SEMARAK ILMU PUBLISHING
202103268166(003316878-P)

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage:
https://semarakilmu.com.my/journals/index.php/applied_sciences_eng_tech/index
ISSN: 2462-1943

Analyzing the Impact of Loss Functions on Dehazing Effectiveness and Unveiling the Discrepancy between Quantitative and Qualitative Results

Ali Murtaza[1,2,*], Uswah Binti Khairuddin[1,2], Ahmad 'Athif Mohd Fauzi[1], Kazuhiko Hamamoto[3]

[1] Centre for Artificial Intelligence and Robotics, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[2] Malaysia-Japan International Institute of Technology, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[3] Graduate School of Information and Telecommunication Engineering, Tokai University, Japan

ARTICLE INFO

Article history:
Received 9 September 2023
Received in revised form 12 December 2023
Accepted 10 February 2024
Available online 25 April 2024

Keywords:
Dehazing; Loss Functions; Computer Vision; SSIM; MS-SSIM; IW-SSIM; MAE

ABSTRACT

The work focuses on the analysis of the impact of loss functions on the effectiveness of a model for dehazing images. Dehazing, the process of removing haze or atmospheric scattering from images, plays a crucial role in various computer vision applications. To enhance the performance of dehazing models, it is essential to examine different loss functions and their variations. In this study, we employ a Generative Adversarial Network (GAN) as our model and evaluate the performance of various loss functions. The primary objective is to assess how well each loss function is capable of dehazing an image, while specifically investigating the influence of various structural similarity index (SSIM) loss variations on the dehazing effectiveness. Our experimental results reveal a notable discrepancy between qualitative and quantitative outcomes. Contradicting the traditional interpretation in literature, our qualitative analysis reveals that the SSIM IQA metric may not be a fully reliable indicator of dehazing effectiveness despite it being viewed to be correlated to human visual perception unlike Mean Square Error and Peak Signal to Noise Ratio Metrics. Moreover, we demonstrate that relying solely on quantitative results may lead to the selection of an inappropriate loss function. This finding emphasizes the significance of qualitative analysis in evaluating the performance of dehazing models. The disparity between quantitative and qualitative results emphasizes the need for newer image assessment metrics in the domain that can effectively bridge this gap. Such metrics should be able to be better correlated with human perception. This research contributes to the field of image dehazing by shedding light on the importance of qualitative analysis in addition to quantitative evaluation. By comprehensively analyzing the impact of variations of SSIM loss functions and their combination with Mean Absolute Error (MAE) loss, we provide valuable insights into enhancing the effectiveness of dehazing models.

1. Introduction

Contaminants in the air, such as haze, fog, and smoke, can have a significant impact on the quality of images captured by cameras. These airborne particles scatter light, leading to a loss of contrast,

* *Corresponding author.*
E-mail address: murtaza@graduate.utm.my

https://doi.org/10.37934/araset.44.1.1221

color saturation, and sharpness in the image. According to [1] this effect is especially pronounced at longer ranges, where the scattered light becomes dominant. Haze, fog, or smoke in images detrimentally affect computer vision applications like object detection and recognition as shown by [2]. In Literature, the process of removing or minimizing the effects of these contaminants is generally referred to as 'Dehazing'. Dehazing techniques aim to remove these effects, improve image quality, and restore lost details, benefiting object detection and recognition systems. This has benefits in tasks like discovery of Archeological structures where quality and quantity of datasets is a significant challenge as mentioned by [3], dehazing can improve datasets by removing clouds, haze, dust and et cetera from the images.

The advent of Deep Learning and its superior ability to deal with images compared to traditional Machine Learning approaches have revolutionized the domain of Dehazing. While significant emphasis has been placed on the development of diverse deep learning models for image dehazing, relatively less attention has been directed towards the loss layer within these models. The choice of loss functions in deep learning is of paramount importance as they play a crucial role in training neural networks and optimizing their performance. Furthermore, the careful selection of an appropriate loss function can significantly enhance the performance of a deep learning model, thereby maximizing its effectiveness. By optimizing the loss function, it becomes possible to achieve improved results without the necessity of constructing an entirely new model to attain a similar level of performance.

One of the most commonly used loss is Mean Square Error Loss or $\mathcal{L}_2$ Loss which has been extensively utilized in dehazing techniques like the works by [4-8]. MSE loss places higher emphasis on larger differences between pixel values, penalizing larger errors more heavily. Since it magnifies the impact of outliers or large error it can lead to over smoothing and blurry reconstruction as shown by [9]. Mean Absolute Error or $\mathcal{L}_1$ Loss performs better than $\mathcal{L}_2$ Loss since it is less sensitive to outliers and hence produces sharper images and better edge preservation. This loss too has been widely used too like in the works of [10,11]. Furthermore work by [11] shows that both $\mathcal{L}_1$ and $\mathcal{L}_2$ loss prevent artifacting.

