Article

Gradient-based metric for the evaluation of image defogging without a ground truth

Gerard deMas-Giménez [1,*,†], Pablo García-Gómez [2], Josep R. Casas [3] and Santiago Royo [1−2]

[1] Centre for Sensors, Instrumentation and Systems Development, Universitat Politècnica de Catalunya (CD6-UPC), Rambla de Sant Nebridi 10, Terrassa 08222, Spain; santiago.royo@upc.edu (S.R.)
[2] Beamagine S.L.; Carrer de Bellesguard 16, Castellbisbal 08755, Spain; pablo.garcia@beamagine.com
[3] Image Processing Group, TSC Department, Universitat Politècnica de Catalunya (UPC), Carrer de Jordi Girona 1-3, Barcelona 08034, Spain; josep.ramon.casas@upc.edu
* Correspondence: gerard.demas@gmail.com;
† Current address: Rambla de Sant Nebridi, 10, 08222 Terrassa, Barcelona, Spain.

Abstract: Fog, haze, or smoke are usual atmospheric phenomena that dramatically compromise the overall visibility of any scene, critically affecting features such as illumination, contrast, and contour detection of objects. The decrease in visibility compromises the performance of computer vision algorithms such as pattern recognition and segmentation, some of them very relevant for decision-making in the fields of security and autonomous vehicles. Several dehazing methods have been proposed. However, to the best of our knowledge, all proposed metrics in the literature compare the defogged image to its ground truth for evaluation of the defogging algorithms, or need to estimate parameters through physical models. This fact hinders progress in the field as obtaining proper ground truth images is costly and time-consuming, and physical parameters greatly depend on the scene conditions. This paper aims to tackle this issue by proposing a gradient-based metric for image defogging evaluation that does not need a ground truth image or a physical model. The proposed metric only requires the original hazy RGB image and the RGB image obtained after the defogging procedure. A comparison of the proposed metric with metrics already used in the NTIRE 2018 defogging challenge is performed to prove its effectiveness in a general situation, showing comparable results to conventional metrics. A Matlab implementation of the proposed metric can be found at the following GitHub repository: https://github.com/GDMG99/Gradient-based-metric-for-image-defogging-without-ground-truth.

Keywords: Image defogging; image evaluation metrics, visual enhancement evaluation; edge detection; deep neural networks; autonomous systems.

Citation: deMas-Giménez, G.; García-Gómez, P.; Royo, S.; Casas, J.R. Gradient-based metric for image defogging evaluation without ground truth. *Preprints* **2022**, *1*, 0. https://doi.org/

1. Introduction

In recent years, there has been a rise and important advances in automated surveillance and autonomous vehicles of different kinds. These systems use cameras as well as other sensors to gather data about their surroundings and process it with image processing algorithms to obtain relevant information about the scene. Nevertheless, these algorithms perform poorly under adverse weather conditions such as fog, smoke, or haze, since they compromise the visibility in images. Other atmospheric scattering media, such as sand or smog, behave similarly. They critically affect illumination, color, contrast, and contours of the scene due to the scattering behavior of the media. Therefore, there is a need to achieve a processing solution that reduces the effect of bad weather conditions, as the reduction of visibility can directly affect the performance of the algorithms involved. The process of developing image processing algorithms for enhancing visibility on images in bad weather conditions is known as defogging or dehazing.

Nowadays there are several approaches to defog an image. On the one hand, active approaches rely on using gated images [1] or polarized light [2,3] to get more information about the scene. Gated imaging requires usually expensive electronics, and polarimetric imaging is difficult to implement in outdoor systems, which is the main target of defogging.

 doi:10.20944/preprints202209.0127.v1

Polarimetric images are also difficult to automatize and to be implemented in autonomous systems, because they usually require estimating physical parameters of the scene [4].

Another common approach to tackle defogging is to apply Deep Neural Networks (DNNs), which have already produced some very promising results. The New Trends in Image Restoration and Enhancement Workshop and Challenges (NTIRE) reflects the advancement in the image defogging field in image and video processing. This workshop proposes challenges in image and video processing in several fields. For instance, homogeneous [5,6] and non-homogeneous [7,8] fog removal were among the topics of interest explored for some years in the workshop. In these challenges, some research groups exploited previous information on the image and tried to evaluate the natural parameters through deep learning techniques [9,10]. Alternatively, other groups took advantage of the generative capability of DNNs, especially with Generative Adversarial Networks (GANs), and used them to directly generate a defogged image from a foggy one without estimating any physical parameter [11–13]. In order to evaluate the effectiveness of the defogging networks, classical computer vision metrics such as the structural similarity index (SSIM), the Peak Signal to Noise Ratio (PSNR), or CIEDE2000 [14] were used to compare the defogged image with a ground truth of the scene. Nevertheless, classical computer vision metrics for evaluation perform poorly when it comes to quantifying an enhancement in the visibility of the scene. Moreover, and as its most important drawback, these metrics need a ground truth image which is not always available.

Obtaining ground truth images in adverse weather conditions is costly, time-consuming and, often, simply unfeasible. In natural conditions fog is a time-variant and complex weather phenomenon. Reproducing the same scene for acquiring images without fog but with equivalent luminance, positioning of the objects, etc, is a very complex task in practice. Thus, research is often based on artificial fog generation in rather controlled environments, usually large-scale fog chambers. However, such artificially generated fog is not fully equivalent to natural fog in terms of homogeneity and distribution [15]. This problem is especially sensitive with DNNs because they need huge datasets to achieve good results and avoid overfitting. Even though there exist defogging DNNs which are trained in an unpaired manner [11], the problem still perseveres when it comes to validation, as the most used evaluation metrics require a ground truth for comparison.

Hence, this work proposes a novel, general-purpose gradient-based metric for evaluating image defogging that needs neither a ground truth image of the scene nor an evaluation of the physical parameters of the image. The proposed metric only relies on the original and defogged images. The proposed metric will be compared for validation with the performance of SSIM on the O-Haze [16] dataset with some results of the NTIRE 2018 defogging challenge [5].

