Haze Removal Algorithm to Enhance Low Light Images

Smt. K. Maheswari, Associate professor Dept. of ECE, Sanskrithi School of Engineering

K. Charan
Dept of ECE
Sanskrithi School of engineering,
Puttaparthi 515134

K. Annamayya
Dept of ECE
Sanskrithi School of engineering
Puttaparthi 515134

K. Naveen
Dept of ECE
Sanskrithi School of engineering
Puttaparthi 515134

K. Manoj Kumar
Dept of ECE
Sanskrithi School of engineering
Puttaparthi 515134

Abstract:

Image captured in hazy, visually degraded visibility caused by foggy atmospheric conditions; it obscures image quality. Instead of producing clear images, by the pixel-based metrics does not guarantee, this updated image is used as input in computer vision of low-level tasks like segmentation. To upgrade this, for image dehazing we are introducing a novel called end-to-end approach, so maintain the generated images visual quality. So here we take a step further to explore possibility of exploiting a network to perform semantic segmentation technique with U-Net. To further improve the perceptual quality of output U-Net going to designed and used in this model.

Keywords: Dehaze, Image Segmentation, U-Net

Introduction:

Haze is formed due to is very fine, widely dispersed particles in atmosphere, in relatively dry air, it gives air an opalescent appearance. Haze is traditionally an atmospheric phenomenon in which smoke, dry particles and other dust particles obscure sky clarity. The World Meteorological Organization manual codes have a classification of horizontal obscuration into categories of fog, ice fog, steam fog, mist, haze, smoke, volcanic ash, dust, sand, and snow. Haze particles having sources like traffic, wildfire, dry weather, industry. For example, an approaching airplane and haze can be appeared brownish or bluish depending on the direction of view with respect to the Sun, while mist become a bluish grey. Haze is formed by dry air and mist is formed by humid air. However, mist droplets are formed by the subsequent of the condensation nuclei(haze particles), such produces of haze are known as "wet haze".

In general, haze is just like a fog in nature when we capture images in this situation. Image losses its actual visibility, changing colors and obscure scenes. Ultimately it degrades the image quality. It is because of scattering and absorption of small particle in atmosphere. This we can refer as “Atmospheric Scattering Model”. This haze is an undesired and problematic to photography and also some applications like Object detection, scene understanding, and surveillance etc. To get better understanding and identify objects in image has to be done. For this haze must be removed from the image, it is not a simple task in computer vision.

$$\boldsymbol{I}(\boldsymbol{x}) = \boldsymbol{J}(\boldsymbol{x})\boldsymbol{t}(\boldsymbol{x}) + \boldsymbol{A}(\mathbf{1} - \boldsymbol{t}(\boldsymbol{x}))$$

Where,
I(x) = Hazy input image
J(x) = Original image

A = Global Atmospheric Lighting
$t(x)$ = Transmission map

A defines the natural atmospheric light across the entire image. $t(x)$ is the amount of light reaches to the object from camera.

It determines as follows:

$$\boldsymbol{t(x) = e^{-\beta d(x)}}$$

where,
β = Scattering Co-efficient (non-negative)
$d(x)$ = Distance between the object and observer

The idea is that the atmosphere scatters light coming from the object before it reaches the camera. The amount of light scattered depends on the atmospheric properties (captured by β) and the distance of the object from the camera captured by $d(x)$. Rest of the light is reaches to the camera. Since β and $d(x)$ are both non-negative, the value of $t(x)$ is in the range [0,1]. 0 refers to no light from the object reaches the camera all light is scattered, and 1 refers to all the light from the object reaches the camera (no scattering). A higher value of β represents an atmosphere that tends to scatter light more. Moreover, If the object moves away from the camera, more light is scattered.

Now, given values of transmission for each pixel and the global atmospheric light, the light received by the camera can be calculated. If $t(x)$ is 1, then the camera sees the light from the object perfectly (no scattering). If $t(x)$ is 0, then the camera sees only the atmospheric light (complete scattering). Else, the camera sees a LI(linear) interpolation between the light from the object and the atmosphere.

