

Enhanced Satellite Image Haze Removal using Deep Learning Convolutional Neural Networks and Multiple Exposure Fusion

Pilli Indira Kumari¹, Duddu Satish Kumar², D.Karthikeyan S³

^{1,2,3}Assistant Professor, Department of ECE, Malla Reddy Engineering College and Management Sciences, Hyderabad, Telangana.

Abstract

Haze removal from satellite images is vital for various outdoor applications. However, existing techniques often lack the necessary knowledge to effectively restore hazy satellite images. They tend to rely on attributes with constant values, resulting in suboptimal dehazing outcomes. This review paper presents a structured overview of well-known haze removal techniques, shedding light on their limitations. To address these drawbacks, this research leverages advanced Deep Learning Convolutional Neural Networks (DLCNN) to enhance satellite image dehazing. The network is trained and tested with Laplacian and Gaussian pyramid-based features, which are used to modify Multiple Exposure Fusion (MEF) properties, enabling precise enhancement. The proposed DLCNN-MEF technique outperforms state-of-the-art approaches in terms of various parameters, including PSNR, SSIM, and MSE values for result comparison.

Keywords: Deep Learning Convolutional Neural Network, Multiple Exposure Fusion, Satellite Image Dehazing, Image Enhancement, PSNR, SSIM, MSE.

1. Introduction

With the fast advance of technologies and the prevalence of imaging devices, billions of digital images are being created every day. Due to undesirable light source, unfavourable weather or failure of the imaging device itself, the contrast and tone of the captured image may not always be satisfactory. In fact, image enhancement algorithms have already been widely applied in imaging devices for tone mapping. For example, in a typical digital camera, the CCD (Charge Coupled Device) or CMOS (Complementary Metal Oxide Semiconductor) array receives the photons passing through lens and then the charge levels are transformed to the original image. Usually, the original image is stored in raw format, with a bit length too big for normal displays.

An underwater image bears poor quality of images due to the nature of the light. When light enters the water, it gets refracted, absorbed and scattered in different directions. Scattering causes the blurring of light and reduces the color contrast. These effects on underwater images are due to the nature of the water. So, image enhancement is the mechanism to process the input image to make it clearly visible as this image enhancement improves the information content and alters the visual impact. Satellite images captured in poor weather conditions may lose their potential information, due to a dirty medium such as particles and water droplets in an atmospheric veil. Thus, these hazy satellite images do not provide enough significant details for future vision applications. Poor illumination decreases the visibility of satellite images. Thus, these satellite images are not suitable for future vision applications such as weather forecasting, radar tracking system, lane detection system, etc. Therefore, these kinds of applications demand haze removal techniques as a preprocessing tool to improve the performance of vision applications under poor environmental conditions. The number of particles available in the atmosphere fluctuates according to the weather conditions. Depending upon the category of the visual belongings, poor weather circumstances are categorized into two types: Steady and dynamic [1]. In steadily poor weather, ingredient droplets are minimum (1–10 μm) and steadily float in the atmosphere. Haze, mist, and fog are examples of steady weather conditions. The illumination effect at a given pixel is because of the collective consequence of the high degree of

droplets within the pixel's solid angle. In dynamic poor weather circumstances, ingredient droplets are 1000 times more (0.1–10 μm) than steady weather [2]. Snow and rain are examples of dynamic weather circumstances. Most vision applications provide poor results in case of weather degraded satellite images [3].

Rest of the paper is organized as follows: Section 2 details about literature survey, section 3 details about the proposed methodology, section 4 details about the results with discussion, and section 5 concludes article with references.

2. Literature Survey

Jiao, W., Jia, X., et al. (2021) [11] Numerical iterative strategy is employed to further optimize the atmospheric light and transmission. Extensive experiments demonstrate that our method outperforms existing state-of-the-art methods on synthetic datasets and real-world datasets. Xu, Zunxiao, et al. (2021) [12] provided constructing data pairs of cloud-free scene images and cloudy scene images, which are highly needed but considerably insufficient in the remote sensing literature. They referred to the overall paradigm consisting of the two fundamental operations as cloudy image arithmetic. We explore the use of the cloudy image arithmetic for the purpose of thin cloud removal. Song, Yuda, et al. (2022) [13] proposed dehaze former, which consists of various improvements, such as the modified normalization layer, activation function, and spatial information aggregation scheme. They train multiple variants of dehaze former on various datasets to demonstrate its effectiveness. Liu, Juping, et al. (2021) [14] classified the current available algorithm into three categories, i.e., image enhancement, physical dehazing, and data driven. The advantages and disadvantages of each type of algorithm are then summarized in detail. Finally, the evaluation indicators used to rank the recovery performance and the application scenario of the RS data haze removal technique are discussed, respectively. In addition, some common deficiencies of current available methods and future research focus are elaborated. Shi, Weipeng, et al. (2022) [15] proposed a reliable framework which is evaluated across various severity levels of fog corruptions. Utilizing HRNet as the backbone to maintain high-resolution representations, they develop a multimodal fusion module (MMF) to exploit the complementary information of lidar and multispectral data.

