



Deep Variational Bayesian Modeling of Haze Degradation Process

Eun Woo Im*

Department of Artificial Intelligence
Hanyang University
Seoul, Republic of Korea
iameuandyou@hanyang.ac.kr

Sungyong Baik

Department of Data Science
Hanyang University
Seoul, Republic of Korea
dsybaik@hanyang.ac.kr

Junsung Shin*

Department of Artificial Intelligence
Hanyang University
Seoul, Republic of Korea
junsung6140@hanyang.ac.kr

Tae Hyun Kim[†]

Department of Computer Science
Hanyang University
Seoul, Republic of Korea
taehyunkim@hanyang.ac.kr

ABSTRACT

Relying on the representation power of neural networks, most recent works have often neglected several factors involved in haze degradation, such as transmission (the amount of light reaching an observer from a scene over distance) and atmospheric light. These factors are generally unknown, making dehazing problems ill-posed and creating inherent uncertainties. To account for such uncertainties and factors involved in haze degradation, we introduce a variational Bayesian framework for single image dehazing. We propose to take not only a clean image and but also transmission map as latent variables, the posterior distributions of which are parameterized by corresponding neural networks: dehazing and transmission networks, respectively. Based on a physical model for haze degradation, our variational Bayesian framework leads to a new objective function that encourages the cooperation between them, facilitating the joint training of and thereby boosting the performance of each other. In our framework, a dehazing network can estimate a clean image independently of a transmission map estimation during inference, introducing no overhead. Furthermore, our model-agnostic framework can be seamlessly incorporated with other existing dehazing networks, greatly enhancing the performance consistently across datasets and models.

CCS CONCEPTS

• Computing methodologies → Computational photography.

KEYWORDS

Image dehazing; Variational Bayesian method; Computer vision; Machine learning; Image processing

*Equal contribution.

[†]Corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
CIKM '23, October 21–25, 2023, Birmingham, United Kingdom
© 2023 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 979-8-4007-0124-5/23/10...\$15.00
<https://doi.org/10.1145/3583780.3614838>

ACM Reference Format:

Eun Woo Im, Junsung Shin, Sungyong Baik, and Tae Hyun Kim. 2023. Deep Variational Bayesian Modeling of Haze Degradation Process. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management (CIKM '23)*, October 21–25, 2023, Birmingham, United Kingdom. ACM, New York, NY, USA, 10 pages. <https://doi.org/10.1145/3583780.3614838>

1 INTRODUCTION

Haze is an atmospheric phenomenon, where airborne particles (e.g., fog, dust, etc.) between the scene and an observer obscure the scene. Such phenomenon causes poor visibility and thereby severely affects the performance of high-level vision tasks, such as semantic segmentation and object detection. The extent of haze effects is determined by how far the scene is and the amount of airborne particles that either attenuate the visibility of the scene or scatter global atmospheric light towards an observer. As such, an atmospheric scattering model [32, 33] formulates haze effects as:

$$I = J \odot t + A \cdot (1 - t), \quad (1)$$

where I and J are an observed hazy image and a scene radiance (i.e. clean haze-free image), and \odot indicates the pixel-wise multiplication. The scalar A denotes global atmospheric light, and t is the transmission map representing the remaining fraction of light that reaches an observer from the scene. In general, t and A are unknown, and thus recovering the clean image J from a given hazy image I is a highly ill-posed and challenging problem.

Based on this physical haze model, early works have imposed constraints with strong assumptions or priors (e.g., hazy regions have higher intensity values than haze-free regions [42] or haze-free regions have at least one color channel with low intensity [16]). Due to such strong priors, these prior-based methods fail to work under scenarios where assumptions do not hold, resulting in poor generalization. To alleviate this, recent data-driven approaches rely on large-scale datasets and the representation power of neural networks to recover clean haze-free images by learning to estimate transmissions [3, 37] or directly learning a mapping of haze-free images [10, 26, 28, 38, 48] or jointly estimate both from hazy images [47, 49]. However, there are inherent ambiguities and uncertainties (e.g., airlight-albedo ambiguity: we cannot tell how much light is from scene radiance or atmospheric light [11]), causing inaccurate estimation of transmission map or haze-free images.

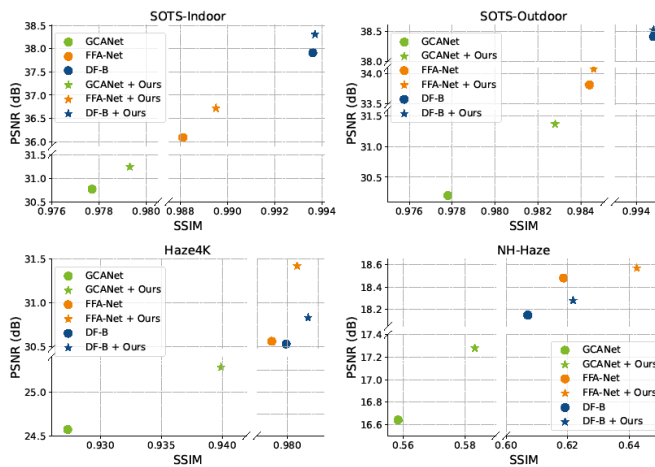


Figure 1: Our variational Bayesian framework is model-agnostic, and consistently improves the performance of existing dehazing neural networks across different benchmark datasets (SOTS [22], Haze4K [27] and NH-Haze [1]) in terms of PSNR and SSIM values. Upward-right movement of the star indicates better restoration.

In this work, instead of focusing on building elaborate and effective network architectures, we shift the attention to modeling uncertainties involved in the single image dehazing problem. To this end, we propose a new variational Bayesian framework, which incorporates the physical model (*i.e.*, atmospheric scattering model) and takes not only hazy image and scene radiance but also transmission as latent variables conditioned on a given hazy image. Then, the variational posteriors of scene radiance and transmission are parameterized by dehazing network and transmission network, respectively. Upon our variational framework, we derive a new objective function that induces synergy between the training of two networks, thereby improving the overall dehazing performance. Note that our framework allows for the joint training of dehazing and transmission networks, without making them dependent on each other. Thus, during inference, a dehazing network can be used independently of a transmission network, introducing no extra overhead. Furthermore, our framework design is model-agnostic, allowing for seamless integration with any dehazing neural network. The contributions of this work can be summarized as follows:

- (1) The proposed method models uncertainties of transmission maps and haze-free images by integrating the Bayesian modeling and data-driven methods.
- (2) Our model-agnostic framework can seamlessly employ any conventional dehazing neural network without any architecture modification.
- (3) Our framework consistently improves the performance of existing methods, including state-of-the-art models, across various benchmark datasets as illustrated in Figure 1.

2 RELATED WORK

In general, the haze effect is dependent on depth (*i.e.*, a deep scene or a distant object produce minimal transmission, therefore resulting in a substantially hazy image). The main challenge of the dehazing

task lies in effectively extracting information on atmospheric light and transmission map merely from the given hazy image. Most of existing dehazing strategies can be categorized into two approaches: prior and learning-based methods.

Prior-based Methods. Early dehazing algorithms mostly depend on Eq. (1) and statistical prior to impose constraints on the solution space. Fattal *et al.* [11] assumed that shading of the object and transmission are statistically uncorrelated over the entire image patch to estimate the transmission map and albedo of the medium. Tan *et al.* [42] proposed to compute transmission map using Markov random field, utilizing the prior that haze-free regions have higher contrast than hazy ones. He *et al.* [16] proposed dark channel prior to estimate transmission maps and atmospheric light, based on the observation that the lowest intensity among the color channels of natural outdoor images is close to zero due to factors, such as shadow or color patterns. Zhu *et al.* [52] proposed a linear model that restores depth maps with assumed color attenuation prior which describes the relationship between the pixel intensity, saturation, and their differences. Berman *et al.* [2] introduced a non-local method with haze-line prior, which assumes that a few hundred distinct colors can successfully approximate the color of haze-free regions, forming compact clusters in RGB space. These methods often fail to work due to strong priors and assumptions.