To overcome the limitations of these losses they are often combined with other losses to improve the performance of a model. One such loss is Structural Similarity Index Measure (SSIM) Loss based on SSIM IQA metric proposed by [12]. Diverging from traditional metrics that primarily account for pixel-level disparities, SSIM is widely regarded as a metric that captures the human visual system's sensitivity to structural distortions. It assesses image quality based on three components: luminance comparison, contrast comparison, and structural comparison. SSIM has become a popular choice for image quality assessment in various image processing applications and is one of the main benchmarks for Hazy Images. Research by [13,14] utilized this loss for Dehazing purposed and demonstrated it preserved details related to Brightness, Contrast and Structure.

Multi-Scale Structural Similarity Index or MS-SSIM proposed by [15] is an extension of SSIM operating on multiple scales to capture both global and local information from an image. By considering multiple scales, MS-SSIM is traditionally thought to better account for the perceptual characteristics of human vision, capturing both fine details and global structural information. It has been used in various image reconstruction/enhancement tasks like [16] denoising, super-resolution and artifact removal which demonstrated superiority of MS-SSIM over SSIM in these domains. Work by [17] further demonstrated that MS-SSIM encouraged the network to learn low to mid-level structures within the image & it was the loss that generated the most pleasing results as perceived by a human. However, when it comes to dehazing, MS-SSIM loss has not been experimented in any other literature. While there has been development of newer loss related in the visual domain over recent years, they tend to be complex, computationally expensive since they often rely on neural networks and lack interpretability. Meanwhile, the work by [18] demonstrated that MS-SSIM performed almost as well as these newer losses while not having the drawbacks related to them.

IW-SSIM (Information Weighted Structural Similarity Index) proposed by [19] is an advanced variant of the SSIM metric that incorporates information theory principles. By assigning different weights to image regions based on their perceptual importance, IW-SSIM provides a more refined assessment of image quality. This approach enhances the sensitivity of the metric to structural details, leading to a more accurate evaluation of image similarity and perceptual fidelity. It too has not been utilized as a loss function in the domain of dehazing.

Our research will evaluate the performance of different loss functions in the context of dehazing by employing three metrics: $\mathcal{L}_1$ Loss, SSIM-Loss, MS-SSIM Loss, IW-SSIM loss $\mathcal{L}_1$ Loss combined with SSIM, $\mathcal{L}_1$ Loss combined with MS-SSIM, and L1 Loss combined with IW-SSIM. We will perform quantitative and qualitative analysis on the results of these losses to measure the effectiveness of each approach in improving dehazing outcomes. Furthermore, we will investigate if quantitative analysis correlates with qualitative analysis, something which has been inadequately investigated in the literature. By comparing the results obtained from these configurations of loss function, we aim to gain insights into the impact of incorporating what are traditionally thought to be perceptual similarity metrics, SSIM, MS-SSIM and IW-SSIM, alongside the traditional $\mathcal{L}_1$ Loss in deep learning-based dehazing models.

2. Methodology

The primary objective of our study is to investigate the influence of the $\mathcal{L}_1$, SSIM, MS-SSIM & IW-SSIM loss functions on the performance of a Deep Learning Model. Specifically, we have selected the Pix2Pix GAN [9] shown in Figure 1 as the Deep Learning Model for our analysis. The Pix2Pix Model in literature uses Adversarial Loss and L1 Loss.

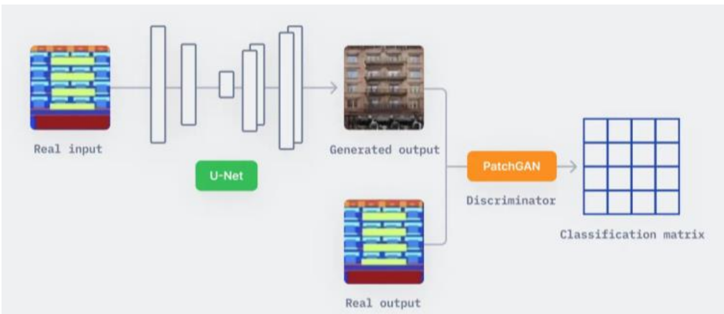

Fig. 1. Pix2Pix GAN Architecture

The model will be trained on O–Haze [20] Database consisting of real world pair of Hazy and Ground Truth Images. Considering the limited availability of images for our study, Data Augmentation by extracting patches of size 256x256 from the images will be employed to expand the dataset for training. Of the original pair of images, 3 will be kept for validation, 5 for testing and the remaining for training. Validation dataset too will have patches extracted from it to augment its size.

