JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

In-depth Investigation of Single Image Dehazing Methods and their variants: An Extensive Analysis

[1]Chintan Dave, [2]Hetal Patel

[1]Research Scholar, [2]Professor

[1]Gujarat Technological University, Ahmedabad, 382424, India

[2]Department of ECE, A. D. Patel Institute of Technology, The Charutar Vidya Mandal (CVM) University, V.V. Nagar, 388120, India

Abstract: Haze formation is caused by tiny particles suspended in the air that scatter light coming from the scene. As light travels through the atmosphere, it can interact with molecules and particles. The tiny particles, like dust, smoke, or water droplets, bend the light rays in different directions. The scattered light from haze particles creates a whitish or greyish veil across the image, reducing the contrast and color fidelity of the images. Many computer vision tasks rely on clear and accurate image data. Haze disrupts this by obscuring objects and reducing their visibility. Image dehazing improves the quality of the image data, allowing computer vision algorithms to perform tasks like object detection, recognition, and scene understanding more effectively. Real-time applications like self-driving cars heavily rely on image dehazing to avoid accidents in bad weather. The paper summarizes the current state-of-the-art single image dehazing methods and classifies them according to their groups. The paper analyze the methods, significant techniques, and models used to improve the image dehazing task. To conclude, the paper provides a comprehensive comparison of various image dehazing techniques. This comparison includes both quantitative and qualitative evaluations of each method's performance. Additionally, the paper highlights the key strengths and limitations associated with each dehazing approach.

Index Terms **- Single image dehazing, supervised learning, unsupervised learning, generative adversarial network, MRF**

I.INTRODUCTION

Single image dehazing is a technique used to recover a clear image from a hazy one, all based on the information within the single hazy image itself. The amount of haze affecting any given part of the image depends on the distance to that part, known as the depth of the object from the viewing point. This unknown depth creates a variable haze concentration throughout the image. Image dehazing finds application in surveillance and security for object detection and tracking in outdoor environments, enhancing the effectiveness of video surveillance cameras. It plays a crucial role in remote sensing applications such as satellite imagery and aerial photography. It can enhance the clarity of images captured from aerial platforms or satellites, enabling better analysis of land cover, environmental changes, and natural disasters. It can also find application in autonomous vehicles, allowing them to make better decisions and navigate safely.

In the early stages, image dehazing techniques were primarily based on simple image processing operations such as histogram equalization [1], contrast stretching [2], and image polarization [3-4]. These methods aim to enhance the visibility of hazy images by adjusting their pixel intensities or contrast. These methods do not give better performance under different environmental conditions. Tan [5] utilized the concept of higher contrast and smooth airlight for a clear image. Markov Random Field (MRF) is used to model the airlight through the contrast maximizing technique and compute the light chromaticity from airlight. The method is intended for contrast enhancement rather than restoration. Fattal [6] uses Independent Component Analysis (ICA) with the assumption that surface shading and transmission map are locally uncorrelated. This method struggles with scenes containing complex lighting conditions and dense haze. Tarel et al. [7] proposed a faster approach using median filtering instead of the computationally expensive soft mating process. However, this method suffers when the depth of the scene varies significantly. He et al. [8] leverage a key observation about outdoor haze-free images that at least one color channel in most local image patches will have very low intensity, known as Dark Channel Prior (DCP). This inherent property allows us to estimate the haze thickness across the image, ultimately leading to a clearer restored image using the atmospheric scattering model. The method fails to produce better results in bright regions like the sky. To overcome this issue, region-wise transmission map estimation is proposed by Yuan et al. [9] to estimate transmission map separately for the sky and non-sky regions using color characteristics of the sky region and the dark channel prior, respectively. The segmentation of a sky region is performed using quad-tree splitting based feature pixel detection, and then the transmission map in the sky region is estimated using color characteristic observation of the sky and edge-preserving guided filtering. Zhu et al. [10] introduced a novel approach known as Color Attenuation Prior(CAP) that establishes a relationship between haze concentration, contrast and brightness within the image.

The emergence of large image datasets [11-13] has revolutionized the field of image dehazing. These extensive datasets, containing vast numbers of hazy and corresponding clear image or transmission map pairs, have accelerated the development of learning-based approaches. By training on these datasets, the methods have learned complex patterns of haze formation, which leads to a dramatic improvement in dehazing performance compared to traditional techniques. CNN based image dehazing