Cherian, Aswathy K., et al. (2021) [16] focused is on summarizing the main advanced approaches used to enhance classification precision. The result shows that the Convolutional Neural Network outperforms all the other traditional classification method. Saxena, Gaurav, et al. (2021) [17] proposed to identify the hazy and non-hazy images which makes all haze removal systems more efficient. Moreover, proposed model also classifies the hazy images based on their densities which can be utilized for selection of configuration parameters of haze removal techniques. Finally, the performance of the proposed model is demonstrated and the valuable identification, as well as classification parameters, are computed and presented. Bai, Yu, et al. (2022) [18] provided a brief review of 2674 related papers in deep learning with remote sensing (DL-RS) from 2014 to 2020. Keywords, publication years, journals, countries, and other essential characteristics of the papers were extracted. Also, they had set up some data items for information collection, such as remote sensing image categories, remote sensing applications, commonly used public datasets, and basic deep learning models. Jiao, Wenjiang, et al. (2021) [19] proposed a mixed iterative model is proposed, which combines a physical model-based method with a learning-based method to restore high-quality clear images, and it has good performance in maintaining natural attributes and completely removing haze. Unlike previous studies, they first divide the image into different regions according to the density of haze to accurately calculate the atmospheric light for restoring haze-free images. Then, dark channel prior and Dehaze Net are used to jointly estimate the transmission to promote the final clear haze-free image that is more like the real scene. Wang, Shan, et al. (2022) [20] proposed a dynamic

mutual enhancement network (DMENet) for haze removal in remote sensing images. It has three major advantages compared with other dehazing algorithms: 1) The proposed DMENet is based on the U-Net architecture to extract features effectively, which is composed of three components, i.e., a multi-scale encoder, a middle transmission layer (MTL), and a dynamic mutual decoder. 2) The dynamic mutual enhancement (DME) module is designed to dynamically integrate multi-level feature maps in a mutual way, which contains the low-level detail information and high-level semantic information respectively. 3) To improve the robustness and generalization performance of the DMENet, the hybrid supervision is built for network training between the restored results and their ground-truth labels, which consists of the pixel-level supervision, patch-level supervision, and image-level supervision. Experimental results on both synthetic datasets and real remote sensing hazy images demonstrate that the proposed DMENet can gain significant progresses over the competing methods.

3. Proposed Methodology

Figure 1 shows the Physical models of natural image dehazing and remote sensing image dehazing. In Figure (a) Natural dehazing and (b) Remove sensing image dehazing (RSID).