In the Pix2Pix Paper the $\mathcal{L}_1$ loss functions were multiplied with $\boldsymbol{\lambda}$ of value 100. Since we will be using SSIM, MS-SSIM and IW-SSIM Loss as well we will need to assign them with a λ value too, in which we chose it to be 100. So, the loss functions we will be investigating are:

$$G^* = \arg\min_{G}\max_{D} L_{cGAN}(G,D) + \boldsymbol{\lambda_1}\mathrm{L}_{L1}(G) \quad (1)$$
$$G^* = \arg\min_{G}\max_{D} L_{cGAN}(G,D) + \boldsymbol{\lambda_1}L_{SSIM}(G) \quad (2)$$
$$G^* = \arg\min_{G}\max_{D} L_{cGAN}(G,D) + \boldsymbol{\lambda_1}L_{MS-SSIM}(G) \quad (3)$$
$$G^* = \arg\min_{G}\max_{D} L_{cGAN}(G,D) + \boldsymbol{\lambda_1}L_{IW-SSIM}(G) \quad (4)$$
$$G^* = \arg\min_{G}\max_{D} L_{cGAN}(G,D) + \boldsymbol{\lambda_1}\mathrm{L}_{L1}(G) + \boldsymbol{\lambda}_2 L_{SSIM}(G) \quad (5)$$
$$G^* = \arg\min_{G}\max_{D} L_{cGAN}(G,D) + \boldsymbol{\lambda_1}\mathrm{L}_{L1}(G) + \boldsymbol{\lambda}_2 L_{MS-SSIM}(G) \quad (6)$$
$$G^* = \arg\min_{G}\max_{D} L_{cGAN}(G,D) + \boldsymbol{\lambda_1}\mathrm{L}_{L1}(G) + \boldsymbol{\lambda}_2 L_{IW-SSIM}(G) \quad (7)$$

Where G^* is our final objective. G represents Generator, D represents Discriminator, L_{cGAN} is Adversarial Loss, L_{L1} is $\mathcal{L}1$ Loss, L_{SSIM} is SSIM Loss, $L_{MS-SSIM}$ is MS-SSIM Loss, $\boldsymbol{\lambda_1}$ and $\boldsymbol{\lambda_2}$ are Loss coefficient which value is set to 100. Mean of each L_{L1}, L_{SSIM}, $L_{MS-SSIM}$ and $L_{IW-SSIM}$ values will be taken for each batch.

Each model will be trained for 25 epochs using Adam Optimizer with Learning Rate of 0.0001, $\beta_1 = 0.5$, $\beta 2 = 0.999$. Batch size of training set is 8 and validation set is 48. Shuffle for training set is set to True, while for validation set, it is False. Our training device is Intel Core i7-8700 CPU @ 3.20GHz, 16GB RAM, NVIDIA GeForce RTX 3070 (GPU Memory 8.0GB).

The model is created using PyTorch deep learning framework and SSIM and MS-SSIM losses calculated using the PyTorch Image Quality (PIQ) Library [21,22]. To ensure result consistency and comparability, we minimize inherent randomness by setting a fixed random seed (29) for Python, the deep learning architecture (e.g., PyTorch), CUDA, and the data loader. Additionally, deterministic convolutional algorithms are employed within the PyTorch framework, guaranteeing consistent outputs for a given input and eliminating randomness associated with non-deterministic implementations.

To comprehensively evaluate the dehazing performance of our model, we acknowledge the necessity of a dual-pronged approach that incorporates both quantitative and qualitative analyses. While quantitative analysis relies on established Image Quality Assessment (IQA) metrics from the literature, it is important to recognize their limitations in fully capturing human visual perception. Therefore, qualitative analysis assumes a vital role in our evaluation process, allowing us to consider subjective factors such as the effectiveness of dehazing, visual clarity, naturalness, and realism of the dehazed images. By combining objective and subjective aspects, our evaluation aims to provide a holistic assessment of dehazing performance.

3. Results

Considering the inherently subjective nature of data, a need for a comprehensive approach to evaluate dehazing performance of our model is recognized. This entails employing both quantitative and qualitative analysis. For Quantitative analysis Image Quality Assessment (IQA) metrics established in literature will be utilized. However, IQA metrics may not fully capture the nuances and subtleties of human visual perception. Therefore, qualitative analysis will also play a crucial role in our evaluation process. This approach allows us to consider factors that are difficult to quantify, such as how well dehazing was performed, the overall visual clarity and realism of the dehazed images. Our goal is to ensure that our evaluation process captures both objective and subjective aspects.