Learning-based Methods. As deep learning technology and large scale open source datasets become increasingly procurable, data-driven learning-based methods have become prevalent [5, 18, 36, 44, 45, 51]. In contrast to prior-based methods, these learning-based methods learn to map hazy images to haze-free images directly in an end-to-end manner. Cai *et al.* [3] proposed DehazeNet, an end to end modeling with CNN, and Ren *et al.* [37] introduced, a multi-scale neural network (MSCNN), both successfully estimate transmission maps. Li *et al.* [21] introduced a new variable dependent on hazy input by using the atmospheric scattering model in Eq. (1), enabled to reconstruct latent clean image by predicting the variable. Zhang *et al.* [47] proposed an edge-preserving loss, multi-level architectures, and introduced a discriminator to estimate transmission map and atmospheric light mutually. Ren *et al.* [38] proposed several pre-processing steps and multi-scale-fusion-based network to learn confidence maps for improved global visibility. Dong *et al.* [7] incorporated generic recursive boosting algorithm in the dense feature fusion model for information preservation and performance improvement. Guo *et al.* [15] added geometrical information to a transformer module and concatenated with a CNN module to increase the local and global connectivity.

Variational Bayesian Modeling. Bayesian modeling allows for the modeling of uncertainties and latent variables that may not be readily apparent from the observed data. While variational inference is a potent tool for approximating intricate probability distributions, its practical application requires careful consideration of prior knowledge. For instance, to tackle denoising problem, Yue *et al.* [46] model noise-free image and its variance as latent variables utilizing a conjugate prior. Wang *et al.* [43] leverage Dirichlet distribution to model blur kernel and deblurred image as latent variables and introduced two inference structures that are independent and dependent on the estimated blur kernel under blur process. In this

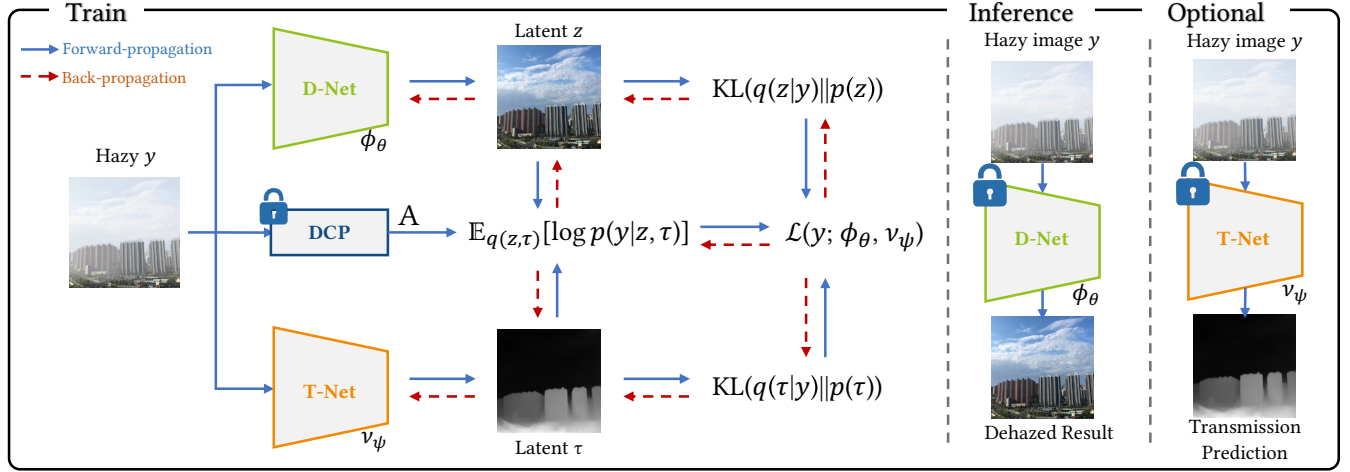


Figure 2: The architecture of proposed variational network. Blue solid lines represent forward process and red dotted lines denote gradient flow in back-propagation. Note that our *D-Net* and *T-Net* are not depending on specific network architectures. In addition, we can only employ *D-Net* to output the haze-free image in the inference stage, hence no additional overhead during inference.

work, we take prior knowledge from the atmospheric scattering model, in which haze degradation involves several factors, such as transmission, that introduce uncertainties. Motivated by the physical model, we employ variational Bayesian modeling and take not only hazy image and haze-free images but also transmission as latent for facilitating training and the modeling of uncertainties.

3 VARIATIONAL HAZE REMOVAL FRAMEWORK

Let \mathbb{D} be a training dataset composed of n triplets (x, y, t) of hazy image $y \in \mathbb{R}^{h \times w \times 3}$, ground truth clean image $x \in \mathbb{R}^{h \times w \times 3}$, and transmission map $t \in \mathbb{R}^{h \times w \times 1}$, respectively. h and w are the height and width of an image in RGB space, respectively. Moreover, we denote latent haze-free image, latent transmission, and atmospheric light as z , τ , and A . We consider the clean image and transmission map as latent variables, and proceed to calculate their posterior distribution based on haze degradation. This work aims to construct a variational function approximation of the posterior given a single hazy image through the Bayesian model including likelihood and priors. Learning the joint distribution of our latent variables can further escalate the performance of the conventional dehazing networks. The details are elaborated as in the following.

3.1 Bayesian Model Construction

Likelihood Model. Based on the atmospheric scattering model in Eq. (1), we start with taking intensity value of a hazy image as latent, which we assume to follow a Gaussian distribution as:

$$y_i \sim \mathcal{N}(z_i \odot \tau_i + A(1 - \tau_i), \sigma^2), \quad (2)$$

where y_i , z_i , and τ_i denote pixel values of a hazy image, haze-free image, and transmission map at a pixel location i respectively. Moreover, $\mathcal{N}(\mu, \sigma^2)$ is the Gaussian distribution with mean μ and variance σ^2 . For the sake of analytical feasibility and the basic properties of the Gaussian distribution that facilitate the parameterization of

latent variables, we model y_i as a Gaussian distribution [31]. Since our training dataset \mathbb{D} includes the ground truth haze-free image and transmission map, we can further take z and τ as latent and train neural networks to estimate their posteriors.

Haze-free Image. In general, optimizing the L_1 loss function encourages the median estimation of the observations rather than mean as with L_2 loss [17, 41]. Therefore, conventional dehazing networks favor L_1 loss variants to L_2 loss [15, 24, 35, 40] and minimize the absolute difference between the ground truth clean and predicted dehazed images during training to produce sharper edges/boundaries while suppressing the noise in homogeneous regions. If regression model errors are assumed to follow a Laplace distribution, then maximum likelihood estimates of the distribution parameters correspond to L_1 regression estimates [30]. Accordingly, we model the haze-free image under data-driven Laplace prior as:

$$z_i \sim \text{Laplace}(x_i, \epsilon_1^2), \quad (3)$$

where x_i is a pixel value of the ground truth clean image at i , $\text{Laplace}(n, \delta^2)$ denotes the Laplace distribution with parameters of location n and scale δ^2 . ϵ_1^2 is the mean absolute deviation from the median of z_i . Given that x_i serves as a reliable prior for z_i , we set a small value to ϵ_1^2 .

