Image De-Hazing Using Deep Learning Approach

Najwa Fadhil Abbas Alhadeethy, Akram M Zeki, Asadullah Shah

najwaalhadeethy80@gmail.com, akramzeki@iium.edu.my, asadullah@iium.edu.my

Malaysia

Abstract – Image de-hazing has always been a challenging task in the domain of Computer vision since the beginning. Sometimes the images that are taken in certain atmospheric conditions turn out to be blurred and hazy, and of low quality because of the adverse environmental conditions, resulting in foggy and hazy capture. This, in turn, creates a lot of trouble to identify or recognize the objects or key items in the captured images. This causes many issues, particularly for the researchers working in the domain of Computer vision, who mainly depend upon the prominence and perceptibility of the captured images as the study data. In this research article, a deep learning algorithm comprising of generative adversarial network architecture has been implemented to de-haze the image. The actual de-hazing methods and processes involve the use of per-pixel loss, which during the analysis period creates a big problem even if there is a difference of only one pixel in the whole of the image or even if the captured images are seemed to be perceptually alike. The function of the perceptual gains the high-level characteristics of the captured images by making use of the pre-trained algorithms on ImageNet, which minimizes or diminishes the challenges that occur in the per-pixel loss function.

Keywords – Image de-hazing, Deep learning algorithm, Convolution Neural network, CNN, Perceptual loss function.

1 Introduction

Image de-hazing is considered to be one of the major problems of the Computer Vision domain. Even though it is not an intricate task to tackle this, yet it is difficult to acquire for all that development and progression in technology assisting this field. The challenging issues of image de-hazing include a wide range of applications in the Computer Vision domain along and also in daily life. This issue is not restricted to only this domain but can be found pertinent or relevant to several other aspects of daily life. The actual problem is the image De-noising which is one of the fragments of a whole set of issues in the area of Image processing [1].

When the light is reflected from the object or any item, it gets scattered by the surroundings prior to reaching the camera. This occurrence of scattering the rays of light within the surroundings is because of the existence of “aerosol particles” in the atmosphere. This occurrence, eventually, upsets the actual image taken by the camera lens [2]. Due to the dust, fog particles, and fumes present in the surrounding or in the atmosphere, the quality of the image gets affected. The images or pictures which are captured in such circumstances lack not only the small little details but also vividness.

All of the images in which clear and radiant details are not present can pose a greater threat if being counted on as a decent source in apps like surveillance or transportation. Therefore, there comes the need to de-haze the image, considering it as the most important task to do.

Aforementioned that image de-hazing is a simple yet problematic process to implement. In all of the processes, figuring out the “transmission map” from the obscure picture is the most challenging one. Prior to the currently existing technologies, this image de-hazing process was used to be done via conventional/traditional computer vision approaches like dark channel [3] or else manual heuristics [4].

All of these approaches were effective enough to remove the haze from the images, but because of the non-linear enigmatic nature of the image haze, it became difficult to bring some major improvements in those approaches using conventional algorithms alone. Thus, the research was turned to machine

DOI: http://doi.org/10.24086/cocos2022/paper.577

learning (ML) techniques in which deep learning algorithms were then used. Recent ML methods are taken into two major categories.

The majority of the researchers mainly count on having the training data that is matched, comprising of both hazy as well as clear views of the location or scene as it is. This is considered to be the major restriction as it becomes difficult to obtain such kind of dataset, thus, making it a challenging task to design and build a model that is generalized for any condition to remove the haze from the image. On the other hand, there is the vast majority of researchers that creates model independent of the training of matched dataset [5]. The current execution of such methods is considered to be highly limited because of slow de-hazing speed as well as intricacy in recuperating extremely fine details.

Taking into consideration all of these issues, a new approach has been proposed in this paper that perfectly fits the category where there is no need to match the training dataset and is anticipated to build the autonomous measure of blurriness in an image to initiate the ML and deep learning (DL) procedure. Thus, for ML, one challenge is to build an appropriate training dataset.

