Universidade Estadual de Campinas
Instituto de Computação

Percy Maldonado Quispe

Unsupervised Deep-Learning Method for Haze Removal Without Paired Images

Método Não Supervisionado de Aprendizado de Máquina Profundo para Remoção de Neblina Sem Imagens Emparelhadas

CAMPINAS
2024

Percy Maldonado Quispe

Unsupervised Deep-Learning Method for Haze Removal Without Paired Images

Método Não Supervisionado de Aprendizado de Máquina Profundo para Remoção de Neblina Sem Imagens Emparelhadas

Dissertação apresentada ao Instituto de Computação da Universidade Estadual de Campinas como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Dissertation presented to the Institute of Computing of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Computer Science.

Supervisor/Orientador: Prof. Dr. Hélio Pedrini

Este exemplar corresponde à versão final da Dissertação defendida por Percy Maldonado Quispe e orientada pelo Prof. Dr. Hélio Pedrini.

CAMPINAS
2024

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

M293u	Maldonado Quispe, Percy, 1997- Unsupervised deep-learning method for haze removal without paired images / Percy Maldonado Quispe. – Campinas, SP : [s.n.], 2024. Orientador: Hélio Pedrini. Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Computação. 1. Aprendizagem não-supervisionada (Aprendizado do computador). 2. Aprendizado Zero-shot. 3. Análise de imagem. 4. Visão por computador. 5. Redes neurais (Computação). I. Pedrini, Hélio, 1963-. II. Universidade Estadual de Campinas. Instituto de Computação. III. Título.

Informações Complementares

Título em outro idioma: Método não supervisionado de aprendizado de máquina profundo para remoção de neblina sem imagens emparelhadas
Palavras-chave em inglês:
Unsupervised learning
Zero-shot learning
Image analysis
Computer vision
Neural networks (Computer science)
Área de concentração: Ciência da Computação
Titulação: Mestre em Ciência da Computação
Banca examinadora:
Hélio Pedrini [Orientador]
César Armando Beltrán Castañón
Helena de Almeida Maia
Data de defesa: 22-03-2024
Programa de Pós-Graduação: Ciência da Computação

Identificação e informações acadêmicas do(a) aluno(a)
- ORCID do autor: https://orcid.org/0009-0007-6391-5370
- Currículo Lattes do autor: http://lattes.cnpq.br/3360189827691469

Universidade Estadual de Campinas
Instituto de Computação

Percy Maldonado Quispe

Unsupervised Deep-Learning Method for Haze Removal Without Paired Images

Método Não Supervisionado de Aprendizado de Máquina Profundo para Remoção de Neblina Sem Imagens Emparelhadas

Banca Examinadora:

- Prof. Dr. Hélio Pedrini
  Instituto de Computação, Universidade Estadual de Campinas (UNICAMP)
- Dra. Helena de Almeida Maia
  Instituto de Computação, Universidade Estadual de Campinas (UNICAMP)
- Prof. Dr. César Armando Beltrán Castañón
  Ingeniería Informática, Pontificia Universidad Católica del Perú (PUCP)

A ata da defesa, assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Campinas, 22 de março de 2024

Vita brevis, ars longa, occasio praeceps,
experimentum periculosum,
iudicium difficile.

(Hippocrates)

Acknowledgments

- First of all, I would like to thank God for guiding me during these two years of study.
- I would like to thank my parents and grandparents for their unconditional love and support given all this time, despite the distance I always felt close and present at this stage of my life.
- I am especially grateful to my advisor Prof. Dr. Hélio Pedrini, for affording me the opportunity to be his student. In addition, his help was very important to finish this research work, since with his knowledge and patience they showed me many times the way to follow with this M.Sc. dissertation.
- A special acknowledgment goes to my dear friend, partner, and girlfriend, Jesamin Zevallos, whose love, patience, and understanding were instrumental during this phase of my life. Her emotional support played a crucial role in the successful execution of this research.
- I extend my thanks to the Institute of Computing at the University of Campinas (IC-UNICAMP) for opening its doors and providing me with the opportunity to advance in my professional career.
- This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Resumo

Neste estudo, abordamos um paradigma fundamental e ainda relativamente pouco explorado no campo das redes neurais artificiais para o desembaçamento não supervisionado de imagens. Ao conceber uma imagem nebulosa pela superposição de várias camadas "mais simples", tais como uma camada de imagem livre de neblina, uma camada de mapa de transmissão e uma camada de luz atmosférica, inspirada no modelo de espalhamento atmosférico, propomos uma abordagem baseada no conceito de desentrelaçamento de camadas. Nosso método, denominado XYZ, representa uma melhora substancial nas métricas de qualidade de imagem, como SSIM e PSNR, bem como BRISQUE, PIQE e NIQE. Este avanço é alcançado por meio da combinação estratégica dos métodos XHOT, YOLY e ZID, capitalizando os pontos fortes individuais de cada um. Um aspecto distintivo e valioso da abordagem XYZ é a sua natureza não supervisionada, o que implica que ela não depende de conjuntos de dados contendo pares de imagens nítidas e desfocadas para treinamento. Isto contrasta com o paradigma tradicional de treinamento profundo, marcando uma inovação no campo da remoção de névoa. Além disso, destacamos dois benefícios fundamentais da abordagem XYZ proposta. Em primeiro lugar, por não ser supervisionada, evita a necessidade de utilizar conjuntos de dados exaustivos que incluem imagens nítidas e desfocadas como referência fundamental. Em segundo lugar, abordamos a questão da neblina a partir de uma perspectiva multifacetada, reconhecendo e desvendando as complexidades inerentes a este fenômeno atmosférico. Esta abordagem em camadas permite uma representação mais precisa e detalhada da cena, melhorando assim a qualidade das imagens sem neblina. Resultados experimentais obtidos para a conjunto de dados RESIDE são comparados com outros métodos da literatura.

Abstract

In this study, we address a fundamental and still relatively less explored learning paradigm in the field of neural networks for image dehazing: the unsupervised dehazing of an image. By conceiving a hazy image as the superposition of several “simpler” layers, such as a haze-free image layer, a transmission map layer, and an atmospheric light layer, inspired by the atmospheric scattering model, we propose an approach based on the concept of layer disentangling. Our method, called XYZ, represents a substantial improvement in image quality metrics, such as SSIM and PSNR as well as BRISQUE, PIQE and NIQE. This advancement is achieved through the strategic combination of the XHOT, YOLY and ZID methods, capitalizing on the individual strengths of each. A distinctive and valuable aspect of the XYZ approach is its unsupervised nature, which implies that it does not rely on data sets containing pairs of clear and hazy images for training. This contrasts with the traditional deep training paradigm, marking an innovation in the field of dehazing. Furthermore, we highlight two fundamental benefits of the proposed XYZ approach. Firstly, being unsupervised, it frees the process from the need to use exhaustive datasets that include clear and hazy images as a fundamental reference. Secondly, we approach the haze issue from a multi-layered perspective, recognizing and unraveling the complexities inherent to this atmospheric phenomenon. This layered approach allows for a more accurate and detailed representation of the scene, thereby improving the quality of haze-free images. Experimental results obtained for the RESIDE dataset are compared with other methods from the literature.

List of Figures

List of Tables

List of Abbreviations and Acronyms

AOD-Net	All-in-One Dehazing Network
ASM	Atmospheric Scattering Model
BRISQUE	Blind/Referenceless Image Spatial Quality Evaluator
CAP	Color Attenuation Prior
CNN	Convolutional Neural Networks
CVPR	Computer Vision and Pattern Recognition
DCP	Dark Channel Prior
FFA	Feature Fusion Attention
GAN	Generative Adversarial Networks
GPU	Graphics Processing Unit
HSTS	Hybrid Subjective Testing Set
HSV	Hue, Saturation, Value
IDBP	Image Dehazing Using Blended Priors
MAFFNet	Multi-Scale Attentive Feature Fusion Network
MSAFF-Net	Multi-Scale Attention Feature Fusion Network
MSE	Mean Squared Error
NIQE	Naturalness Image Quality Evaluator
PIQE	Perception-based Image Quality Evaluator
PSNR	Peak Signal to Noise Ratio
RESIDE	Realistic Single Image Dehazing
RGB	Red, Green, Blue
SOTS	Synthetic Objective Testing Set
SSIM	Structural Similarity Index Measure
VAE	Variational Autoencoder
XHOT	X-Zero Shot Method
XYZ	XHOT + YOLY + ZID (Our Haze Removal Method)
YOLY	You Only Look Yourself Method
ZID	Zero-Shot Image Dehazing
ZSL	Zero-Shot Learning Method

Contents

Chapter 1

Introduction

In challenging weather situations, such as those affected by fog and haze, image quality suffers significant degradation due to the influence of particles suspended in the atmosphere. These particles scatter light, attenuating the intensity of the light reflected in the scene. The scattered atmospheric light mixes with the light captured by the camera, generating noticeable changes in the contrast and color of the image.

The absorption caused by particles suspended in the environment directly affects the quality of the images captured. This phenomenon, evident on hazy days, has an adverse impact on photographic practice. The contrast of the image decreases, and the colors undergo perceptible alterations. Simultaneously, the textures and contours of objects in the scene become blurred. Figure 1.1 illustrates the disparity in pixel histograms between images with and without the presence of haze.

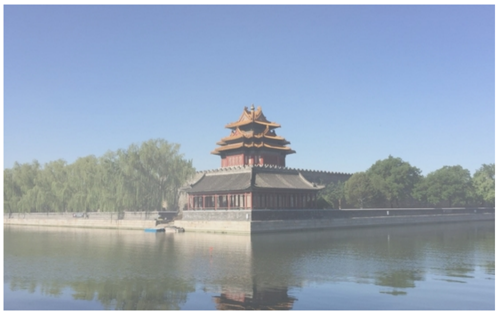
This degradation in image quality not only constitutes a visual hindrance, but also negatively impacts fundamental computer vision tasks such as object detection and image segmentation. In this context, low-quality inputs can significantly degrade the performance of meticulously designed models, highlighting the urgent need for effective dehazing solutions in image processing.

In the critical task of estimating global atmospheric light and per-pixel transmission coefficients in atmospheric dispersion models, the haze removal community has developed several methodologies. These can be classified into two main categories: assumption-based prior methods and learning-based methods.

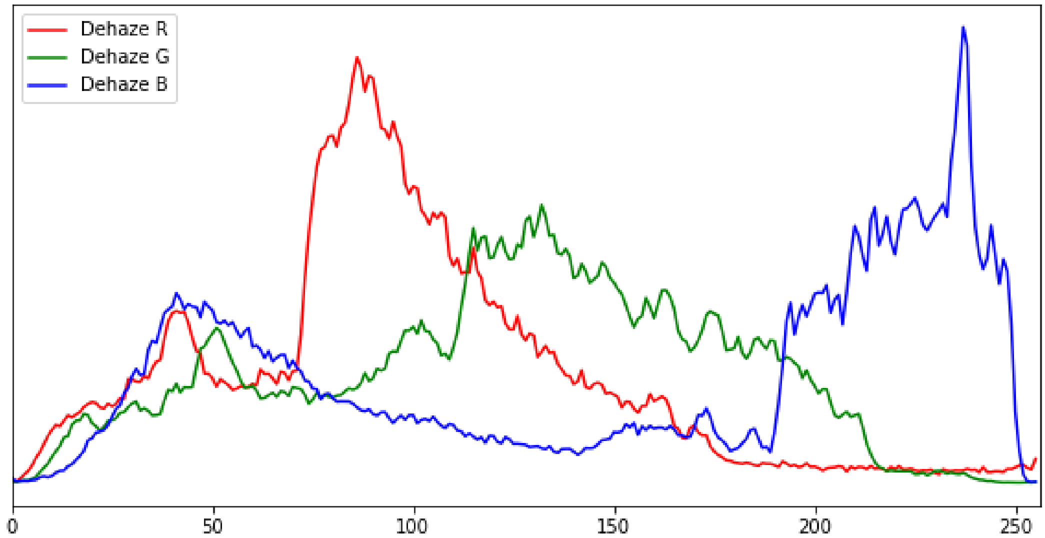
Previous methods, based on assumptions derived from images, have proven effective in certain contexts. For example, Tan [48] proposed to maximize the local contrast of the image to eliminate haze, based on the premise that clean images tend to exhibit higher contrast. In a similar vein, Berman et al. [3] focused on dehazing, assuming that the colors in a dehazed image can be approximated by a reduced set of colors. Despite the notable achievements of these methods, the quality of haze removal largely depends on the agreement between the data distribution used and reality.

To overcome the limitations imposed by prior assumptions, a promising approach has been the application of deep neural networks. These methods not only detect and remove haze, but also directly learn atmospheric dispersion parameters from training data. A significant example is the work developed by Cai et al. [6], who proposed a convolutional neural network designed for this purpose. However, it is crucial to note

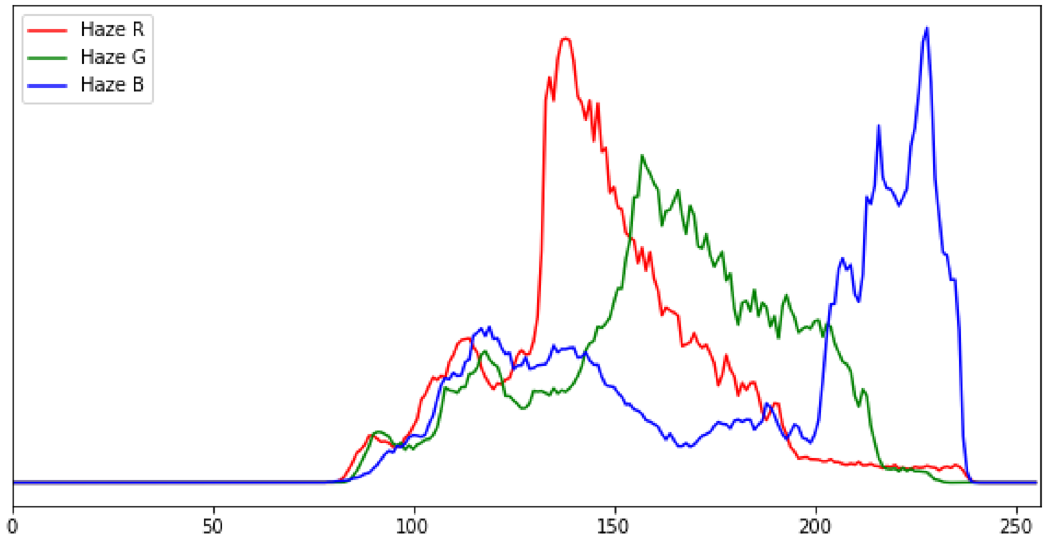
(a) Clear image



(b) Haze image



(c) Histogram of clear image



(d) Histogram of haze image

Figure 1.1: Histograms for clear and haze images. This figure illustrates the difference in pixel histograms between images with and without the presence of haze.

that such methods, by requiring an extensive set of clean and haze image pairs, fall into the category of supervised learning.

This shift toward deep neural network-based approaches presents considerable potential to improve the ability to remove haze more accurately and widely. As we evolve toward more complex models and more advanced training techniques, the door opens to solutions that can adapt more robustly to various atmospheric conditions and complexities. The fusion of knowledge derived from previous methods and the expressive power of neural networks offers a promising horizon to address significant challenges in the field of haze removal.

1.1 Context and Motivation

The massive proliferation of digital cameras both in the consumer market and in specialized detection systems has accentuated the relevance of dehazing in outdoor images. This process has evolved to become an essential component in various scientific fields and cutting-edge applications, from astronomy and medical sciences to surveillance, agronomy, archeology and environmental studies, as you can see some examples in Figure 1.2.

Figure 1.2: Applications with use of images. Applications range from astronomy and medicine to autonomous vehicles and security cameras, among others.