3.1 Quantitative Analysis

In dehazing research, commonly employed IQA metrics include MSE, PSNR, and SSIM. Our analysis involved resizing five test images to 2048x2048 and extracting 16 patches from each. IQA metrics were computed for each patch, and their average and standard deviation were calculated for the entire set. The results are summarized in Table 1. The MSE score measures pixel-level dissimilarity rather than perceptual differences, PSNR assesses detail preservation, and SSIM evaluates structural preservation. In literature, SSIM is generally regarded as an indicator of dehazing effectiveness.

From Table 1 it can be seen that L_{cGAN}+ L_{L1} outperforms all metrics based on MSE score. However, it ranks low when it comes to PSNR and SSIM score. In fact, it is the worst by some margin when it comes to SSIM indicating it does not preserve structures well. L_{cGAN}+ L_{L1} + L_{SSIM} however is ranked the highest by SSIM and PSNR metrics, while it is very narrowly beaten by L_{cGAN}+ L_{L1} for the first position in the MSE metric. Another trend that can be seen is that without L1 LSSIM, LMS-SSIM and LIW-SSIM perform the worst in the metrics but combining them with L1 loss boosts their performance significantly so much so that they outperform the individual losses. Interestingly MS-SSIM is outperformed by SSIM both individually and when they are combined with L1 loss. A reason for that might be that its added complexity leads to a higher number of parameters to calibrate. So, there is a possibility that finetuning MS-SSIM's parameters might improve its performance over SSIM as has been proven by research in other domains. IW-SSIM's performance is only improved one rank in each of the metrics by the addition of L1 loss, unlike SSIM & MS-SSIM Loss which were amongst the worst performers when alone but became the highest performers when combined with L1.

Table 1
MSE, PSNR SSIM average values and standard deviation for each loss function. Ranking performed based on average values

Losses	MSE	Std Dev	Rank	PSNR	Std Dev	Ranking	SSIM	Std Dev	Rank
$L_{cGAN} + L_{L1}$	0.0062	0.0118	1	21.52	3.91	4	0.7275	0.0984	7
$L_{cGAN} + L_{SSIM}$	0.0082	0.1723	7	20.95	4.09	6	0.7441	0.1020	5
$L_{cGAN} + L_{MS-SSIM}$	0.0082	0.0152	7	20.62	3.96	7	0.7350	0.1009	6
$L_{cGAN} + L_{IW-SSIM}$	0.0072	0.0127	5	21.30	3.95	5	0.7472	0.0939	4
$L_{cGAN} + L_{L1} + L_{SSIM}$	0.0063	0.0146	2	21.98	3.93	1	0.7716	0.0949	1
$L_{cGAN} + L_{L1} + L_{MS-SSIM}$	0.0064	0.0123	3	21.76	4.21	2	0.7521	0.0993	2
$L_{cGAN} + L_{L1} + L_{IW-SSIM}$	0.0066	0.0130	4	21.62	4.09	3	0.7457	0.0983	3

A comparison between IQA metrics of whole test images (resized to 256x256) is made in Table 2. The results largely follow the trend from Table 1 with various SSIM related losses being the best performers when combined with L1 Loss. The only new significant finding is that MS-SSIM drops down to being the 4th best performer when it comes to SSIM score. However, the difference is marginal between it and the ones above it. Figure 2 shows cumulative ranking by combining rankings from Table 1 and Table 2 for easier interpretability.

Table 2
MSE, PSNR SSIM values for each loss function across 5 images. Ranking performed based on average values

Comparing MSE Values							
Losses	Image 1	Image 2	Image 3	Image 4	Image 5	Average Value	Rank
$L_{cGAN} + L_{L1}$	0.0025	0.0133	0.0024	0.0036	0.0126	0.0069	5
$L_{cGAN} + L_{SSIM}$	0.0057	0.0131	0.0027	0.0042	0.0069	0.0065	4