Considering this problem, the method proposed in this paper is to make use of the "unmatched" training dataset that would help in improving the whole process of image de-hazing, which will be circumvented on the whole rather than making an attempt to solve directly. The proposed approach is based on two networks assemble, one is the discriminator network, and the second is the de-hazing network. This assures the process of fine-tuning, which eventually is considered to be the significant contribution since it countenances with the proposed approach to acquire the outcome that would be comparable or at least exceed the state-of-the-art performance with a knowingly smaller data set.

2 Literature Review

There are a wide variety of approaches that can be deployed to remove the haze from the images. The haze removal (or image de-hazing) process assists the researchers in several other applications that include surveillance or transportation and thus, is considered to be one of the most significant tasks. Sometimes, when the image is taken in severe weather conditions, it becomes degraded or blurry because of the atmospheric scattering/ scattering. With the help of methods involved in the image de-hazing process, this blurriness can be removed. Based on the no. of inputs, these image de-hazing approaches are as follows:

- Single Image De-hazing
- Multiple Image De-hazing
- Polarizing filter Image De-hazing

In the multiple-image de-hazing approach, several images are taken of the same location or surrounding under atmospheric circumstances. This approach analyzes several parameters on the basis of the images being taken at various conditions of the atmosphere in order to perform the process of de-hazing the image.

In the polarization filtering approach, like the multiple-image approach, a number of images are taken as input data. However, these images are not required to be captured at different conditions of weather or atmosphere, but diverse polarization filters are smeared over multiple images or single images to put on or simulate changing weather/ atmospheric conditions [3].

In the single image de-hazing approach, no additional information is required to be acquired from different images taken at distinct weather conditions. The process of single de-hazing of the image further involves two sub-categories, i.e., learning-based approach and prior-based approach [6].

The prior approach makes use of the "statistical information/ data" to estimate the transmission gap of the images. Usually,

DOI: http://doi.org/10.24086/cocos2022/paper.577

the output of this process gets affected because of the wrong assumptions being made by the researchers. Other approaches like dark-channel prior as well as color-attenuation prior are considered to be working well for image de-hazing. However, both of these prior-based approaches fail if the assumptions being made are not correct [3].

The learning Approach makes use of machine learning or deep learning algorithms to acquire the de-hazing in images. In addition to that, the convolutional neural network is also used under learning methods to learn the depth of the atmospheric conditions as well as to learn about the characteristics of haze to gain the image de-hazing in an effective manner and up to the mark. However, intrinsically, all these algorithms of learning have to estimate and analyze the transmission gap, and then the "atmospheric scattering model" is used to get to the final objective or goal (de-hazed image) [7].

Mathematically, the image de-hazing model can be represented as:

$$I\,(a) = J(a)\mathrm{t}(a) + A(1 - t(a)) \quad (1)$$

Where;

I = Hazy RGB image

J = Original De-Hazed Image

A = Atmospheric Light

t = transmission coefficient

In the proposed approach, the "generative network adversarial (GAN)" approach has been implemented that makes use of the encoder-decoder network in order to acquire a clear image [6]. The concept of the GAN approach was first developed by Ian Goodfellow, which proved to be effective in many various de-hazing processes. Furthermore, the GAN architecture involves two subcategories of the neural network model, which are as follows:

- Generator
- Discriminator

In this architecture, the generator gets a random value as the input, which is then passed thru the network, and as output is produced in the required dimension. Usually, the generator in GAN architecture creates the new data samples; on the other hand, discriminators recognize the originality or fakeness of the data. Figure 1 below describes the process of operation of both generator and discriminator architecture.