algorithms like Dehazenet [14], Multi-Scale Convolutional Neural Network (MSCNN) [15] and All-in-One Dehazing Network (AOD-Net) [16] establish the pioneering work in the haze removal process using CNN. Image dehazing has shown enhanced performance with the application of Generative Adversarial Networks (GANs) [17]. The network proposed in [18-20] used GAN to estimate the transmission map from the input image, and an atmospheric scattering model utilized to recover the clear image. The network present in [21-24] used GAN to directly estimate the clear image from the hazy image without estimating any intermediate parameters. Zhang et al. [19] introduced a novel image dehazing approach using a densely connected pyramid structure. This network jointly estimates the transmission map and atmospheric light to recover a clear image through an atmospheric scattering model. To further enhance the results, the authors incorporated a joint discriminator that takes both the estimated transmission map and the dehazed image as input. Notably, the encoder part of the generator leverages the initial layers of a pre-trained DenseNet-121 [25] model for improved efficiency. It is noticed that there is a major change in Y-plane compared to Cb-plane and Cr-plane when image is affected by the haze. Based on that observation, [26-27] utilized the Y-plane of a clear and hazy image after conversion from RGB to YCbCr-color space. After conversion, the hazy image Y-plane is given to the network, and it is trained to generate the Y-plane similar to a clear image. The generated output is combined with the CbCr-plane and transformed to the RGB-plane to obtain a dehaze image. Zhu et al. [28] proposed a unique network named DehazeGAN, which transforms the atmospheric scattering model into a unique generative adversarial network by applying the concept of differential programming. Zhang et al. [29] used the concept of [16] to estimate the single variable K, which integrates the transmission map and atmospheric light into the single variable. The proposed FAMED-Net architecture used a multi-scale encoder and fusion module to restore the clear image. Attention based image dehazing networks FFA-Net [30] and GridDehazeNet [31] show significant improvement in terms of qualitative and quantitative performance with color restoration. Guo et al. [32] first time utilize the transformer module with CNN in image dehazing, which not only gives the haze location but also gives density at different locations. The architecture has an encoder-decoder structure with a feature modulation module as a sandwich block that combines the features of the encoder and transformer before feeding them to the decoder block. Zhen et al. [33] introduce UVM-Net, which merges the strengths of CNN for capturing local details and State Space Sequence Models to understand the long-range dependencies in the image.

II. BACKGROUND

The hazy image formation is represented by the atmospheric scattering model [5], [6], [34-36] given by:

$$\mathrm{I}(x,y) = \mathrm{J}(x,y) * t(x,y) + \mathrm{A} * \big(1 - t(x,y)\big) \qquad (2.1)$$

Where *I(x, y)* is the hazy image, *J(x, y)* is the scene radiance, *t(x, y)* is a transmission map, and *A* is the atmospheric constant. The first part of Equation 2.1 is called direct transmission, and the additive term is known as airlight. The objective is to retrieve the scene radiance from the haze image. To achieve this, we need to estimate the t(x, y) and A and invert the model expressed in the Equation 2.1. In a homogeneous environment, *t(x, y)* can be expressed as:

$$t(x,y) = e^{-\beta * d(x,y)} \qquad (2.2)$$

Where β is the scattering coefficient of the atmosphere and depth *d(x, y)* gives the object distance from the camera point of view. This equation tells us that scene radiance decreases exponentially as the depth of the image increases. Consequently, if we know the scene depth for every point, we could directly estimate the transmission map. As hazy image is known to us, haze-free image can be obtained by estimating *t(x,y)* and *A*.

III. IMAGE DEHAZING METHODS

Single image dehazing methods tackle the challenge of restoring visibility and clarity in images obscured by atmospheric haze. These methods play a crucial role not only in enhancing image quality but also as a vital preprocessing step for various computer vision applications. Single image dehazing methods can be broadly classified into two main categories, as shown in Figure 1. (A) Image enhancement and prior-based methods; and (B) Learning-based methods. The state-of-the-art methods proposed by Tan [5], Fattal [6], He et al. [8], and Zhu et al. [10] laid the foundation for image dehazing and fall under the categories of image enhancement and prior based methods. These methods showed significant improvement in the resultant image before the learning-based approaches are introduced. The availability of large datasets has played a crucial role in the learning-based approach. By analyzing large amounts of data, this learning base approach shows dramatically improved ability to identify complex patterns and generate the desired output. As the networks are trained using dataset images, their weights are modified according to the generated loss