$L_{cGAN} + L_{MS-SSIM}$	0.0379	0.0176	0.0023	0.0054	0.0050	0.0136	7
$L_{cGAN} + L_{IW-SSIM}$	0.0227	0.0046	0.0023	0.0052	0.0047	0.0079	6
$L_{cGAN} + L_{L1} + L_{SSIM}$	0.0020	0.0134	0.0037	0.0042	0.0059	0.0058	3
$L_{cGAN} + L_{L1} + L_{MS-SSIM}$	0.0021	0.0090	0.0023	0.0031	0.0043	0.0042	1
$L_{cGAN} + L_{L1} + L_{IW-SSIM}$	0.0029	0.0088	0.0027	0.0032	0.0049	0.0045	2
Comparing PSNR Values							
Losses	Image 1	Image 2	Image 3	Image 4	Image 5	Average Value	Rank
$L_{cGAN} + L_{L1}$	25.42	16.13	24.75	23.90	17.58	21.56	6
$L_{cGAN} + L_{SSIM}$	21.79	17.42	24.50	23.08	19.15	21.19	7
$L_{cGAN} + L_{MS-SSIM}$	23.46	16.05	25.33	22.09	21.10	21.61	4
$L_{cGAN} + L_{IW-SSIM}$	25.91	12.42	25.63	22.59	21.49	21.61	4
$L_{cGAN} + L_{L1} + L_{SSIM}$	26.40	17.29	23.30	23.04	20.66	22.14	3
$L_{cGAN} + L_{L1} + L_{MS-SSIM}$	26.23	18.54	25.82	24.57	21.80	23.39	1
$L_{cGAN} + L_{L1} + L_{IW-SSIM}$	24.79	19.36	24.70	24.53	21.85	23.05	2
Comparing SSIM Values							
Losses	Image 1	Image 2	Image 3	Image 4	Image 5	Average Value	Rank
$L_{cGAN} + L_{L1}$	0.8310	0.6099	0.8485	0.8172	0.5987	0.7411	7
$L_{cGAN} + L_{SSIM}$	0.8234	0.7897	0.8877	0.8490	0.6829	0.8065	2
$L_{cGAN} + L_{MS-SSIM}$	0.8404	0.7191	0.8524	0.8221	0.6552	0.7778	5
$L_{cGAN} + L_{IW-SSIM}$	0.8566	0.6135	0.8663	0.8373	0.6803	0.7708	6
$L_{cGAN} + L_{L1} + L_{SSIM}$	0.8746	0.7637	0.8797	0.8516	0.6770	0.8093	1
$L_{cGAN} + L_{L1} + L_{MS-SSIM}$	0.8576	0.7049	0.8718	0.8460	0.6762	0.7913	4
$L_{cGAN} + L_{L1} + L_{IW-SSIM}$	0.8517	0.7283	0.8617	0.8504	0.6686	0.7921	3

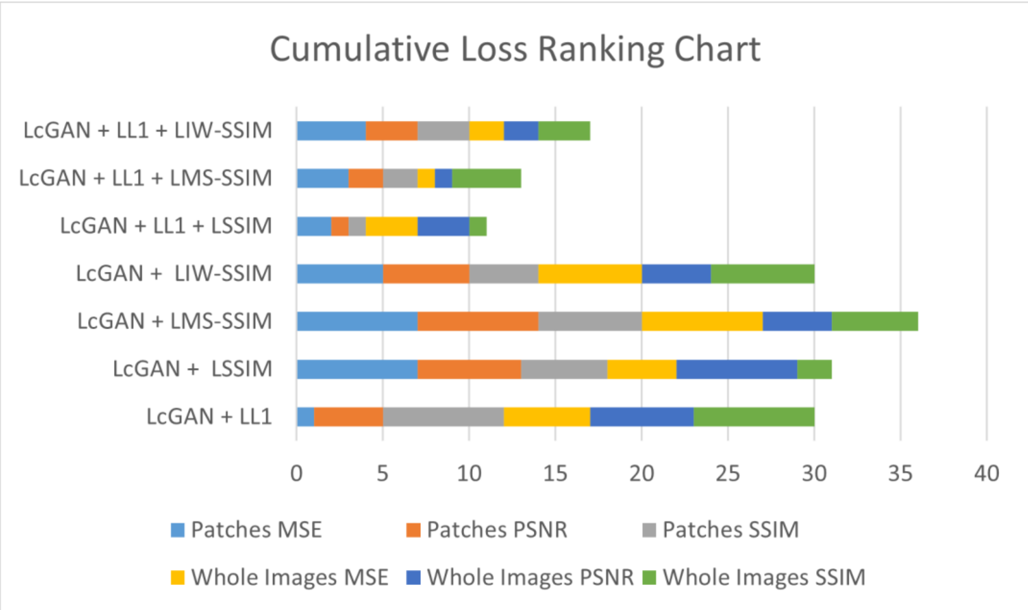

Fig. 2. Cumulative ranking of the loss functions. Lower is better

3.2 Qualitative Analysis

Qualitative analysis is essential alongside quantitative analysis when evaluating dehazing performance. Unlike quantitative measures, qualitative analysis provides insights that quantitative measures cannot capture. It enables visual perception assessment, artifact detection, and facilitates comparative analysis between algorithms. Additionally, research has demonstrated that IQA metrics often fail to accurately predict human perception, sometimes resulting in seemingly good results despite lingering haze. Therefore, conducting qualitative analysis is vital to fully understand the effectiveness of dehazing.