Transmission Map. In this work, we assume that the atmosphere is homogeneous in the scene as in previous arts [6, 33]. Under this assumption, the scene radiance is exponentially attenuated [42] and the transmission map can be formulated with scattering coefficient β and depth map d [6, 34] as:

$$t = e^{-\beta \cdot d}. \quad (4)$$

For reason similar to Eq. (2), we can model the probability of scene depth as a normal distribution. When the logarithm of a variable is normally distributed, then the variable has log-normal distribution. In addition, as in haze-free image modeling, a large number of transmission maps t from the training data provide a

strong data-driven prior to our latent transmission map τ . Therefore, we model the latent transmission map at pixel location i as follows:

$$\tau_i \sim \text{Lognormal}(-\beta d_i, \varepsilon_2^2), \quad (5)$$

where d_i denotes the pixel value of the latent depth map, $\text{Lognormal}(m, b)$ is lognormal distribution with parameters of scale m and scatter b . Notably, $-\beta d_i$ and ε_2^2 are equal to $\ln t_i$ and $\varepsilon_d^2 \beta^2$, which shall be also a small value, similar to that of ε_1^2 . We can compute $-\beta d_i$ as:

$$-\beta d_i = \log \left(\frac{y_i - A}{x_i - A} \right). \quad (6)$$

Moreover, to simplify the overall framework, we deal with σ^2 , ε_1^2 , ε_2^2 as hyper-parameters rather than latent variables, each of which controls the uncertainty of its associated variable.

Atmospheric Light. As for atmospheric light A , we adopt atmospheric light estimation result from the dark channel prior (DCP) [16] under the assumption of homogeneous atmosphere. As several conventional datasets provide the ground truth A , we can also treat A as a latent and train a model to estimate it in our Bayesian framework. However, we empirically observed that the difference between dehazing results with the ground truth A and the atmospheric light obtained by DCP are insignificant. Therefore, we do not explicitly model the atmospheric light as latent and reduce the complexity of our Bayesian modeling by using the atmospheric light results by DCP in our work.

To be specific, the dark channel is defined as the morphological minimum filtered values among the RGB channels. The most haze-opaque region in the image can be detected by collecting the top 0.1% of brightest pixels in the dark channel [42]. The one with the highest intensity in the corresponding hazy y is selected as A .

3.2 Variational Formulation of Posterior

We aim to infer the posterior of the latent variables by merging the Bayesian models in Eqs. (2-5). The direct estimation of the true posterior of the latent variables z and τ solely from a single hazy image y (i.e., $p(z, \tau|y)$) is computationally infeasible. Therefore, we construct a variational surrogate distribution $q(z, \tau|y)$ to approximate $p(z, \tau|y)$. Following the mean field assumption, we partition the variables into independent parts (i.e., assume the conditional independence between two variables z , and τ):

$$q(z, \tau|y) = q(z|y)q(\tau|y). \quad (7)$$

Using Eqs. (3) and (5) with an assumption that surrogate distribution $q(z|y)$ and $q(\tau|y)$ have scale parameter ε_1^2 and scatter parameter ε_2^2 respectively, we formulate the variational posterior as:

$$\begin{aligned} q(z|y) &= \prod_i \text{Laplace}(\phi_\theta(y)_i, \varepsilon_1^2), \\ q(\tau|y) &= \prod_i \text{Lognormal}(\log v_\psi(y)_i, \varepsilon_2^2), \end{aligned} \quad (8)$$

where $\phi_\theta(\cdot)$, and $v_\psi(\cdot)$ are neural networks that are trained to estimate the posterior distribution parameters of latent variable z , and τ , conditioned on the input hazy image y . Since our proposed framework focuses on modeling the posterior of latent variables, without any assumption on the form of $\phi_\theta(\cdot)$, and $v_\psi(\cdot)$, our framework is model-agnostic. In particular, ϕ_θ , which we call *D-Net*, can be any existing dehazing networks (e.g., [4]), which is a neural

network with parameters θ trained to estimate the haze-free image. Similarly, v_ψ , named as *T-Net*, is an auxiliary neural network with parameters ψ to estimate the transmission map from a given input hazy image. Note that there is no additional overhead during inference as we can estimate a haze-free image with only *D-Net*, and the transmission map is optionally obtainable as shown in Figure 2.

3.3 Variational Lower Bound

With the functional parameterization of the variational posterior, we can optimize the trainable parameters θ , and ψ to maximize the posterior probability. To do so, we decompose the marginal log-likelihood and obtain the variational lower bound. For notational simplicity, we use ϕ_i and v_i rather than $\phi_\theta(y)_i$ and $v_\psi(y)_i$, then the marginal log-likelihood is given by,

$$\begin{aligned} \log p(y; z, \tau) &= \mathbb{E}_{q(z, \tau|y)} [\log p(y)] \\ &= \iint q(z, \tau|y) \log \left(\frac{p(y, z, \tau)}{p(z, \tau|y)} \right) dzd\tau \\ &= \iint q(z, \tau|y) \log \left(\frac{p(y|z, \tau)p(z)p(\tau)}{p(z, \tau|y)} \right) dzd\tau \\ &= \iint q(z, \tau|y) \log \left(\frac{p(y|z, \tau)p(z)p(\tau)}{q(z, \tau|y)} \frac{q(z, \tau|y)}{p(z, \tau|y)} \right) dzd\tau \\ &= \iint q(z, \tau|y) \log \left(\frac{p(y|z, \tau)p(z)p(\tau)}{q(z, \tau|y)} \right) dzd\tau \\ &\quad + \iint q(z, \tau|y) \log \left(\frac{q(z, \tau|y)}{p(z, \tau|y)} \right) dzd\tau \\ &= \mathbb{E}_{q(z, \tau|y)} \left[\log \left(\frac{p(y|z, \tau)p(z)p(\tau)}{q(z, \tau|y)} \right) \right] + \text{KL}(q(z, \tau|y) \| p(z, \tau|y)) \\ &\equiv \mathcal{L}(y; \phi_\theta, v_\psi) + \text{KL}(q(z, \tau|y) \| p(z, \tau|y)). \end{aligned} \quad (9)$$

where $\text{KL}(\cdot \| \cdot)$ computes the Kullback–Leibler (KL) divergence of two distributions, and $\mathcal{L}(y; \phi_\theta, v_\psi)$ is the variational lower bound which can be combined with Eqs. (7) and (8) as follows:

$$\begin{aligned} \mathcal{L}(y; \phi_\theta, v_\psi) &= \mathbb{E}_{q(z, \tau|y)} \left[\log \left(\frac{p(y|z, \tau)p(z)p(\tau)}{q(z, \tau|y)} \right) \right] \\ &= \mathbb{E}_{q(z, \tau|y)} [\log p(y|z, \tau)] - \mathbb{E}_{q(z, \tau|y)} \left[\log \left(\frac{p(z)p(\tau)}{q(z|y)q(\tau|y)} \right) \right] \\ &= \mathbb{E}_{q(z, \tau|y)} [\log p(y|z, \tau)] - \text{KL}(q(z|y) \| p(z)) - \text{KL}(q(\tau|y) \| p(\tau)). \end{aligned} \quad (10)$$

and each term in Eq. (10) can be calculated analytically as follows:

$$\begin{aligned} &\mathbb{E}_{q(z, \tau|y)} [\log p(y|z, \tau)] \\ &= \sum_{i=1}^{hw} \left\{ -\frac{\log 2\pi\sigma^2}{2} - \frac{(y_i - (\phi_i v_i + A(1 - v_i)))^2 + \sigma^2}{2\sigma^2} \right\}, \end{aligned} \quad (11)$$

$$\begin{aligned} &\text{KL}(q(z|y) \| p(z)) \\ &= \sum_{i=1}^{hw} \left\{ \exp \left(-\frac{|\phi_i - x_i|}{\varepsilon_1^2} \right) + \frac{|\phi_i - x_i|}{\varepsilon_1^2} - 1 \right\}, \end{aligned} \quad (12)$$

and

$$\text{KL}(q(\tau|y) \| p(\tau)) = \sum_{i=1}^{hw} \left\{ \frac{1}{2\varepsilon_2^2} (\log v_i - \log t_i)^2 \right\}. \quad (13)$$

Note that all terms in Eq. (10) are differentiable, and we can train the network parameters θ and ψ over the given training dataset \mathbb{D} by optimizing the following objective function:

$$\min_{\theta, \psi} -\mathcal{L}(y; \phi_{\theta}, v_{\psi}). \quad (14)$$

3.4 Learning with Variational Lower Bound

By minimizing the final objective function in Eq. (14) through conventional back-propagation without using the reparameterization trick [20], we can train the parameters of networks ϕ_{θ} and v_{ψ} and estimate the posterior of latent variables z and τ as illustrated in Figure 2. Notably, the roles of three terms composing the total objective can be explained as follows. The first term represents the likelihood of the observed hazy images and is responsible for encouraging cooperation between dehazing and transmission networks based on the Eq. (1), which describes the relationship between the three latent variables: hazy image, haze-free image, and transmission. The second (Eq. (12)) and the third (Eq. (13)) terms act as regularization, making the posterior distribution close to prior distribution. Therefore, two separate networks ϕ_{θ} and v_{ψ} can complement each other with the aid of the joint term, and they are simultaneously trained by simulating the physical haze degradation process.