The paper is organized as follows. The next section overviews the current state-of-the-art of defogging evaluation metrics, and presents several proposals that tackle the problem of obtaining the ground truth images of natural fog scenes. Secondly, we present our method, a gradient-based metric for evaluating image defogging that can directly measure the improvement in the output defogged image with respect to the original hazy RGB image without the need of any ground truth or the estimation of any physical parameter. Afterwards, we compare our metric with the currently used SSIM algorithm on the O-Haze dataset [16] applied to some defogging results of the NTIRE 2018 defogging challenge [5] to prove its effectiveness.

2. State of the art

The problem of evaluating the visibility of a scene without having any reference beyond the original fogged RGB image has been of interest in the past years due to the complexity of obtaining reliable ground truth images of fogged scenes. Within this section, we briefly review different approaches used for evaluation of defogging algorithms. We can divide the evaluation methods into three groups [17]. The first two are called full-reference image quality assessment (FR-IQA) and no-reference image quality assessment

 doi:10.20944/preprints202209.0127.v1

(NR-IQA). The first group, FR-IQA, needs a ground truth image to evaluate quantitatively the defogging result. This is the case of SSIM and PSNR. On the contrary, NR-IQA metrics either do not need a reference or do not use a fog-free ground truth image for comparison. Our proposed metric explained in Section 3 falls into this category. The third group simulates hazy images from clear images based on Koschmieder's law [18] and then employs FR-IQA metrics to evaluate dehazing algorithms.

Hautière *et. al.* [19] and Pormeleau *et. al.* [20] presented different NR-IQA methods to evaluate the attenuation coefficient of the atmosphere by means of a single camera on a moving vehicle. Nevertheless, their method cannot be used as metric for a general single image visibility evaluator, as Pormeleau *et. al.* needed multiple images of the scene and Hautière *et. al.* required a road and the sky to be present in the scene.

A different NR-IQA method was presented by Liu *et. al.* [21] and consisted on the analysis of the histogram of the image on the HSV colourspace. First, the image is converted from the RGB colourspace to HSV. Then, fog detection is achieved by analyzing different features of the histogram of each channel: Hue (H), Saturation (S), and Value (V). They stated that the overall value of the three channels decreased due to scattering resulting from the fog, so the distribution was modified in presence of fog. Feature extraction of each histogram was performed by adding the values of the pixels of the image and normalizing to the number of pixels different of 0 in the channel. After that, a classification into different visibility categories was done by comparing the results obtained from the histogram with some empirical values. Even though Liu *et. al.* claimed good results with this method there is certain subjectivity in the choice of values of the thresholds for the classification.

Li *et. al.* [22] compared the results of two FR-IQA (SSIM and PSNR) with two NQ-IQA methods (spatial-spectral entropy-based quality - SSEQ) [23] and blind image integrity notator using DCT statistics (BLIINDS-II) [24]). However, their results do not bring to a general conclusion about which IQA method has a better judgement. Besides, BLIINDS-II [24] is based on the statistical behavior of a group of 100 people, so there is inherent subjectivity in the metric.

Also, Choi *et. al.* [25] presented a reference-less prediction of perceptual fog density and perceptual image defogging based on natural scene statistics and fog-aware statistical features. Their proposed model, Fog Aware Density Evaluator (FADE), predicts the visibility of a foggy scene from a single image without reference to a corresponding fog-free image and without being trained on human-rated judgments. FADE only makes use of measurable deviations from statistical regularities observed in natural foggy and fog-free images. Apart from that, they present a single image defogging network called DEFADE. More recently, Chen *et. al.* [26] presented a visibility detection algorithm of a single fog image based on the ratio of wavelength residual energy. Nevertheless, their algorithm uses the transmisivity map, which is obtained by estimating certain atmospheric parameters.

Other approaches have been trying to fix the method using edge detection metric evaluation [27], as we propose. However, they are mostly focused on the evaluation of the edge detection method rather in an improvement of the visibility of a scene by gradient comparison. Moreover, these metrics require a ground truth edge image for a proper evaluation.

Currently, the most used metric in defogging challenges is SSIM [28]. This well-known metric takes into account different aspects of an image and directly compares them with a sample image. SSIM basically focuses on contrast, luminance and structure. In fact, these are some of the most affected image features when fog is present in a scene. Nevertheless, defogging techniques, usually, do not try to completely recreate the original image but rather to produce an enhancement in the visibility of the fogged image by adjusting luminance, contrast, and other aspects of the scene. This could lead to a defogging procedure being heavily punished for not being similar enough to its ground truth even if the defogging results are good. Still, the main drawback of the metric for defogging evaluation is the need of a ground truth. As mentioned earlier, obtaining a ground truth

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 September 2022 doi:10.20944/preprints202209.0127.v1

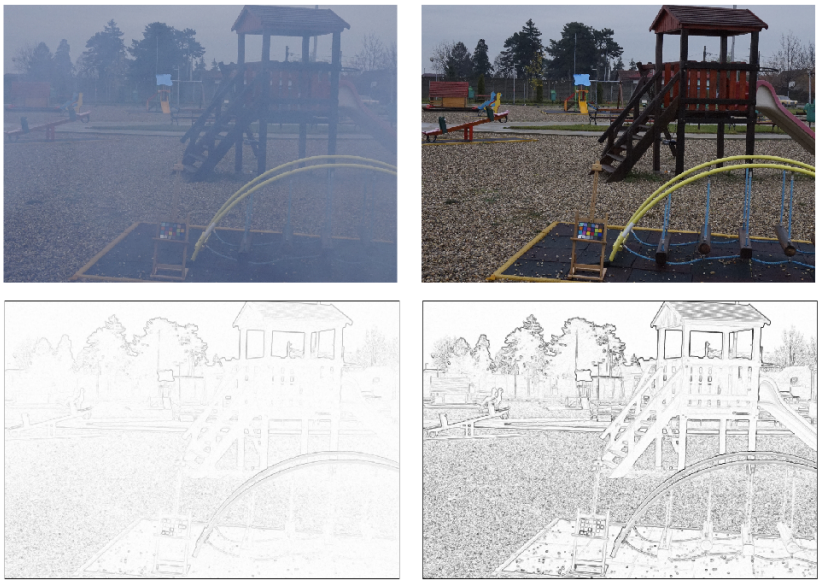

Figure 1. Gradient comparison between a fogged image (left) and its fog-free ground truth (right). Both colour images are presented on top with their associated edge images below.

image of a natural foggy scene is difficult, and the issue becomes more relevant when DNNs are introduced.