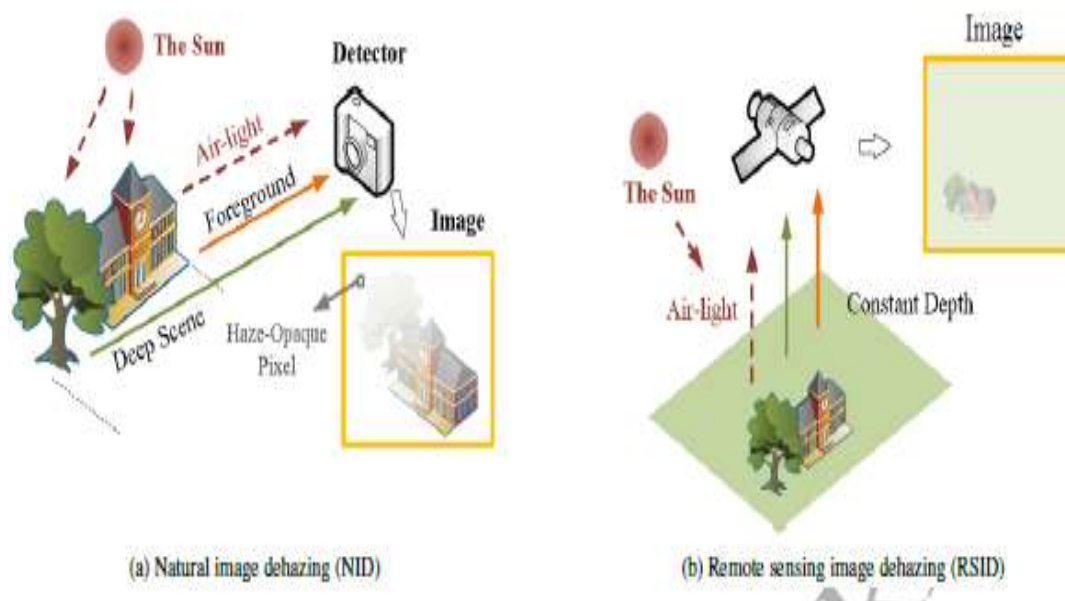


Figure 1. Physical models of natural image dehazing and remote sensing image dehazing.

3.1 DLCNN ARCHITECTURE

To overcome the outcomes of the related works in this proposed method a novel architecture is utilized by considering the network depth as 10. This network contains several layers. This proposed method improves more visibility of the enhanced haze images. The Figure 2 shows the block diagram of the network which is used for proposed implementation.

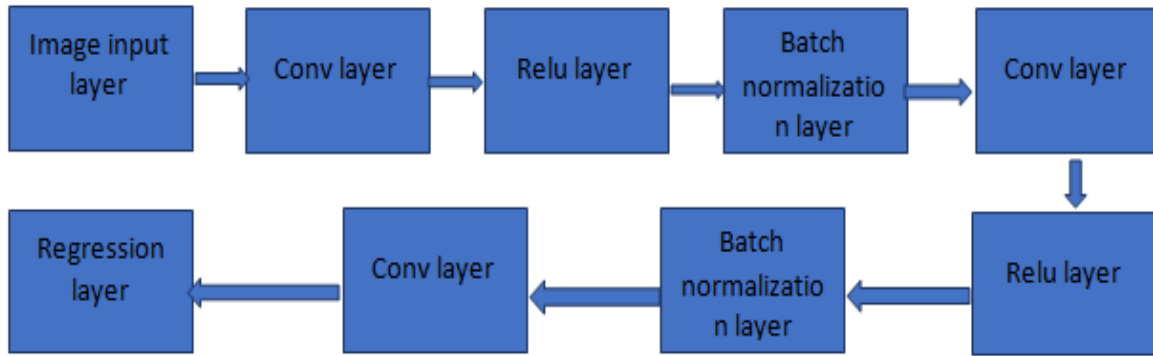


Figure 2. DL-CNN Network architecture.

The first layer of this network is image Input Layer having image data. Applies data normalization by considering this input data into the network. To extract the features from input layer convolutional layer is used. By taking the image and kernel as inputs this layer prevent the feature values from neighborhood pixels by adding a bias to the input of the convolutional layer along with the size of the kernel or filter. Filters are used to improve the operations including such as blur, sharpen and edge detection. Then we are using ReLU (rectified linear unit) layer as activation where the input value is set as zero which is less than zero by applying the threshold set-up before applying the pooling layer. After applying the convolutional layer, next layer will be the batch normalization layer in order to increase the network training process correspondingly by reducing the activations and gradients. The network initialization is applied to drop the sensitivity. This layer sustains the network as steadier, and it decreases the over fitting problems which is occurs due to the. By calculating the loss and accuracy which are called as activation metrics we can estimate the over fitting problem. This is main problem in the neural network. And also, we can observe the over fitting whenever the validation loss is more the training loss will be reduced. Then, it calculates the normalized activations as

$$x_i^{\wedge} = \frac{x_i - \mu_{\beta}}{\sqrt{\sigma_{\beta}^2 + \epsilon}} \dots\dots (1)$$