Furthermore, σ^2 , ε_1^2 and ε_2^2 can be interpreted as not only uncertainty of each variable but also the importance of the associated term. For instance, the importance of the KL divergence between $q(z|y)$ and $p(z)$ increases as ε_1^2 approaches to zero.

4 EXPERIMENTAL RESULTS

4.1 Experimental Setting

Datasets. We conducted our experiments on both synthetic and real-world datasets. We utilize the RESIDE [22] and Haze4K [27] as synthetic datasets, and the NH-Haze [1] and Fattal evaluation set [12] for real-world datasets. The RESIDE benchmark comprises synthetic hazy images along with their corresponding clean images captured in both indoor and outdoor scenarios. The Synthetic Objective Test Set (SOTS) is used to evaluate the performance of the models on RESIDE dataset. The indoor training set (ITS) of RESIDE benchmark consists of 13990 generated hazy images from 1399 clean images. The outdoor training set (OTS) of RESIDE benchmark includes a total of 313950 hazy images generated by using the collected real outdoor images. Haze4K is constructed by generating 4000 hazy images with randomly sampled atmospheric light A and scattering coefficient β from 500 clean indoor images in NYU-Depth [39] and 500 outdoor images in OTS. NH-Haze contains 55 paired images of real-world haze scenes.

Implementation. For our *D-Net*, we can employ any conventional dehazing networks, and we use GCANet [4], FFA-Net [35], and a recent state-of-the-art network DehazeFormer-B [40] as our baseline dehazing networks. For our *T-Net*, we use GCANet with a clamping activation function on the output layer. For fair comparison, we follow all training and evaluation strategies of the baselines (e.g., total epoch, optimizer, etc.) and our Bayesian framework is implemented based on the officially available code of each baseline.

In the case of real-world NH-Haze dataset, total training epoch is set to 300 using the official train-test split. As the NH-Haze train

set lacks the ground truth transmission map, we estimated the map using a clean and hazy image pair while assuming $A = 1$ from Eq. (1) (i.e., $t = (I - A)/(J - A + \epsilon) \in \mathbb{R}^{h \times w \times 3}$ with $\epsilon = 10^{-6}$ for numerical stability).

Notably, as our baseline networks originally employ either L_1 or L_2 loss functions, we do not employ additional objective functions (e.g., adversarial loss [8, 9, 23, 50], contrastive loss [44], and perceptual loss [25, 36]) for fair comparisons. We empirically determine $\sigma^2 = 10^{-5}$, $\varepsilon_1^2 = 10^{-6}$, and $\varepsilon_2^2 = 10^{-5}$. Our source code is publically available.¹

4.2 Performance Evaluation

To evaluate the performance of the proposed Bayesian framework, we compare the dehazing results with and without using the proposed framework both on synthetic and real-world haze datasets.

Results on Synthetic Datasets. Table 1 presents the dehazing results on three different datasets (SOTS-Indoor, SOTS-Outdoor, and Haze4K), comparing with DCP [16], BCCR [29] CAP [52], NLD [2], DehazeNet [3], AOD-Net [21], MSBDN [7], and DeHamer [15]. Notably, for our baseline GCANet [4], FFA-Net [35], and DehazeFormer-B [40], we provide two sets of PSNR and SSIM values: the scores reported in their original manuscripts and reproduced numbers in our experiments which are within the parentheses. As shown, our proposed method integrated with DehazeFormer-B obtains the highest metric scores in every domain, except for the PSNR on Haze4K, where FFA-Net + Ours performs the best.

Moreover, in Figure 4, we present visual comparisons of our method with baseline models on SOTS-Indoor test set. We see that our method produces clear images with less artifacts.

Results on Real-world Datasets. In Table 2, the evaluation results in terms of PSNR, SSIM, and LPIPS obtained from the NH-Haze test set [1] are provided. We compare with DCP, CAP, MSBDN, and our baselines. Our proposed method, when combined with FFA-Net, achieved the best performance in every metric. Notably, we observed an improvement in performance of over 0.28 dB in PSNR and 0.021 in SSIM on average.

We compare our method against baseline methods on the Fattal evaluation set, as demonstrated in Figure 5, where all networks were trained on the Haze4K dataset. It can be observed that models trained with our method outperform each baseline, and effectively removing haze while producing more vivid colors with less artifacts. The results demonstrate the effectiveness of our framework in removing depth-independent haze, with manually calculated transmission map from atmospheric scattering model.

User Study Results. Due to the lack of ground truth clean images or object annotations in the Fattal evaluation set [12], as well as the need for further evaluation of the qualitative aspects, we conducted a user study. The details are described as follows. First, we randomly selected six images each from SOTS-indoor, SOTS-outdoor [22], Haze4K test set [27], and Fattal evaluation set [12]. For each hazy image, we randomly chose a pair of dehazing results from one of the three base models (GCANet [4], FFA-Net [35], and DehazeFormer-B [40]) and corresponding enhanced model by our

¹<https://github.com/imeunu/Variational-Dehazing-Networks>

Methods	SOTS-Indoor		SOTS-Outdoor		Haze4K	
	PSNR \uparrow	SSIM \uparrow	PSNR \uparrow	SSIM \uparrow	PSNR \uparrow	SSIM \uparrow
DCP [16]	16.62	0.818	19.13	0.815	14.01	0.760
BCCR [29]	17.04	0.785	15.51	0.791	-	-
CAP [52]	18.97	0.815	18.14	0.759	16.32	0.782
NLD [2]	17.29	0.777	17.97	0.819	-	-
DehazeNet [3]	19.82	0.821	24.75	0.927	19.12	0.840
AOD-Net [21]	20.51	0.816	24.14	0.920	17.15	0.830
MSBDN [7]	33.67	0.985	33.48	0.982	22.99	0.850
DeHamer [15]	36.63	0.988	35.18	0.986	27.28	0.956
GCANet [4]	30.23 (30.77)	0.9764 (0.9777)	- (30.20)	- (0.9778)	- (24.57)	- (0.9273)
GCANet + Ours	31.25	0.9793	31.37	0.9828	25.28	0.9399
FFA-Net [35]	36.39 (36.09)	0.9886 (0.9881)	33.57 (33.81)	0.9839 (0.9844)	- (30.56)	- (0.9787)
FFA-Net + Ours	36.72	0.9895	34.07	0.9846	31.42	0.9808
DehazeFormer-B [40]	37.84 (37.91)	0.9936 (0.9936)	34.95 (38.41)	0.9843 (0.9948)	- (30.53)	- (0.9799)
DehazeFormer-B + Ours	38.31	0.9937	38.52	0.9948	30.83	0.9817

Table 1: The PSNR(dB), SSIM comparison of image dehazing methods on different synthetic data benchmarks. The numbers within parenthesis represent reproduced results. Baseline dehazing networks are GCANet [4], FFA-Net [35], and basic DehazeFormer [40]. + Ours indicates baseline networks trained with the proposed Bayesian framework. The best values are indicated as bold text.