3. Methodology

In this section we introduce our proposed gradient-based metric for image defogging without the need of a ground truth. We thoroughly explain every step to validate a defogging procedure without a ground truth reference using our method. The reader can find a Matlab implementation of the gradient-based metric algorithm on the following GitHub repository: https://github.com/GDMG99/Gradient-based-metric-for-image-defogging-without-ground-truth.

As Fig. 1 shows, the main effect that hazy weather has on a scene is decreased luminance and contrast, that dramatically reduce the contours and textures of the scene. Maintaining defined contours in adverse weather conditions is key to proper object recognition and segmentation, which are the basis of several applications. The visibility metric we present is based on gradient detection for image defogging evaluation. Our approach compares the gradient of the foggy image to the gradient of its defogged counterpart, i. e. after the defogging procedure is done. Hence, there is no need for a ground truth. Besides that, our method does not need to estimate any atmospheric parameter, which is difficult from a single RGB image and, in general, requires the sky to be present in the image.

Thus, as a first step, we need to obtain the derivative of both images (original and defogged), as can be seen in Fig. 1. There are several well-known image processing operators to perform such procedure. Some of the most used are Canny [29], Roberts, Prewitt, and Sobel [30]. For our method, we used the Sobel edge detector [31] due to its simplicity. The horizontal and vertical derivatives are obtained by respectively applying the horizontal and vertical kernels on the image, as shown in Eq. 1,

$$F_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \circledast I \quad ; \quad F_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \circledast I, \tag{1}$$

where F_x and F_y are the corresponding horizontal and vertical derivatives of the image I. The image integrating all gradients is retrieved following

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 September 2022 doi:10.20944/preprints202209.0127.v1

$$F = \sqrt{F_x^2 + F_y^2}. \tag{2}$$

Note that in any image, most of the pixels do not represent an edge, yielding small values in the processed gradient image. This can be appreciated in Fig. 1 where the white pixels that represent null or negligible gradients are dominant in the image. Hence, we define a threshold value for the gradient values in order to differentiate the gradients of interest from the background (white).

Gradient images are normalized to one, so our proposed threshold value is to use 5% of the maximum edge value present in the image to still keep all the relevant information related to edges while disregarding the background data. This threshold will be further discussed after presenting Eq. 4. Such 5% value could be optimized for a different dataset if needed. After obtaining the derivative of each image, we perform the relative difference between the gradient images of the fogged and its defogged counterpart pixel by pixel, as stated in Eq. 3,

$$RD(x,y) = \begin{cases} \dfrac{d_e(x,y) - f_e(x,y)}{f_e(x,y)} & d_e(x,y), f_e(x,y) > \text{threshold} \\ 0 \quad \text{otherwise} \end{cases}, \tag{3}$$

where $RD(x,y)$ is the relative difference computed at pixel (x,y), $d_e(x,y)$ is the defogged gradient image and $f_e(x,y)$ is the fogged gradient image. Let us take a moment to analyze the "*relative difference image*" so obtained. This image has the same dimensions as both input images. Each pixel stores the relative difference between the corresponding pixel of both input gradient images. If the value of a pixel in the relative difference image is positive, the strength of the gradients in the defogged image has improved, i. e. the gradient value in the defogged image is larger than the gradient value in the original image. Otherwise, if the value of a pixel in the relative difference image is negative, the strength of the gradient has decreased after the defogging algorithm. Therefore, the value of the difference quantifies the improvement in gradient strength obtained after the defogging process.

Once we compute the relative difference image, we perform the histogram of said image excluding the background pixels of the image, that is, the null values corresponding to those pixels below the threshold value. Fig. 2 presents the resulting histogram of the defogged image from the previous Fig. 1. The vast majority of edges in this image are better defined when fog is not present on the scene because of the defogging algorithm, as we would expect. Negative values close to 0 in the histogram correspond to some regions that have not been really affected by fog, or that in such areas the defogging process has introduced small variations in the gradient strength for these regions. These pixels, however, are quite residual compared to the rest. Note that positive pixels can reach values as large as 6, meaning a 6-fold improvement in the gradient strength.

At this point, the strategy of the gradient-based metric becomes clear. However, we still need a scalar value to quantify the enhancement of the defogging procedure consistent with the information that can be graphically observed in the histogram presented in Fig. 2. There are several options to obtain this numerical value. Our proposal consists on calculating the weighted ratio between the positive part of the histogram and the whole one. Mathematically,

$$R = \frac{\sum_{i=0}^{\infty} r_i^+ \cdot h(r_i^+) - \sum_{i=0}^{\infty} |r_i^-| \cdot h(r_i^-)}{\sum_{i=0}^{\infty} r_i^+ \cdot h(r_i^+) + \sum_{i=0}^{\infty} |r_i^-| \cdot h(r_i^-)} \tag{4}$$

where $r_i^{\pm}$ is the value of the relative difference, either positive or negative, and $h(r_i^{\pm})$ corresponds to the histogram value of $r_i^{\pm}$, so the total counts on the gradient image of such a value. R can take values from -1 to 1, being 1 when all the gradients have been enhanced and -1 when the defogging procedure has worsened every gradient of the image. The weighted character of the metric is used to strengthen the gradient that have been greatly

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 September 2022 doi:10.20944/preprints202209.0127.v1