Figure 3 displays a list of Hazy Images, Ground Truth Images and their dehazed versions generated via different loss functions. The first 5 images are patches extracted from the test set while the last two images are whole images from test set (Images 1 and 5 from Table 2).

After evaluating the patches, the dehazing performance reveals a favorable outcome, wherein all loss combinations demonstrate good results with no significant variations among them. Notably, the preservation of intricate details remains prominent across the tested methods. However, in terms of color fidelity, IW-SSIM and MS-SSIM (with and without L1) exhibit a closer resemblance to the Ground Truth images, while the remaining methods exhibit a subtly warmer color representation.

The effectiveness of dehazing images is most evident in image (g), which initially received a low score. Notably, the most successful dehazing techniques were SSIM, IW-SSIM, and MS-SSIM, whereas the remaining images still retained some level of haze, particularly when L1 was used alone, resulting in the worst performance. This pattern suggests that L1 leads to inadequate dehazing performance when combined with SSIM-based loss functions or when alone, despite its demonstrated improvement in quantitative performance as indicated in Table 2. To further bolster our hypothesis, we examine image (f). Although the overall performance is superior in this image, the best dehazing outcomes are once again achieved by SSIM, IW-SSIM, and MS-SSIM when not combined with L1. Conversely, the L1 function proves to be the least effective once more. Consequently, our initial hypothesis drawn from image (g) is further reinforced, indicating that L1 loss results in inferior perceptual performance, despite its potential to enhance quantitative measurements.

This observation emphasizes the significance of incorporating qualitative analysis alongside quantitative analysis in the field of dehazing since our analysis between the two has produced contradicting results. The images with higher SSIM scores had more haze than ones with lower SSIM scores. It becomes evident that relying solely on present quantitative measurements can be misleading and may not accurately reflect how humans perceive image quality. While several studies have suggested a correlation between SSIM and human perception in our study on dehazing these findings are challenged which validates the findings of [23] whose work suggested that SSIM was not closely related to human perception and in its essence was just a statistical measure. Therefore, it is crucial to develop or incorporate new full-reference IQA metrics specifically tailored for evaluating dehazing quality. The existing metrics demonstrate inadequacy in capturing the nuanced perceptual aspects of dehazing, highlighting the need for improved evaluation methods that better align with human perception.