Method	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow
DCP [16]	12.72	0.4419	0.5168
BCCR [29]	13.12	0.4831	-
CAP [52]	12.88	0.4929	0.6223
NLD [2]	12.23	0.4823	-
AOD-Net [21]	15.31	0.4584	0.5121
MSBDN [7]	17.34	0.5566	0.5026
DeHamer [15]	17.91	0.5781	0.4816
GCANet [4]	16.64	0.5583	0.4356
GCANet+Ours	17.28	0.5832	0.4057
FFA-Net [35]	18.48	0.6186	0.3694
FFA-Net + Ours	18.57	0.6424	0.3415
DehazeFormer-B [40]	18.15	0.6070	0.4192
DehazeFormer-B + Ours	18.28	0.6217	0.4066

Table 2: The PSNR(dB), SSIM and LPIPS results on NH-Haze test set [1].

approach. To further validate the performance of perceptual quality in a real-world scenario, we selected images containing distinct objects on Fattal evaluation set and utilized Yolov5 [19] to detect the objects in the dehazed image pair. Finally, 18 raters were asked to vote on 18 dehazed result pairs that appeared more visually convincing, and 6 object detection output pairs with more precise bounding boxes and accurate classification of the object.

As summarized in Figure 3, the pie charts (a) and (b) indicate that our proposed method consistently generates more visually pleasing clean images than the baseline models. In addition, pie chart (c) demonstrates that our framework produces dehazed images with superior perceptual quality. Therefore, the results of our user study clearly demonstrate the effectiveness of our proposed method in improving the quality of dehazed images.

4.3 Object Detection Application

We further assess the quality of the estimated haze-free images by evaluating how much the haze removal improves a downstream

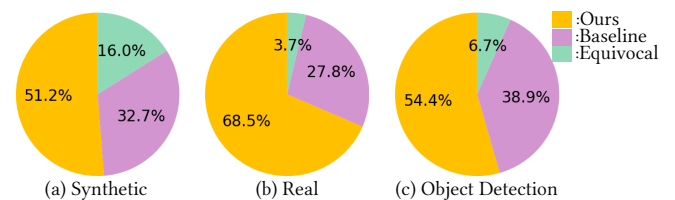


Figure 3: User study results.

YOLOv5 [19]	mAP50 \uparrow	mAP50-95 \uparrow
Hazy images	68.5	45.2
Ground truth clean images	93.0	69.6
GCANet [4]	76.4	53.7
GCANet + Ours	78.6	55.3
FFA-Net [35]	75.2	53.3
FFA-Net + Ours	78.3	56.6
DehazeFormer-B [40]	87.3	63.2
DehazeFormer-B + Ours	87.6	63.7

Table 3: Object detection results with YOLOv5 [19]. Mean average precision scores larger than 0.5 overlap IOU (mAP50) and 0.5~0.95 overlap IOU (mAP50-95) on the KITTI Haze dataset [7] are reported.

task: object detection in this work. We perform experiments with YOLOv5 [19] as an object detector on KITTI Haze dataset [7], which is synthesized based on KITTI detection dataset [13], following the dataset generation algorithm of RESIDE dataset [22] with depth estimation method [14]. The quality of dehazed images is evaluated with how much detection accuracy improves in terms of mean average precision.

Table 3 summarizes the detection performance on hazy images, ground truth clean images (upper bound), estimated clean images by the baselines GCANet [4], FFA-Net [35], DehazeFormer-B [40], and ours applied to each model, where all models are trained with RESIDE outdoor dataset [22]. The object detector clearly benefits

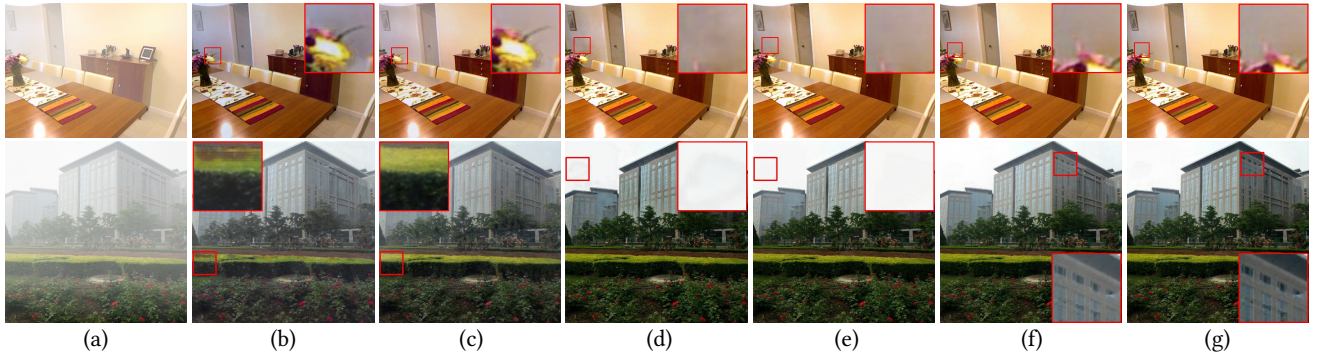


Figure 4: Visual comparisons between the baseline models and our enhanced models on the SOTS dataset [22]. (a) Input hazy image. (b) GCANet [4]. (c) GCANet + Ours. (d) FFA-Net [35]. (e) FFA-Net + Ours. (f) DehazeFormer-B [40]. (g) DehazeFormer-B + Ours. Best viewed on high-resolution display.

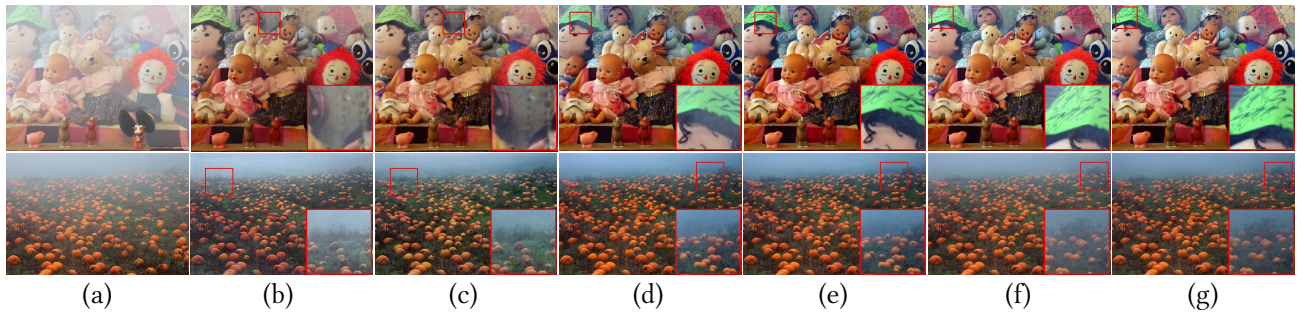


Figure 5: Visual comparisons of image dehazing methods on Fattal evaluation set [12]. (a) Hazy input image. (b) GCANet [4]. (c) GCANet + Ours. (d) FFA-Net [35]. (e) FFA-Net + Ours. (f) DehazeFormer [40]. (g) DehazeFormer + Ours.

Configuration	Joint	Loss	PSNR \uparrow	SSIM \uparrow
GCANet [4]	\times	L_2	24.57	0.9273
GCANet + Modified [49]	\checkmark	L_2	24.97	0.9153
GCANet + Ours	\checkmark	Eq. (14)	25.28	0.9399

Table 4: Comparison with different joint training method. The terms Joint and Loss indicate whether the training strategy is joint or not and the training objective, respectively. We compare original GCANet, GCANet with a modified joint training strategy from [49] (GCANet + Modified [49]) and GCANet with our final Bayesian framework (GCANet + Ours). The PSNR and SSIM results are evaluated on Haze4K [27] test set.

from the haze removal, and is greatly improved by using the dehazed images, and we observe that our framework shows consistent improvement over baselines. Moreover, Figure 6 presents the qualitative results of object detection both on synthetic dataset [7] and on real-world dataset [12] and demonstrates that our framework not only improves the dehazing performance but also allows detection module to recognize distant objects with more clear image, while enhancing the confidence.