Visual data, fundamental to the understanding and analysis of the human brain, occupy a central position in the hierarchy of perception. About one-third of the cortical area of the human brain is dedicated exclusively to processing visual data [13]. Consequently, the sharpness and clarity of images become crucial aspects for various tasks in the field of image processing, exercising a direct influence on the quality of the interpreted information.

In the specific context of vehicular systems, the ability of cameras to generate clear images becomes of outstanding importance, especially in adverse weather conditions. Visibility affected by haze can seriously compromise the safety and effectiveness of these systems, underscoring the urgency of advanced haze removal solutions to ensure optimal performance in a wide range of challenging atmospheric conditions.

In this scenario, research and development in haze removal techniques through the use of artificial intelligence, specifically through unsupervised neural networks, is placed at the vanguard. This approach not only seeks to improve the visual quality of images. In this convergence between advanced technology and fundamental practical needs, a vast field opens up to explore and refine methods that not only transform the way we interpret and use visual data, but also set standards in haze removal in challenging environments.

The combination of artificial intelligence and unsupervised neural networks in this context promises a significant advance in improving the quality of images and the adaptability of vision systems in adverse atmospheric conditions.

1.2 Problem Characterization

Although supervised learning methods have achieved remarkable levels of performance in haze removal, they face significant challenges, one of them being the need for large pairs of haze and clean images for training. This requirement is usually met by artificially synthesizing haze images through physical models with predetermined parameters and the corresponding clean image. However, as Golts et al. [11] pointed out, synthesized databases tend to lack the informational richness and consistency present in real datasets. This gap raises the urgency of developing unsupervised approaches.

In practice, obtaining a large-scale dataset with the desired ground truth becomes a formidable challenge due to variations in the scene and other factors such as lighting. Consequently, most methods choose to initially collect clean images and then synthesize the corresponding blurred images using an atmospheric dispersion model with predefined parameters. However, when applying models trained on synthetic datasets to real-world blurred photos, the challenge of domain switching arises, as synthetic blurred images may lack the consistency and information needed to address real-world situations. Against this backdrop, the creation of an deep neural network is presented as a key solution to overcome the mentioned problems.

The development of a deep neural network capable of simultaneously addressing the challenges of operating unsupervised and without the need for extensive training is envisioned. This direction, so far little explored, seeks not only to mitigate the limitations derived from supervised methods, but also to address the specific problems of domain switching and lack of coherence in synthetic data. Progress in this direction will not only improve the effectiveness of haze removal, but will also significantly contribute to the autonomy and adaptability of these models in real-world environments, marking a shift in the convergence of unsupervised neural networks and haze removal in images.

1.3 Main Challenges

Haze removal using unsupervised neural networks presents several key challenges that require careful attention to achieve effective results. Some of the main challenges include:

- Lack of Labeled Datasets: Unlike supervised approaches, where pairs of haze and unhazed images are required for training, unsupervised neural networks must learn directly from unlabeled data. This poses a fundamental challenge of not having ground truth to guide the training process.

- Varied Atmospheric Complexity: Atmospheric conditions can vary significantly, from light haze to dense haze situations. The ability of an unsupervised neural network to adapt and generalize effectively to these variations is a significant challenge, as haze features can be highly complex and contextual.

- Data Heterogeneity: Visual data can be highly heterogeneous in terms of settings, lighting, and weather conditions. The network must be able to handle the diversity of input data and learn relevant patterns that enable effective dehazing in a wide range of environments.

- Unsupervised Learning Transfer: Achieving effective learning transfer in unsupervised environments is a challenge. The network must learn useful features without relying on explicit supervision, which requires strategies to capture relevant information without a dataset specifically labeled for the dehazing task.

- Evaluate Removal Quality: Effectively measuring the quality of haze removal without the existence of labeled pairs of clear and haze images can be difficult. Objective and subjective metrics are needed that evaluate the visual fidelity and perceptual improvement of the network without relying on ground truth.

- Robustness to Unusual Conditions: Unsupervised neural networks must be robust to extreme or unusual atmospheric conditions. The ability to deal with unpredictable situations, such as heavy haze or sudden changes in lighting, is an essential challenge for practical applicability.

Addressing these challenges effectively is essential to advance the successful application of unsupervised neural networks in dehazing and improving the visual quality of images in challenging environments.

1.4 Objectives and Research Questions

The fundamental objectives of this work focus on proposing an approach based on unsupervised learning, recognizing the intrinsic need for this approach due to the scarcity of pairs of real images with and without haze. Additionally, a comprehensive comparative analysis of a diverse set of haze removal methods is sought to be carried out. This analysis will not only address the individual capabilities of each method, but will also explore the advantages and disadvantages of their application both individually and in combination.

Specific Objectives

To achieve the main objectives of this work it is necessary to achieve some specific objectives:

1. Development of a Robust Unsupervised Method: Design and develop an unsupervised learning-based dehazing method that can learn autonomously from unlabeled data. This involves creating an unsupervised neural network model capable of capturing the inherent complexities of haze images.

2. Comparative Analysis: Conduct a detailed comparative analysis between the proposed method and a variety of existing approaches for haze removal. Evaluate performance in terms of visual quality, effectiveness and computational efficiency, considering relevant metrics for each aspect.

3. Combination of Methods: Investigate and analyze how the combination of different methods can enhance or mitigate their respective limitations. Explore synergies and possible improvements by applying multiple approaches together.

Ultimately, the success of this work will depend on the ability to comprehensively address these specific objectives, thereby providing a valuable contribution to the field of haze removal using unsupervised learning-based approaches.

Research Questions

In this section, we present some research questions that motivated our thesis proposal:

1. How can existing unsupervised dehazing models be improved and optimized using advanced deep learning techniques?

2. What is the impact of neural network architecture on the performance of unsupervised dehazing models?

3. How do accurate and relevant evaluation metrics influence the performance of unsupervised haze removal models in real-world conditions?

1.5 Contributions

As fundamental contributions, we propose a deep learning model aimed at removing haze in images, addressing both outdoor and indoor environments. Our approach is based on the use of unsupervised neural networks, which is an approach to overcome the limitations associated with the reliance on paired datasets. This strategic approach is based on the need to address real-life conditions, where obtaining pairs of clear and haze images is impractical and, in many cases, extremely tedious.

Additionally, we plan to conduct extensive experimental evaluation using challenging dataset, captured in real-world environments. This evaluation will not only measure the effectiveness of our approach in removing haze under diverse conditions, but will also support the robustness and generalizability of the model to varied and challenging scenarios. This experimental analysis will become an essential component to validate and demonstrate the effectiveness of our model in real-world situations, thus consolidating its contributions in the field of haze removal.

1.6 Publications

The following papers have been published from the development of this research work:

- **P. Maldonado-Quispe**, H. Pedrini. *Image Dehazing Using a Simple Convolutional Autoencoder.* 9th IEEE Latin American Conference on Computational Intelligence (LA-CCI). Recife-PE, Brazil, pp. 1-6, October 29 - November 01, 2023.
- **P. Maldonado-Quispe**, H. Pedrini. *XYZ Unsupervised Network: A Robust Image Dehazing Approach.* 19th International Conference on Computer Vision Theory and Applications (VISAPP). Rome, Italy, pp. 500-507, February 27-29, 2024.

1.7 Text Organization

Chapter 2 presents the fundamental concepts that serve as the basis for our research. Additionally, we address related work that is of relevance to our research and perform a comparison in three areas of study linked to our work: (i) *Assumptions*, we examine the existing literature on methods that employ prior information or assumptions for haze removal; (ii) *Supervised Learning*, we review the literature related to methods using supervised neural networks, which implies the need for paired images for haze removal; (iii) *Unsupervised Learning*, we explore the literature related to dehazing methods that dispense with paired images, that is, those that only require the image with the presence of haze.

Chapter 3 describes in detail our proposed methodology, which is the XYZ unsupervised approach [29]. This approach is distinguished by not requiring an extensive dataset of paired data for subsequent training and evaluation. The XYZ methodology comprises the combination of three unsupervised neural networks (XHOT [29], YOLY [25] and ZID [24]), capitalizing on the particular advantages of each of them. Furthermore, we provide detailed information about the dataset used to evaluate our approach. In this case, we use the RESIDE dataset, thus ensuring a reliable comparison with various existing methods in the literature.

Chapter 4 presents the results obtained by applying our methodology to a widely used dataset. We carry out a comprehensive evaluation that covers both quantitative and qualitative aspects. For quantitative evaluation, we employ Full-Reference metrics such as PSNR and SSIM, as well as Non-Reference metrics such as BRISQUE, NIQE, and PIQE, given the unsupervised nature of the problem. Additionally, a detailed discussion on the visual results for the qualitative evaluation is included. For a more rigorous quantitative evaluation, metrics are presented that compare our results to leading state-of-the-art methods.

Chapter 5 presents our contributions and conclusions from our research work. In addition, we describe some directions for future studies. Finally, Appendix A includes some relevant information about the RESIDE dataset.

Chapter 2

Background

The purpose of this chapter is to provide the essential conceptual foundations and review relevant related works in the context of this research, as well as the key methods used in its development.

Section 2.1 describes the fundamental concepts in the haze removal process. Some concepts present in this section are Atmospheric Scattering Model (ASM), Dark Channel Prior (DCP), Color Attenuation Prior (CAP), among other.

Section 2.2 provides a comprehensive introduction to Machine Learning and Deep Learning and their specific application in the field of haze removal. The unsupervised approach is highlighted, where the need for paired or labeled data becomes unnecessary. Likewise, we focus especially on Zero-Shot Learning, a Deep Learning paradigm that is characterized by requiring a minimum amount of data and even just one piece of data for the training of unsupervised neural networks.

Section 2.3 presents a classification of related works into three different categories: (i) Assumption, (ii) Supervised Learning and (iii) Unsupervised Learning. Assumption-based methods use prior information present in the image, such as color or saturation, among other visible characteristics. Moreover, we perform a detailed analysis and comparison of methods using supervised neural networks, which imperatively require paired images (with hazy and without hazy). On the other hand, we explore unsupervised methods, which dispense with paired data, alleviating the need for paired datasets, which in many cases are impractical to obtain.

This section also introduces and analyzes the methods YOLY [25] and ZID [24]. The essence of both approaches and how they are applied in the haze removal process is explored. It is important to note that both YOLY and ZID are distinguished by not depending on extensive datasets, since they are based on the philosophy of Zero-Shot Learning.

Section 2.4 highlights the relevance of the REalistic Single Image DEhazing (RESIDE) [23] dataset, widely recognized for its quality in evaluating large-scale haze removal methods. This dataset is characterized by its diversity and is divided into two test subsets to address different environments and challenges. The Synthetic Objective Testing Set (SOTS) subset, consisting of 500 images with haze, focuses on indoor environments, while the Hybrid Subjective Testing Set (HSTS) subset is designed for outdoor evaluation, including 10 images of synthetic haze and 10 images that simulate a realistic approach.

This diversification allows for a more thorough and robust assessment of haze removal methods under a variety of conditions.

2.1 Fundamental Concepts

In this section, we describe fundamental concepts in the process of haze removal, as well as the relevant characteristics present in it.

2.1.1 Dehazing Process

Haze is a natural phenomenon (Figure 2.1) that can be roughly explained by the ASM. McCartney [30] proposed a basic ASM to explain the principle of haze formation. Narasimhan and Nayar [34] and Nayar and Narasimhan [35] extended and developed ASM, which is widely used today. ASM provides a solid theoretical foundation for imaging range studies. The formula is:

$$I(x) = J(x)\,t(x) + A(1 - t(x)) \tag{2.1}$$

where x is the coordinate value and A stands for global atmospheric light. In various documents, A is sometimes referred to as air light or ambient light. In this study, A is given as atmospheric light. For ASM-based haze removal methods, A is generally unknown. $I(x)$ represents a haze image and $J(x)$ represents a clean scene. For most common models, $I(x)$ is the input and $J(x)$ is the desired output. Moreover, $t(x)$ represents the average transmission map.

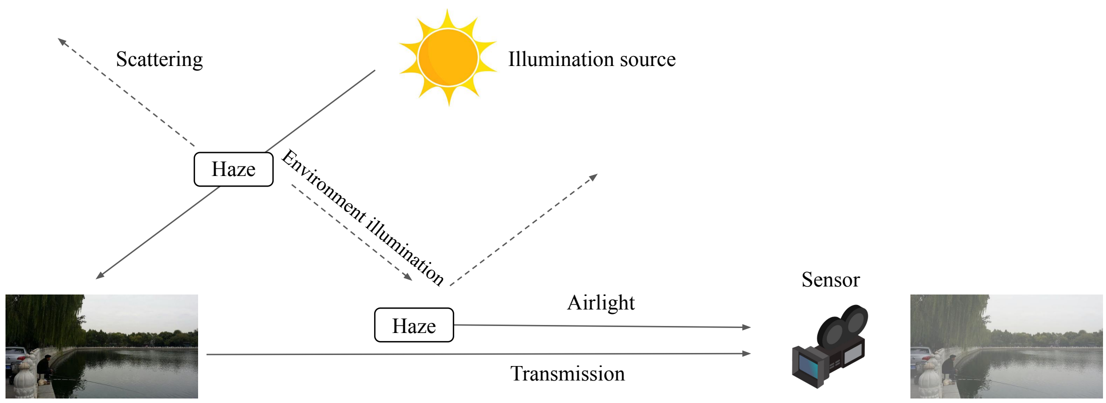


Figure 2.1: Atmospheric Scattering Model (ASM). Extracted from the work developed by McCartney [30].

2.1.2 Haze Features

In the context of haze removal, it is essential to understand the key characteristics associated with the haze phenomenon. Two pertinent features are outlined as follows:

1. Dark Channel Prior (DCP) is based on extensive study of haze-free outdoor images. Most of the haze-free patches contain at least one color channel with some pixels

with very low, almost zero intensity values. The smallest color in a particular area of pixels is called the dark channel [15]:

$$D(x) = \min_{y \in \Omega_r(x)} \left(\min_{c \in \{r,g,b\}} I^c(y) \right), \tag{2.2}$$

where I^c is an RGB color channel of I and $\Omega_r(x)$ is a local patch centered at x with the size of $r \times r$. The dark channel feature has a high correlation to the amount of haze in the haze image, and is used to estimate the transmission map directly.

2. Color Attenuation Prior (CAP) arises from the observation pointed out by Zhu et al. [60], where they infer that the effect of white light on the observed values is additive. In the case of atmospheric light, an increase in image brightness is observed at the expense of a decrease in saturation. This background information provides an opportunity to use the difference between brightness and saturation to estimate haze concentration. Since the concentration of the haze increases as the depth of the scene changes, there is a positive correlation between the depth of the scene and the concentration of the haze. This relationship can be expressed as follows:

$$d(x) = \theta_0 + \theta_1 V(x) + \theta_2 S(x) + \epsilon(x), \tag{2.3}$$

where x represents the position within the image, d corresponds to the depth of the scene, V is the component of brightness of the image with haze, S is the component of saturation, θ_0, θ_1 and θ_2 are the unknown linear coefficients and finally $\epsilon(x)$ represents a random error of the model.

2.2 Machine Learning

In the field of image processing, Machine Learning [4, 5, 33] represents a crucial discipline that has revolutionized the way we approach the improvement and interpretation of visual content. This approach stands as a powerful tool that allows algorithms to learn intrinsic patterns and complex relationships in large data sets, making it easier to automate previously challenging tasks.

In essence, Machine Learning is a branch of artificial intelligence that enables machines to learn without direct human intervention. Instead of being programmed with specific rules, machines use algorithms that adjust and evolve, refining their performance as they are exposed to more data. This paradigm is particularly relevant in image processing, where visual complexity and variability of conditions require adaptive approaches.