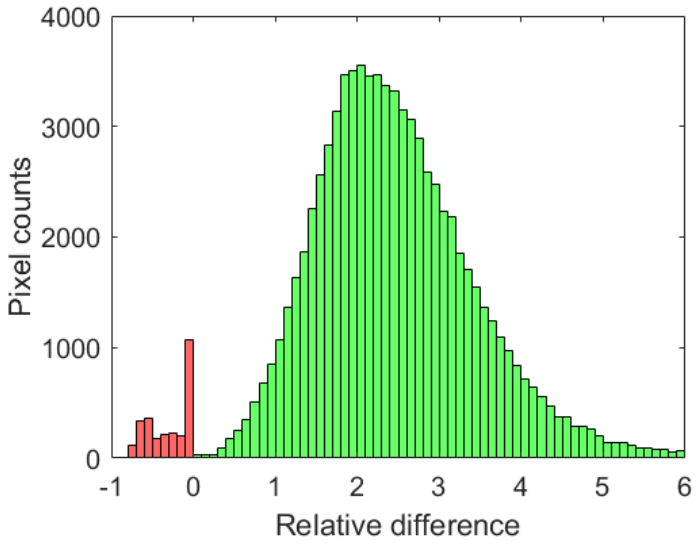

Figure 2. Histogram of the relative difference between the gradient images of the ground truth and the fogged image shown in Fig. 1. The positive values of the histogram are shown in green whereas negative values are shown in red.

improved or worsened. If we compute the proposed metric value for the example images shown in Fig. 1 we get $R = 0.9732$. This is a reasonable result since we are comparing a fogged image directly with its fog-free ground truth, mimicking an ideal defogging algorithm.

As we previously mentioned, the threshold value in Eq. 3 has been empirically fixed. We noted that setting a threshold above 8% disregards low intensity gradients, and decreases the value of the metric. On the other side, a threshold smaller than 5% considers a gradient almost any variation due to noise. We shall recall the reader that this threshold was introduced to separate the background pixels, where there are no contours, edges, or textures, from the low intensity edges.

An additional remark must be made. As previously discussed, DNNs, and especially GANs, are nowadays used to tackle defogging. GANs are very useful when it comes to generating new data that resembles the data distribution it has learned from. This means that these networks tend to generate new features in the images, which leads to new contours producing better results in our metric even if the defogging is poor.

These situations may occur with images lacking edge information. Under this condition, two scenarios could happen. First, the original haze-free might does not have any contours. In this case, fog will not be a problem since no information would be hidden due to fog. Moreover, once the defogging procedure is done the resulting image will be very similar to the original hazy one because there is no element on the scene that needs to be improved. Second, the original haze-free scene has contours, but the fog is so dense that there is no visibility. This is a more delicate case since there are elements in the image that could be improved. Nevertheless, no realistic defogging method could recover any information under such conditions. Any contour generated under extremely low visibility can in practice be considered a “ghost” object as long as it is appearing in the image from nothing.

In our opinion, generating these "*ghost*" features in the image should directly discard the defogging method. Defogging is especially useful to increase the performance of object detection and image segmentation, which will ultimately execute an action in an autonomous vehicle. Executing an action due to a "*ghost*" feature could be extremely dangerous. So our metric works under the premise that no new features are added to the defogged image during the defogging procedure, and only already existing features are highlighted.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 September 2022 doi:10.20944/preprints202209.0127.v1

A complete algorithm for the metric computation is presented in Algorithm 1.

Algorithm 1: Algorithm to compute the gradient-based metric for image defogging without ground truth.

- Obtain the gradient images of both Fogged and Defogged images:

$$F_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \circledast I \quad ; \quad F_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \circledast I$$

$$F = \sqrt{F_x^2 + F_y^2}$$

- Obtain the *relative difference image* from the gradient images:

$$RD(x,y) = \begin{cases} \dfrac{d_e(x,y) - f_e(x,y)}{f_e(x,y)} & d_e(x,y), f_e(x,y) > \text{threshold} \\ 0 \quad \text{otherwise} \end{cases}$$

- Perform the histogram of the *relative difference image*;
- Obtain the ratio of enhancement:

$$R = \frac{\sum_{i=0}^{\infty} r_i^+ \cdot h(r_i^+) - \sum_{i=0}^{\infty} |r_i^-| \cdot h(r_i^-)}{\sum_{i=0}^{\infty} r_i^+ \cdot h(r_i^+) + \sum_{i=0}^{\infty} |r_i^-| \cdot h(r_i^-)}$$

4. Results and discussion

To validate our proposed metric, we tested it on the O-Haze dataset [16]. This dataset was used in the NTIRE 2018 challenge [5]. It consists of 45 outdoor scenes. Each fogged scene has its ground truth counterpart, so there is a total of 90 images in the dataset. The example images shown in Fig. 1 are also part of this dataset. We used our metric to compare the results of some groups who participated in the challenge. During the challenge, the groups received 35 fogged images, with their respective ground truth for training their networks. They also received 5 more images for validation purposes and 5 more for testing, which were evaluated by the jury. Again, the last ten images had their respective ground truths delivered.

As mentioned above, the metrics used for evaluation in the NTIRE 2018 challenge were SSIM and PSNR, calculated relative to the ground truth image. Fig. 3 shows several examples from the O-Haze dataset. The defogged images have 800 pixels of height or width at most whereas both the ground truth and the original hazy images have greater resolutions so we resized them to match the dimensions of the defogged image, to enable proper comparison. The resize method used was the bi-cubic algorithm. After resizing, we calculated both the SSIM value and our proposed contour metric. Numerical values are shown in Table 1. The classification according to their ranking can be seen in Fig. 4, where worst and best values of each metric are plotted in red and green, respectively.

Table 1. Mean over the 45 images of the O-Haze [16] dataset of SSIM (with ground truth) and our proposed metric (without ground truth). The best and worst performing results are colored in green and red respectively for each metric.