4.4 Ablation Study

Joint optimization and Bayesian modeling. we performed ablation study in order to validate the effectiveness of our Bayesian

modeling, we compared our framework with the slightly modified version of the joint training framework proposed by Zhang *et al.* [49]. Specifically, the estimated joint transmission map is concatenated to the input of the dehazing module, facilitating the joint training of the dehazing and transmission estimation module as illustrated in Figure 7. Note that for a fair comparison, we have excluded the adversarial network, adversarial loss, and perceptual loss, and employed GCANet [4] for both modules.

To verify the effectiveness of the joint optimization strategy and our Bayesian modeling, we compare these with the modified version of joint training framework introduced in Zhang *et al.* [49]. Specifically, the estimated transmission map is concatenated to the input of the dehazing module to jointly train the dehazing and transmission estimation networks. We utilize GCANet for dehazing and transmission estimation module for a fair comparison and train with the ground truth clean image and transmission map using the L_2 objective. This configuration (GCANet + Modified [49]) allows joint training, but does not take into account the haze degradation process. From the comparison result, it is observed that our final model (GCANet + Ours) using the joint optimization with our Bayesian framework outperforms the best. We analyze that the likelihood term (*i.e.*, Eq. (11)) in the proposed objective is responsible for leveraging the relationships between latent variables and uncertainty, resulting in our method’s superior performance in comparison with the modified joint method, which lacks this

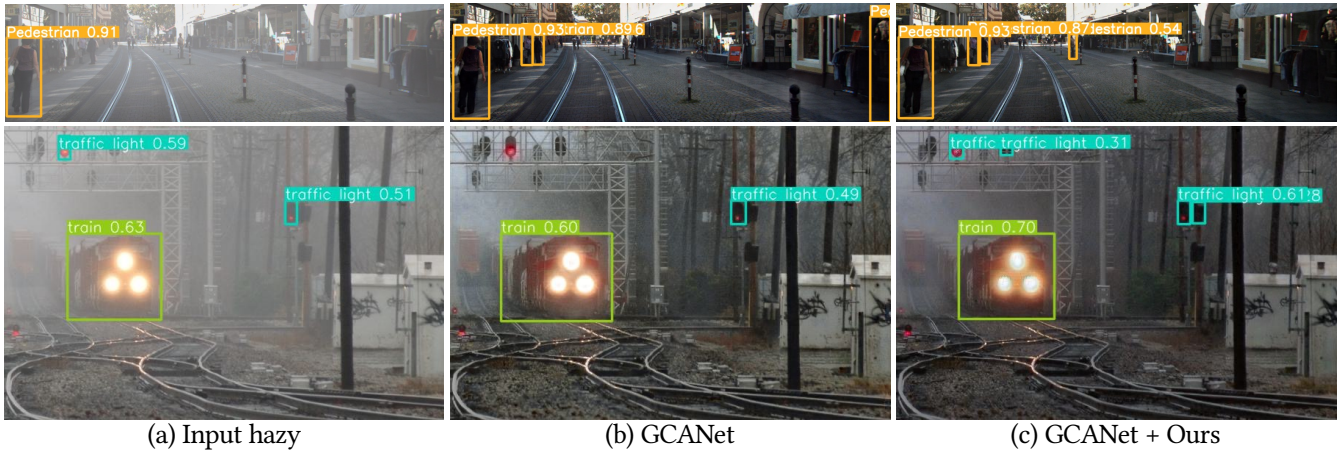


Figure 6: Object detection results by Yolov5 [19] on estimated clean images. Top to bottom: Results on the KITTI Haze dataset [7] and the Fattal evaluation set [12]. (a) Detection results from input hazy images. (b) Detection results with GCANet. (c) Detection results with GCANet + Ours. Best viewed on high-resolution display.

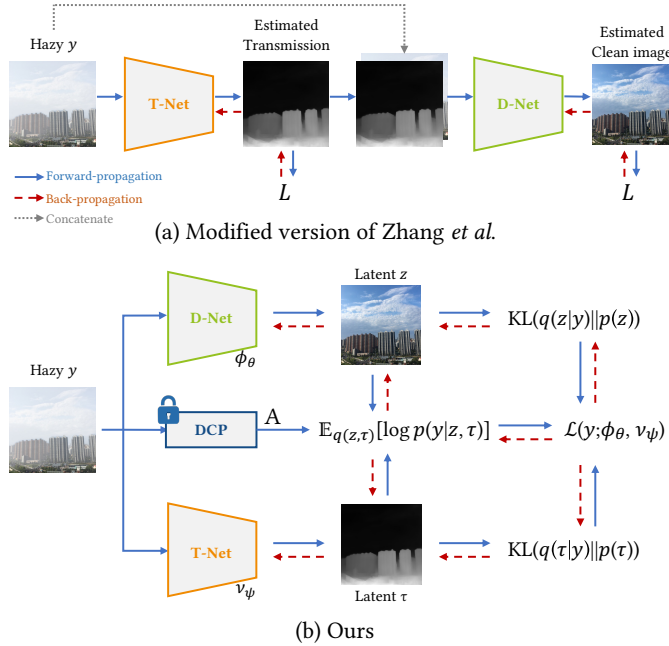


Figure 7: Illustration of the modification made to the framework from Zhang et al. [49].

term. In addition, joint optimization helps to utilize transmission information, thus both contribute to performance gain.

Prior Distribution. We further justify our choice of prior distributions for z and τ by performing ablation study on the prior distributions. As reported in Table 5, we perform ablations based on GCANet backbone, replacing each prior distribution with Gaussian distributions of variance $(\sigma^2, \epsilon_1^2, \epsilon_2^2)$. In the results, Laplace and Log-normal distribution tend to yield better performance. Notably, using the Laplace distribution for z provides higher SSIM values than Gaussian distribution, corroborating our motivation that Laplace encourages z to learn the latent of sharp boundaries.

4.5 Discussion

Role of T-Net. *T-Net* serves as an auxiliary branch in our framework, designed to estimate the scale parameter of the Lognormal distribution modeling the latent variable τ , which is not utilized in the inference phase. However, we can further inspect the reason behind the output of *D-Net* by examining the output of *T-Net*. For instance, the dehazing network may prioritize the restoration of more vivid colors in areas where lower transmission is estimated.

Influence of Model Capacities. We studied how the size of each module of our framework could affect the other by conducting experiments on Haze4K [27] dataset. We estimate the impact of the *T-Net* on the performance of *D-Net* by evaluating the ones we used for benchmarking Haze4K dataset. Note that the architecture for *T-Net* is fixed in this experiment. On the other hand, to evaluate the impact of *D-Net* performance on the *T-Net* structure, we modified the number of filters in the hidden layer of GCANet. Specifically, we fixed the architecture of *D-Net* as GCANet and employed another GCANet as *T-Net* with 48 and 96 filters per layer, respectively, while the original structure has 64 filters per layer. The estimated transmission maps and their evaluation metrics (mean squared error (MSE) and SSIM) on the Haze4K test set are presented in Figure 8 and Table 6, respectively. Although the same network architecture is used for *T-Net*, improving *D-Net* architecture from a simple one (e.g., GCANet) to advanced ones (FFA-Net, DehazeFormer-B) tends to improve the transmission accuracy. Likewise, we observe that the performance of *D-Net* improves as the capacity of *T-Net* increased as reported in Table 7. In other words, *T-Net* improves as *D-Net* architecture improves, owing to our proposed framework that allows cooperation between two branches. Thus, improving *D-Net* better facilitates the training of *T-Net* or vice versa.