Within the image processing, Machine Learning plays a crucial role in tackling challenging tasks such as dehazing. In contexts where the presence of haze significantly degrades image quality, Machine Learning models can learn to discern specific patterns associated with haze and apply corrections automatically.

There are two main approaches in applying Machine Learning to image processing [7, 43, 44]. The first, known as supervised learning, involves training models using labeled datasets, which contain pairs of images with and without haze. This allows models to learn the relationship between visual characteristics and the presence of haze, extrapolating that knowledge to new images.

On the other hand, unsupervised learning excels in situations where obtaining pairs of images is difficult or impractical in real-world environments. Unsupervised learning models can learn directly from unlabeled data, capturing subtle patterns without relying on ground truth. This approach is especially valuable in image processing under challenging atmospheric conditions, such as haze.

Therefore, the integration of Machine Learning into image processing represents a significant evolution, enabling notable advances in dehazing and other complex visual tasks. The ability of models to learn and adapt to the variability of atmospheric conditions contributes to a substantial improvement in the visual quality of images, making this discipline a cornerstone in Haze Removal research.

2.2.1 Deep Learning

Deep Learning [12, 14, 16, 47] has significantly transformed the image processing landscape, standing out for its inherent ability to learn complex representations and perform tasks automatically from raw data. This approach is based on deep neural network architectures, such as Convolutional Neural Networks (CNN) and Generative Adversarial Networks (GAN), which have proven effective in various applications, including haze removal.

Deep neural networks are capable of learning high-level patterns and features in visual data, allowing us to address the haze problem more effectively than traditional methods. In the specific context of haze removal, deep learning offers a substantial advantage by allowing models to automatically and hierarchically learn the complexities associated with adverse atmospheric conditions.

In the words of Ren et al. [40], “deep learning models can automatically and efficiently capture the intrinsic features of images affected by haze, thereby improving the detail recovery ability under challenging atmospheric conditions”. The ability of these networks to discern spatial patterns and discriminative features gives them the ability to decompose image information, separating light transmission and atmospheric light.

It is essential to highlight that the unsupervised approach to deep learning is particularly valuable in the context of haze removal. Unlike supervised methods that require pairs of clear and haze images for training, deep learning models can generalize to various atmospheric conditions without relying on specific datasets.

Ren et al. [40] highlighted this advantage by noting that “deep learning's ability to generalize to various atmospheric conditions makes it particularly valuable in real-world situations, where conditions can vary widely”.

In summary, Deep Learning is presented as an essential tool in haze removal, providing a unique ability to automatically and efficiently learn the characteristics associated with adverse atmospheric conditions. Its unsupervised approach and its ability to discern complex visual patterns position it as a key technology for improving the visual quality of images affected by haze.

2.2.2 Unsupervised Learning

In the field of image processing [17, 37, 38, 41], Unsupervised Learning is emerging as an innovative paradigm that revolutionizes the way algorithms can address visual complexity without relying on explicit labeled data. As Bishop [4] and Goodfellow et al. [12] highlight, this approach is distinguished by its ability to learn intrinsic patterns and underlying structures directly from unlabeled datasets.

Unsupervised learning is presented as a valuable alternative in image processing, particularly in scenarios where obtaining labeled image pairs is challenging. Rather than relying on ground truth, unsupervised learning models [2] have the ability to extract subtle patterns from unlabeled data, which is essential for realistic situations where the availability of annotated data is limited.

Neural networks, in particular, have proven to be effective tools in the application of unsupervised learning in image processing. As LeCun et al. [21] highlighted, architectures such as GANs have revolutionized the generation of realistic images without the need for explicit labels. In the context of haze removal, this implies that models can directly learn the complexities of atmospheric dispersion and apply corrections autonomously.

By avoiding the need for labels, unsupervised learning in image processing, as addressed by Goodfellow et al. [12], aligns perfectly with the challenging nature of atmospheric phenomena such as haze. The ability of these models to capture relevant visual features without reliance on previously labeled data opens new possibilities in improving visual quality in real-world situations.

Unsupervised learning represents a revolutionary approach in image processing, especially dehazing. Its ability to learn autonomously and adaptively without requiring annotated information offers a promising path to advance the visual quality of images affected by challenging atmospheric conditions.

2.2.3 Zero-Shot Learning

Zero-Shot Learning (ZSL) [46, 51] in image processing represents an advanced paradigm that challenges traditional machine learning conventions. This approach allows models to generalize and recognize classes without requiring specific training examples, which is essential in situations where the availability of annotated data is limited.

In essence, ZSL is based on the transfer of knowledge from previously known classes to new classes during inference. This is achieved through techniques such as semantic representation, where visual concepts are mapped to semantic vectors, allowing models to generalize to classes not seen during training [36].

The applications of ZSL in image processing are diverse and promising. Lampert et al. [20] highlighted its usefulness in scenarios where the identification of new classes is essential, such as in challenging atmospheric conditions where the haze can introduce unexpected visual phenomena. The ability of ZSL models to adapt to classes not considered during training becomes a valuable resource in changing and challenging contexts.

In conclusion, ZSL emerges as a transformative approach in image processing, providing an innovative response to the need to generalize to new classes without specific training data. Its ability to transfer knowledge, as evidenced by some research works [20, 36, 46], offers a unique and adaptable perspective in solving complex visual challenges, such as dehazing.

2.3 Related Work

Many researchers are trying to recover sharp, high-quality scenes from haze images. Table 2.1 presents a classification of the outstanding approaches in haze removal, according to three main features: Assumptions, Supervised and Unsupervised Learning. Before the widespread use of deep learning in computer vision tasks, haze removal algorithms were mainly based on some previous assumptions [15] and the ASM [30]. The processing sequences of these rule-based statistical methods are easy to interpret. However, they can fail when faced with complex real-world scenarios. For example, the popular dark channel before [15] (DCP, CVPR 2009 Best Paper) does not handle empty regions well.

The works [6, 27, 28, 58] are inspired by deep learning and combines ASM with CNN to estimate physical parameters. Quantitative and qualitative experimental results show that deep learning can help predict physical parameters in a supervised manner. Wang et al. [50], on the other hand, propose the use of an attention-convolutional module.

Due to advancements in deep learning and convolutional neural networks, numerous neural network methodologies have emerged for haze removal. A majority of these approaches rely on paired samples to facilitate supervised learning during network training. For instance, DehazeNet[6] is an end-to-end haze removal model that estimates the transmission map, estimating a clear image from a hazy input. AOD-Net [22] integrates the ASM with deep learning, simplifying calculations and reducing the number of variables for restoring clear images. MSBDN-DFF [9] adopts a U-Net architecture with dense feature fusion for multi-scale enhanced dehazing, addressing spatial information preservation issues in U-Net. FFA [39] expands CNN capabilities by applying attention mechanisms to different features and pixels. MAFFNet [56] leverages multi-scale attention feature fusion, utilizing U-Net features to transmit shallow information to deep features while assigning varying weights through pixel and channel attention modules.

MSAFF-Net [26] is a compact multi-scale attention feature fusion network designed for end-to-end single image haze removal, considering regions with haze-related features through channel and multi-scale spatial attention modules. TSDN [55] adopts a two-step dehazing network with intra-domain and constrained inter-domain adaptations, subdividing synthetic domain distributions into subsets and identifying optimal subsets via loss-based supervision. DCNet [54] employs Net-U and Net-D sub-networks to progressively obtain a haze-free image in a coarse-to-fine manner.

Following this, Liu et al. [27] and Zheng et al. [58] demonstrated that end-to-end supervised dehazing networks can be implemented independently of the ASM. Thanks to the powerful feature extraction capability of CNN, these non-ASM-based dehazing algorithms can achieve similar accuracy as ASM-based algorithms.

ASM-based and non-ASM-based supervised algorithms have shown impressive performance. However, they often require synthetic paired images that are inconsistent with real-world haze images. While these supervised methods have shown competitive results

Table 2.1: Haze removal methods available in the literature, categorized by the type of approach employed.

Methods	Main idea	Type
Tan [48]	Maximizing the local contrast	Assumption
Berman et al. [3]	Colors could approximate a few colors	
He et al. [15]	Estimating the ASM	
Cai et al. [6]	Learning of $t(x)$	Supervised Learning
Zhang and Patel [57]	Learning $t(x)$ and A	
Zhu et al. [59]	Generative Adversarial Network	
Yang et al. [53]	Generative Adversarial Network	
Liu et al. [27]	Convolutional Neural Network	
Zheng et al. [58]	Convolutional Neural Network	
Maldonado-Quispe and Pedrini [28]	Convolutional Neural Network	
Wang et al. [50]	Attention Convolutional Module	
Zhang and Wu [56]	Multi-scale Attention Feature Fusion	
Lin et al. [26]	Multi-scale Attention Feature Fusion	
Ju et al. [18]	Multiple Prior Constrain	
Engin et al. [10]	Unsupervised domain transfer	Unsupervised Learning
Golts et al. [11]	Learning without haze-free images	
Chen et al. [8]	Generative Adversarial Networks	
Wang et al. [52]	Generative Adversarial Networks	
Li et al. [24]	Zero-Shot learning	
Li et al. [25]	Zero-Shot learning	

on synthetic datasets based on objective metrics, a significant challenge arises when transitioning these models to real-world scenarios. Acquiring paired hazy/clear image samples in real-world conditions is nearly impossible, leading to suboptimal performance when applied to actual haze images in practical applications. This limitation has fueled the exploration of unsupervised approaches [10, 11, 24, 25], aiming to enhance the adaptability and effectiveness of dehazing models across diverse and challenging environments.

A noteworthy approach involves introducing constraints at different algorithmic stages using ASMs. Zhu et al. [59] pioneered the integration of GANs into a dehazing method based on the ASM. This novel implementation introduces new constraints through a discriminator, breaking away from the conventional reliance on paired samples. Disent-GAN [53] generates cyclic images based on the ASM model, incorporating cycle losses to achieve dehazing. Due to the asymmetry in features between hazy and clear images, defining their relationship accurately using weak constraints alone poses challenges for GAN-based methods.

Furthermore, Chen et al. [8] designed a method based on two stages, which eliminates haze using DCP and subsequently optimizes the results using existing features between the transmission and depth map. Wang et al. [52] proposed a method completely independent of real haze-free images. However, GAN-based networks have great complexity. Agrawal and Jalal [1] proposed a dehazing method that combines superpixel transformations with nonlinear transformations. Superpixels enhance the accuracy of transmission map estimation, reducing artifacts and better preserving the image structure.

The authors of IDBP [18] developed a robust haze removal technique based on the ASM model. This technique includes an atmospheric light estimation module and a multi-prior constraint module that integrates non-local/local priors and global constraints. In these methods, the parameters of the ASM model are reconstructed in the image domain, reducing the convergence of different information. However, it is important to note that the estimation error resulting from assumptions in the ASM model remains an inevitable challenge in these approaches.

Data-driven unsupervised dehazing methods have achieved impressive performance. Unlike models that require sufficient data to perform network training, Li et al. [24] proposed a neural network dehazing process that only requires a single example. They further reduced the dependence of the parameter learning process on data by combining the advantages of unsupervised learning and zero-shot learning.

Methods for unsupervised dehazing that are data-driven have demonstrated outstanding results. In the approaches ZID and YOLY developed by Li et al. [24] and Li et al. [25] respectively, a dehazing based on neural networks uses a single example, in contrast to methods that require sufficient data to perform network training. By combining the benefits of unsupervised learning and zero-shot learning, the authors further decreased the dependence of the parameter learning process on data.

Supervised haze removal methods have achieved excellent results. However, this requires paired data, which is difficult to obtain in the real world. For outdoor scenes with grass, water, or moving objects, it is difficult to guarantee that two images taken on a clear, cloudy day have exactly the same content. That is the reason why the works developed by Engin et al. [10], Golts et al. [11], Li et al. [24, 25] explored unsupervised dehazing algorithms.

In the next subsections, we present the methods used in this work, which include YOLY and ZID.

2.3.1 YOLY

Considering a haze image, represented as $I(x)$, the central purpose is to restore the image without haze $J(x)$ without making use of information beyond what is contained in the image itself. The essence of this method is based on breaking down x into three sub-networks, as illustrated in Figure 2.2.

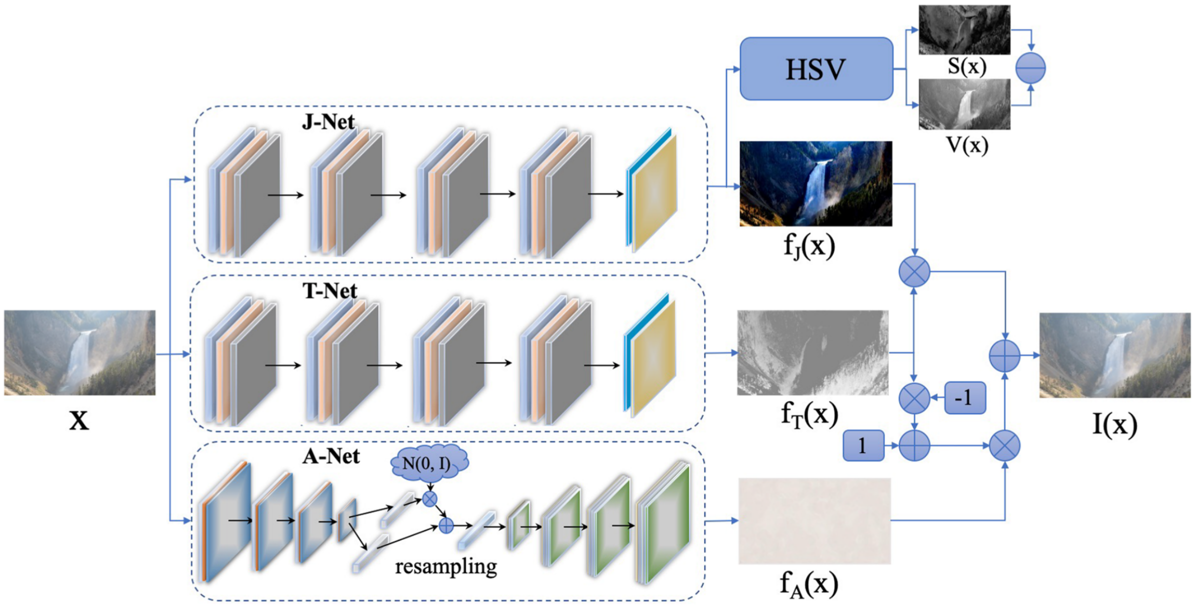


Figure 2.2: The architecture of YOLY. Extracted from the work developed by Li et al. [25].

To be more precise, YOLY simultaneously channels x through three main networks: the first is designed to estimate $J(x)$, the second to estimate the transmission map $T(x)$, and the third focuses on estimating atmospheric light A. Subsequently, the results of these networks are further combined to reconstruct x in the upper layer, making use of the atmospheric scattering model (Equation 2.1).

Loss Function: Therefore, the model as a whole learns in an unsupervised way. In summary, the goal is to minimize the following loss function:

$$L_{rec} = ||I(x) - x|| \tag{2.4}$$

The cleared image $J(x)$ is obtained by combining the outputs generated by the three sub-networks, as expressed in Equation 2.1. The loss function L_{rec} was designed to regulate the performance of the system as a whole, encompassing both the individual sub-networks and the reconstruction of the haze image $I(x)$ after calculating its components. More precisely, this loss function monitors and guides the disentangling process, and this is achieved by incorporating the haze generation process.

In addition, YOLY proposes a new loss function taking into account the HSV color space, which arises based on the observation made by Zhu et al. [60], which indicates that the difference between brightness and saturation is close to zero in the haze-free zones.

To make use of this previous information, they propose the following equation regarding the prediction of $J(x)$.

$$L_J = ||V(J(x)) - S(J(x))|| \tag{2.5}$$

In addition, YOLY proposes a set of loss functions for training the A-Net sub-network, including a regularization function to avoid overfitting.