Where x_i is the input and μ_{β} is the mean and variance σ_{β}^2 . These two parameters are measured after the training process only. If we consider the pre-trained network then the mini batch values are applied for mean and variance. In order to consider the mean as zero and variance as one the activation is computed as

$$y_i = \gamma x_i^{\wedge} + \beta \dots\dots (2)$$

Where γ and β are called as the training parameters. These two are stored as trained mean and variance values. Finally, we are using the regression layer in this architecture. This layer is used to predict the output from the feature values of the input. To estimate the half, mean square error problems regression layer were utilized. Where, t_i is the output target, y_i is the network prediction and number of responses is denoted as R and the half, mean square error for single observation is given as

$$MSE = \sum_{i=1}^R \frac{(t_i - y_i)^2}{R} \beta \dots\dots (3)$$

To calculate the half, mean square error for image-to-image regression network where H, W, C are the height, width and number of channels and the loss function is given as

$$loss = \frac{1}{2} \sum_{\rho=1}^{HWC} (t_i - y_i)^2 \dots\dots (4)$$

For regression problems it will help to increase the network training process. This architecture is easy to implement and gives the more accuracy of training process.

3.2 PROPOSED METHOD

Haze Image contrast enhancement is a method which is used to increase the visibility of images. Here, the intensity of the pixel grey level is modified based on a function. Intensity based methods are of the form:

$$I_o(x, y) = f(I(x, y)) \dots\dots (5)$$

The original image is $I(x, y)$, the output image is $I_o(x, y)$ and f is the DLCNN transformation function. Intensity based methods transmute the grey levels over the entire image. Even after the transformation pixels with same grey levels throughout the image remain same. Contrast stretching is a generally used method that falls into this group. Figure 4 represents the proposed method of Haze removal with contains the White balance based brightness preserving DLCNN-MEF approach as its operational function $f(I(x, y))$.

The step wise operation of proposed method as follows:

Step 1: Haze images corrupted by lighting conditions may lose the visibility of the scene considerably. Enhancing approach is not available to remove entire haze effects in degraded images. The proposed algorithm includes deriving two inputs from the single degraded image and it recovers colour and visibility of the entire image. Colour correction is applied to the image after the fusion process. In the proposed algorithm, two inputs are derived from a single degraded input. The first derived input is obtained by white balancing the input. This step aims at removing chromatic casts in the image. More attention is given to red channel of the image as it attenuates more in haze. The second derived input I_2 can enhance those regions having low contrast. It is obtained by contrast stretching the median filtered input image. Therefore, the first derived input I_1 avoids colour casts, and the second derived input improves the uneven contrast. The important features of these derived inputs are computed using different weight maps. The different weight maps derived in the proposed model are exposedness, saliency, Laplacian, and colour cast.