We believe this is because *D-Net* and *T-Net* are complementary in that they are jointly trained to minimize the objective function. Specifically, more accurate estimation of either z or τ results in a lower loss value, allowing for more accurate gradient computation. Thus, transmission estimation module can improve with dehazing

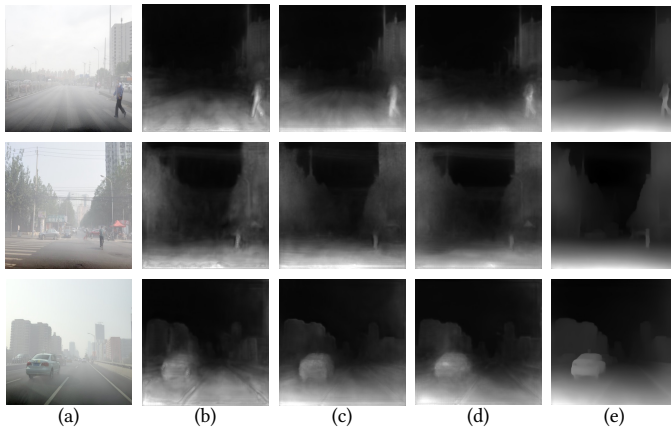


Figure 8: Visualization of transmission map produced by our T -Net utilizing our framework on Haze4k dataset [27]. Note that each T -Net has the same architecture, but jointly trained with different D -Net architectures in our framework. We denote the architecture of each branch as the combination of D -Net + T -Net. (a) Hazy image. (b) GCANet + GCANet. (c) FFA-Net + GCANet. (d) DehazeFormer-B + GCANet. (e) Ground truth transmission map.

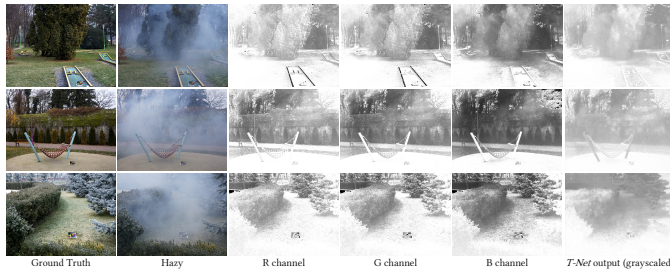


Figure 9: Visualization of the transmission map implemented as explained in Section 4.5. The prior assumption is violated as the ratio differs by RGB channels.

module. The results also allude to the efficacy of our joint optimization of clean haze-free image and transmission map, which are related by our proposed Bayesian framework and objective.

Implementation for NH-Haze Dataset. Unlike synthetic datasets, the assumption of the atmosphere scattering model in Eq. (1) (*i.e.*, $I = J \odot t + A \cdot (1 - t)$) may not hold true in real-world hazy images due to several violations of its underlying assumptions. One of the most critical violations is the significant difference in interpolation ratio among RGB channels as depicted in Figure 9. Note that the hazy image is represented through interpolation between the pixel value of the clean image and atmospheric light, while the transmission being the ratio between them. Another violation is that the atmospheric light is not close to 1, leading to the interpolation ratio being outside the range between 0 and 1. These violations make it challenging to model the haze degradation process with the atmosphere scattering model and integrate it into our framework. However, we can mitigate the adverse effects of these assumption violations through setting $t = (I - A)/(J - A + \epsilon)$ and $A = 1$. By doing so, we can model different interpolation ratios between 0

z distribution	τ distribution	PSNR \uparrow	SSIM \uparrow
Gaussian	Gaussian	30.96	0.9748
Gaussian	Lognormal	31.22	0.9772
Laplace	Gaussian	31.16	0.9771
Laplace	Lognormal	31.25	0.9793
Baseline (GCANet [4])		30.77	0.9777

Table 5: The PSNR(dB), SSIM results of GCANet [4] + Ours on SOTS-Indoor dataset [22] according to changing the prior models.

Model	GCANet [4]	FFA-Net [35]	DF-B [40]
MSE \downarrow	0.028	0.019	0.020
SSIM \uparrow	0.711	0.770	0.747

Table 6: The MSE, SSIM results of inferred transmission map on Haze4K test set [27]. Both metrics improve as the performance of D -Net on Haze4K benchmark increase.

# of filters	48	64	96
PSNR \uparrow	24.72	25.28	25.74
SSIM \uparrow	0.9349	0.9399	0.9447

Table 7: The performance of D -Net in terms of PSNR and SSIM on Haze4K test set. Both metrics improved as the capacity of T -Net increased.

and 1 for each RGB channel under this assumption. As a result, this implementation enables modeling the haze degradation process on real scenarios, including depth-independent scenarios, and integrating the atmosphere scattering model into our framework. Finally, as detailed above, D -Net can leverage the outputs of T -Net to enhance its dehazing performance.

5 CONCLUSION

This work is founded on the motivation that there are inherent uncertainties that make the single image dehazing problem challenging. To alleviate this problem, we propose to formulate a variational Bayesian framework for single image dehazing. Incorporating the atmospheric scattering model, we handle uncertainties involved in estimating transmission and haze-free images. In particular, we take transmission and haze-free images as latent variables and use neural networks to parameterize the approximate posterior distribution of these joint latent variables. Our framework provides consistent performance improvement across various models and numerous datasets.

ACKNOWLEDGEMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00156, Fundamental research on continual meta-learning for quality enhancement of casual videos and their 3D metaverse transformation) and Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2020-0-01373, Artificial Intelligence Graduate School Program(Hanyang University))

