> Wed, May 24, 2023 at 10:03 PM
Reply-To: Vishnu Sharma <vishnu.sharma@galgotiacollege.edu>
To: Deepanker Rawat <Deepanker.cse.rawat@gmail.com>

Dear Deepanker Rawat,
Delhi Technological University

Greetings from ICAC3N-23 ...!!!

Congratulations....!!!!!

On behalf of the 5th ICAC3N-23 Program Committee, we are delighted to inform you that the submission of "Paper ID- 773 " titled " A Comparative Study on Single Image Dehazing using Deep Learning-Based Techniques " has been accepted for presentation and further publication with IEEE at the ICAC3N- 23 subject to incorporate the reviewers and editors comments in your final paper. All accepted papers will be submitted to IEEE for inclusion into conference proceedings to be published on IEEE Xplore Digital Library.

For early registration benefit please complete your registration by clicking on the following Link:
https://forms.gle/8e6RzNbho7CphnYN7 on or before 31 May 2023.

Registration fee details are available @ https://icac3n.in/register.
https://drive.google.com/file/d/1RGu6i4eGUI5B07zOfgRmDRfyjOhOhiC6/view?usp=share_link

You can also pay the registration fee by the UPI. (UPI id - icac3n@ybl) or follow the link below for QR code:
https://drive.google.com/file/d/1Ry-sF0apvy_0zUjM8lNW02lZgWAsE0YD/view?usp=sharing

You must incorporate following comments in your final paper submitted at the time of registration for consideration of publication with IEEE:

Reviewers Comments:
The theme and title of the article "A Comparative Study on Single Image Dehazing using Deep Learning-Based Techniques" is suitable and appropriate for publication.
Paper is well written and technically sound.
Overall flow of paper is very good.
English language is satisfactory.

Editor Note:
1. All figures and equations in the paper must be clear. Equation and tables must be typed and should not be images.
2. Final camera ready copy must be strictly in IEEE format available on conference website www.icac3n.in.
3. Transfer of E-copyright to IEEE and Presenting paper in conference is compulsory for publication of paper in IEEE.
4. If plagiarism is found at any stage in your accepted paper, the registration will be cancelled and paper will be rejected and the authors will be responsible for any consequences. Plagiarism must be less then 20% (checked through Turnitin). However the author will be given fair and sufficient chance to reduce plagiarism.
5. Change in paper title, name of authors or affiliation of authors will not be allowed after registration of papers.
6. Violation of any of the above point may lead to rejection of your paper at any stage of publication.
7. Registration fee once paid will be non refundable.

If you have any query regarding registration process or face any problem in making online payment, you can Contact @ 8168268768 (Call) / 9467482983 (Whatsapp/UPI) or write us at icac3n23@gmail.com.

Regards:
Organizing committee
ICAC3N – 2023

turnitin

Similarity Report ID: oid:27535:36498590

PAPER NAME

Fusion_based_GAN_for_single_image_dehazing.pdf

WORD COUNT

9841 Words

CHARACTER COUNT

55134 Characters

PAGE COUNT

49 Pages

FILE SIZE

4.5MB

SUBMISSION DATE

May 30, 2023 1:33 PM GMT+5:30

REPORT DATE

May 30, 2023 1:34 PM GMT+5:30

15% Overall Similarity

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

- 9% Internet database
- 10% Publications database
- Crossref database
- Crossref Posted Content database
- 9% Submitted Works database

Excluded from Similarity Report

- Bibliographic material