2.3.2 ZID

The approach followed by ZID is similar to YOLY [25], in terms of deinterlacing the problem into capable simpler ones. However, ZID differs from YOLY in two fundamental aspects. First, a distinction is made in terms of the loss function used. Specifically, ZID proposes a loss similar to that used in DCP for J-Net training, while YOLY is based on observing the HSV color space. In addition to this, ZID introduces a smooth regularization in the outputs of T-Net and A-Net, as opposed to YOLY, which only applies this regularization in A-Net. Second, network architectures vary. ZID adopts a structure analogous to the U-Net architecture, in contrast to YOLY, which is based on a non-degenerate architecture. Figure 2.3 shows in general the architecture adopted by ZID.

Figure 2.3: The framework of ZID. Extracted from the work developed by Li et al. [24].

2.4 Dataset

We are set to conduct comprehensive experiments leveraging a dataset recognized for its prowess in large-scale haze removal, namely the REalistic Single Image DEhazing dataset, abbreviated as RESIDE[1]. This dataset, meticulously curated for its diverse and

[1] https://sites.google.com/view/reside-dehaze-datasets/reside-standard

challenging nature, has proven instrumental in evaluating the efficacy of dehazing methodologies [23]. To provide a snapshot of the dataset's characteristics, we present an overview in Table 2.2, delineating two distinctive test subsets: SOTS and HSTS.

The SOTS subset encompasses 500 haze images captured in indoor settings. Notably, these images are synthesized using a sophisticated physical model, incorporating meticulously tuned parameters. This synthesis process ensures a diverse representation of hazing scenarios, enriching the dataset with an abundance of indoor atmospheric conditions. On the other hand, the HSTS subset is tailored for outdoor evaluation and comprises a curated collection. It consists of 10 synthetic haze images, intelligently generated to simulate a spectrum of real-world outdoor hazing scenarios. Additionally, HSTS incorporates 10 real-world blur images, meticulously captured across diverse scenes, further enhancing the dataset's realism and relevance to practical dehazing scenarios. To evaluate our approach to haze removal, it is important to note that we dispense with the training subset and rely solely on the validation set to perform relevant evaluations.

Table 2.2: Overview of the RESIDE dataset.

RESIDE	
Subset	**Images**
Indoor Training Set	13,990
Synthetic Objective Testing Set	500
Hybrid Subjective Testing Set	20

Figures 2.4 and 2.5 illustrate some samples of the dataset, which are divided in terms of haze inside of a closed environment (indoor: Figure 2.4) or outside (outdoor: Figure 2.5). It is worth noting that the images on SOTS are manually generated by the physical model to represent the haze effect. A more detailed description can be found in Appendix A.

2.5 Final Remarks

This chapter aimed to present the background context of our research, where we present the fundamental concepts to understand our approach and the evaluations carried out and review the research related to our study. After reviewing the state of the art, our focus will be on applying Zero-Shot Learning and Unsupervised Learning, thus avoiding dependency on paired data sets. Our strategy leverages the benefits of integrating different sources of prior information, such as Dark Channel Prior and Color Attenuation Prior, together and in a coordinated way. The chapter established a solid foundation for our research work.

The following chapter focuses on a detailed description of our proposed method, considering each module that composes it, as well as a description of the steps involved in each of these modules, providing a complete description of our proposed method.

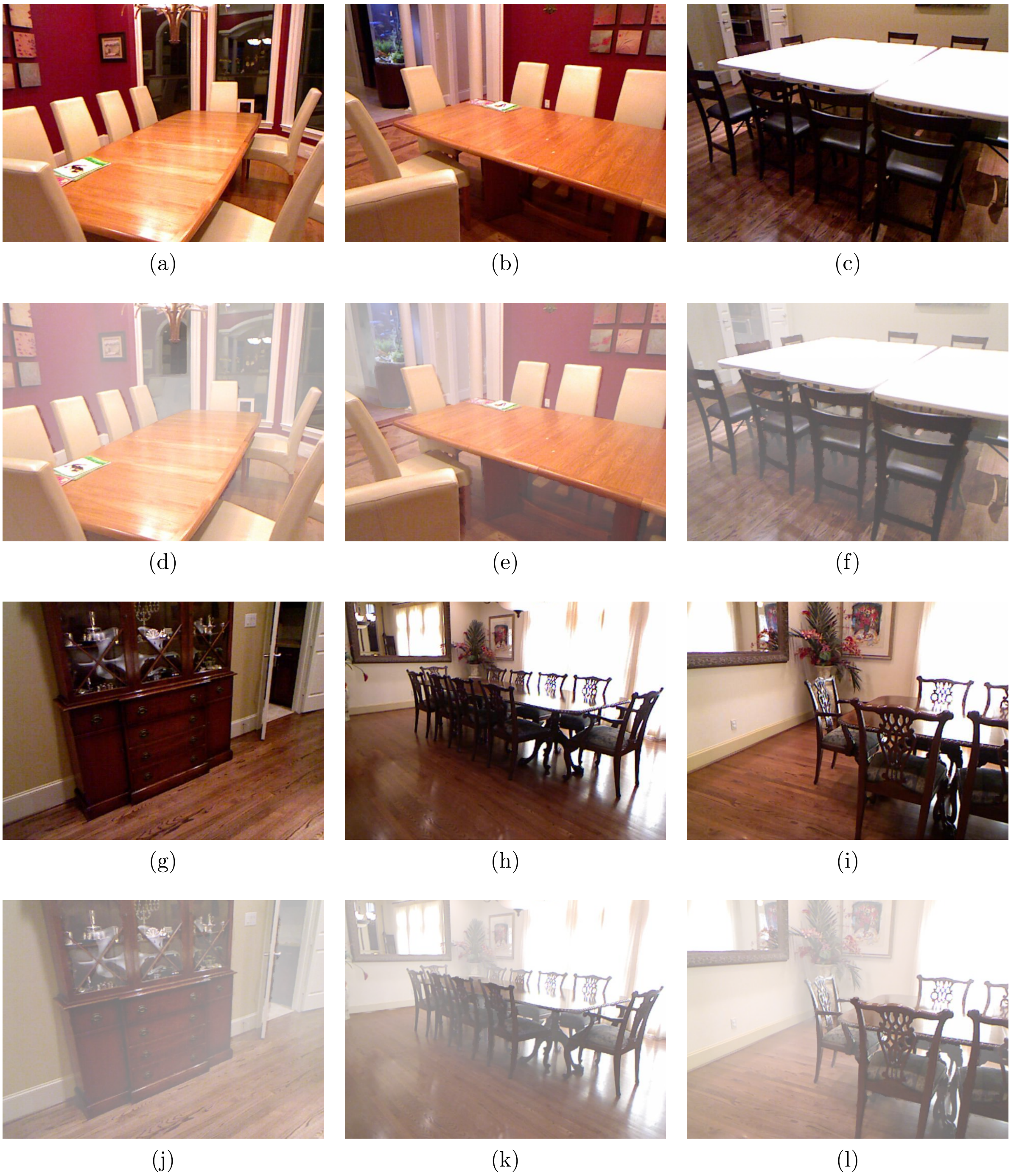


Figure 2.4: SOTS indoor images: The first and third row of images, (a), (b), (c), (g), (h) and (i), show haze-free samples in a closed environment, for example, inside a house. The images in the second and fourth row, (d), (e), (f), (j), (k) and (l), show samples with haze.


(a)


(b)


(c)


(d)


(e)


(f)


(g)


(h)


(i)


(j)


(k)


(l)

Figure 2.5: SOTS outdoor images: The first and third row of images, (a), (b), (c), (g), (h) and (i), show haze-free outdoor samples in a street or environment itself that reflect the effects of nature. The images in the second and fourth row, (d), (e), (f), (j), (k) and (l), show images with haze, fog, and other effects.

Chapter 3

Proposed Method

This chapter presents a detailed description of our main approach, named XYZ, to haze removal. We propose a method based on supervised neural networks, thus addressing the elimination of haze without relying on paired data, which is a distinctive advantage of our approach. We further describe XHOT, which is our first approach to haze removal.

Section 3.1 details the design and evolution of two innovative unsupervised learning methods for haze removal: (i) XHOT [29] and (ii) XYZ [29]. The XHOT method presents a simplified strategy for haze removal, contrasting with the complexity of the ZID method. In addition, it introduces our main approach, XYZ, which capitalizes on the strengths of the XHOT, YOLY and ZID methods. XYZ, by amalgamating these techniques, achieves a synergy that significantly enhances the model's ability to address various challenges associated with haze removal in images. This approach reflects the search for a comprehensive and effective solution in the landscape of unsupervised learning for the improvement of images affected by haze.

Section 3.2 addresses the comprehensive process of evaluating results obtained through our approaches, XHOT and XYZ. This process involves an evaluation that will range from Full-Reference metrics to Non-Reference metrics. Full-Reference metrics, such as PSNR and SSIM, will provide a detailed quantitative assessment, allowing a rigorous comparison with other approaches and revealing performance in terms of fidelity and structural similarity. On the other hand, Non-Reference metrics, such as BRISQUE, NIQE and PIQE, will offer an additional perspective by evaluating image quality more closely to human perception, considering factors such as naturalness and visual perception. This comprehensive approach to evaluation reflects our commitment to a complete and accurate understanding of XYZ's performance in the critical task of haze removal.

3.1 Unsupervised Approach

In this section, we delve into the evolution of our proposed methodologies, a fusion of three unsupervised learning techniques. Firstly, we introduce XHOT (Subsection 3.1.1), rooted in the foundational principles expounded in prior investigations, notably drawing inspiration from the notable works of [25] and [11]. Subsequently, our exploration extends to YOLY, influenced by the groundbreaking research articulated by Li et al. [25], and

ZID, shaped by the insightful findings presented by Li et al. [24].

The remarkable efficacy of both YOLY and ZID is underscored by their performance across key metrics such as SSIM and PSNR, firmly establishing them as benchmarks within the challenging RESIDE dataset.

Furthermore, it is crucial to highlight a pivotal characteristic shared by all the methodologies detailed henceforth - the absence of a requirement for paired image datasets during the training process. This unique attribute accentuates the adaptability of each method, autonomously addressing the challenge of haze removal for individual images. Notably, this not only underscores their practical utility in real-world scenarios but also positions them as versatile solutions free from the constraints of paired image dependencies.

In the following subsections, we present and analyze our two developed approaches, named XHOT and XYZ.

3.1.1 XHOT Approach

The conceptual genesis of the XHOT network [29] is rooted in the exigency to forge an effective haze removal solution, deliberately deviating from the conventional dependence on paired data inherent in unsupervised learning methodologies. Acknowledging the intricate nature of haze removal and its decomposability into simpler constituent elements, our strategic approach pivots on the synergy derived from the fusion of multiple elementary layers.

This architectural decision, as elucidated in prior research [24], stems from the recognition that a hierarchical combination of these fundamental layers can significantly enhance the efficacy and efficiency of haze removal processes.

In essence, our network's architecture leverages the collaboration of these simple layers, each contributing to the overall task of haze removal. This modularized strategy allows for adaptability to diverse atmospheric conditions and varying degrees of haziness. By dissecting the complexity of haze removal into manageable components, the XHOT network is designed to navigate the challenges posed by real-world scenarios, where paired data may be scarce or unavailable.

Furthermore, the reliance on unsupervised learning principles aligns with the pragmatic need to develop solutions that do not necessitate paired data for training. This choice emphasizes the network's capacity to autonomously learn and adapt to diverse haze patterns, ensuring its versatility across a spectrum of environmental conditions. In summary, the XHOT network represents a purposeful stride towards an effective and efficient unsupervised haze removal solution, leveraging the power of modular simplicity to navigate the complexities of real-world atmospheric conditions.

In tackling this challenge, we have devised three neural networks, among which two are convolutional neural networks meticulously crafted to ascertain the optimal values for $J(x)$ and $T(x)$, as delineated in Figure 3.1. However, when it comes to computing the atmospheric light, a distinctive component of the image content, we have elected to adhere to the methodology put forth by earlier research, particularly by Li et al. [24] and Li et al. [25]. Their work leveraged a Variational Autoencoder (VAE) to estimate the atmospheric light A.

This strategic decision stems from recognizing the demonstrated efficacy of these methodologies in the specific task of atmospheric light estimation. By harnessing convolutional neural networks to discern the parameters $J(x)$ and $T(x)$, we aim to enhance our model's capacity to accurately and efficiently capture the fundamental features of haze-affected images.

The utilization of a dedicated convolutional neural network for the estimation of each component is pivotal in addressing the inherent complexity of the dehazing task. This modular approach allows for a more precise and specialized focus on each aspect of the process, thereby optimizing the overall performance of the XHOT network in haze removal from images.

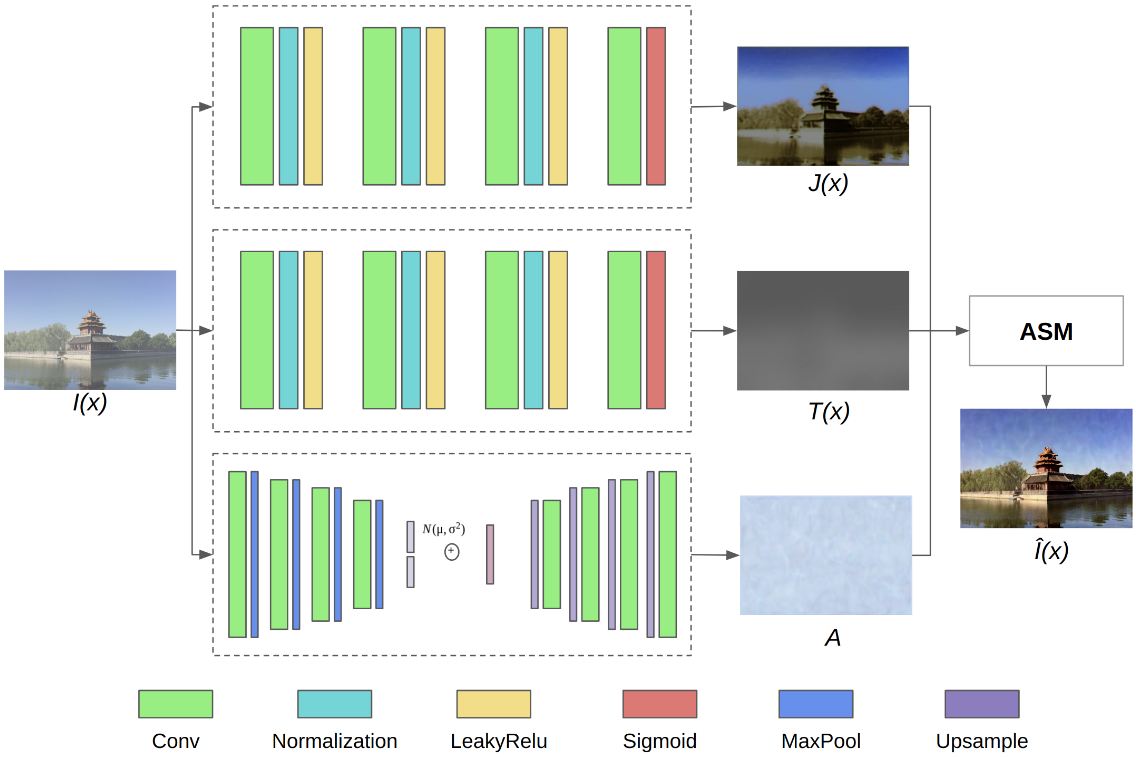


Figure 3.1: Architecture of our proposed XHOT, in which we have three sub-networks to calculate the variables $J(x)$, $T(x)$ and A, respectively.