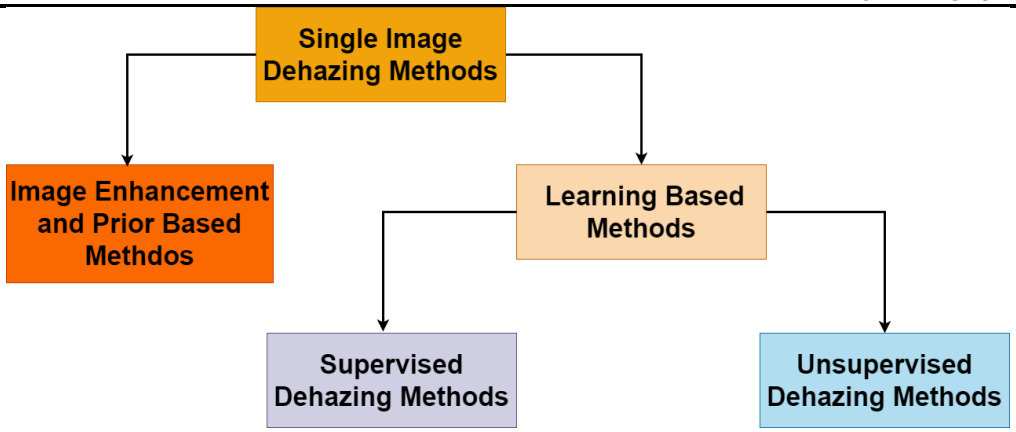


Figure 1. Classification of Single Image Dehazing Methods

between the predicted output of the learning network and the ground truth image. The method proposed in [14-16], [19], [21], [26], [28], [30], [37] used a supervised learning approach in which a network is trained using a hazy image and its corresponding ground truth image or transmission map to predict the desired output. An unsupervised learning approach eliminates the need for a pre-existing dataset containing paired hazy and clean images or a transmission map for training. The methods proposed in [24], [38-42] fall under the unsupervised image dehazing methods, where methods have either used unpaired images from the dataset of clear and hazy images or used prior information from the input image to train the network. The following section outlines the most notable and effective methods for single image dehazing, which have revolutionized the field by providing novel approaches to enhancing image quality in single hazy images.

3.1 Image Enhancement and Prior Based Methods for Image Dehazing

This section focuses on prior-based methods that recovered the haze-free image using physical scattering models or image enhancement techniques with various observations and assumptions. The state-of-the-art methods proposed by Tan [5], Fattal [6], He et al. [8], and Zhu et al. [10] have demonstrated substantial enhancements in resultant image quality prior to the introduction of learning-based approaches.

Tan [5] utilized the concept of higher contrast and smooth airlight for a clear image. Markov Random Field (MRF) is used to model the airlight through contrast maximizing technique and compute the light chromaticity of the airlight. In this proposed method, the data cost and smoothness cost for each pixel build up a complete MRF. After optimization with the inference method, the airlight is estimated, which is used to compute direct attenuation that represents enhanced scene visibility. The method is intended for contrast enhancement rather than restoration. The proposed method creates a halo effect near depth discontinuities because of patch operation. The actual value of the atmospheric constant is not known, and because of that, oversaturation is observed, but the visibility of the image is considerably improved.

In the method proposed by Fattal [6], it is assumed that lighting variations and the transmission are not directly related to each other. The author split the unknown image J into two components. First is surface albedo (R), which refers to the coefficient of reflectance. The second component, the shading factor (l), gives the amount of light coming to the view point. R is further divided into two components, which are parallel and orthogonal to atmospheric light. The transmission map is estimated by using the projection of the input image onto atmospheric light and the surface albedo, which is orthogonal to atmospheric light. The Independent Component Analysis (ICA) method is followed to find the parameter. The author performs the noise estimation present in the image and applies statistical extrapolation to withstand large errors. Images produced by this method are clear and have regained their contrast. Some of the drawbacks of this method are as follows: Its performance highly depends on the quality of the input image, as a statistical property is used to estimate the parameters. Insufficient signal-to-noise ratio will lead this method to failure, and it will also not work for images containing dense fog.

He et al. [8] proposed a dehazing method that was based on statistics of the outdoor haze-free images. This method performs very well on outdoor hazy images, except for the image region having a scene radiance similar to an airlight. The authors found that in an outdoor clear image, one channel of the RGB plane has a very low value. This channel is named Dark Channel. This dark channel has a good correlation with the transmission map. The concept of a dark channel is expressed as follows [8]:

$$\mathrm{J}^{\mathrm{dark}}(\mathrm{p}) = \min_{\mathrm{c}\epsilon\{\mathrm{r,g,b}\}} \left(\min_{n\epsilon\Omega(p)} J^{c}(q) \right) \qquad (3.1)$$

Where *Jc* is a color channel of *J* and *Ω(p)* denotes a patch centered at pixel p. Based on the observation of a clear outdoor image, except for some conditions like the sky, Equation 3.1 can be written as:

$$J^{dark} \rightarrow 0 \qquad (3.2)$$

This observation is referred to as a dark channel prior (DCP). The intensity of the dark channel provides an idea of haze density in a particular area. Utilizing the DCP and Equation 2.1, the transmission map can be represented as [8]:

$$\tilde{t} = 1 - \omega * \min_{c\epsilon\{r,g,b\}} \left(\min_{n\epsilon\Omega(p)} \frac{I^c(q)}{A^c} \right) \quad (3.3)$$

Here, the local patch was assumed to be constant. Equation 3.3 is derived by taking the minimum operation on both sides after normalizing Equation 2.1. The second term in Equation 3.3 is the dark channel of the normalized hazy image. In bright regions like the sky, the dark channel value is very similar to A, and because of that second term of Equation 3.3 goes to near 0 which leads to wrong estimation. The soft matting process was introduced further, which managed to capture sharp edges with discontinuities in the depth map. From the dark channel, 0.1% of the brightest pixels are selected, and among these, a pixel having the highest intensity in the input image is selected as atmospheric light. The reformulation of Equation 2.1 to find *J(x,y)* can be expressed as:

$$\mathrm{J}(x,y) = \frac{(I(x,y) - A)}{max(t(x,y), t_0)} + \mathrm{A} \quad (3.4)$$

t_0 is a lower bound of the transmission map t with a constant value of 0.1. It is observed that a moderate patch size is also essential for this method, as it directly affects the output by creating a variance in depth. This method outperforms the previous state-of-the-art methods. However, it fails to handle bright regions like the sky, where dark channel intensity is near the atmospheric light.