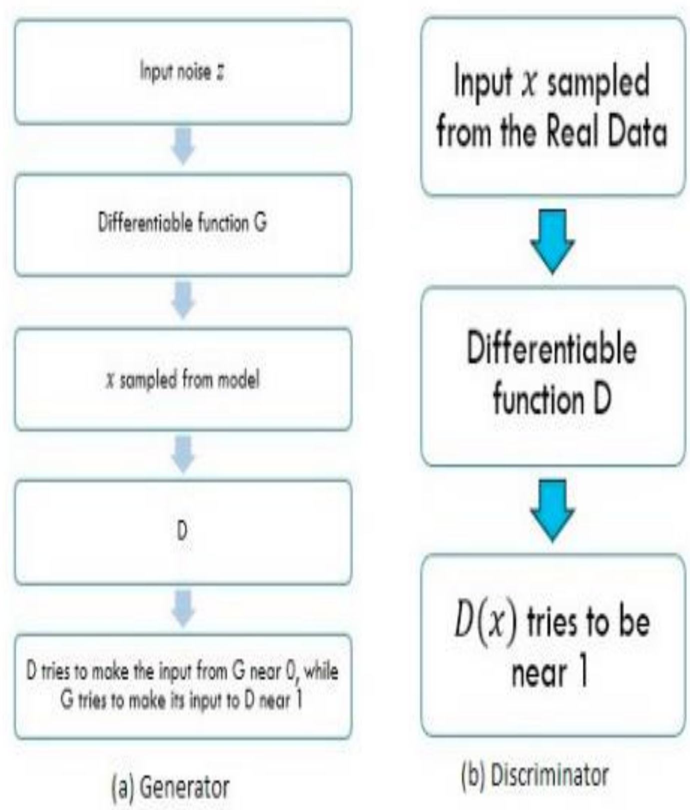

Figure 1. Generator and Discriminator Model *[6]*

2.1 Related Work

This section describes the core techniques and algorithms that have already been used in the past for the image de-hazing process.

In [8], the Siamese Network algorithm has been introduced to tackle the customer signature verification issue. In a research article, the signature of the customer or client is installed in a

DOI: http://doi.org/10.24086/cocos2022/paper.577

network accompanied by the reference signature to find out that either the signatures are forgery or genuine. For this purpose, an approach of a Neural network algorithm comprising of two identical networks has been proposed that extracts the learning features out of the input data and, as an output, sends the metric of resemblance or connection among the two signatures. A cosine vector having the potential to extract the feature vectors has been exploited as a "similarity metric." The actual value of using this approach is to determine the resemblance or make a comparison between the two dissimilar samples to be done on the basis of learned features, irrespective of classifying them into different classes. Since then, the algorithm of the Siamese network has been exploited to determine the verification issues comprising of pedestrian re-identification [9], signature verification [8], and/or reduction of data dimension [10].

In the deep learning (DL) domain, the "Convolutional Neural Network (CNN)" algorithm has been considered to be of great interest for the past several decades. This technique has come out to be exploited on broad terms along with the previously existing de-hazing techniques that used to be deployed in the 1980s. The algorithm of CNN has been driven by the influx of cost-effective GPUs, approaches that provide support for training the deep network algorithms, as well as large datasets in such a way that the actual value of the ML algorithm has been realized. A large number of the currently existing research can be found in [11], where the significant turning point of using this latest technique can be observed in the deep learning domain.

In [3], the dark channel Prior approach has been used in order to resolute the transmission map (t) of the provided image. Even though this approach is said to be effective in various manners but it has certain limitations. High contrast surfaces like white objects in any location (or bright sky) are taken as haze or blurry via the DCP approach. In [2], an attempt has been made to enhance the performance of such approaches with the augmented recovery of J to gain more precise outcomes for high "textured image patches." In [12], an alternative analytical method has been demonstrated where the color relationship among the cluster pixels has been used in order to resolve the actual un-hazed image.

In [13], a deep learning algorithm has been used, demonstrating to be an effective platform for the image de-hazing technique development. The hierarchal nature of DL algorithms has extracted the relevant characteristics of haze, providing a number of significant enhancements in the model. DL techniques and algorithms have proved to be a versatile solution for a wide range of image types, lightening, and scene distances, along with bringing enhancement in processing speed. In addition to that, in [14], DL has been used that has also used the prior approach. For this reason, the DL model has come out to be an accepted standard for all of the recently developed de-hazing algorithms.

3 Proposed Methodology and Implementation

3.1 Data Augmentation

While training the ML models, parameters were tuned essentially such that input could be mapped to an output. The main aim of the model was to minimize the loss of data while processing. Recent neural networks have these parameters in millions. In order to acquire a large amount of data, there comes no need to go and collect new data; on the contrary, data augmentation can be used by creating smaller changes in currently available images like flips, translation, rotation, etc. On a less number of available data, data augmentation can be applied, and if there is a large number of data, augmentation can be applied on randomly opted images. Furthermore, with the deployment of augmentation, neural networking can be averted to learn the unnecessary patterns, eventually enhances the overall algorithmic performance.