Architectures: Three sub-networks were constructed to estimate the values for $J(x)$, $T(x)$ and A:

- J-Net: This sub-network is designed as a non-degenerative convolutional neural network, meaning it preserves the dimensions of the input. Comprising three convolutional blocks, each conducts a convolution with a kernel size of 5, followed by batch normalization, and finally, a LeakyReLU activation function with a slope of 0.01. At the conclusion of the third block, we apply a convolution along with a Sigmoid activation function to normalize the output between 0 and 1. The resulting output of this sub-network is a 3-channel image representing the haze-free image $J(x)$, which will play a crucial role in guiding the training process.

The architectural choices for J-Net are grounded in principles aimed at maintaining information integrity and effective feature extraction throughout the network. The non-degenerative nature ensures that the spatial dimensions of the input are conserved, preserving crucial details during the processing. The LeakyReLU activation function contributes to introducing non-linearity, allowing the network to capture complex relationships within the data.

Additionally, the use of batch normalization enhances training stability by normalizing the inputs at each layer. The final application of the Sigmoid activation function ensures that the output is appropriately scaled between 0 and 1, aligning with the characteristics of the haze-free image. This output, denoted as $J(x)$, serves as a vital reference for the training process, guiding the network towards generating dehazed images that closely resemble the ground truth.

- T-Net: This sub-network adopts the identical neural network architecture as J-Net, yet introduces a key distinction in the output. The output of this neural network is a single-channel image, representing the transmission map $T(x)$ within the given image. The utilization of the same neural network structure for T-Net as in J-Net is strategic, leveraging the network's capacity to capture and understand complex features related to haze removal. By focusing on the transmission map, which signifies the attenuation of light due to haze, T-Net plays a critical role in estimating the transmission properties across the image.

  The decision to have a single-channel output for the transmission map aligns with the inherent characteristics of the transmission map, which is a scalar value representing the degree of haze in different regions of the image. This design choice ensures that the output is tailored to the specific information content required for transmission map estimation, streamlining the learning process.

  In summary, while T-Net shares its architectural foundation with J-Net, its distinct focus on generating a transmission map provides a complementary piece of information necessary for the overall dehazing process. Together, J-Net and T-Net form a framework, each contributing specialized outputs that collectively guide the XHOT network in effectively removing haze from images.

- A-Net: This sub-network consists of a Variational Autoencoder, since the variable A is not related to the content of the image, similar to the work by Li et al. [24], it is assumed that A is sampled from a latent Gaussian distribution, and so the problem becomes a variational inference [19].

Loss Function: In the training process of our unsupervised model, a pivotal aspect lies in the careful design of an effective loss function that guides the neural network to learn the underlying features of haze removal. For this purpose, we leverage a comprehensive loss function, denoted as L_{XHOT}, which combines the loss functions from both J-Net and A-Net components. The formulation is depicted in Equation 3.1:

$$L_{\text{XHOT}} = L_J + L_A, \tag{3.1}$$

where L_J is the loss function between the input x and the result of generating haze $I(x)$, this value is calculated taking into account the 3 variables predicted following Equation 2.1. Then, we can define L_J as:

$$L_J = \text{MSELoss}(I(x), x). \tag{3.2}$$

In addition, the loss function for the A-Net sub-network is expressed as:

$$L_A = L_H + L_{KL}, \tag{3.3}$$

where L_H is the loss function MSELoss between A and A_H, A_H is the pre-calculated value of the image with haze x using Dark Channel Prior [15]. Finally, the function L_{KL} is given following Equation 3.4, where KL() denotes the Kullback-Leibler divergence between two distributions:

$$\begin{aligned} L_{KL} &= \text{KL}(N(\mu_z, \sigma_z^2)||N(0,1)) \\ &= \frac{1}{2m}\sum_i \left((\mu_{z_i})^2 + (\sigma_{z_i})^2 - 1 - \log(\sigma_{z_i})^2\right). \end{aligned} \tag{3.4}$$

3.1.2 XYZ Approach

In the subsequent sections, we delve into the intricacies of our second proposed approach, marked by the seamless integration of the strengths derived from the XHOT [29], YOLY [25], and ZID [24] methods. The distinctiveness of this approach lies in its ability to harness the unique advantages presented by each unsupervised learning method, fostering a synergy that elevates the overall performance of the dehazing model.

A noteworthy aspect is the preservation of the individual loss functions associated with the XHOT, YOLY, and ZID methods. These loss functions serve as pivotal components guiding the training process of our proposed robust model. By retaining these distinct loss functions, we capitalize on the specialized capabilities encoded within each method, ensuring that the model benefits from the subtle insights provided by XHOT, YOLY, and ZID.

The integration of multiple unsupervised methods introduces a level of adaptability and versatility to our proposed approach. Each method contributes a unique perspective and set of features, collectively enhancing the model's capacity to handle diverse atmospheric conditions and image complexities. This synergistic fusion of methodologies is poised to yield a dehazing model that not only surpasses individual benchmarks but also exhibits a robustness essential for real-world deployment.

Architecture

Illustrated in Figure 3.2, our methodology leverages insights from the works developed Li et al. [24] and Li et al. [25]. As described in Section 2.3, our devised methods are strategically designed to disentangle the intricate effects of haze into more computationally manageable layers.

This deinterlacing process becomes feasible through the application of Equation 2.1. The atmospheric dispersion model, comprising three unknown variables, forms the core of many haze removal techniques, centering their efforts on elucidating these critical values.

In alignment with this principle, our approach capitalizes on the disentanglement facilitated by XYZ [29] to compute the haze-free image. To achieve this, our methodology features two sub-network groups alongside an unsupervised neural sub-network. The gJ-Net group is dedicated to computing the clear image $J(x)$ and, subsequently, training our neural network with its attendant loss function. Concurrently, the gT-Net subgroup is entrusted with computing the transmission map $T(x)$. Lastly, the A-Net sub-network takes on the responsibility of estimating the atmospheric light embedded within the image.

This strategic division into sub-network groups not only underscores the complexity of haze removal but also highlights the multifaceted nature of our approach. By breaking down the dehazing process into these specialized components, our methodology attains a comprehensive understanding of the atmospheric conditions, enabling the model to generate superior results. The coordinated efforts of these sub-networks collectively contribute to the potency of our unsupervised neural network, setting the stage for enhanced dehazing performance in diverse scenarios.

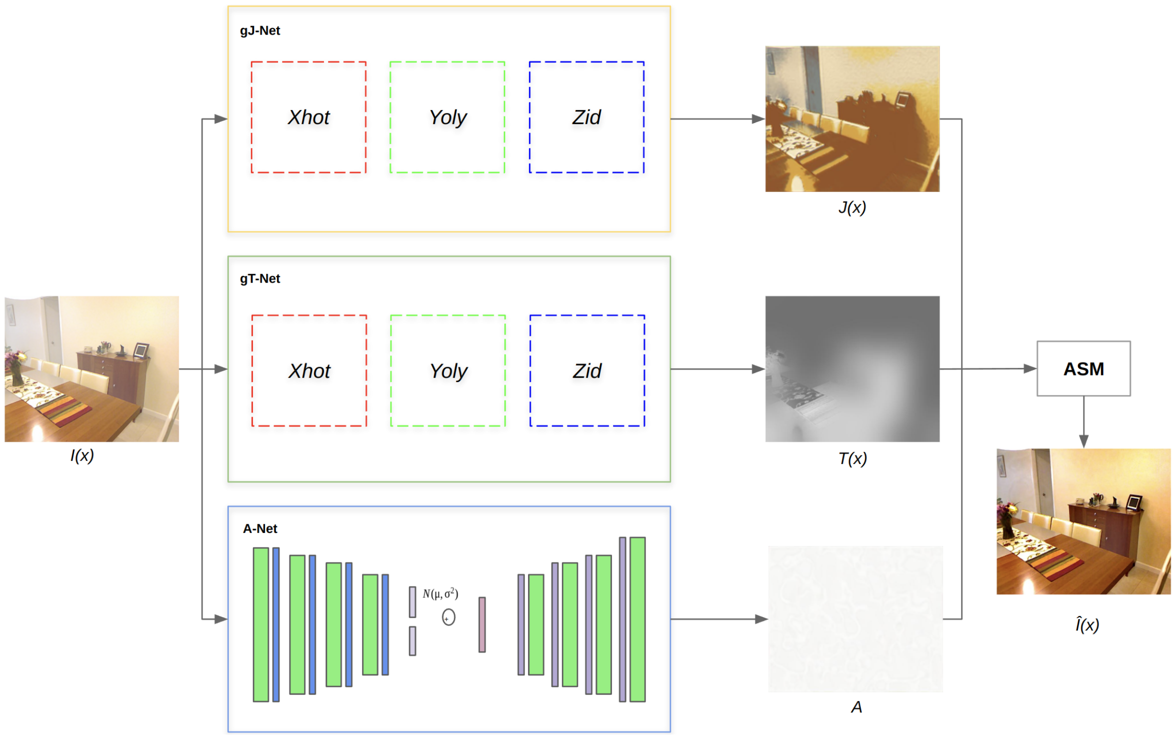


Figure 3.2: The architecture of our proposed XYZ approach. XYZ includes two groups of sub-networks: gJ-Net group of sub-networks (XHOT, YOLY and ZID) for image estimation clean $J(x)$, gT-Net group of sub-networks for transmission map estimation $T(x)$. In addition, XYZ maintains layer disentanglement for the estimate of A, which is unique for the two groups described previously.

- **gJ-Net**: The gJ-Net group (Figure 3.2 (a)) emerges as a synergistic integration of the XHOT, YOLY, and ZID methods, strategically fusioning the strengths of each approach. Drawing inspiration from XHOT, which employs a non-degenerative neural network akin to YOLY, this fusion serves a dual purpose. First, it aids in the preservation of critical details essential for generating a high-quality haze-free image. This includes the retention of intricate object shapes within the haze image while maintaining the original dimensions. YOLY, in its contribution, leverages the HSV color space to optimize the training of the unsupervised network. This innovative approach is inspired by the findings of Zhu et al. [60], who observed that in haze-free regions, the disparity between brightness and saturation tends to be close to zero.

Further enriching the gJ-Net group, ZID introduces a novel perspective by incorporating a loss function based on Dark Channel Prior Loss [11]. In tandem, ZID adopts a degenerative neural network of the U-Net architecture, complete with skip connections. The U-Net structure is particularly instrumental in capturing and reconstructing intricate features from various scales of the input, aligning with the subtle nature of haze removal. This strategic combination of architectures, along with the incorporation of the HSV color space and Dark Channel Prior Loss, reflects the improvement in our gJ-Net group. The fusion of these diverse methodologies empowers the neural network to adeptly navigate the challenges posed by haze, resulting in a holistic and robust haze removal mechanism.

- **gT-Net:** In parallel with the gJ-Net architecture, this group (Figure 3.2 (b)) mirrors the three previously introduced methods, differing solely in the output of each method. Regardless of the specific method employed, the common output is the transmission map associated with the image under processing. This design choice enhances the flexibility and adaptability of the overall framework, allowing for seamless integration of diverse methodologies while maintaining a unified focus on the transmission map, a pivotal element in the dehazing process.

  The versatility embedded in this sub-network design not only streamlines the integration of distinct methodologies but also emphasizes the significance of the transmission map as a shared representation across the methods. By maintaining consistency in this crucial component, the group ensures a coherent and unified approach to haze removal, capitalizing on the unique strengths of each method while preserving a common thread that unifies their contributions. This unified output, despite methodological differences, lays the foundation for a more cohesive and robust dehazing mechanism within the overarching neural network architecture.

- **A-Net:** Within this critical sub-network, the primary objective is the computation of the atmospheric light inherent in an image (Figure 3.2 (c)), irrespective of its content. The presumption is made that this atmospheric light arises independently and exists within a latent space shaped by a Gaussian distribution. To tackle this subtle task, a strategic choice is made to employ a dedicated neural network, specifically a VAE, as elucidated in Section 3.1.1.

  The rationale behind opting for a VAE lies in its capacity to encapsulate the complex characteristics of the atmospheric light within a probabilistic latent space. VAEs offer a principled approach to modeling this latent space, allowing for the extraction of meaningful representations. By leveraging the probabilistic nature of the latent space, the neural network within this sub-network becomes adept at capturing the inherent variability associated with atmospheric light across diverse images.

This deliberate choice reflects a subtle understanding of the complexities involved in estimating atmospheric light, acknowledging the need for a sophisticated model that can navigate the inherent uncertainties in this crucial component of haze removal.

Loss Function

To train our model composed of three unsupervised methods, we use together a loss function taking into account the output of each method as described in Equation 3.5:

$$L_{\mathrm{XYZ}} = L_{\mathrm{XHOT}} + L_{\mathrm{YOLY}} + L_{\mathrm{ZID}} \tag{3.5}$$

where L_{XHOT} is the loss function for our XHOT method, L_{YOLY} is the loss function for YOLY and finally L_{ZID} is the loss function for ZID, we decided to give the same weight to each loss function as the three guide our training, taking into account the assumptions made, such as the case of the HSV color space, Dark Channel Prior Loss, in addition to the XHOT loss function, which turns out to be an improvement of Golts et al. [11] and Li et al. [25].

3.2 Evaluation

In the training phase of our model, consisting of three unsupervised methods, a critical aspect lies in the formulation of an effective loss function that can adequately guide the optimization process. The synergy of these three methods, namely XHOT [29], YOLY [25], and ZID [24], necessitates a comprehensive loss formulation to capture the nuanced aspects of haze removal. We introduce a unified loss function that strategically considers the output contributions from each method, synthesizing them into a cohesive optimization objective.

The loss function, as depicted in Equation 3.5, reflects a holistic approach that accommodates the distinctive characteristics of each method. This formulation is designed to strike a balance between preserving crucial details, such as object shapes and scene-specific features, while enhancing overall image clarity and transmission map accuracy. It is essential to underscore the significance of this unified loss function in training a model that leverages the strengths of each method to collectively achieve superior dehazing performance.

Equation 3.5 incorporates a multi-term structure that accounts for the contributions of XHOT, YOLY, and ZID methods individually. This allows the model to discern the relative importance of each method's output in the overall optimization process. Moreover, the formulation considers factors such as spatial consistency, color preservation, and contrast enhancement, ensuring a comprehensive evaluation of the dehazing model's performance.

By unifying the loss function, we create a training framework that maximizes the benefits of each unsupervised method. This approach not only facilitates efficient convergence during training but also enhances the adaptability of the model to diverse environmental conditions, contributing to the robustness and generalization capability of our proposed haze removal solution.

3.2.1 Full-Reference Quality Metrics

In the area of haze removal, accurate evaluation of the quality of the algorithms' outputs plays a crucial role in measuring their effectiveness. Full References Quality Metrics constitute an essential approach in this context. These metrics focus on the direct comparison between processed images and dehazed reference images, thus providing a comprehensive and objective evaluation of the fidelity and accuracy of dehazing. By considering the totality of the information present in the reference image, allowing a detailed and precise evaluation of the effectiveness of the algorithms in adverse atmospheric conditions. That is the reason why the evaluation of our approach includes both PSNR and SSIM metrics.

1. **Peak Signal to Noise Ratio (PSNR)**: Indicates the proportion of the image signal's maximum value to its noise. The higher the value, the less noise interferes with the image and the better the image quality. It is a commonly used metric in the field of image processing, including haze removal. It is used to evaluate the quality of an image by measuring the relationship between the signal (image information) and noise (any distortion or degradation present). The PSNR is expressed in decibels (dB) and calculated by the formula 3.6:

$$\begin{aligned} \mathrm{MSE}(X,Y) &= \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} (X(i,j) - Y(i,j))^2 \\ \mathrm{PSNR} &= \log \left[\frac{(2^N - 1)^2}{\mathrm{MSE}} \right], \end{aligned} \tag{3.6}$$

   where MSE (Mean Squared Error) represents the average square difference between the pixels of the original (X) and the degraded (Y) image. H, W correspond to height and width of the images, and N corresponds to bits per pixel (8).