Zhu et al. [10] proposed a linear model to predict scene depth. Author utilize the fact of a direct correlation between scene radiance and both saturation and contrast. A rise in haze concentration leads to a proportional increase in the influence of airlight and a corresponding reduction in scene radiance. The denser the haze, the stronger the influence of airlight on hazy image formation. The relation between them is expressed as [10]:

$$d(x) \propto c(x) \propto v(x) - s(x) \quad (3.5)$$

Where d is the scene depth, c is the concentration of haze, v is the brightness, and s is the saturation of the scene. This represents the Color Attenuation Prior (CAP). To achieve depth restoration, the authors leverage the HSV color space. Within the HSV space, they define an angle α as the projection of image I between the saturation (s) and value (v) channels. This angle, ranging from 0 to 90 degrees, correlates with the scene depth. Building upon this observation, the authors in their work designed a specific linear model with edge preserving property [10].

$$d = \theta_0 + \theta_1 v(x) + \theta_2 s + \epsilon(x) \quad (3.6)$$

where θ_0, θ_1, and θ_2 are the unknown coefficients that are estimated by the supervised learning method. A random variable, $\epsilon(x)$, accounts for potential errors in the model's predictions. By taking the gradient of Equation 3.5, the authors demonstrate the method's ability to preserve edges while recovering depth information, even near sharp depth transitions in the image. To address the issue of white objects, or those with high brightness and low saturation, scene depth is estimated by considering each pixel's neighborhood. Finally, a technique called guided image filtering is employed to refine the resulting depth map. The haze-free image is recovered by inverting Equation 1 for scene radiance. This method outperforms the previous state-of-the-art methods proposed in [8], [7] in terms of performance and running time. The major drawback of this model is that when the haze is dense, it fails to give accurate results. The method assumes a constant scattering coefficient (β) of 1. However, this value varies depending on the severity of the haze.

3.2 Learning Based Methods for Image Dehazing

Machine learning-based methods use convolutional neural networks or recently widely used generative adversarial networks for single image dehazing. There are two approaches to single image dehazing. In the first approach, the estimation of transmission is achieved by CNN, and the atmospheric constant is estimated by the traditional method mentioned in DCP [8]. These estimated transmission map and atmospheric constant are utilized in the atmospheric scattering model to recover haze free image. In the second end-to-end approach, the network directly gives a haze-free image without utilizing the scattering model. Recently, the GAN based model has shown significant improvement in quantitative and qualitative performance, with the dehaze image very close to the ground truth image. But on the other hand, the GAN-based model finds it hard to train and suffers from non-convergence, mode collapse, and training time issues compared to CNN-based methods. As shown in Fig. 1, learning-based methods are mainly classified into two categories. In supervised learning, a paired dataset of hazy images and their ground truth images, or transmission maps, is utilized to train the network parameters. An unsupervised learning approach eliminates the need for paired datasets of clear and hazy image or their transmission maps. The subsections 3.2.1 and 3.2.2 provide a brief overview of supervised and unsupervised learning-based methods, respectively, for single image dehazing.

3.2.1 Supervised Single Image Dehazing Methods

Cai et al. [14] proposed an end-to-end single image dehazing network called Dehazenet, in which layers of the convolutional neural network are specially designed to embody priors. Layers of Maxout units are specifically used for feature extraction; they extract almost all the haze relevant features. By using a novel bilateral rectified linear unit called BReLU, the quality of the output image has been improved. Dehazenet comprises four specially designed layers that provide sequential operations to generate the transmission map from the hazy image patch. The first layer extracts the haze relevant feature from the hazy image patch of size 16x16 through a convolution filter of size 5x5. Inspired by extremum processing, the Maxout unit transformed it into a new feature map through dimensionality reduction. Inspired by GoogLeNet [43], the second layer consists of 16 parallel filters with different scales of 3x3, 5x5, and 7x7 for multi-scale mapping. In the third layer, spatial invariance is achieved

by the max pool operation using a filter size of 7x7 to obtain the local extremum operation. In the fourth layer, nonlinear regression is achieved through a novel bilateral rectified linear unit called BReLU, which has a better convergence rate than ReLU and also makes a significant improvement in performance. This activation function is intended to bound the transmission map in a given range. The estimated transmission map and atmospheric constant, whose value is set to 1, are used in the atmospheric scattering model to obtain a haze-free image. This method is more robust compared to the previously proposed methods in case of different haze conditions. In terms of quality, the image recovered from Dehazenet is not oversaturated and retains the color of the sky region in the dehazing process.