	He *et. al.*	Meng *et. al.*	Fattal *et. al.*	Bermann *et. al.*	Cai *et. al.*	Ren *et. al.*	Ancuti *et. al.*
SSIM	0.399	0.498	0.441	0.545	0.433	0.519	**0.573**
Ours	0.911	0.872	0.769	0.944	**0.747**	0.889	**0.975**

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 September 2022 doi:10.20944/preprints202209.0127.v1

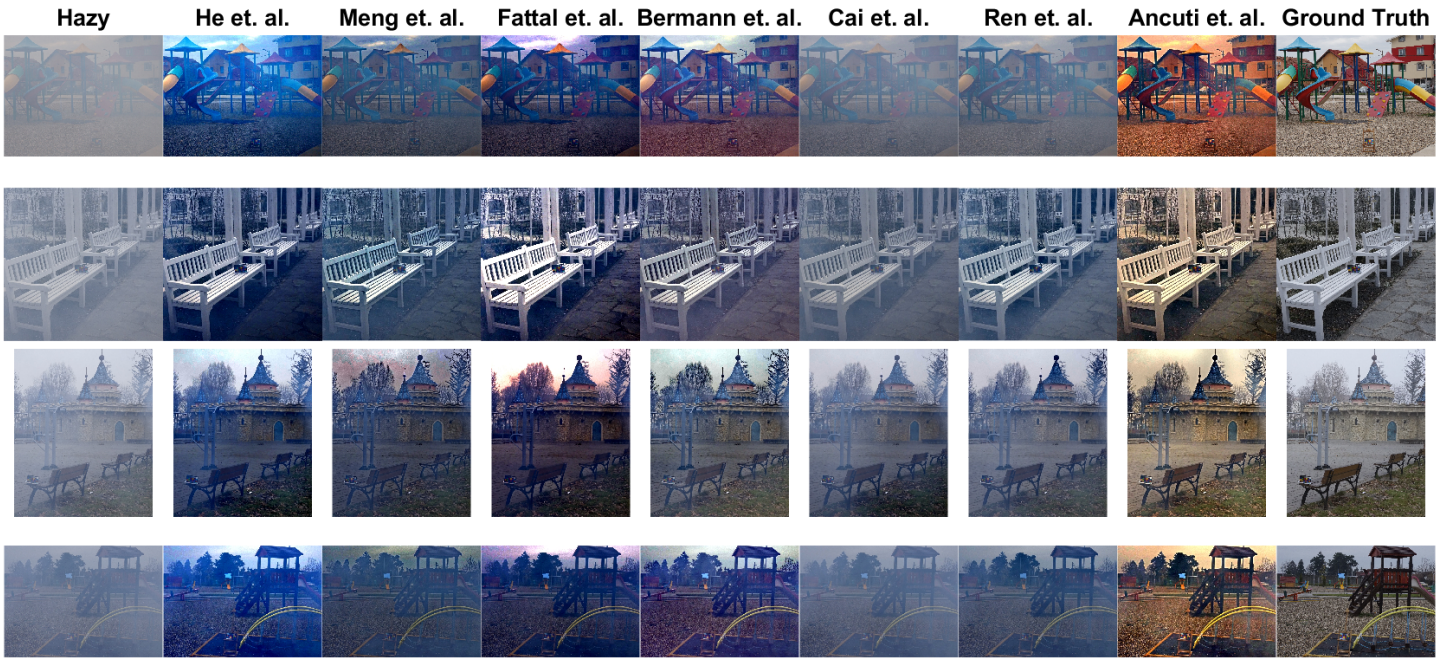

Figure 3. Several examples from the O-Haze dataset. From left to right, the hazy scene, He *et al.* [32], Meng *et. al.* [33], Fattal *et. al.* [34], Bermann *et. al.* [35], Cai *et. al.* [36], Ren *et. al.* [37], Ancuti *et. al.* [38] and the ground truth.

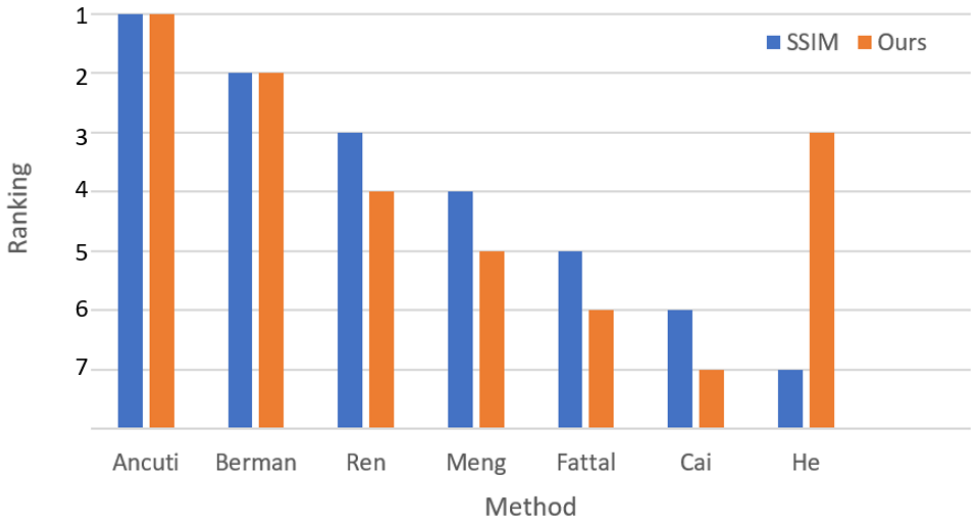

Figure 4. Classification of the mean over the 45 images of the O-Haze dataset for SSIM (with ground truth) and our proposed metric (without ground truth).

 doi:10.20944/preprints202209.0127.v1

As it can be seen, both metrics agree on the ranking of the different methods. Both SSIM and our proposal rank Ancuti as the best defogging method followed by Bermann. Ren, Fattal and Meng dispute the middle positions in both metrics. Cai performs poorly in both metrics. However, there is one significant discrepancy between the metrics, which becomes evident in the case of He. According to its SSIM value, He is the worst, despite according to our metric, it climbs up to third place. Let's discuss this issue by looking at Fig. 3. When it comes to defogging and, especially, to differentiating objects, He's results are visibly better than Meng's, Cai's or even Ren's. Nevertheless, all previous groups are ahead of them when SSIM is applied. This can be explained by looking at the colors of each image and comparing them to the ground truth. The color aberration introduced by He is measured by SSIM as a bad defogging method. On the contrary, our metric strictly considers one of the most affected features by fog, the edges of objects, leading to a more reasonable position of He's defogging method even without the need of a ground truth comparison.