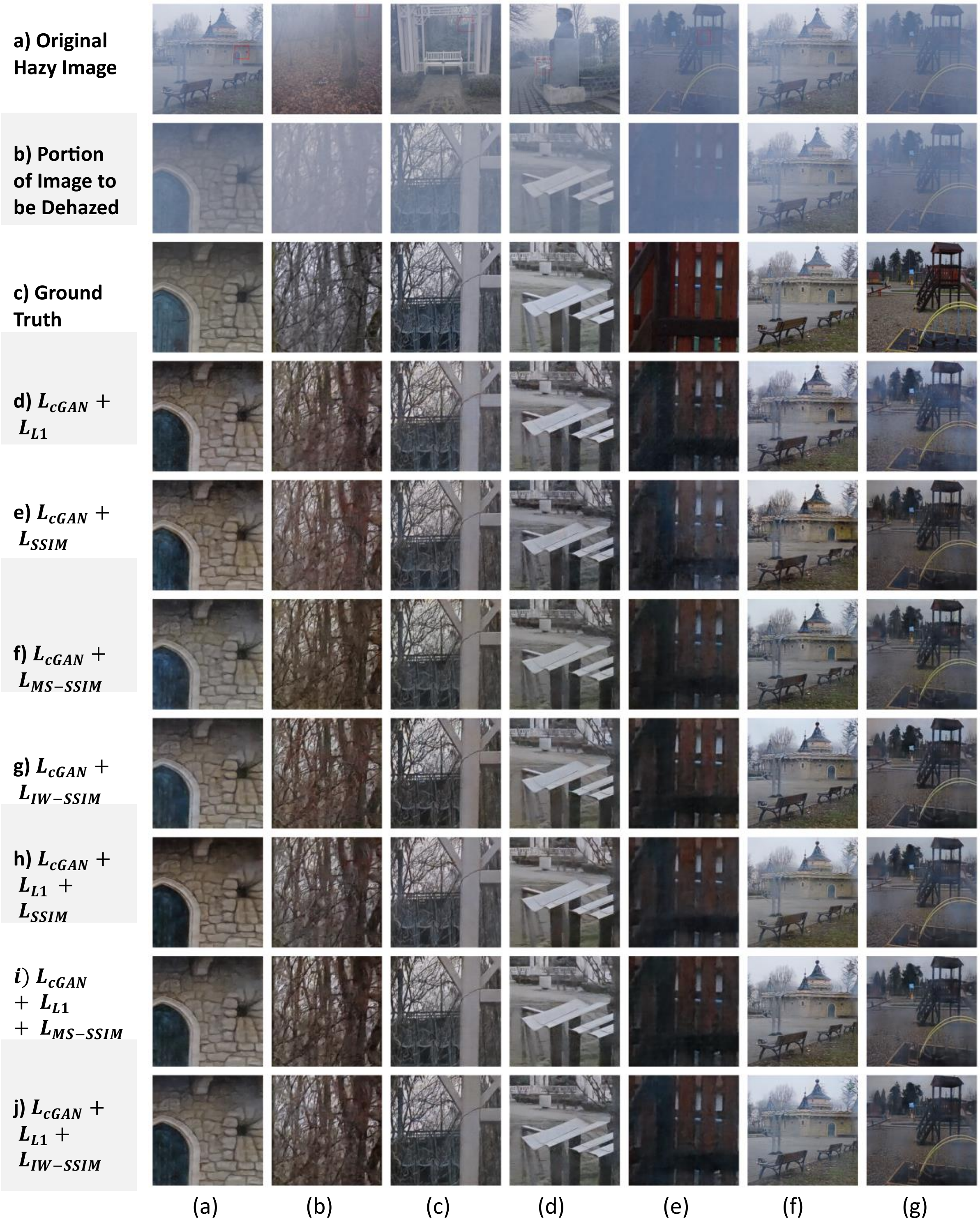

Fig. 3. Dehazing results. Images (a) to (e) are extracted patches while (f) & (g) are whole images

4. Conclusion

In conclusion, our research examined the impact of loss function selection on dehazing performance. We found that there is a lack of correlation between quantitative and qualitative results in dehazing evaluation, cautioning against relying solely on quantitative measures. Specifically, our investigation revealed that utilizing L1 loss improves quantitative performance but leads to poorer qualitative outcomes. Furthermore, our findings indicate that SSIM scores do not show a strong correlation with dehazing effectiveness as traditionally thought and relying on it can give misleading results. This highlights the need for the development or integration of new IQA metrics in the domain of Dehazing to bridge the gap between quantitative and qualitative assessments. Furthermore, our work emphasizes the importance of incorporating the need for a thorough qualitative analysis instead of solely relying on quantitative results to measure the performance of dehazing since quantitative results can lead to misleading results. However, it is important to note that our study was limited to a single dataset, and in the future, we plan to conduct further research to validate our findings. Overall, our research emphasizes the importance of the choice of loss functions, complexity of evaluating dehazing performance and the importance of considering qualitative aspects in assessment in addition to quantitative results.

Acknowledgement
This research was not funded by any grant.