Zhang et al. [19] proposed a method called Densely Connected Pyramid Dehazing Network (DCPDN), which is a joint discriminator based generative adversarial network having an edge-preserving densely connected encoder-decoder structure with a multi-level pyramid pooling module. A dense block is used in the network to maximize the information flow, and the pyramid pooling structure refines the feature learning process. Generator consists of encoder-decoder architecture in which a few layers of Densenet [25] architecture are incorporated in the encoder part with pre-defined weights. At the end, multilevel pooling of the output feature map with pool sizes of 1/4, 1/8,1/16, and 1/32 is used, which is upsampled to the original output feature size. The outputs from the multilevel pooling after resizing are concatenated with the output feature map. The discriminator is trained jointly with the estimated transmission map and dehazed image. DCPDN is trained by an overall loss function that includes an edge-preserving loss to overcome the issue of blurred results due to inaccurate estimation of the transmission map. The total loss includes edge-preserving loss, L2 loss, joint discriminator loss, and dehazing loss. There are multiple losses considered to train the model, which hardens the tuning of parameters for different losses. This method preserves a sharper contour with less color distortion, and the output is closer to the ground truth image. On the real-world dataset, this method also provides visually pleasing results.

Liu et al. [31] proposed ”GridDehazeNet” with attention mechanisms incorporated in the form of a grid. It contains three serial modules named pre-processing, backbone, and post-processing. The first module transforms the input hazy image into sixteen feature maps, representing the hazy image from different angles of view. The backbone is a refined GridNet [44] architecture in the form of a 3x6 matrix with different scales (16, 32, 64) in each row with 5 residual dense blocks. The learned features of each scale are transferred through the column after up/down sampling. The post-processing module, with the same structure as the first module, is designed to nullify the effects of artifacts created by the backbone network. The weight used to fuse the features of different scales is trainable, as all of them are not of the same importance. The L1 loss and perceptual loss are utilized to train the network parameters.

Qin et al. [30] designed ”FFA-Net” with residual connection followed by an attention mechanism to solve the issue of varying haze concentration in the image. The residual connections are employed to pass the low frequency haze free feature and concentrate on hazy features related to high frequencies. These features are transferred to the channel attention module to retrieve the global information, as each channel has different weighted information. After that, it passed to the pixel attention module to learn the haze level information at the pixel level. There is a major improvement observed in the PSNR value compared to the state-of-the-art methods.

3.2.2 Unsupervised Single Image Dehazing Methods

Golts et al. [38] proposed ”Deep DCP” with an unsupervised approach, which eliminates the need for a paired dataset of clear and hazy image or transmission map. It utilize the dark channel prior loss to estimate the refined transmission map. The dark channel prior is applied to the hazy image to obtain the raw transmission map. The objective is to minimize the DCP energy function so that network parameters can be learned to generate a refined transmission map. The predicted transmission map and raw transmission map is utilize in the DCP energy function to minimize the smooth loss and data loss. There are 6 blocks with 2 convolution layers and 1 dilated convolution layer, with an exponentially increasing dilated rate to increase the receptive field size. The residual connection from the input of each block to the output helps to propagate the low level information in the estimation of the refined transmission map. It is a novel approach to integrate the concept of DCP to train the network parameters.

Li et al. [24] proposed a workaround solution for the unavailability of real world paired hazy and clear image datasets. The authors proposed the novel ”USID-Net,” which learns from the unpaired image dataset to give a haze-free image directly without estimating any intermediate parameters. The direct estimation of the haze-free image has a less inference time enabling it to be used in real time applications. The architecture utilizes the concept of CycleGAN [45] with integrating encoders before the generator. The architecture consists of two encoders, one OctEncoder, two generators, and two discriminators, in which the combined encoder and generator block is named the multi-scale feature attention module. The encoder job is to find out content-related information, and the OctEncoder job is to find out haze details from the given input hazy image. Like FFA-Net [30] channel-wise and pixel wise attention mechanisms are developed to learn global details and local haze level details, respectively.

IV. COMPARISON OF STATE-OF-THE-ARTS METHODS EXPERIMENTAL RESULTS:

4.1 Quantitative Analysis:

Table 1 shows the quantitative analysis for state-of-the-art methods categorized according to prior, supervised, and unsupervised methods. The PSNR and SSIM are used as the evaluation parameters on the SOTS indoor and SOTS outdoor RESIDE datasets. The quantitative performance of DCP and CAP is lower due to its reliance on specific prior information. From the table, it is observed that the learning-based method has shown significant improvement in performance compared to prior based methods. In the early stages of single image dehazing using machine learning, DehazeNet achieved superior results by surpassing the limitations of prior-based methods. The advancements in attention-based approaches like GridDehazeNet and FFA-Net have shown dramatic changes in PSNR and SSIM. FFA-Net achieves the highest PSNR and SSIM for both the indoor and outdoor datasets. In terms of

SSIM, GRIDDehazeNet performance is comparable to FFA-Net. In the unsupervised learning based approach, USID showed better results on the indoor image dataset, while Deep DCP exhibited better results on outdoor image datasets.