DOI: http://doi.org/10.24086/cocos2022/paper.577

3.2 Network Architecture

Network architecture includes the base layer of the network, ReLU activation, pooling, dropout, and normalization. The details of each are described below.

3.2.1 Base Layer of Network

The convolutional layers are considered to be the basic building blocks of CNN. When images are dealt with, creating a complete connection with neurons becomes impossible. DL algorithm puts more focus on a square patch of the picture and then passes by the filter having the same size. The whole of the images are enclosed by moving the filter, and eventually, activation maps are generated. In further steps, the output of the neurons is computed by the conventional layer (CL).

3.2.2 ReLU Activation

In the Image de-hazing process, ReLU activation is defined as the non-linear function for activation as shown in equation 1 and represented in Figure 2 below.

$$R(a) = \max(0, a) \quad (2)$$

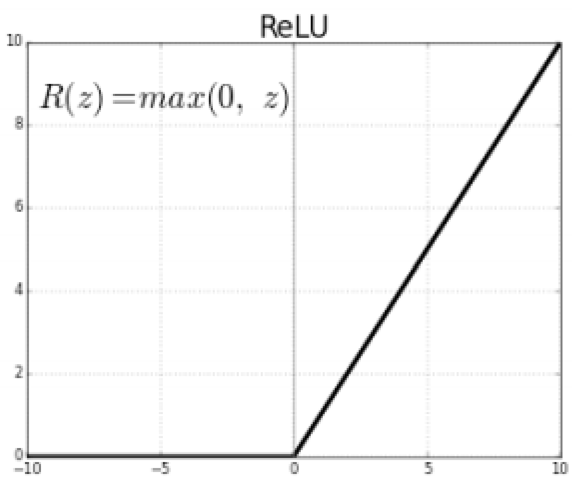

Figure 2. ReLU Activation

3.2.3 Pooling Layer

The pooling layer is implanted among the successive CL in the DL algorithm in a periodic manner. It minimizes the no. of parameters as well as computations appearing in a network and helps in lessening the over-fitting. The usually used pooling layer dimensions have been shown in Figure 3 below:

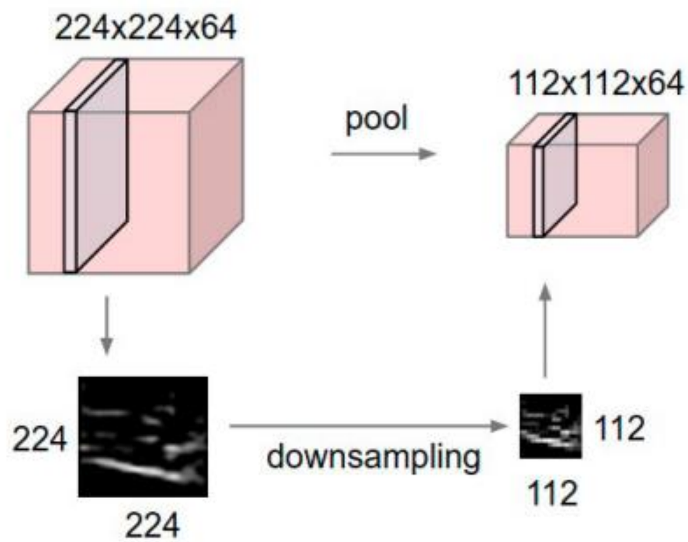

Figure 3. Pooling Layer

3.2.4 Dropout

In order to avoid the model over-fitting, methods such as early stopping of training were introduced. Moreover, computing the average result of several networks or training of deep learning algorithm data was not suitable enough. Therefore, a method named, Dropout [7], was introduced, that refers to drop several neural units out of the network. This process or dropout is completed haphazardly or randomly, where each of the units is reserved with 0.5 probability, in most cases.