   In the context of haze removal, a higher PSNR indicates lower degradation and higher fidelity in the processed image. However, it should be noted that the PSNR only measures the difference pixel by pixel, as it does not consider more complex perception characteristics.

2. **Structural Similarity Index Measure (SSIM)**: It quantifies how closely two images match up structurally, taking into account detail, edge, and contour information. The more haze is eliminated, the less the restored image's structure differs from the original image, and the higher the SSIM value is because the comparison and analysis are between the processed image and the original haze image. The value of SSIM approaches one if the amount of haze eliminated is lower since it

is more comparable to the original haze image and their structural similarities are greater.

$$\text{SSIM}(X,Y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \tag{3.7}$$

where μ_x and μ_y correspond to mean of X and Y, respectively, σ_x and σ_y correspond to variance of X and Y, respectively, and σ_{xy} corresponds to covariance between X and Y.

Unlike simple metrics such as PSNR, SSIM takes into account visual perception and image structure. Evaluate the structural similarity between a reference image and a processed image

3.2.2 Non-Reference Quality Metrics

Non-Reference Quality Metrics play an essential role in evaluating the effectiveness of dehazing algorithms without relying on clear reference images. This approach is particularly valuable in real-world situations, where obtaining haze-free images may be difficult or even impossible.

These metrics focus on the intrinsic evaluation of the processed images, using internal properties and specific image characteristics to determine the quality of dehazing. By dispensing with reference images, Non-Reference Quality Metrics adjust to practical scenarios, offering an objective evaluation based on the intrinsic information of the cleared images. This approach is crucial for measuring the effectiveness of algorithms in real-world conditions, where the availability of reference images may be limited.

1. **Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE)**: It is an image quality metric that focuses on evaluating image quality distortion without requiring a reference image. Developed by Mittal et al. [31], BRISQUE is based on the analysis of natural image statistics to identify perceptual distortions.

   BRISQUE's approach is to extract statistical features from an image and compare them to a reference model that has been built from a diverse set of natural images. These characteristics include the magnitude and orientation of the gradients, high-frequency energy, and pixel distribution.

   The BRISQUE quality score is calculated by comparing the features extracted from the image under test with those of the reference model. A lower score indicates a better quality image, while a higher score suggests greater distortion.

2. **Naturalness Image Quality Evaluator (NIQE)**: It is a metric designed to evaluate the quality of images completely blindly Mittal et al. [32], that is, without the need for a quality reference. The fundamental purpose of NIQE is to measure the perceptual naturalness of an image, evaluating the statistical discrepancy between the statistics of the original image and the statistics of an ideally natural image. NIQE uses second-order statistics to evaluate image texture, seeking to capture how the image deviates from natural statistical properties.

The NIQE formula involves extracting features based on image quality statistics and comparing them to a reference model. Its performance has proven effective in evaluating the quality of natural images and has found applications in various fields, including image enhancement and dehazing.

3. **Perception-based Image Quality Evaluator (PIQE)**: PIQE stands as a non-reference image quality evaluation method designed for real-world images, employing a perception-based approach. The mean subtraction contrast normalization coefficient is utilized to derive the image quality score [49]. Unlike some methods that rely on supervised learning for quality prediction, PIQE is an unsupervised approach, eliminating the need for a learning model.

   PIQE draws inspiration from key principles governing human perception of image quality. First and foremost, human visual attention tends to focus strongly on salient points and spatially active regions within an image. PIQE leverages this insight by estimating distortion primarily in spatially prominent areas [49]. Additionally, PIQE recognizes that local quality at the block/patch level significantly contributes to overall human perception of image quality. To address this, the distortion level is calculated at the local block level, typically with a size of $n \times n$ where $n = 16$.

3.3 Final Remarks

This chapter has delineated our innovative approaches, XHOT and XYZ, designed with the specific purpose of addressing haze removal without relying on paired images. By adopting an unsupervised approach, XYZ leverages the inherent information contained in haze images to learn and improve autonomously.

The next chapter will delve into the practical application of our XHOT and XYZ method, highlighting the results obtained by implementing it in the RESIDE dataset. Through this detailed application, XYZ capabilities and performance will be revealed in real-world scenarios, providing a comprehensive view of its effectiveness in haze removal under various conditions.

Chapter 4

Experimental Results

This chapter describes and discusses the results obtained by applying our XYZ approach [29] to the RESIDE dataset, a widely recognized resource in the field of haze removal.

Section 4.1 provides a detailed overview of the experimental setup where the experiments that tested the effectiveness of our approaches XHOT and XYZ were conducted. The conditions, parameters and variables relevant to understanding the environment in which the results were evaluated are described.

Section 4.2 focuses on an exhaustive quantitative and qualitative comparison of the results obtained in our experimental environment. For quantitative comparison, the metrics presented in Chapter 3 are used, allowing an objective and detailed evaluation of the performance of our XHOT and XYZ approaches compared to other methods in haze removal. In addition, qualitative analyses are presented that highlight visual aspects and specific details of the images processed by our method. This comprehensive approach provides a complete understanding of the effectiveness of XYZ in various evaluation dimensions.

4.1 Experimental Setup

Running experiments is essential to evaluate the effectiveness and robustness of the proposed approaches in dehazing using unsupervised neural networks. The experimental setup was carried out with the aim of providing reliable and comparable results. The configuration used in the experiments is detailed below:

- Dataset: For the purpose of evaluating haze removal, we chose to use the REalistic Single Image DEhazing dataset. This data set has gained recognition in the community due to its breadth, diversity and relevance in the evaluation of haze removal methods. This choice provides a robust and representative environment to test the effectiveness of our approaches, ensuring significant results and comparable with research standards in this field.

- Implementation: Unsupervised models, such as XHOT, ZID [24], and YOLY [25], along with our proposed combination XYZ, were implemented using the PyTorch deep learning framework. Each algorithm underwent an independent training and evaluation process, with specific configurations adapted to the characteristics of each model. This choice of framework and individualized approach ensures robust implementation and accurate evaluation of each method in the context of haze removal.

- Hardware and Software: The experiments were conducted in a computer environment equipped with a high-performance GPU, in particular an NVIDIA GeForce RTX 3070 graphics card was used to speed up both the training and evaluation process. In addition, it leveraged the capabilities of deep learning libraries provided by PyTorch, combined with scientific development environments such as Jupyter Notebooks. This configuration ensures the efficiency and precision needed to address the challenges of haze removal in images.

- Evaluation Metrics: To quantitatively evaluate the quality of the output images, standard metrics such as SSIM and PSNR were used. These metrics provide an objective measure of the structural similarity and fidelity of the restored images compared to the original images. Furthermore, we use non-referenced metrics such as BRISQUE, NIQE and PIQE for qualitative evaluation, based on the fact that obtaining real images without haze is complex and in many cases impossible.

4.2 Comparison

Evaluation of haze removal algorithms is a critical task to understand their performance and applicability in real-world scenarios. In this section, we will conduct a comprehensive qualitative and quantitative comparison of different haze removal approaches, aiming to provide a comprehensive view of their strengths and limitations.

4.2.1 Quantitative Results

In this section, we present the results when applying our method to deinterlace complex layers into simpler layers, in comparison to the methods proposed by Li et al. [24] and Li et al. [25]. It should be noted that the methods with which we are carrying out the comparison make use of the defocusing of haze deinterlacing in simpler layers.

Full-Reference

In the analysis of the results, we evidence notable observations that provide valuable perspectives on the performance of our two approaches in comparison with the ZID and YOLY methods, reflected in Table 4.1.

Although our XHOT approach does not exceed the ZID and YOLY methods in terms of SSIM and PSNR metrics, it is essential to highlight the progress made in haze removal. This result suggests that, while not leading in all metrics, XHOT remains a valuable and effective choice to address the challenge of haze removal. The ability to provide substantial improvements, even without reaching the leadership position, underscores its relevance in specific contexts and application scenarios.

Table 4.1: Results for XHOT, YOLY, ZID and XYZ methods applied to the SOTS and HSTS datasets. The results are shown in relation to SSIM and PSNR metrics.

Network	SOTS Indoor		HSTS Outdoor	
	SSIM↑	PSNR↑	SSIM↑	PSNR↑
ZID [24]	0.815	18.313	**0.851**	21.650
YOLY [25]	0.807	17.950	0.832	**22.217**
XHOT [29]	0.803	17.860	0.829	21.430
XYZ [29]	**0.818**	**18.530**	0.846	21.680

It is important to note that the effectiveness of XHOT can be manifested more prominently in certain particular contexts or conditions such as those presented by Li et al. [23], where its particular characteristics prove to be beneficial. The ability to deliver significant improvements, even in specific scenarios, highlights the versatility and applicability of XHOT in various haze removal situations.

Our XYZ approach, in contrast, distinguishes itself by improving the results on the SOTS dataset, which is specifically composed of images captured in closed environments. Although it still fails to outperform the ZID and YOLY methods in general metrics, the optimization capability in specific environments highlights the adaptability of XYZ to particular conditions. This achievement highlights the importance of considering different contexts and underscores the approach's ability to excel in specific scenarios.

It is essential to note that the effectiveness of XYZ can be manifested more prominently in particular situations, where its particular characteristics prove to be beneficial. The adaptability of XYZ to specific environments suggests that it could be a preferred option in particular conditions or specialized applications, where the ability to excel in specific scenarios is crucial. In situations where the dataset consists of images that capture the random and authentic nature of the haze, rather than synthetically generated images, our approach demonstrates significant improvements. Specifically, we observed notable improvements in images with high levels of reflected light, while results may be less effective in images with low levels of lighting.

It is important to note that, while our techniques may not exceed the state of the art in all metrics, they rank as the second method with the best results. This achievement consolidates our approaches as valuable and competitive contributions in the field of haze removal. The relevance of the improvements achieved and the ability to compete with leading methods underline the effectiveness of our strategies in solving this challenging problem. This success highlights the prominent position of our methods and their ability to offer meaningful solutions in the context of haze removal.

Non-Reference

The benchmarking of the XHOT, YOLY, ZID and XYZ methods used non-referenced image quality metrics, such as BRISQUE, NIQE and PIQE. These metrics play a crucial role in providing an objective assessment of the quality of the output images of each method, without relying on reference images. Below is the description of the results in terms of the SOTS subset containing interior images.

In terms of BRISQUE, the ZID method showed promising results in achieving a significant reduction in the distortion of the output images, closely followed by XHOT. On the other hand, YOLY and XYZ show acceptable results, but with a slightly higher score in the BRISQUE metric, indicating greater distortion compared to YOLY and XYZ. As for NIQE, ZID and XHOT stand out by achieving a substantial improvement in image quality, being consistently superior to YOLY and XYZ. This metric highlights the ability of ZID and XHOT methods to reduce distortion and improve image clarity by getting rid of haze effectively. In relation to PIQE, it is evident that XHOT and ZID exhibit superior performance, keeping distortion scores low. Although YOLY and XYZ show acceptable results, they do not exceed the quality offered by XHOT and ZID, highlighting the effectiveness of the latter in eliminating haze.

In summary, the evaluation with non referenced metrics BRISQUE, NIQE and PIQE suggests that ZID, XHOT are the most effective methods in eliminating haze, achieving a substantial improvement in image quality compared to YOLY and XYZ. These results support the effectiveness of our XYZ approach in solving the challenge of haze removal.

Exploring the performance of methods in outdoor environments reveals a slightly different picture. The Table 4.2 highlights that our main approach XYZ exhibits superior performance compared to other methods in both BRISQUE and PIQE metrics. These results point to the effectiveness of XYZ in eliminating haze in outdoor environments, highlighting its ability to maintain low levels of distortion, as indicated by the BRISQUE and PIQE metrics. This superior performance suggests that XYZ achieves significant improvement in image quality, even in the presence of additional challenges associated with outdoor environments.

This finding reinforces the versatility and robustness of our main approach, XYZ, by addressing haze removal in various scenarios, both indoor and outdoor. The ability to maintain consistent performance across different environments underscores the applicability and potential of XYZ to deliver significant improvements in image quality under various atmospheric conditions.

Table 4.2: Results for XHOT, YOLY, ZID and XYZ methods applied to the SOTS and HSTS datasets. The results are shown in relation to BRISQUE, NIQE and PIQE metrics.

Network	SOTS Indoor			HSTS Outdoor		
	BRISQUE↓	NIQE↓	PIQE↓	BRISQUE↓	NIQE↓	PIQE↓
ZID [24]	**39.011**	**5.015**	**34.224**	20.565	**5.024**	10.907
YOLY [25]	40.621	5.074	35.426	20.015	5.279	10.593
XHOT [29]	39.719	5.037	35.065	22.500	5.570	12.205
XYZ [29]	40.085	5.108	35.168	**19.706**	5.515	**10.032**

In conclusion, this detailed analysis provides a balanced view of the results, highlighting significant advances, areas of improvement and the overall relevance of our approaches in the landscape of haze removal using unsupervised neural networks.

4.2.2 Qualitative

In this section, we carefully examine the visual results obtained by applying our layer deinterlacing methods in outdoor and indoor environments. Figures 4.1 and 4.3 show the results for XHOT, ZID [24], YOLY [25], and XYZ algorithms in outdoor environments. open air and closed, respectively. These visual representations provide an intuitive assessment of the effectiveness of each method in terms of color and detail.

Close observation of Figure 4.1 reveals that the XHOT technique presents notable improvements, especially in areas such as the sky. This advance is attributed to the decision not to rely directly on a Dark Channel Prior approach, in contrast to ZID, which exhibits significant artifacts due to its DCP-based approach. In this context, XYZ shows significant improvement in attempting to make a more consistent distinction, although some artifacts remain in non-reference metrics. Additionally, Figure 4.2 shows real-world images from the HSTS dataset, where XYZ manages to effectively reduce haze in most of the image, although it does not eliminate it completely.

In the case of closed environments, Figure 4.3 shows that the ZID method stands out in the third image, preserving color details in areas such as the floor. However, XYZ demonstrates superior performance in higher illumination images, as seen in the first, second, and fifth images.