Table 1. Quantitative result analysis of single image dehazing methods

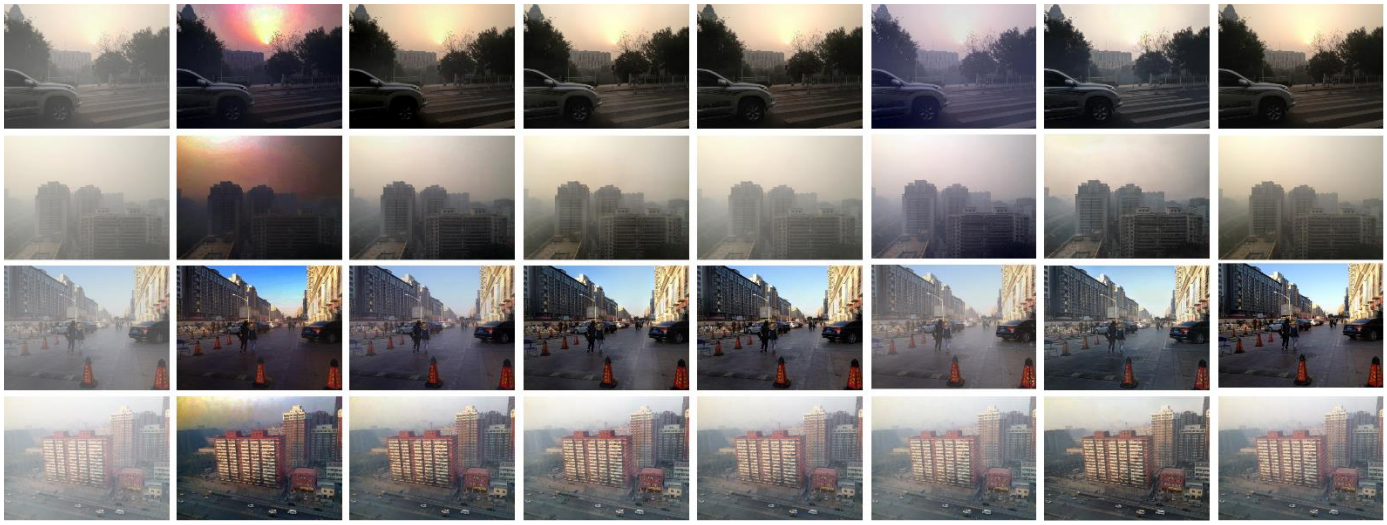
Type	Methods	SOTS-indoor		SOTS-outdoor	
		PSNR	SSIM	PSNR	SSIM
Prior	DCP	20.15	0.872	17.56	0.822
	CAP	19.05	0.8364	18.12	0.757
Supervised	AOD-Net	19.04	0.8504	22.71	0.9112
	DehazeNet	21.14	0.8472	22.57	0.863
	GFN	22.3	0.88	21.25	0.84
	GRID-DehazeNet	32.16	0.9836	30.86	0.9819
	FFA-Net	36.39	0.9886	33.57	0.984
Unsupervised	DeepDCP	19.25	0.832	24.08	0.933
	USID	21.41	0.8947	23.89	0.919

4.2 Qualitative Analysis:

Fig. 2 and 3 show the qualitative analysis of DCP [8], DehazeNet [14], GridDehazeNet [31], FFA-Net [30], Deep DCP [38] and USID-Net [24] on indoor and outdoor RESIDE dataset images, respectively. The performance comparison in dense haze conditions is depicted in the first, third, and fourth rows of Fig. [1]. DehazeNet and Deep DCP can struggle to completely remove haze from images. The second last column of Fig. [1] and Fig. [2] shows that USID-Net introduces a shading effect in dehaze image. The second column of Fig. 3 illustrates the darkening effect and distortion in sky region of DCP on a haze-free image. From the

Hazy image DCP DehazeNet GRIDDehaze Net FFA-Net Deep DCP USID-Net Ground Truth

Figure 2. Qualitative result analysis of SOTS indoor images. Prior based method: DCP, Supervised Methods: DehazeNet, GRIDDehazeNet, FFA-Net, Unsupervised Methods: Deep DCP, USID-Net



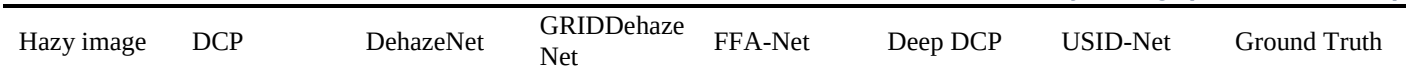


Figure 3. Qualitative result analysis of SOTS outdoor images. Prior based method: DCP, Supervised Methods: DehazeNet, GRIDDehazeNet, FFA-Net, Unsupervised Methods: Deep DCP, USID-Net

analysis, it is observed that DCP produces a better result except in bright regions like the sky. For light haze conditions, the output image looks more saturated. DehazeNet gives better color restoration compared to DCP but does not remove haze completely from dense haze areas. GRIDDehazeNet and FFA-Net surpass the quality of dehaze with the previous state-of-the-art methods. GRIDDehazeNet produces the haze-free image close to ground truth with finer details. FFA-Net is remarkably effective at restoring image details and achieving true color information.