In Fig. 5, we present a comparison between SSIM and our metric by showing some examples of the relative difference image histogram and the defogged result for the images corresponding to different defogging methods in Fig. 3. The figures in the last row represent the relative difference image. To ease interpretation, background is painted in white, positive edge values in green and negative ones in red. The intensity of the edges is conserved so darker regions express little difference between the fogged and defogged images. An important feature to consider is that the better the defogging method, the more similarities can be found between the histogram of the defogged image and the ground truth, having a larger positive area under the curve when our metrics value is closer to one. Also, our metrics values in this example agree with what we can observe: Ancuti's method performs a better defogging job than Meng's and Cai's. However, the same thing cannot be said about the SSIM evaluation. Moreover, according to SSIM, Cai's and Meng's resulting defogged images are worse than the original hazy image even though they visibly perform a good defogging task. Again, this proves that SSIM might not be a proper metric for image defogging evaluation.

5. Conclusions

We have proposed a gradient-based metric for image defogging that does not need a ground truth image, and measures the improvement on gradient strength on the defogged image without estimating any atmospheric parameter. We have also reviewed several state-of-the-art defogging techniques and metrics for evaluation. Finally, we compared our proposed metric with the current metric used in defogging challenges, SSIM, through the O-Haze dataset. We compared the similarities and discrepancies of both metrics and concluded that the proposed metric properly measures visual enhancement of image defogging without any reference other than the original RGB fogged scene. This metric further enables progress in the defogging field because, in particular, it enables fast validation of defogging DNNs with unpaired fog-fog free datasets. Additionally, other reference-less edge-sensitive image processing tasks like blind deblurring[39] and blind super-resolution[40] might use this metric for IQA evaluation as well. Based on its good results proved in this paper, proper adjustments on the metric's algorithm might broaden its use for other low-vision tasks like the above-mentioned.

Author Contributions: Conceptualization, Gerard deMas-Giménez, Pablo García-Gómez and Santiago Royo; methodology, Gerard deMas-Giménez and Pablo García-Gómez; software, Gerard deMas-Giménez; formal analysis, Gerard deMas-Giménez; investigation, Gerard deMas-Giménez and Pablo García-Gómez; resources, Pablo García-Gómez; data curation, Gerard deMas-Giménez; writing—original draft preparation, Gerard deMas-Giménez; writing—review and editing Santiago Royo, Josep R. Casas and Pablo García-Gómez; supervision, Santiago Royo and Josep R. Casas. All authors have read and agreed to the published version of the manuscript.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 8 September 2022 doi:10.20944/preprints202209.0127.v1