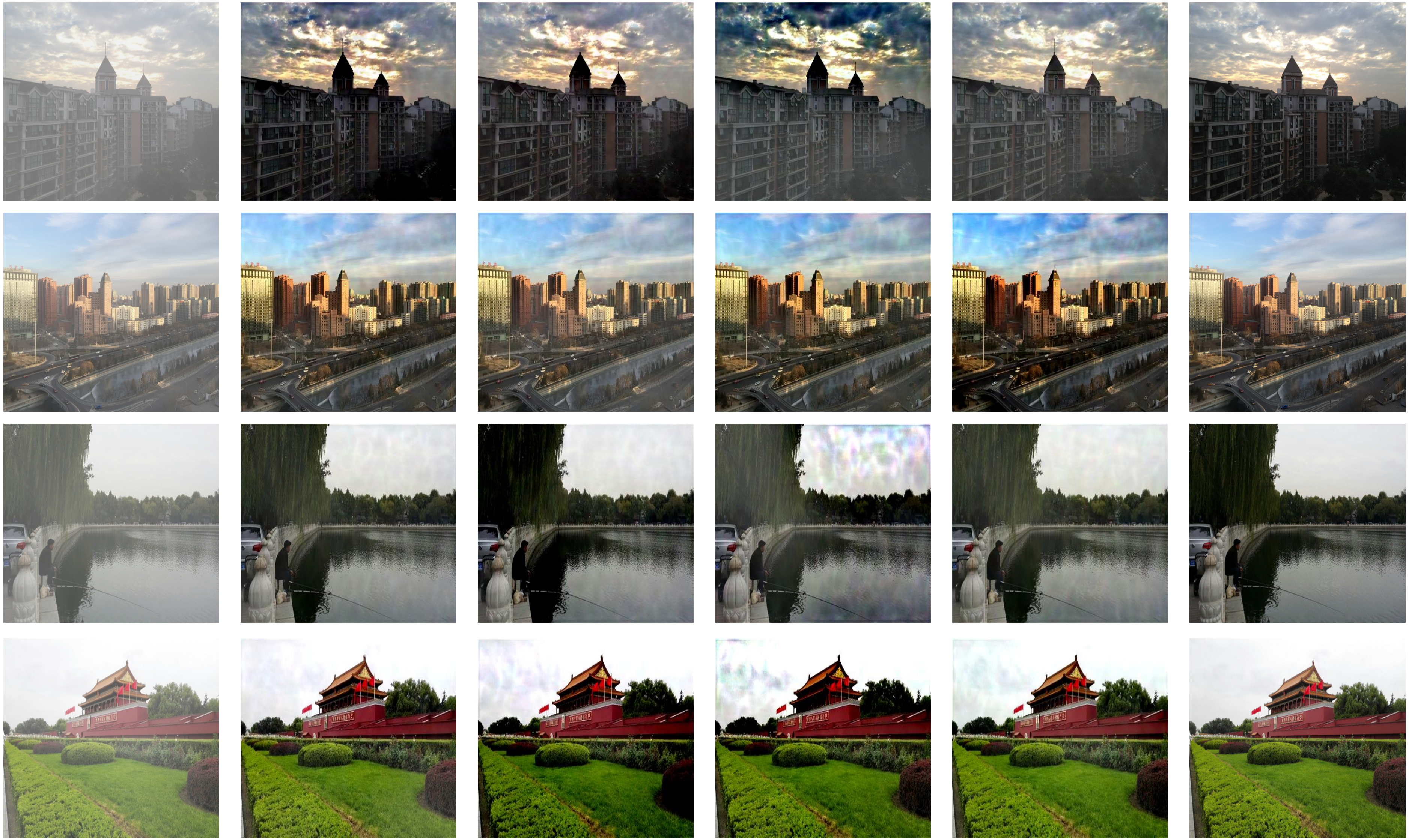
These visual results highlight the strengths and limitations of each method in specific contexts. While XHOT shows excellent results in outdoor environments by avoiding DCP artifacts, XYZ manages to effectively mitigate haze in indoor environments with varied lighting conditions. The choice of the most appropriate method will depend on the specific application context.

4.3 Final Remarks

This chapter immerses the reader in the implementation and application of our proposed method, XYZ. It provides a description covering each phase of our two approaches from its inception to its conclusion. Every aspect of the process is examined, from data collection to the combination of various methods, highlighting the advantages inherent in each. In addition, a comprehensive assessment is presented that addresses several relevant aspects.

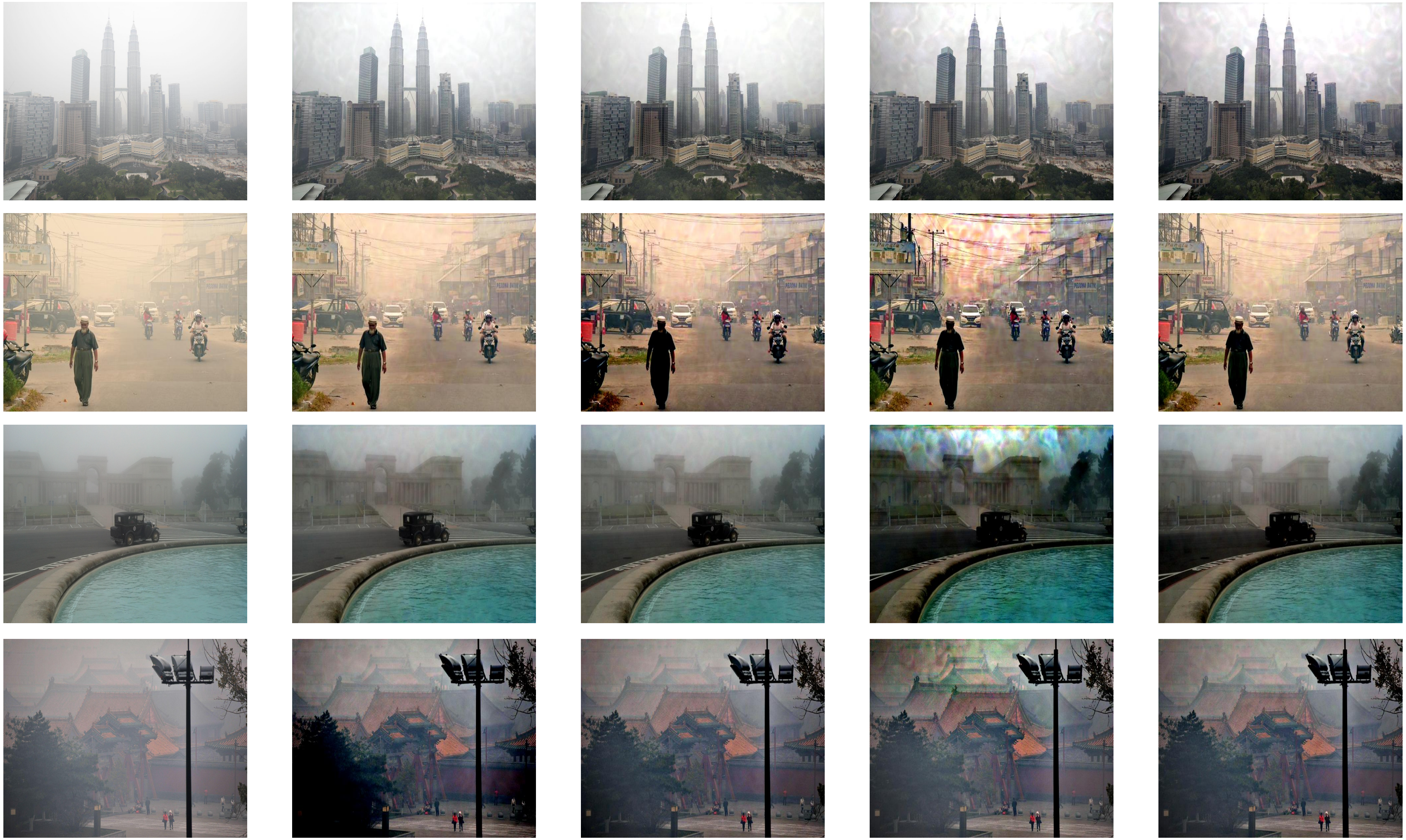
The chapter specifically details the images selected for evaluation, highlighting their importance in the validation and analysis of our approaches. By going deeper into each stage of the implementation process, we offer a complete overview that allows you to understand the workflow and strategic decisions made during the XYZ application.

The next chapter will consolidate these contributions and present key findings from this research. In addition, promising directions for future research will be explored, providing a comprehensive view of the impact and continuing potential of our approaches to haze removal.



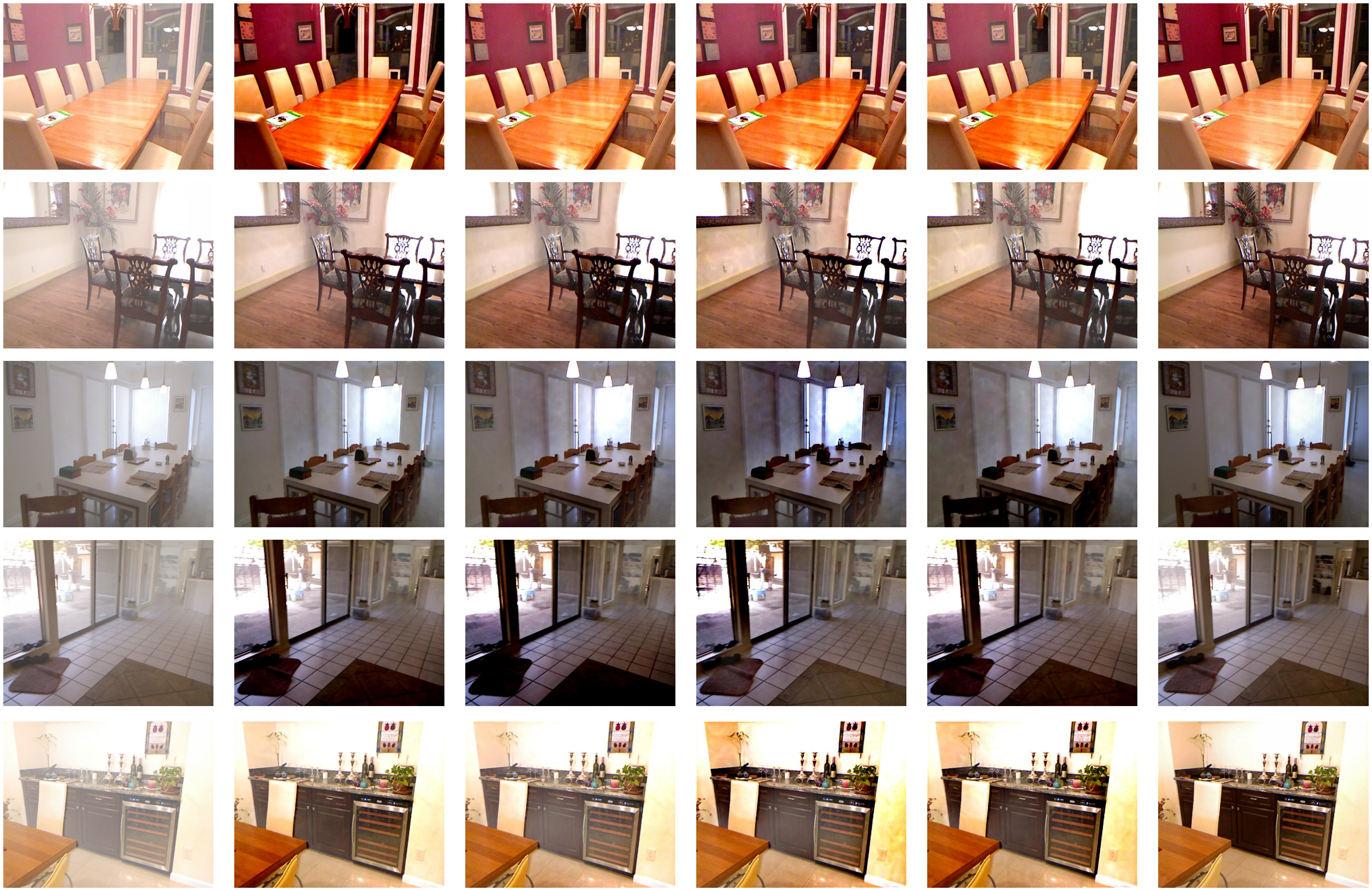
(a) Haze (b) XHOT (c) YOLY (d) ZID (e) XYZ (f) Ground Truth

Figure 4.1: Qualitative comparisons on HSTS Outdoor dataset for different methods.



(a) Haze (b) XHOT (c) YOLY (d) ZID (e) XYZ

Figure 4.2: Qualitative comparisons on HSTS Real World for different methods.



(a) Haze (b) XHOT (c) YOLY (d) ZID (e) XYZ (f) Ground Truth

Figure 4.3: Qualitative comparisons on SOTS Indoor dataset for different methods.

Chapter 5

Conclusions

This chapter presents the conclusions of this research work around of our two methods (XHOT and XYZ) and their evaluation with different images. In addition, the contributions arising from this study and future studies are highlighted.

Section 5.1 is dedicated to detailing the significant achievements made throughout our study. We provide a detailed analysis of the achievements made in each phase of the research, highlighting both the theoretical and practical aspects that have enriched our work. In addition, the research questions initially set out in Chapter 1 are addressed, providing informed answers and evaluating the contribution of our approach to haze removal.

In Section 5.2, the contributions derived from the development of this research are discussed. A detailed analysis of the results obtained and their relevance in the general context of the elimination of haze. In addition, Section 5.3 explores possible directions for future studies, based on the findings and results of this research. These future studies point to opportunities to improve and expand our method, providing a forward-looking vision oriented towards continuous progress in the field.

5.1 Synthesis of Achievements

In this work, we present an approach to haze removal based on disentanglement of complex layers into simpler layers. We have developed an unsupervised method, called XHOT, which is simple and lightweight. While it is important to note that, to date, this method does not surpass the leading approaches in the state of the art in unsupervised haze removal. Additionally, we propose a group disentanglement approach (our main method) using three unsupervised methods: XHOT, YOLY and ZID.

Our main approach, called XYZ, represents an effective strategy that combines the advantages of these three individual methods. XYZ results show significant improvements in image quality metrics, such as SSIM, PSNR, BRISQUE, NIQE, and PIQE, supporting its effectiveness in haze removal. In other words, both approaches presented in this study address the challenge of haze removal in an unsupervised manner. This is especially valuable as we overcome the limitation of the lack of real-world paired images and eliminate the need to train with an extensive dataset. Our methods address the haze problem individually for each image, which represents a significant advance in unsupervised haze removal.

We presented three research questions in Chapter 1, and we provide the following answer relying on the research conducted in this dissertation:

1. *How can existing unsupervised dehazing models be improved and optimized using advanced deep learning techniques?* To improve and optimize unsupervised haze removal models using advanced deep learning techniques, multifaceted strategies can be implemented. First, it seeks to enhance the architecture of the model by using deeper neural networks, with the aim of capturing more effectively the complex characteristics present in images with haze. The incorporation of spatial attention mechanisms allows the model to focus on crucial areas, thus improving the quality of haze removal. In addition, the careful design of loss functions contributes to enhance the visual quality of the resulting images.

   It is essential to note that this research focuses on the elimination of haze from a single image, alleviating the need for a large dataset of paired data. Combining multiple models and exploring advanced hyperparameter settings, along with post-processing techniques, form a comprehensive approach to optimizing unsupervised haze removal models. This approach makes it possible to advance the effectiveness and applicability of models, positioning them as valuable tools in practical and diverse scenarios.

2. *What is the impact of neural network architecture on the performance of unsupervised dehazing models?*

   The impact of neural network architecture on the performance of unsupervised haze removal models is crucial and determinant. The choice of architecture directly influences the model's ability to capture and understand the complex features present in haze images. Incorporating specific layers, such as spatial attention mechanisms, can improve the model's ability to focus on critical areas, thus improving the quality of haze removal. It is essential to consider the complexity of the environment and the diversity of haze conditions when selecting or designing the architecture, as this directly affects the overall adaptability and performance of the model in real-world situations.

3. *How do accurate and relevant evaluation metrics influence the performance of unsupervised haze removal models in real-world conditions?*

   Developing more accurate and relevant evaluation metrics to measure the performance of unsupervised haze removal models under real-world conditions involves considering the complexity of human vision and variability of atmospheric conditions. It is essential to design metrics that accurately reflect human visual perception, incorporating aspects such as sharpness, color fidelity and the preservation of important details. In addition, the assessment should specifically address real-world conditions, where haze may vary in density and distribution. Metrics that integrate the model's ability to improve visibility in different types of scenes, both indoor and outdoor, will be more effective. In this research, we used three well-defined metrics in the literature to evaluate our approach.

5.2 Contributions

This work achieved significant results. Our main contribution was to develop a method that allows haze removal using unsupervised neural networks. The results we obtained demonstrate that our method allows us to eliminate the existing haze in images without compromising its coherence and consistency. In summary, we present the key contributions of this research work as follows:

- The robust XYZ method for haze removal using unsupervised neural networks is a key contribution to any technique that makes use of images taken by digital cameras, from security applications, medical applications, among others.
- XYZ adheres to an unsupervised paradigm, eliminating the need to use paired datasets, reflecting a more realistic approach to dehazing in real-world situations.
- XYZ combines the XHOT, YOLY and ZID methods, taking advantage of the individual strengths of each. It maintains the specific advantages of each method, such as preserving details, exploring the HSV color space, and using loss functions based on DCP Loss.