V. CONCLUSION:

The paper investigates major state-of-the-art single image dehazing methods and classifies them according to their approach to the single image dehazing. The enhancement and prior-based methods fail to generate satisfactory results in cases of dense haze and different lighting conditions. DCP came up with pioneering work and was used as the basic building block of many image dehazing methods. DehazeNet cleverly embedded the prior in deep architecture but degraded the output in cases of dense haze. In supervised learning, GRIDDehazeNet and FFA-Net architectures with attention mechanisms find a way to concentrate on haze density at the pixel level to generate the dehaze output nearer to the ground truth image. In an unsupervised approach, Deep DCP and USID-Net demonstrate a novel idea of dehazing without using any paired dataset. Despite unsupervised learning, the results are quite good for outdoor image datasets. In terms of speed, networks like GRIDDehazeNet, FFA-Net, and USID-Net that directly restore the haze-free image from the input haze image require less inference time compared to methods like DehazeNet, DCPDN, and Deep DCP that calculate the transmission map and restore the image using an atmospheric scattering model.

REFERENCES:

[1] Z. Xu, X. Liu, and N. Ji, “Fog Removal from Color Images using Contrast Limited Adaptive Histogram Equalization,” 2nd International Congress on Image and Signal Processing, pp. 1–5, 2009.
[2] S. G. Narasimhan and S. K. Nayar, “Contrast Restoration of Weather Degraded Images,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 6, pp. 713–724, 2003.
[3] Y. Y. Schechner, S. G. Narasimhan, and S. K. Nayar, “Instant Dehazing of Images Using Polarization,” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 325–332, 2001.
[4] S. Shwartz, E. Namer, Y. Y. Schechner, T. Israel, and I. Technology, “Blind Haze Separation,” IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2, pp. 1984–1991, 2006.
[5] R. Tan, “Visibility in Bad Weather,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8, 2008.
[6] R. Fattal, “Single Image Dehazing,” ACM Transactions on Graphics, vol. 27, no. 3, p. 1, 2008.
[7] J.-P. Tarel and N. Hautiere, “Fast Visibility Restoration from a Single Color or Gray Level Image,” 2009 IEEE 12th International Conference on Computer Vision, pp. 2201–2208, 2009.
[8] K. He, J. Sun, and X. Tang, “Single Image Haze Removal Using Dark Channel Prior,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 12, pp. 2341–2353, 2011.
[9] H. Yuan, C. Liu, Z. Guo, and Z. Sun, “A Region-wised Medium Transmission Based Image Dehazing Method,” IEEE Access, vol. 5, pp. 1735–1742, 2017.
[10] Q. Zhu, J. Mai, L. Shao, and S. Member, “A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior,” IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 3522–3533, 2015.
[11] C. Ancuti, C. O. Ancuti, and C. D. Vleeschouwer, “D-HAZY: A Dataset to Evaluate Quantitatively Dehazing Algorithms,” IEEE International Conference on Image Processing, pp. 2226–2230, 2016.
[12] 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, vol. 28, no. 1, pp. 492–505, 2019.
[13] N. Silberman, D. Hoiem, P. Kohli, and R. Fergus, “Indoor Segmentation and Support Inference from RGBD Images,” In European Conference on Computer Vision, pp. 746–760, 2012.
[14] 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, vol. 25, no. 11, pp. 5187–5198, 2016.
[15] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M. H. Yang, “Single Image Dehazing via Multi-scale Convolutional Neural Networks,” European Conference on Computer Vision, pp. 154–169, 2016.
[16] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, “Aod-net: All-in-One Dehazing Network,” IEEE International Conference on Computer Vision, pp. 4770–4778, 2017.
[17] A. C. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D.Warde-Farley, S. Ozair and Y. Bengio, “Generative Adversarial Nets,” In Advances in Neural Information Processing Systems (NIPS), pp. 2672–2680, 2014.
[18] Y. Pang, J. Xie, and X. Li, “Visual Haze Removal by a Unified Generative,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 29, no. 11, pp. 3211–3221, 2019.
[19] H. Zhang and V. M. Patel, “Densely Connected Pyramid Dehazing Network,” The IEEE Conference on Computer Vision and Pattern Recognition, pp. 3194–3203, 2018.
[20] R. Mondal, S. Santra, and B. Chanda, “Image Dehazing by Joint Estimation of Transmittance and Airlight using Bi Directional Consistency Loss Minimized FCN,” IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1033–1041, 2018.