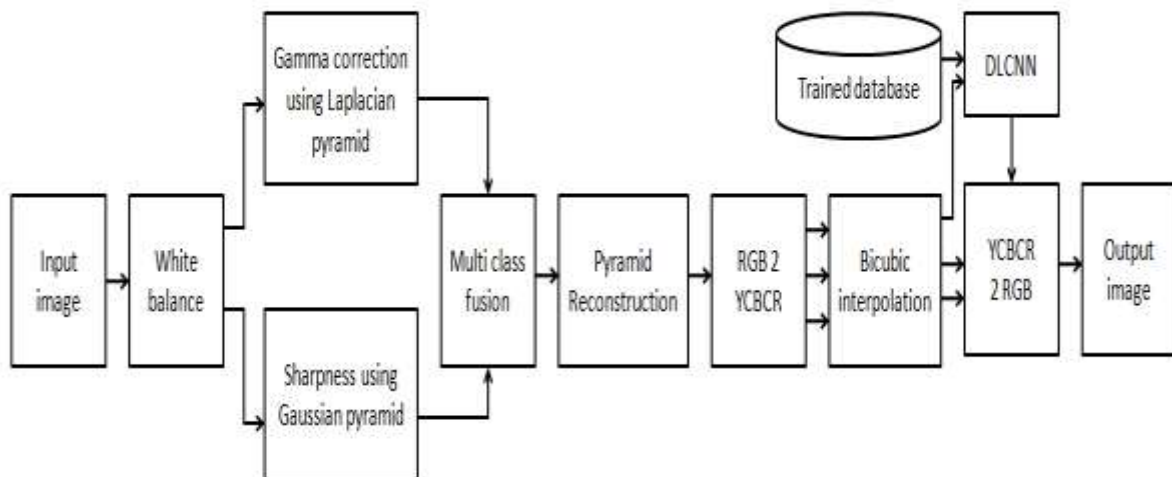


Figure 3: Proposed DLCNN with MEF block diagram

Step 2: The first output will be applied as input to the gamma correction using Laplacian pyramid operation. The weight maps play a critical role in the outcome of the final fused result. The weight maps generated should have a non-negative value. The weight maps are some measures of the input image. It represents the finer details of the image. The finer details from each image have to be extracted out and fuse them together to form the enhanced image. The weight maps are to be designed carefully to extract the details. New weight maps were used which showed better results in enhancing degraded haze images. The different weight maps are Laplacian, saliency, colour cast and exposedness. It is the measure of visibility of each pixel such that it gives higher values to high visibility areas and smaller values to the remaining areas. The main aim of this weight map is to assign a high value to edges and textures. Laplacian filter is a better choice for edge preserving. Laplacian of an image accentuates the regions having abrupt or rapid intensity change. The second derivative measurements are very delicate to noise. Since the images were noise reduced, further noise reduction is not needed. Even though it preserves edges, this weight map doesn't give the details of contrast since it can't discriminate the flat, valley or rapid regions.

Step 3: The Second input will be applied as input to the sharpening using Gaussian pyramid operation. In an haze image, all pixels will not be exposed. Pixels are commonly better exposed when they have normalised values close to the average value of 0.5. This weight avoids an over or underexposed look by constraining the result to match the average luminance. a similar weight in the context of tone mapping. The exposedness weight map is expressed as a Gaussian-modelled distance to the average normalised range value (0.5). One of the main challenges in haze removal is that the objects within the image lose their visibility and hence discrimination of the objects from the background scene becomes difficult. The quality that makes the object distinctive relative to its neighbours is known as saliency. It is based on the concept of centre-surround contrast wherein a saliency map is developed so that the contrast of the main object of interest is enhanced. Another main advantage of this approach is that the mid values of the image are not affected while applying this method.

Step 4: Multi-Scale Fusion technique is used to improve the visibility of poorly illuminated areas by using multi-scale fusion-based strategy [22]. The fundamental principle of this method is extracting the best features from the given input image and then to apply the estimated perceptual based qualities called weight maps to them and finally fuses them together to form the enhanced output. In the fusion-based strategy, two inputs derived from the single image will be by white balancing and by median filtering followed by contrast stretching. White balancing is used to reduce colour casting from the individual input image; estimation of four weight maps is done. One is exposedness weight map, which measures the exposedness of pixels not exposed to haze constraints. Laplacian weight map assigns high values to textures as well as edges. Colour cast weight map, introduced newly, increases red channel value thereby reducing colour cast. Saliency weight map measures the amount of discernible concerning other pixels. These weight maps are computed and applied to both the derived inputs. This method is a per pixel implementation. Fusion is a process of combining the relevant information from a set of input images into a single image, where the resultant fused image will be more informative and complete than the input images.

Step 5: Pyramid reconstruction is the process of reconstruction of the Gaussian and Laplacian pyramid outcomes and results the pyramid reconstruction output respectively.

Step 6: The Pyramid reconstruction output is applied to RGB2YCBCR operation. It will generate the output as Luminance(Y), chromium blue (CB) and chromium red (CR) outcomes respectively.

Step 7: Perform the bicubic interpolation operation individually on Y, CB and CR outcomes. These bicubic interpolations will results the perfect region of interest respectively.

Step 8: The detailed operation of DLCNN explained in section 3.1. The Luminance output of the bicubic interpolation is applied to implement Mean brightness; we used morphological based operation, which is computationally not exhaustive. We are essentially extracting pixel block centered also referred to as patch at $\Omega(x,y)$. We determine the minimum value for each block. Hence, we get three values corresponding to each colour for every block of pixels. From these three minimum intensity values we chose the most minimum value and replace it at the center location of the processed patch $\Omega(x,y)$. This step is repeated till the entire image is operated upon. Finding the minimum value for a pixel block in a grayscale image is same as carrying out a morphological operation. In this case, we can separately apply this operation on individual colour channels corresponding to H, S and I. This step is then followed with finding the minimum out of the three colour planes for any structuring element. Inspired from mean channel prior, we derive the modified red color priority depth map image, which is given as

$$I_I^{mean}(x, y) = \min_{\gamma \in R} (\min_{y \in \Omega(x,y)} (I_R^n(x, y))) \dots (6)$$

where $\Omega(x,y)$ is the structuring element, n corresponds to the number of partitions, I corresponds to the intensity color channel. This depth map gives the pictorial estimate of the presence of haze in an image and is useful for estimating the mean transmission map.