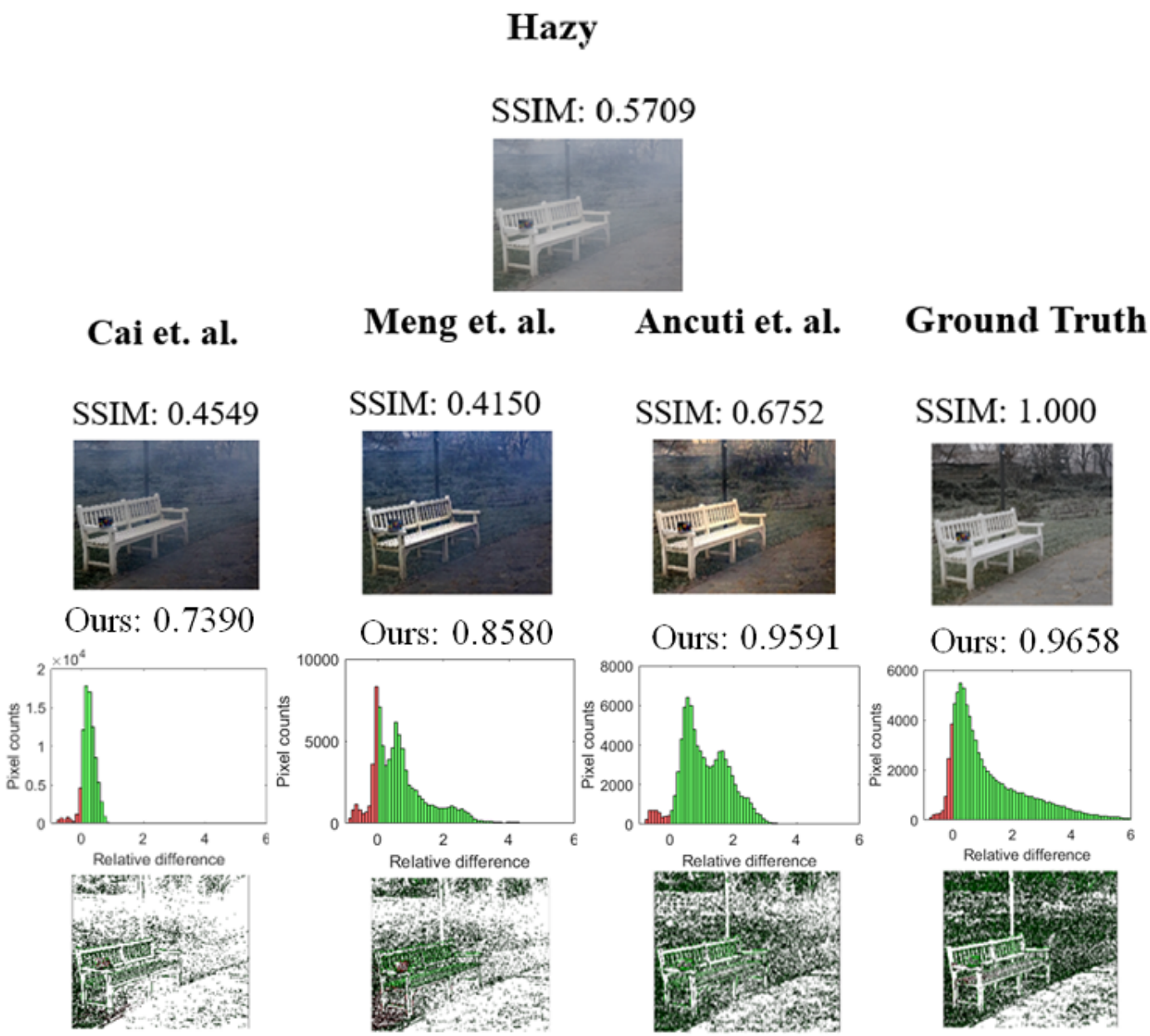

Figure 5. Comparison between SSIM and our metric on different defogging models (by columns). The first two rows correspond to the original hazy image and the defogging results. The second row corresponds to the relative difference image histogram, where positive values are represented in green and negative ones in red. The last row corresponds to the relative difference image. The white points are the background, green points are positive edge difference values, and red points are negative ones. The intensity of the difference is conserved.

Funding: This research was partially supported by MICINN and Agencia Estatal de Investigación thorugh projects PID2020-119484RB-I00 and PDC2021-121038-I00

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The O-Haze dataset, as well as the different defogging results cited on the text, have been publicly obtained at the following site https://data.vision.ee.ethz.ch/cvl/ntire18//o-haze/.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

DNNs	Deep Neural Networks
NTIRE	New Trends in Image Restoration and Enhancement
SSIM	Structural Similarity Index
PSNR	Peak Signal to Noise Ratio
FR-IQA	Full-Reference Image Quality Assessment
NR-IQA	No-Reference Image Quality Assessment
SSEQ	Spatial Spectral Entropy-based Quality
BLIINDS-II	Blind Image Integrity Notator usinge Dct Statistics
FADE	Fog Aware Density Evaluator
GANs	Generative Adversarial Networks

References

1. Gruber, T.; Julca-Aguilar, F.; Bijelic, M.; Heide, F. Gated2Depth: Real-Time Dense Lidar From Gated Images. In Proceedings of the The IEEE International Conference on Computer Vision (ICCV), 2019.
2. Schechner, Y.Y.; Narasimhan, S.G.; Nayar, S.K. Polarization-based vision through haze. Applied Optics **2003**, 42. https://doi.org/10.1364/ao.42.000511.
3. Peña-Gutiérrez, S.; Ballesta-Garcia, M.; García-Gómez, P.; Royo, S. Quantitative demonstration of the superiority of circularly polarized light in fog environments. Optics Letters **2022**, 47. https://doi.org/10.1364/ol.445339.
4. Li, X.; Han, Y.; Wang, H.; Liu, T.; Chen, S.C.; Hu, H. Polarimetric Imaging Through Scattering Media: A Review. Frontiers in Physics **2022**, 10. https://doi.org/10.3389/fphy.2022.815296.
5. Ancuti, C.; Ancuti, C.O.; Timofte, R.; Van Gool, L.; Zhang, L.; Yang, M.H.; Patel, V.M.; Zhang, H.; Sindagi, V.A.; Zhao, R.; et al. NTIRE 2018 Challenge on Image Dehazing: Methods and Results. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018, pp. 1004–100410. https://doi.org/10.1109/CVPRW.2018.00134.
6. Ancuti, C.O.; Ancuti, C.; Timofte, R.; Gool, L.V.; Zhang, L.; Yang, M.H.; Guo, T.; Li, X.; Cherukuri, V.; Monga, V.; et al. NTIRE 2019 image dehazing challenge report. 2019, Vol. 2019-June. https://doi.org/10.1109/CVPRW.2019.00277.
7. Ancuti, C.O.; Ancuti, C.; Vasluianu, F.A.; Timofte, R.; Liu, J.; Wu, H.; Xie, Y.; Qu, Y.; Ma, L.; Huang, Z.; et al. NTIRE 2020 challenge on nonhomogeneous dehazing. 2020, Vol. 2020-June. https://doi.org/10.1109/CVPRW50498.2020.00253.
8. Ancuti, C.O.; Ancuti, C.; Vasluianu, F.A.; Timofte, R.; Fu, M.; Liu, H.; Yu, Y.; Chen, J.; Wang, K.; Chang, J.; et al. NTIRE 2021 nonhomogeneous dehazing challenge report. 2021. https://doi.org/10.1109/CVPRW53098.2021.00074.
9. Ancuti, C.O.; Ancuti, C.; Hermans, C.; Bekaert, P. A fast semi-inverse approach to detect and remove the haze from a single image. 2011, Vol. 6493 LNCS. https://doi.org/10.1007/978-3-642-19309-5_39.
10. Galdran, A.; Alvarez-Gila, A.; Bria, A.; Vazquez-Corral, J.; Bertalmío, M. On the Duality Between Retinex and Image Dehazing. CoRR **2017**, abs/1712.02754, [1712.02754].
11. Engin, D.; Genc, A.; Ekenel, H.K. Cycle-dehaze: Enhanced cyclegan for single image dehazing. 2018, Vol. 2018-June. https://doi.org/10.1109/CVPRW.2018.00127.
12. Dong, Y.; Liu, Y.; Zhang, H.; Chen, S.; Qiao, Y. FD-GAN: Generative adversarial networks with fusion-discriminator for single image dehazing. 2020. https://doi.org/10.1609/aaai.v34i07.6701.
13. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, 2017. https://doi.org/10.48550/ARXIV.1703.10593.
14. Sharma, G.; Wu, W.; Dalal, E.N. The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations. Color Research and Application **2005**, 30. https://doi.org/10.1002/col.20070.
15. Duthon, P.; Colomb, M.; Bernardin, F. Fog classification by their droplet size distributions: Application to the characterization of Cerema's platform. Atmosphere **2020**, 11. https://doi.org/10.3390/atmos11060596.