3.3 Loss Function

DL algorithms efficiently work by learning the weight of every neuron by realizing more on elevating the "objective function." Here, the loss function signifies the performance of the deployed model in comparison to the expected outcome. It is better to reduce the loss function as much as possible. The single loss function performs for all types of data. It all mainly relies on optimizing function and data gathered, as well as the DL algorithm. In the proposed work, the objective is to learn the relation among the two different sets of samples X and Y that need to be trained. The two functions are A: $X \rightarrow Y$ and

DOI: http://doi.org/10.24086/cocos2022/paper.577

B: Y → X. These have their discriminators as D_Y and D_X. The data is distributed in the following way, as shown in equation 3:

$$a \sim pdata\ (a)\ and\ b \sim pdata(b) \qquad (3)$$

Loss functions are of two types, i.e., Adversarial loss and cycle consistency loss.

The main objective regarding the adversarial loss function for G is shown in Equation 4 below:

$$LGAN(G, DY, X, Y)\ =\ \mathrm{E}b\ \sim pdata(b)[logDY(b)] + \mathrm{E}a \sim pdata(a)[log(1\ -\ DY(G(a)))] \qquad (4)$$

Here, G attempts to reduce the loss; on the contrary, D attempts to increase it. Thus, the final objective comes out to be minGmaxDYLGAN.

In addition to the adversarial loss, there is another one known as "cycle consistency loss" which can be represented as shown in equation 5 below:

$$Lcyc(G, F) = Ea{\sim}pdata(a)[||F(G(a) - a||1] \quad + \quad Eb{\sim}pdata(b)[||G(F(b) - b||1] \qquad (5)$$

3.4 Implementation

In order to implement this proposed work, Python3 along with DL libraries were used using Google Cloud Platform. A virtual machine along with 16 GB RAM and vCPU and Nvidia-K80 GPU was deployed. The development of the program was done by making use of Pytorch.

For the proposed work, the generator was designed where the CNN algorithm took the input image dimensions as $256x256x3$. These were then passed via stride 2 CL. After that, these images were down-sampled through the whole set of CL and were surmounted through a long residual block series. The output was passed through a group of altered convolutions so that they could be resampled, at the end are passed through the convolutional layer (CL). It is the same as the one that was used in neural transfer.

During the implementation process, a discriminator was used that acted as a classifier to differentiate among the input data (images) on receiving from generators to find out the original one. If the classification process is done correctly, the generator tries to learn from this and makes a fool of the discriminator. In the proposed work, "PatchGAN" classifier network architecture was used for creating the discriminator, which is also known as "Markovian Discriminator" [13]. It comprises a group of Conv-BatchNorm-LeakyRelu layers. The proposed network took a 256x256x3 image as the input and delivered one value that represented the class of the image, rather than capturing two input images as of "PatchGAN" discriminatory. The proposed network was trained for generator images and real images.

The training of the input data was prepared by limiting the batch size to that of the Adam Solver. 0.0002 as the learning rate was kept for the first 50 epochs, and linear decay was set to 0 for the next epochs. The proposed model was trained for 70 epochs having more than 1000 input images from the dataset of NYU depth. Here, the cyclic loss was increased to about 10x by keeping $\lambda = 10$. Then the two values of Adam's solver, i.e., $\beta1$ and $\beta2$, were set to 0.5 and 0.9, respectively.

4 Results and Findings

The collected dataset was comprised of several pictures of the indoor environment that were recorded in D and RGB format. In the data set, there was a total of 1449 pairs that were labeled comprising RGB and D type images that were then synchronized as well as labeled for each of the input images. The images were considered to be part of the D-HAZY image dataset. The depth map was not used in the proposed approach. Some of the sample images from the dataset are shown in Figure 4 below. The first columns images are blurred and hazy, while the last three columns are clear and haze-free images corresponding to the first ones. Qualitative results were obtained using the proposed approach even with the images captured in an adverse hazy environment.

DOI: http://doi.org/10.24086/cocos2022/paper.577

Figure 4. From Hazy, blurred images to clear, haze-free images

While processing the images, the quality of the images was degrading continuously. Therefore, some metrics were exploited to measure the quality of the image. So, for the proposed approach, PSNR and SSIM were used to analyze the image quality being generated at the output.