The output of the mean brightness operation applied to DLCNN specification. DLCNN specification used the mean for dividing the image input into two different parts. The sub images first part consists of pixels values that are up to mean, and the second part consists of pixels that are higher than mean of original image. After these two different layered ranges which don't overlap were attained. After these two sub-parts i.e., sub- layers were equalized by layers specification. From the results it was found that h DLCNN specification was proficient of preserving the actual brightness to level when input feature of the image had quasi-symmetrical distribution near to its mean. In this technique, the mean intensity of image's every pixel with range of feature from 0 to M-1 is presented by input mean. The 1st feature consists of pixels that are from zero to mean and the 2nd consists of from mean+1 to M-1. DLCNN specification has been implemented to these two sub images separately, and after that both equalized images are integrated. DLCNN specification can improve input image and can be utilized for user electronic by preserving the mean brightness. It will be represented in mathematical form that DLCNN specification much considerably preserves the images' mean brightness than usual DLCNN by improving the contrast. Consequently, it gives more enhancements which can be used in user electronic devices. Assume that mean of image X is X_m which is given as:

$$X_m \in \{X_0, X_1, \dots, X_{l-1}\} \dots (7)$$

On the basis of mean, the decomposition of image is done into 2 sub images i.e. X_L and X_U

In this , X_L consists of :

$$(X_0, X_1, \dots, X_m) \dots (8)$$

And X_U consist of:

$$(X_{m+1}, X_{m+2}, \dots, X_{L-1}) \dots (9)$$

Same as DLCNN, in this, the CDF is utilized in the form of transform function as :

$$\in \{X_0, X_1, \dots, X_{l-1}\} \dots (10)$$

And

$$f_U(x) = X_{m+1} + (X_{L-1} - X_{m+1})C_U(x) \quad \dots\dots\dots (11)$$

On the basis of above equation, the equalization of the divided sub images is performed separately and ensuing sub images' composition comprises the DLCNN specification output i.e., finally, the Histogram specification output image Y is given as:

$$Y = f_L(X_L) \cup f_U(X_U) \quad \dots\dots\dots (12)$$

Step 9: The adaptive histogram equalization is utilized only to remove the noise over the homogeneous areas. In order to overcome this issue MEF is proposed because it applies each and every pixel of the image. For generation of contrast transform function there are different types of distributions including such as Rayleigh, uniform, and exponential. Rayleigh distribution is the most useful for enhancing the haze images. This algorithm is only applied for smaller regions or areas throughout the image. To eliminate the boundaries by using the bilinear interpolation the tiles are combined. To prevent the over saturation area in case of homogeneous regions we are using the contrast factor. The MEF produces the enhanced results in some times without using the clip limit (contrast factor). In this proposed algorithm we are decreasing the network depth that is 10. And the number of layers which we are used is 10 layers. By decreasing the depth size, the performance of the network is improved and gives the better enhancement results. Finally, DLCNN output is concatenated with other CB and CR components and results the output as the RGB version of contrast enhanced colour image.

4. Results and Discussions

The proposed method is implemented using Matlab 2018a software simulation on various types of haze test images such as low light images, dusty environment images, hazy images and general purposed images. The proposed method performs outstanding enhancement and gives the better qualitative and quantitative evaluation compared to the state of art approaches. The enhancement results of the proposed implementation by using a novel architecture for haze images is shown and discussed.



(a) (b) (c) (d) (e)

Figure 4. (a) satellite input images (b) ID-CNN [12] (c) IDERS [20] (d) TME- MOF [15] and (e) proposed Dehazed image

Table 1. Performance evaluation

Parameters	ID-CNN [12]	IDERS [20]	TME- MOF [15]	Proposed method
PSNR (dB)	21.45	31.33	40.45	55.89
SSIM	0.67	0.71	0.85	0.96
MSE	0.48	0.098	0.082	0.034

5. Conclusion

In this work, the DLCNN algorithm is implemented for enhancing the haze images by decreasing the depth of the network to 10. A novel architecture is designed for this execution. This design enhances the haze images more by utilizing contrast limited adaptive histogram equalization algorithm. In processing of image, role of haze removal is very significant. This paper gives the implementation of on MEF based DLCNN techniques for enhancing the image quality by preserving the brightness and by enhancing contrast of image. The multi-layer deep learning methods give the extended enhancement because various layers are used analyse the each and every pixel region. If the low intensity regions are identified, then they are enhanced by the pre-trained DLCNN network model. The proposed model also responsible for the smoothness enhancement, brightness control, saturation, hue and intensity adjustments to the standard levels. For enhancing the light, edge difference loss levels are calculated. The MEF method responsible for controlling all the parameters, for this purposed mean preserving brightness levels are developed from the DLCNN method. The feature specific preserving values capable of correcting the all errors with detailed enhancement, by using the deep learning-based brightness preserved histogram specification-controlled histogram equalization techniques for enhancing the image quality gives the outstanding results compared to the existing WB-MUV approach. This work can be extended to implement the satellite image processing applications for haze removal.