REFERENCES

- [1] Codruta O. Ancuti, Cosmin Ancuti, and Radu Timofte. 2020. NH-HAZE: An Image Dehazing Benchmark With Non-Homogeneous Hazy and Haze-Free Images. In *CVPRW*.
- [2] Dana Berman, Shai Avidan, et al. 2016. Non-local image dehazing. In *CVPR*.
- [3] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. 2016. Dehazenet: An end-to-end system for single image haze removal. *IEEE TIP* 25, 11 (2016), 5187–5198.
- [4] Dongdong Chen, Mingming He, Qingnan Fan, Jing Liao, Liheng Zhang, Dongdong Hou, Lu Yuan, and Gang Hua. 2018. Gated Context Aggregation Network for Image Dehazing and Deraining. *WACV* (2018).
- [5] Zeyuan Chen, Yangchao Wang, Yang Yang, and Dong Liu. 2021. PSD: Principled synthetic-to-real dehazing guided by physical priors. In *CVPR*.
- [6] Fabio Cozman and Eric Krotkov. 1997. Depth from scattering. In *CVPR*.
- [7] Hang Dong, Jinshan Pan, Lei Xiang, Zhe Hu, Xinyi Zhang, Fei Wang, and Ming-Hsuan Yang. 2020. Multi-Scale Boosted Dehazing Network With Dense Feature Fusion. In *CVPR*.
- [8] Yu Dong, Yihao Liu, He Zhang, Shifeng Chen, and Yu Qiao. 2020. FD-GAN: Generative adversarial networks with fusion-discriminator for single image dehazing. In *AAAI*.
- [9] Yixin Du and Xin Li. 2019. Recursive image dehazing via perceptually optimized generative adversarial network (POGAN). In *CVPRW*.
- [10] Akshay Dudhane, Harshjeet Singh Aulakh, and Subrahmanyam Murala. 2019. Ri-gan: An end-to-end network for single image haze removal. In *CVPR*.
- [11] Raanan Fattal. 2008. Single image dehazing. *ACM MM* 27, 3 (2008), 1–9.
- [12] Raanan Fattal. 2014. Dehazing using Color-Lines. *ACM Transaction on Graphics* 34, 13, Issue 1.
- [13] Andreas Geiger, Philip Lenz, and Raquel Urtasun. 2012. Are we ready for autonomous driving? the kitti vision benchmark suite. In *CVPR*.
- [14] Clément Godard, Oisin Mac Aodha, Michael Firman, and Gabriel J. Brostow. 2019. Digging into Self-Supervised Monocular Depth Prediction. (2019).
- [15] Chun-Le Guo, Qixin Yan, Saeed Anwar, Runmin Cong, Wenqi Ren, and Congyi Li. 2022. Image Dehazing Transformer with Transmission-Aware 3D Position Embedding. In *CVPR*.
- [16] Kaiming He, Jian Sun, and Xiaoou Tang. 2010. Single image haze removal using dark channel prior. *IEEE TPAMI* 33, 12 (2010), 2341–2353.
- [17] Peter J Huber. 1964. Robust Estimation of a Location Parameter. *Ann. Math. Statist.* 35, 4 (1964), 73–101.
- [18] Xingyu Jiang, Hongkun Dou, Chengwei Fu, Bingquan Dai, Tianrun Xu, and Yue Deng. 2022. Boosting Supervised Dehazing Methods via Bi-level Patch Reweighting. In *ECCV*.
- [19] Glenn Jocher, Alex Stoken, Jirka Borovec, NanoCode012, ChristopherSTAN, Liu Changyu, Laughing, tkianai, Adam Hogan, lorenzomammanna, yxNONG, AlexWang1900, Laurentiu Diaconu, Marc, wanghaoyang0106, ml5ah, Doug, Francisco Ingham, Frederik, Guilhen, Hatoviv, Jake Poznanski, Jiacong Fang, Lijun Yu, changyu98, Mingyu Wang, Naman Gupta, Osama Akhtar, PetrDvoracek, and Prashant Rai. 2020. *ultralytics/yolov5: v3.1 - Bug Fixes and Performance Improvements*. <https://doi.org/10.5281/zenodo.4154370>
- [20] Diederik P Kingma and Max Welling. 2013. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114* (2013).
- [21] Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. 2017. Aodnet: All-in-one dehazing network. In *ICCV*.
- [22] Boyi Li, Wenqi Ren, Dengpan Fu, Dacheng Tao, Dan Feng, Wenjun Zeng, and Zhangyang Wang. 2018. Benchmarking single-image dehazing and beyond. *IEEE TIP* 28, 1 (2018), 492–505.
- [23] Runde Li, Jinshan Pan, Zechao Li, and Jinhui Tang. 2018. Single Image Dehazing via Conditional Generative Adversarial Network. In *CVPR*.
- [24] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. 2017. Enhanced deep residual networks for single image super-resolution. In *CVPR*.
- [25] Xiaohong Liu, Yongrui Ma, Zhihao Shi, and Jun Chen. 2019. GridDehazeNet: Attention-Based Multi-Scale Network for Image Dehazing. In *ICCV*.
- [26] Xing Liu, Masanori Suganuma, Zhun Sun, and Takayuki Okatani. 2019. Dual residual networks leveraging the potential of paired operations for image restoration. In *CVPR*.
- [27] Ye Liu, Lei Zhu, Shunda Pei, Huazhu Fu, Jing Qin, Qing Zhang, Liang Wan, and Wei Feng. 2021. From synthetic to real: Image dehazing collaborating with unlabeled real data. In *ACM MM*.
- [28] Kangfu Mei, Aiweng Jiang, Juncheng Li, and Mingwen Wang. 2018. Progressive feature fusion network for realistic image dehazing. In *ACCV*.
- [29] Gaofeng Meng, Ying Wang, Jiangyong Duan, Shiming Xiang, and Chunhong Pan. 2013. Efficient image dehazing with boundary constraint and contextual regularization. In *Proceedings of the IEEE international conference on computer vision*. 617–624.
- [30] Gregory P Meyer. 2021. An alternative probabilistic interpretation of the huber loss. In *CVPR*.
- [31] Kevin P Murphy. 2012. *Machine learning: a probabilistic perspective*. MIT press.
- [32] Srinivasa G Narasimhan and Shree K Nayar. 2000. Chromatic framework for vision in bad weather. In *CVPR*. IEEE.
- [33] Srinivasa G Narasimhan and Shree K Nayar. 2002. Vision and the atmosphere. *ICCV* (2002).
- [34] Shree K Nayar and Srinivasa G Narasimhan. 1999. Vision in bad weather. In *ICCV*.
- [35] Xu Qin, Zhilin Wang, Yuanhao Bai, Xiaodong Xie, and Huizhu Jia. 2020. FFA-Net: Feature fusion attention network for single image dehazing. In *AAAI*.
- [36] Yanyun Qu, Yizi Chen, Jingying Huang, and Yuan Xie. 2019. Enhanced pix2pix dehazing network. In *CVPR*.
- [37] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang. 2016. Single image dehazing via multi-scale convolutional neural networks. In *ECCV*.
- [38] Wenqi Ren, Lin Ma, Jiawei Zhang, Jinshan Pan, Xiaochun Cao, Wei Liu, and Ming-Hsuan Yang. 2018. Gated fusion network for single image dehazing. In *CVPR*.
- [39] Nathan Silberman, Derek Hoiem, Pushmeet Kohli, and Rob Fergus. 2012. Indoor segmentation and support inference from rgbd images. In *ECCV*.
- [40] Yuda Song, Zhuqing He, Hui Qian, and Xin Du. 2022. Vision Transformers for Single Image Dehazing. *arXiv preprint arXiv:2204.03883* (2022).
- [41] Stephen M Stigler. 1986. *The history of statistics: The measurement of uncertainty before 1900*. Harvard University Press.
- [42] Robby T Tan. 2008. Visibility in bad weather from a single image. In *CVPR*. IEEE.
- [43] Hui Wang, Zongsheng Yue, Qian Zhao, and Deyu Meng. 2021. A Deep Variational Bayesian Framework for Blind Image Deblurring. *arXiv preprint arXiv:2106.02884* (2021).
- [44] Haiyan Wu, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. 2021. Contrastive learning for compact single image dehazing. In *CVPR*.
- [45] Yang Yang, Chaoyue Wang, Risheng Liu, Lin Zhang, Xiaojie Guo, and Dacheng Tao. 2022. Self-Augmented Unpaired Image Dehazing via Density and Depth Decomposition. In *CVPR*.
- [46] Zongsheng Yue, Hongwei Yong, Qian Zhao, Deyu Meng, and Lei Zhang. 2019. Variational denoising network: Toward blind noise modeling and removal. *NeurIPS* (2019).
- [47] He Zhang and Vishal M Patel. 2018. Densely connected pyramid dehazing network. In *CVPR*.
- [48] He Zhang, Vishwanath Sindagi, and Vishal M Patel. 2018. Multi-scale single image dehazing using perceptual pyramid deep network. In *CVPR*.
- [49] He Zhang, Vishwanath Sindagi, and Vishal M Patel. 2019. Joint transmission map estimation and dehazing using deep networks. *IEEE Transactions on Circuits and Systems for Video Technology* 30, 7 (2019), 1975–1986.
- [50] Dong Zhao, Long Xu, Lin Ma, Jia Li, and Yihua Yan. 2020. Pyramid global context network for image dehazing. *IEEE TCSVT* 31, 8 (2020), 3037–3050.
- [51] Zhuoran Zheng, Wenqi Ren, Xiaochun Cao, Xiaobin Hu, Tao Wang, Fenglong Song, and Xiuyi Jia. 2021. Ultra-high-definition image dehazing via multi-guided bilateral learning. In *CVPR*.
- [52] Qingsong Zhu, Jiaming Mai, and Ling Shao. 2015. A fast single image haze removal algorithm using color attenuation prior. *IEEE TIP* 24, 11 (2015), 3522–3533.