5 Conclusion

Under certain circumstances, the quality of the captured image gets low, and it seems to be blurred and hazy because of the atmospheric conditions and the addition of loss function affecting the image equality. In this paper, a deep learning (DL) approach has been proposed that generates clear and haze-free data pictures and images out of the blurry and hazy images without any type of human intervention. For this purpose, at first, the loss function was reduced from the input images for which cyclic consistency loss architecture was used that significantly helped to improve the output of the images and made sure that the output images do not come out to be in any other form (in any degrading quality). From the qualitative and quantitative results, it can be seen that the proposed approach of implementing the DL algorithm for image de-hazing has proved to be effective in comparison to several other methodologies and current models. It has also been observed that perceptual consistency values of loss that has been acquired from VGG 19 were much better than VGG 16.

As future work, it can be recommended that the proposed approach can be modified by adding more datasets such as 0-HAZE, I-HAZE, and RESIDE to extend the capability of the proposed algorithm.

References

C. Hodges, M. Bennamoun, and H. Rahmani, "Single image dehazing using deep neural networks," *Pattern Recognit. Lett.*, vol. 128, pp. 70–77, Dec. 2019, DOI: 10.1016/j.patrec.2019.08.013.

C. Chen, M. N. Do, and J. Wang, "Robust Image and Video Dehazing with Visual Artifact Suppression via Gradient Residual Minimization," in *Computer Vision – ECCV 2016*, Oct. 2016, pp. 576–591. DOI: 10.1007/978-3-319-46475-6_36.

[3] M. K. Das, P. Dinulal, G. Koshy, and P. Simon, "Image Dehazing using Improved Dark Channel Prior and Relativity of Gaussian," *Procedia Comput. Sci.*, vol. 165, pp. 442–448, 2019, DOI: 10.1016/j.procs.2020.01.004.

[4] V. Kehar, V. Chopra, B. K. Singh, and S. Tiwari, "Efficient Single Image Dehazing Model Using Metaheuristics-Based Brightness Channel Prior," *Math. Probl. Eng.*, vol. 2021, May 2021, DOI: 10.1155/2021/5584464.

[5] Z. Gu, M. Ju, and D. Zhang, "A Single Image Dehazing Method Using Average Saturation Prior," *Math. Probl. Eng.*, vol. 2017, Mar. 2017, DOI: 10.1155/2017/6851301.

[6] R. R. Choudhary, K. K. Jisnu, and G. Meena, "Image DeHazing Using Deep Learning Techniques," *Procedia Comput. Sci.*, vol. 167, pp. 1110–1119, 2020, DOI: 10.1016/j.procs.2020.03.413.

[7] Y. Liu, J. Pan, J. Ren, and Z. Su, "Learning Deep Priors for Image Dehazing," p. 9.

[8] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, and R. Shah, "Signature Verification using a 'Siamese' Time Delay Neural Network," p. 8.

[9] X. Deng, B. Ma, H. Chang, S. Shan, and X. Chen, "Deep Second-Order Siamese Network for Pedestrian Re-

identification," in *Computer Vision – ACCV 2016*, Nov. 2016, pp. 321–337. DOI: 10.1007/978-3-319-54184-6_20.

[10] R. Hadsell, S. Chopra, and Y. LeCun, "Dimensionality Reduction by Learning an Invariant Mapping," in *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2 (CVPR'06)*, New York, NY, USA, 2006, vol. 2, pp. 1735–1742. DOI: 10.1109/CVPR.2006.100.

[11] "ImageNet classification with deep convolutional neural networks | Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1." https://dl.acm.org/doi/abs/10.5555/2999134.2999257 (accessed Sep. 21, 2021).

[12] D. Berman, T. Treibitz, and S. Avidan, "Non-local Image Dehazing," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 1674–1682. DOI: 10.1109/CVPR.2016.185.

[13] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "DehazeNet: An End-to-End System for Single Image Haze Removal," *IEEE Trans. Image Process.*, vol. 25, no. 11, pp. 5187–5198, Nov. 2016, DOI: 10.1109/TIP.2016.2598681.

[14] D. Yang and J. Sun, "Proximal Dehaze-Net: A Prior Learning-Based Deep Network for Single Image Dehazing," in *Computer Vision – ECCV 2018*, vol. 11211, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham: Springer International Publishing, 2018, pp. 729–746. DOI: 10.1007/978-3-030-01234-2_43.

DOI: http://doi.org/10.24086/cocos2022/paper.577