[21] S. Ki, H. Sim, J. Choi, S. Y. Kim, S. Seo, S. Kim, and M. Kim, "Fully End-to-End learning based Conditional Boundary Equlibrium GAN with Receptive Field Sizes Enlarged for Single Ultra-High Resolution Image Dehazing," IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 930–937, 2018.
[22] H. Sim, S. Ki, J.-S. Choi, S. Seo, S. Kim, and M. Kim, "High-Resolution Image Dehazing with respect to Training Losses and Receptive Field Sizes," IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 1025–1032,2018
[23] R. Li, J. Pan, Z. Li, and J. Tang, "Single Image Dehazing via Conditional Generative Adversarial Network," The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 8202–8211, 2018.
[24] J. Li, Y. Li, L. Zhuo, L. Kuang, and T. Yu, "USID-Net: Unsupervised Single Image Dehazing Network via Disentangled Representations," IEEE Transactions on Multimedia, 2022.
[25] G. Huang and K. Q. Weinberger, "Densely Connected Convolutional Networks," IEEE Conference on Computer Vision and Pattern Recognition, pp. 2261–2269, 2017.
[26] A. Wang, W. Wang, J. Liu, and N. Gu, "AIPNet : Image-to-Image Single Image Dehazing with Atmospheric Illumination Prior," IEEE Transactions on Image Processing, vol. PP, no. c, p. 1, 2018.
[27] A. Dudhane and S. Murala, "RYF-Net: Deep Fusion Network for Single Image Haze Removal," IEEE Transactions on Image Processing, vol. 29, pp. 628–640, 2020.
[28] H. Zhu, X. Peng, V. Chandrasekhar, L. Li, and J.-h. Lim, "DehazeGAN : When Image Dehazing Meets Differential Programming," pp. 1234–1240, 2017.
[29] J. Zhang and D. Tao, "FAMED-Net : A Fast and Accurate Multi-scale End-to-end Dehazing Network," vol. 14, no. 8, pp. 1–13, 2018.
[30] X. Qin, Z. Wang, Y. Bai, X. Xie, and H. Jia, "FFA-Net: Feature Fusion Attention Network for Single Image Dehazing," Thirty-Fourth AAAI Conference on Artificial Intelligence, vol. 34, pp. 11908—-11915, 2020.
[31] J. Chen, "GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing", ICCV, 2019.
[32] C. Guo, Q. Yan, S. Anwar, R. Cong, W. Ren, and C. Li, "Image Dehazing Transformer with Transmission-Aware 3D Position Embedding," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 2022-June, pp. 5802–5810, 2022.
[33] Z. Zheng and C. Wu, "U-shaped Vision Mamba for Single Image Dehazing," arXiv preprint arXiv:2402.04139, 2024.
[34] E. Cozman, F. and Krotkov, "Depth from Scattering," Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 801–806, 1997.
[35] S. G. Narasimhan and S. K. Nayar, "Vision and the atmosphere," International Journal of Computer Vision, vol. 48, no. 3, pp. 233–254, 2002.
[36] S. G. Narasimhan and S. K. Nayar, "Chromatic framework for vision in bad weather," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 598–605, 2000.
[37] W. Ren, L. Ma, J. Zhang, J. Pan, X. Cao, W. Liu, and M.-H. Yang, "Gated Fusion Network for Single Image Dehazing,IEEE Conference on Computer Vision and Pattern Recognition, pp. 3253–3261, 2018.
[38] A. Golts, D. Freedman, and M. Elad, "Unsupervised Single Image Dehazing Using Dark Channel Prior Loss," IEEE Transactions on Image Processing, vol. 29, no. October, pp. 2692–2701, 2020.
[39] 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, vol. 129, no. 5, pp. 1754–1767, 2021.
[40] B. Li, Y. Gou, J. Z. Liu, H. Zhu, J. T. Zhou, and X. Peng, "Zero-Shot Image Dehazing," IEEE Transactions on Image Processing, vol. 29, no. 8,pp. 8457–8466, 2020.
[41] D. Engin, A. Genc¸, and H. K. Ekenel, "Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing," The IEEE Conference on Computer Vision and Pattern Recognition (Work- shops, pp. 938–946, 2018.
[42] X. Yang, Z. Xu, and J. Luo, "Towards Perceptual Image Dehazing by Physics-based Disentanglement and Adversarial Training," 32nd dt, pp. 7485–7492, 2018.
[43] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 07-12-June, pp. 1–9, 2015.
[44] D. Fourure, R. Emonet, E. Fromont, D. Muselet, A. Tremeau, and C. Wolf, "Residual conv-deconv grid network for semantic segmentation," British Machine Vision Conference 2017, BMVC 2017, 2017.
[45] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," 2017 IEEE International Conference on Computer Vision, pp. 2242–2251, 2017.