References

- [1]. Sharma, Neeraj, Vijay Kumar, and Sunil Kumar Singla. "Single image defogging using deep learning techniques: past, present and future." Archives of Computational .Methods in Engineering 28 (2021): 4449-4469
- [2]. He, Zhijie, et al. "Remote Sensing Image Dehazing Based on an Attention Convolutional Neural Network." IEEE Access 10 (2022): 68731-68739.
- [3]. Jiang, Bo, et al. "Deep dehazing network for remote sensing image with non-uniform haze." Remote Sensing 13.21 (2021): 4443
- [4]. Liu, Shuchao, Huajun Shi, and Zhan Guo. "Research on defog algorithm of single remote sensing image based on deep learning." CIBDA 2022; 3rd International Conference on Computer Information and Big Data Applications. VDE, 2022.
- [5]. Xiao, Jinsheng, et al. "Single image dehazing based on learning of haze layers." Neurocomputing 389 (2020): 108-122..

- [6]. Chen, Zhaojie, et al. "Nonuniformly Dehaze Network for Visible Remote Sensing Images." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2022.
- [7]. Pazhani, A. Azhagu Jaisudhan, and S. Periyanyagi. "A novel haze removal computing architecture for remote sensing images using multi-scale Retinex technique." *Earth Science Informatics* 15.2 (2022): 1147-1154.
- [8]. Meng, Jun, et al. "Learning architecture for dehazing of aerial remote sensing images." *Multimedia Tools and Applications* (2022): 1-17.
- [9]. Kalra, Abhi, et al. "A new deep [2] He, Zhijie, et al. "Remote Sensing Image Dehazing Based on an Attention Convolutional Neural Network." *IEEE Access* 10 (2022): 68731-68739.
- [10]. Bi, Guoling, et al. "Haze removal for a single remote sensing image using low-rank and sparse prior." *IEEE Transactions on Geoscience and Remote Sensing* 60 (2021): 1-13.
- [11]. Jiao, W., Jia, X., Liu, Y., Jiang, Q., & Sun, Z. (2021). Single image mixed dehazing method based on numerical iterative model and DehazeNet. *Plos one*, 16(7), e0254664.
- [12]. Xu, Zunxiao, et al. "Cloudy image arithmetic: A cloudy scene synthesis paradigm with an application to deep-learning-based thin cloud removal." *IEEE Transactions on Geoscience and Remote Sensing* 60 (2021): 1-16.
- [13]. Song, Yuda, et al. "Vision transformers for single image dehazing." *arXiv preprint arXiv:2204.03883* (2022).
- [14]. Liu, Juping, et al. "A review of remote sensing image dehazing." *Sensors* 21.11 (2021): 3926.
- [15]. Shi, Weipeng, et al. "Land Cover Classification in Foggy Conditions: Robust Models." *IEEE Geoscience and Remote Sensing Letters* 19 (2022): 1-5. *Toward*
- [16]. Cherian, Aswathy K., and E. Poovammal. "Classification of remote sensing images using CNN." *IOP Conference Series: Materials Science and Engineering*. Vol. 1130. No. 1. IOP Publishing, 2021.
- [17]. Saxena, Gaurav, and Sarita Singh Bhadauria. "Haze Identification and Classification Model for Haze Removal Techniques." *Advances in Intelligent Computing and Communication: Proceedings of ICAC 2020*. Springer Singapore, 2021.
- [18]. Bai, Yu, et al. "Deep learning in different remote sensing image categories and applications: status and prospects." *International Journal of Remote Sensing* 43.5 (2022): 180.
- [19]. Jiao, Wenjiang, et al. "Single image mixed dehazing method based on numerical iterative model and DehazeNet." *Plos one* 16.7 (2021): e0254664.
- [20]. Wang, Shan, and Libao Zhang. "Dynamic Mutual Enhancement Network for Single Remote Sensing Image Dehazing." *2022 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2022.