References
[1] Fattal, Raanan. "Single image dehazing." *ACM transactions on graphics (TOG)* 27, no. 3 (2008): 1-9. https://doi.org/10.1145/1360612.1360671
[2] Feng, Yue. "A survey on video dehazing using deep learning." In *Journal of Physics: Conference Series*, vol. 1487, no. 1, p. 012018. IOP Publishing, 2020. https://doi.org/10.1088/1742-6596/1487/1/012018
[3] Jamil, Amirah Hanani, Fitri Yakub, Azizul Azizan, Shairatul Akma Roslan, Sheikh Ahmad Zaki, and Syafiq Asyraff Ahmad. "A review on Deep Learning application for detection of archaeological structures." *Journal of Advanced Research in Applied Sciences and Engineering Technology* 26, no. 1 (2022): 7-14. https://doi.org/10.37934/araset.26.1.714
[4] Fazlali, Hamidreza, Shahram Shirani, Mike McDonald, Daly Brown, and Thia Kirubarajan. "Aerial image dehazing using a deep convolutional autoencoder." *Multimedia Tools and Applications* 79 (2020): 29493-29511. https://doi.org/10.1007/s11042-020-09383-7
[5] Li, Boyi, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. "Aod-net: All-in-one dehazing network." In *Proceedings of the IEEE international conference on computer vision*, pp. 4770-4778. 2017. https://doi.org/10.1109/ICCV.2017.511
[6] Fu, Hui, Bin Wu, and Yanhua Shao. "Multi-feature-based bilinear CNN for single image dehazing." *IEEE Access* 7 (2019): 74316-74326. https://doi.org/10.1109/ACCESS.2019.2920537
[7] Kumar, Prasen, Sharma Priyankar, and Jain Arijit Sur. "Scale-aware conditional generative adversarial network for image dehazing." In *2020 IEEE Winter conference on applications of computer vision (WACV)*, pp. 2344-2354. IEEE, 2020. https://doi.org/10.1109/WACV45572.2020.9093528
[8] 'Convolutional Autoencoder For Single Image Dehazing | IEEE Conference Publication | IEEE Xplore'. Accessed: May 15, 2023 [Online]. Available: https://doi.org/10.1109/ICIP.2019.8803478
[9] Isola, Phillip, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-image translation with conditional adversarial networks." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1125-1134. 2017. https://doi.org/10.1109/CVPR.2017.632
[10] Wu, Haiyan, Yanyun Qu, Shaohui Lin, Jian Zhou, Ruizhi Qiao, Zhizhong Zhang, Yuan Xie, and Lizhuang Ma. "Contrastive learning for compact single image dehazing." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10551-10560. 2021. https://doi.org/10.1109/CVPR46437.2021.01041
[11] Yang, Dong, and Jian Sun. "Proximal dehaze-net: A prior learning-based deep network for single image dehazing." In *Proceedings of the european conference on computer vision (ECCV)*, pp. 702-717. 2018. https://doi.org/10.1007/978-3-030-01234-2_43
[12] Wang, Zhou, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. "Image quality assessment: from error visibility to structural similarity." *IEEE transactions on image processing* 13, no. 4 (2004): 600-612. https://doi.org/10.1109/TIP.2003.819861
[13] Zhu, Lijun, Weibo Wei, Zhenkuan Pan, Lianshun Ji, Jintao Song, and Jinhan Li. "Semi-supervised learning dehazing algorithm based on the OSV model." *IET Image Processing* 17, no. 3 (2023): 872-885. https://doi.org/10.1049/ipr2.12679
[14] Li, Shaoyi, Jian Lin, Xi Yang, Jun Ma, and Yifeng Chen. "BPFD-Net: enhanced dehazing model based on Pix2pix framework for single image." *Machine Vision and Applications* 32, no. 6 (2021): 124. https://doi.org/10.1007/s00138-021-01248-9

[15] Wang, Zhou, Eero P. Simoncelli, and Alan C. Bovik. "Multiscale structural similarity for image quality assessment." In *The Thrity-Seventh Asilomar Conference on Signals, Systems & Computers, 2003*, vol. 2, pp. 1398-1402. Ieee, 2003. https://doi.org/10.1109/ACSSC.2003.1292216

[16] Zhao, Hang, Orazio Gallo, Iuri Frosio, and Jan Kautz. "Loss functions for image restoration with neural networks." *IEEE Transactions on computational imaging* 3, no. 1 (2016): 47-57. https://doi.org/10.1109/TCI.2016.2644865

[17] Dosovitskiy, Alexey, and Thomas Brox. "Generating images with perceptual similarity metrics based on deep networks." *Advances in neural information processing systems* 29 (2016). https://doi.org/10.48550/arXiv.1602.02644

[18] Ding, Keyan, Kede Ma, Shiqi Wang, and Eero P. Simoncelli. "Comparison of full-reference image quality models for optimization of image processing systems." *International Journal of Computer Vision* 129, no. 4 (2021): 1258-1281. https://doi.org/10.1007/s11263-020-01419-7

[19] Wang, Zhou, and Qiang Li. "Information content weighting for perceptual image quality assessment." *IEEE Transactions on image processing* 20, no. 5 (2010): 1185-1198. https://doi.org/10.1109/TIP.2010.2092435

[20] Ancuti, Codruta O., Cosmin Ancuti, Radu Timofte, and Christophe De Vleeschouwer. "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 workshops*, pp. 754-762. 2018. https://doi.org/10.1109/CVPRW.2018.00119

[21] S. Kastryulin, J. Zakirov, D. Prokopenko, and D. V. Dylov, 'PyTorch Image Quality: Metrics for Image Quality Assessment'. arXiv, Aug. 31, 2022. Accessed: May 19, 2023 [Online]. Available: https://doi.org/10.48550/arXiv.2208.14818

[22] S. Kastryulin, D. Zakirov, and D. Prokopenko, 'PyTorch Image Quality: Metrics and Measure for Image Quality Assessment'. 2019 [Online]. Available: https://github.com/photosynthesis-team/piq

[23] J. Nilsson and T. Akenine-Möller, *Understanding SSIM*. arXiv.org, 2020. Accessed: Jun. 13, 2023 [Online]. Available: https://research.nvidia.com/publication/2020-07_Understanding-SSIM