5.3 Future Work

The XYZ method has made significant progress in dehazing by layer disentangling and combining unsupervised approaches. However, there are opportunities for future research that could further improve its performance and address specific challenges. Some promising areas of research could include:

- Conducting a thorough analysis of the hyper-parameters of the XYZ method represents an essential strategy to optimize its performance in various haze conditions and scenarios. This process involves careful exploration and adjustment of key parameters, with the aim of improving the model's ability to adapt to different degrees of haze density and variations in ambient lighting.
- Exploring more advanced network architectures or variants is presented as a crucial strategy to improve the ability of layer disentangling and generalization of the model in haze removal tasks. The constant advancement in the field of deep learning offers opportunities to adopt more sophisticated network structures, such as residual neural networks or care architectures, which can more effectively capture the complex spatial and feature relationships present in haze images.
- Extending the assessment of the XYZ method to a wider range of environmental conditions is a crucial step in validating its robustness and adaptability in real-world environments. This involves considering variations in lighting, different types

of haze, and various scene configurations. Addressing these aspects will provide a more complete understanding of the method's ability to address challenges in real-world scenarios.

- Exploring additional techniques to attenuate or eliminate residual artifacts in XYZ-generated images is an essential step in improving the method's visual fidelity. These artifacts may arise due to the complexity of the scenes or the presence of dense haze. A potential strategy involves implementing specific post-processing, such as adaptive filters or detail restoration techniques. Moreover, refining the loss functions used during training could help minimize the appearance of artifacts.

- Conducting a thorough assessment of the impact of haze removal through the XYZ method on the performance of computer vision algorithms is essential. This analysis should address critical aspects, such as object detection and facial recognition, to understand how improved image quality directly affects fundamental visual processing tasks.

- The exploration of the potential of GANs is presented as a strategic direction to raise the capacity of the XYZ model in the generation of high quality and realistic images. The addition of GANs could improve the visual fidelity of undone images by introducing a competition between the generator and the discriminator.

Bibliography

[1] S. C. Agrawal and A. S. Jalal. Dense Haze Removal by Nonlinear Transformation. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(2):593–607, 2022. 28

[2] Y. Bengio, A. C. Courville, and P. Vincent. Unsupervised Feature Learning and Deep Learning: A Review and New Perspectives. *CoRR*, abs/1206.5538, 2012. 25

[3] D. Berman, T. Treibitz, and S. Avidan. Non-Local Image Dehazing. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1674–1682, 2016. 14, 27

[4] C. M. Bishop. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, Berlin, Heidelberg, 2006. 23, 25

[5] A. Burkov. *Machine Learning Engineering*, volume 1. True Positive Incorporated Montreal, QC, Canada, 2020. 23

[6] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao. DehazeNet: An End-to-End System for Single Image Haze Removal. *IEEE Transactions on Image Processing*, 25(11): 5187–5198, Nov. 2016. 14, 26, 27

[7] F. Camastra and A. Vinciarelli. *Machine Learning for Audio, Image and Video Analysis: Theory and Applications*. Springer, 2015. 23

[8] X. Chen, Y. Li, C. Kong, and L. Dai. Unpaired Image Dehazing With Physical-Guided Restoration and Depth-Guided Refinement. *IEEE Signal Processing Letters*, 29:587–591, 2022. 27, 28

[9] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, and M.-H. Yang. Multi-Scale Boosted Dehazing Network With Dense Feature Fusion. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2154–2164, 2020. 26

[10] D. Engin, A. Genc, and H. Ekenel. Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pages 938–9388, Los Alamitos, CA, USA, June 2018. IEEE Computer Society. 27, 28

[11] A. Golts, D. Freedman, and M. Elad. Unsupervised Single Image Dehazing Using Dark Channel Prior Loss. *IEEE Transactions on Image Processing*, 29, 2020. 17, 27, 28, 34, 40, 41

[12] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, 2016. 24, 25

[13] S. Hagen. The mind's eye, 2012. URL `https://www.rochester.edu/pr/Review/V74N4/pdf/0402_brainscience.pdf`. Accessed on March 27, 2024. 16

[14] M. Hassaballah and A. I. Awad. *Deep Learning in Computer Vision: Principles and Applications*. CRC Press, 2020. 24

[15] K. He, J. Sun, and X. Tang. Single Image Haze Removal using Dark Channel Prior. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1956–1963, 2009. 23, 26, 27, 38

[16] K. Huang, A. Hussain, Q.-F. Wang, and R. Zhang. *Deep Learning: Fundamentals, Theory and Applications*, volume 2. Springer, 2019. 24

[17] B. Jähne. *Digital Image Processing*. Springer Science & Business Media, 2005. 25

[18] M. Ju, C. Ding, W. Ren, and Y. Yang. IDBP: Image Dehazing Using Blended Priors Including Non-Local, Local, and Global Priors. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(7):4867–4871, 2022. 27, 28

[19] D. Kingma and M. Welling. Auto-Encoding Variational Bayes. *arXiv preprint arXiv:1312.6114*, pages 1–14, 12 2014. 37

[20] C. H. Lampert, H. Nickisch, and S. Harmeling. Attribute-Based Classification for Zero-Shot Visual Object Categorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(3):453–465, 2014. 25

[21] Y. LeCun, Y. Bengio, and G. Hinton. Deep Learning. *Nature*, 521(7553):436–444, May 2015. 25

[22] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng. AOD-Net: All-in-One Dehazing Network. In *IEEE International Conference on Computer Vision*, pages 4780–4788, 2017. 26

[23] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang. Benchmarking Single-Image Dehazing and Beyond. *IEEE Transactions on Image Processing*, 28(1): 492–505, 2019. 21, 31, 47, 62, 63

[24] B. Li, Y. Gou, J. Z. Liu, H. Zhu, J. T. Zhou, and X. Peng. Zero-Shot Image Dehazing. *IEEE Transactions on Image Processing*, 29:8457–8466, 2020. 9, 20, 21, 27, 28, 30, 35, 37, 38, 41, 45, 46, 47, 48, 49

[25] B. Li, Y. Gou, S. Gu, J. Z. Liu, J. T. Zhou, and X. Peng. You Only Look Yourself: Unsupervised and Untrained Single Image Dehazing Neural Network. *International Journal of Computer Vision*, pages 1–14, Mar. 2021. 9, 20, 21, 27, 28, 29, 30, 34, 35, 38, 41, 45, 46, 47, 48, 49

[26] C. Lin, X. Rong, and X. Yu. MSAFF-Net: Multiscale Attention Feature Fusion Networks for Single Image Dehazing and Beyond. *IEEE Transactions on Multimedia*, 25:3089–3100, 2023. 26, 27

[27] X. Liu, Y. Ma, Z. Shi, and J. Chen. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. In *International Conference on Computer Vision*, 2019. 26, 27

[28] P. Maldonado-Quispe and H. Pedrini. Image Dehazing Using a Simple Convolutional Autoencoder. In *IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, pages 1–6, Recife-PE, Brazil, 2023. 26, 27

[29] P. Maldonado-Quispe and H. Pedrini. XYZ Unsupervised Network: A Robust Image Dehazing Approach. In *19th The International Conference on Computer Vision Theory and Applications (VISAPP)*, volume 3, pages 694–701, Rome, Italy, Feb. 2024. 20, 34, 35, 38, 39, 41, 45, 47, 48

[30] E. McCartney. Optics of the Atmosphere – Scattering by Molecules and Particles. *IEEE Journal of Quantum Electronics*, 14(9):698–699, 1978. 22, 26

[31] A. Mittal, A. K. Moorthy, and A. C. Bovik. No-Reference Image Quality Assessment in the Spatial Domain. *IEEE Transactions on Image Processing*, 21(12):4695–4708, 2012. 43

[32] A. Mittal, R. Soundararajan, and A. C. Bovik. Making a "Completely Blind" Image Quality Analyzer. *IEEE Signal Processing Letters*, 20(3):209–212, 2013. 43

[33] K. P. Murphy. *Probabilistic Machine Learning: An Introduction*. MIT Press, 2022. 23

[34] S. Narasimhan and S. Nayar. Contrast Restoration of Weather Degraded Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(6):71–724, 2003. 22

[35] S. Nayar and S. Narasimhan. Vision in Bad Weather. In *Seventh IEEE International Conference on Computer Vision*, volume 2, pages 820–827, 1999. 22

[36] M. Palatucci, D. Pomerleau, G. Hinton, and T. M. Mitchell. Zero-shot learning with semantic output codes. In *22nd International Conference on Neural Information Processing Systems*, NIPS'09, page 1410–1418, Red Hook, NY, USA, 2009. Curran Associates Inc. 25

[37] J. R. Parker. *Algorithms for Image Processing and Computer Vision*. John Wiley & Sons, 2010. 25

[38] I. Pitas. *Digital Image Processing Algorithms and Applications*. John Wiley & Sons, 2000. 25

[39] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia. FFA-Net: Feature Fusion Attention Network for Single Image Dehazing. *AAAI Conference on Artificial Intelligence*, 34 (7):11908–11915, Apr. 2020. 26

[40] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang. Single Image Dehazing via Multi-Scale Convolutional Neural Networks. In B. Leibe, J. Matas, N. Sebe, and M. Welling, editors, *European Conference on Computer Vision*, pages 154–169, Cham, 2016. Springer International Publishing. 24

[41] J. C. Russ. *Introduction to Image Processing and Analysis*. CRC Press, 2006. 25

[42] D. Scharstein and R. Szeliski. High-Accuracy Stereo Depth Maps Using Structured Light. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, volume 1, pages I–I, 2003. 62

[43] N. Sebe. *Machine Learning in Computer Vision*, volume 29. Springer Science & Business Media, 2005. 23

[44] J. Sen, S. Mehtab, and A. Engelbrecht. *Machine Learning: Algorithms, Models and Applications*. BoD–Books on Demand, 2021. 23

[45] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus. Indoor segmentation and support inference from rgbd images. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, and C. Schmid, editors, *European Conference on Computer Vision*, pages 746–760, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. 62, 63

[46] R. Socher, M. Ganjoo, C. D. Manning, and A. Y. Ng. Zero-shot learning through cross-modal transfer. In *26th International Conference on Neural Information Processing Systems - Volume 1*, NIPS'13, page 935–943, Red Hook, NY, USA, 2013. Curran Associates Inc. 25

[47] E. Stevens, L. Antiga, and T. Viehmann. *Deep Learning with PyTorch*. Manning Publications, 2020. 24

[48] R. T. Tan. Visibility in Bad Weather from a Single Image. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2008. 14, 27

[49] N. Venkatanath, D. Praneeth, C.-B. Maruthi, S. S. Channappayya, and S. S. Medasani. Blind Image Quality Evaluation Using Perception Based Features. In *Twenty First National Conference on Communications*, pages 1–6, 2015. 44

[50] E. Wang, S. Shu, and C. Fan. CNN-based Single Image Dehazing via Attention Module. In *IEEE 5th International Conference on Automation, Electronics and Electrical Engineering*, pages 683–687, 2022. 26, 27

[51] W. Wang, V. W. Zheng, H. Yu, and C. Miao. A Survey of Zero-Shot Learning: Settings, Methods, and Applications. *ACM Transactions on Intelligent Systems and Technology*, 10(2):1–37, 2019. 25

[52] Y. Wang, X. Yan, D. Guan, M. Wei, Y. Chen, X.-P. Zhang, and J. Li. Cycle-SNSPGAN: Towards Real-World Image Dehazing via Cycle Spectral Normalized Soft Likelihood Estimation Patch GAN. *IEEE Transactions on Intelligent Transportation Systems*, 23(11):20368–20382, 2022. 27, 28

[53] X. Yang, Z. Xu, and J. Luo. Towards Perceptual Image Dehazing by Physics-Based Disentanglement and Adversarial Training. In *Thirty-Second AAAI Conference on Artificial Intelligence and Thirtieth Innovative Applications of Artificial Intelligence Conference and Eighth AAAI Symposium on Educational Advances in Artificial Intelligence*, New Orleans, Louisiana, USA, 2018. AAAI Press. 27, 28

[54] W. Yi, L. Dong, M. Liu, Y. Zhao, M. Hui, and L. Kong. DCNet: Dual-Cascade Network for Single Image Dehazing. *Neural Computing and Applications*, 34(19): 16771–16783, Oct 2022. 26

[55] X. Yi, B. Ma, Y. Zhang, L. Liu, and J. Wu. Two-Step Image Dehazing with Intra-Domain and Inter-Domain Adaptation. *Neurocomputing*, 485:1–11, 2022. 26

[56] C. Zhang and C. Wu. Multi-Scale Attentive Feature Fusion Network for Single Image Dehazing. In *International Joint Conference on Neural Networks (IJCNN)*, pages 1–7, 2022. 26, 27

[57] H. Zhang and V. M. Patel. Densely Connected Pyramid Dehazing Network. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3194–3203, 2018. 27

[58] Z. Zheng, W. Ren, X. Cao, X. Hu, T. Wang, F. Song, and X. Jia. Ultra-High-Definition Image Dehazing via Multi-Guided Bilateral Learning. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16180–16189, 2021. 26, 27

[59] H. Zhu, X. Peng, V. Chandrasekhar, L. Li, and J.-H. Lim. DehazeGAN: When Image Dehazing Meets Differential Programming. In *27th International Joint Conference on Artificial Intelligence*, pages 1234–1240. AAAI Press, 2018. 27, 28

[60] Q. Zhu, J. Mai, and L. Shao. A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior. *IEEE Transactions on Image Processing*, 24(11):3522–3533, 2015. 23, 29, 40

Appendix A

RESIDE Dataset

A comprehensive overview of RESIDE can be found in Table A.1. Notably, RESIDE stands as the inaugural and only systematic evaluation encompassing various dehazing algorithms, assessing them across multiple criteria on a shared large-scale benchmark notable absence in the existing literature.

Table A.1: Structure of RESIDE Standard and RESIDE-β.

RESIDE Standard			
Subset	**Images**	**Type**	**Annotations**
Indoor Training Set	13,990	Synthetic	No
Synthetic Objective Testing Set	500	Synthetic	No
Hybrid Subjective Testing Set	20	Real	No
RESIDE-β			
Subset	**Images**	**Type**	**Annotations**
Indoor Training Set	72,135	Synthetic	No
Real-world Task-driven Testing Set	4,322	Real	Yes

The RESIDE training set incorporates 13,990 synthetic hazy images, derived from 1,399 clear images originating from well-established indoor depth datasets, namely NYU2 [45] and Middlebury stereo [42]. For each clear image, Li et al. [23] synthesize 10 hazy counterparts, allowing for an optional split of 13,000 for training and 990 for validation. To introduce variability, Li et al. [23] set atmospheric lights A by uniformly selecting each channel from $[0.7, 1.0]$, and β is chosen uniformly at random from $[0.6, 1.8]$. Consequently, the dataset comprises paired clean and hazy images, wherein a single clean ground truth image can yield multiple pairs with hazy counterparts generated under distinct A and β parameters.

The RESIDE testing set consists of the Synthetic Objective Testing Set (SOTS) and the Hybrid Subjective Testing Set (HSTS), strategically designed to represent diverse evaluation perspectives. SOTS includes 500 indoor images from NYU2 [45] (non-overlapping with training images), and follows the same synthesis process as the training data to generate hazy images. Li et al. [23] intentionally created challenging dehazing scenarios for testing, such as scenes with substantial haze added to white scenes. HSTS comprises 10 synthetic outdoor hazy images generated similarly to SOTS, along with 10 real-world hazy images collected from outdoor scenes, providing a comprehensive dataset for human subjective evaluation.

Figure A.1 illustrates some examples of images extracted from subsets in RESIDE Standard and RESIDE-β.

(a) RESIDE Standard

(b) RESIDE-β

Figure A.1: Samples of images from subsets in RESIDE Standard and RESIDE-β.