16. Ancuti, C.O.; Ancuti, C.; Timofte, R.; Vleeschouwer, C.D. O-HAZE: a dehazing benchmark with real hazy and haze-free outdoor images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, NTIRE Workshop, 2018, NTIRE CVPR'18.
17. Zhao, S.; Zhang, L.; Huang, S.; Shen, Y.; Zhao, S. Dehazing Evaluation: Real-World Benchmark Datasets, Criteria, and Baselines. IEEE Transactions on Image Processing **2020**, 29. https://doi.org/10.1109/TIP.2020.2995264.
18. Middleton, W.E.K.; Twersky, V. Vision Through the Atmosphere. Physics Today **1954**, 7. https://doi.org/10.1063/1.3061544.
19. Hautiére, N.; Tarel, J.P.; Lavenant, J.; Aubert, D. Automatic fog detection and estimation of visibility distance through use of an onboard camera. Machine Vision and Applications **2006**, 17. https://doi.org/10.1007/s00138-005-0011-1.
20. Pomerleau, D. Visibility estimation from a moving vehicle using the Ralph vision system. 1997. https://doi.org/10.1109/itsc.1997.660594.
21. Liu, C.; Lu, X.; Ji, S.; Geng, W. A fog level detection method based on image HSV color histogram. 2014. https://doi.org/10.1109/PIC.2014.6972360.
22. Li, B.; Ren, W.; Fu, D.; Tao, D.; Feng, D.; Zeng, W.; Wang, Z. Benchmarking Single-Image Dehazing and beyond. IEEE Transactions on Image Processing **2019**, 28. https://doi.org/10.1109/TIP.2018.2867951.
23. Liu, L.; Liu, B.; Huang, H.; Bovik, A.C. No-reference image quality assessment based on spatial and spectral entropies. Signal Processing: Image Communication **2014**, 29. https://doi.org/10.1016/j.image.2014.06.006.
24. Saad, M.A.; Bovik, A.C.; Charrier, C. Blind image quality assessment: A natural scene statistics approach in the DCT domain. IEEE Transactions on Image Processing **2012**, 21. https://doi.org/10.1109/TIP.2012.2191563.
25. Choi, L.K.; You, J.; Bovik, A.C. Referenceless Prediction of Perceptual Fog Density and Perceptual Image Defogging. IEEE Transactions on Image Processing **2015**, 24. https://doi.org/10.1109/TIP.2015.2456502.
26. Chen, Z.; Ou, B. Visibility Detection Algorithm of Single Fog Image Based on the Ratio of Wavelength Residual Energy. Mathematical Problems in Engineering **2021**, 2021. https://doi.org/10.1155/2021/5531706.
27. Magnier, B.; Abdulrahman, H.; Montesinos, P. A review of supervised edge detection evaluation methods and an objective comparison of filtering gradient computations using hysteresis thresholds. Journal of Imaging **2018**, 4. https://doi.org/10.3390/jimaging4060074.
28. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing **2004**, 13. https://doi.org/10.1109/TIP.2003.819861.
29. Canny, J. A Computational Approach to Edge Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence **1986**, PAMI-8. https://doi.org/10.1109/TPAMI.1986.4767851.
30. Chaple, G.N.; Daruwala, R.D.; Gofane, M.S. Comparisions of Robert, Prewitt, Sobel operator based edge detection methods for real time uses on FPGA. 2015. https://doi.org/10.1109/ICTSD.2015.7095920.
31. Vincent, O.; Folorunso, O. A Descriptive Algorithm for Sobel Image Edge Detection. 2009. https://doi.org/10.28945/3351.
32. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. IEEE Transactions on Pattern Analysis and Machine Intelligence **2011**, 33. https://doi.org/10.1109/TPAMI.2010.168.
33. Meng, G.; Wang, Y.; Duan, J.; Xiang, S.; Pan, C. Efficient image dehazing with boundary constraint and contextual regularization. 2013. https://doi.org/10.1109/ICCV.2013.82.
34. Fattal, R. Dehazing using color-lines. ACM Transactions on Graphics **2014**, 34. https://doi.org/10.1145/2651362.
35. Berman, D.; Treibitz, T.; Avidan, S. Non-local Image Dehazing. 2016, Vol. 2016-December. https://doi.org/10.1109/CVPR.2016.185.
36. Cai, B.; Xu, X.; Jia, K.; Qing, C.; Tao, D. DehazeNet: An End-to-End System for Single Image Haze Removal. IEEE Transactions on Image Processing **2016**, 25, 5187–5198. https://doi.org/10.1109/tip.2016.2598681.
37. Ren, W.; Liu, S.; Zhang, H.; Pan, J.; Cao, X.; Yang, M.H. Single image dehazing via multi-scale convolutional neural networks. 2016, Vol. 9906 LNCS. https://doi.org/10.1007/978-3-319-46475-6_10.
38. Ancuti, C.O.; Ancuti, C. Single image dehazing by multi-scale fusion. IEEE Transactions on Image Processing **2013**, 22. https://doi.org/10.1109/TIP.2013.2262284.
39. Nan, Y.; Quan, Y.; Ji, H. Variational-EM-Based Deep Learning for Noise-Blind Image Deblurring. 2020. https://doi.org/10.1109/CVPR42600.2020.00368.
40. Yin, G.; Wang, W.; Yuan, Z.; Ji, W.; Yu, D.; Sun, S.; Chua, T.S.; Wang, C. Conditional Hyper-Network for Blind Super-Resolution With Multiple Degradations. IEEE Transactions on Image Processing **2022**, 31, 3949–3960. https://doi.org/10.1109/tip.2022.3176526.