Literature Survey:

[Yu Li] In research, haze model is introduced, this haze removed via atmospheric scattering model, transmission map from the haze/snowy input images. the Input is blur/haze image,by using optimization it can be separated as haze free and haze images. Here, thes main aim of this to find the transmission map and atmospheric light. The drawback in this model was it does not give better output.

[Pei and Lee] They introduced a method called Color Transfer Method. This method gives better results for only night-time haze images. An input hazy image, air light color transfers into grey. After, the resulted output image is refined using dark channel prior method. It estimates the transmission and eliminates the haze and uses the bilateral filter to enhance low light. This model limits that haze-free image looks un-natural because of color transfer.

[BolunCai] They proposed AOD-Net. It is an end-to-end framework. AOD-Net is simplest as compared to other proposed methods. In this model, it minimizes the MSE between hazy and dehazed image. This method limits that light passing through emulsion on image is not reflects, it absorbs by a layer.

Proposed Work:

To recover the haze free image from hazy image, it is necessary to estimation of transmission map of an input image. For this, an atmospheric scattering model is taken consideration in the model. In this, a network can be trained by mapping between hazy images and the scene radiance.

Semantic Segmentation: Semantic Segmentation is a classification method in deep learning and computer vision/graphics. It is a most effective technique to extract features from an image by segmenting an image that refers to objects. This semantic segmentation labels the objects in image. It helps to get the transmission map at every pixel and also get standard information present in the image.

Figure 1Input hazy image

U-Net is an end to end trained network. Trained network consists of approximately 10000 hazy and non-hazy reference images. In this work, we are using this trained U-Net to objectifying the image and overlaying objects in image.

Network Architecture:

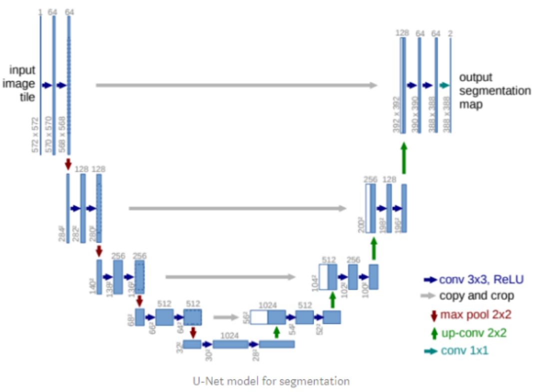

Figure 2: U-Net Architecture

U-Net consists of encoder and decoder that used for image segmentation. This network consists of 3 parts, they are down-sampling, bottleneck and up-sampling. In U-Net, there are cascaded Convolutional and pooling layers along with an activation layer. The down-sampling section having multi-scale blocks where each blocks having two 3X3 convolutional layers with normalization and activation function. The extracted feature map followed by 2X2 max pooling. For every block the image resolution becomes half of its actual resolution simultaneously feature maps are doubled. The encoded features move to the bottleneck to capture both local and global contextual features.

Figure 3: Segmented Image

The output of this bottleneck is sent to decoder to perform up-sampling. The up-sampling is similar opposite process to the down-sampling. While up-sampling, extracted features are concatenated to that particular block therefore the image re-gain its original resolution. Except to the last block, for every up-scaling activation function is applied.

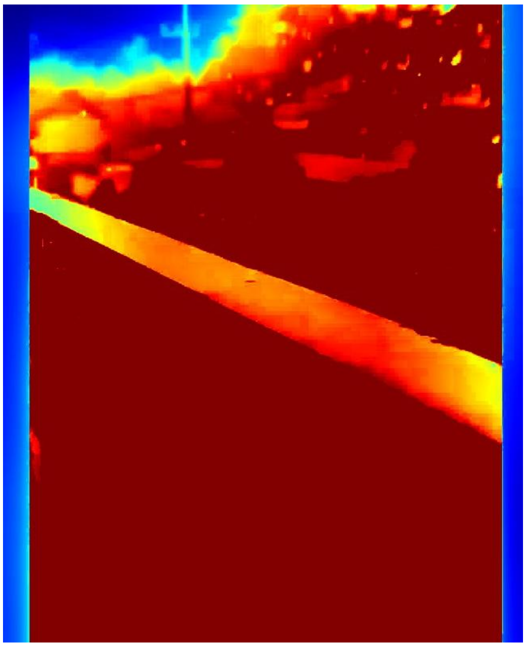

Figure 4: Transmission map

Batch Normalization:

Normalization is a pre-processing data tool, without disturbing its shape, the numeric data converts to common scale. It is a process to make neural networks faster and more stable through adding extra layers in a deep neural network. Output of previous layer is taken as input to next layer and performing the normalizing operations.

Normalization is a process that transforming the data to have a mean zero and standard deviation one. In this step we have our batch input from layer h, first, we need to calculate the mean of this hidden activation.

$$\boldsymbol{\mu} = \frac{\mathbf{1}}{\boldsymbol{m}}\sum \boldsymbol{h_i}$$

Where,

m represents the number of neurons at layer h.

Once we have meant at our end, the next step is to measure the standard deviation of the hidden activations.

$$\boldsymbol{\sigma} = \left[\frac{\mathbf{1}}{\boldsymbol{m}}\sum (\boldsymbol{h_i} - \boldsymbol{\mu^2})\right]$$

Further, as we already measured the mean and standard deviation. We will normalize the hidden activations using these values. For this, we will take differences of mean from every input and divide the whole valuewith thesum of standard deviation and the smoothing term (ε).

The smoothing term (ε) this stops divisible by zero and ensures the numerical stability within the operation.

$$\boldsymbol{h_{i(norm)}} = \frac{(\boldsymbol{h_i} - \boldsymbol{\mu})}{\boldsymbol{\sigma} + \boldsymbol{\varepsilon}}$$

Rescaling Offset:

The re-scaling and offsetting of the input take place. Here two components of the Batch Normalization come into the picture, γ and β. These parameters are used for re-scaling (γ) and shifting (β) of the vector having values from the previous operations.

$$\boldsymbol{h_i} = \boldsymbol{\gamma h_{i(norm)}} + \boldsymbol{\beta}$$

During the training neural network ensures the optimal values of γ and β are used. That will activate the perfect normalization of each batch.

ReLU Activation Function:

Using weights and biases as inputs, performing linear transformation to all the neurons. Even this linear transformation is simple, but this network is less powerful to analyzing complex date. Ultimately, the neural network becomes linear regression model.

In this work, we are using an activation function I.e., the Rectified Linear Unit (ReLU). This ReLU function does not activate all the neurons at a time; this is the reason most popularly in neural networks.The neurons will only be deactivated if the output of the linear transformation is less than 0.

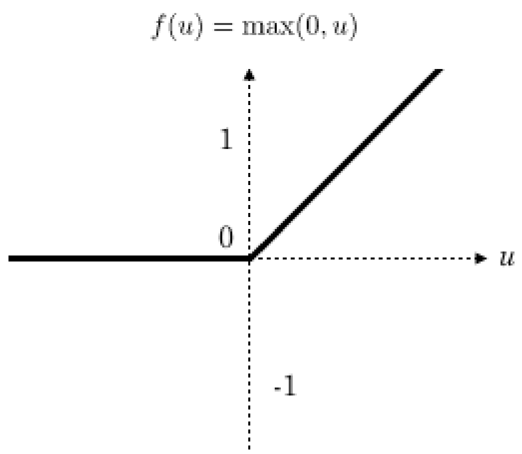

Figure 5: ReLU Activation Function

Experimental Work:

Training Dataset: Images in the training dataset had different sizes, therefore images had to be resized before using it as input to our model.

Square Images were resized to the shape 256x256 pixels. Rectangular images were resized to 256 pixels on their shortest side, then the middle 256x256 square was cropped from the image. The mean pixel values were estimated from the training dataset, one for each of the red, green, and blue channels of the color images.

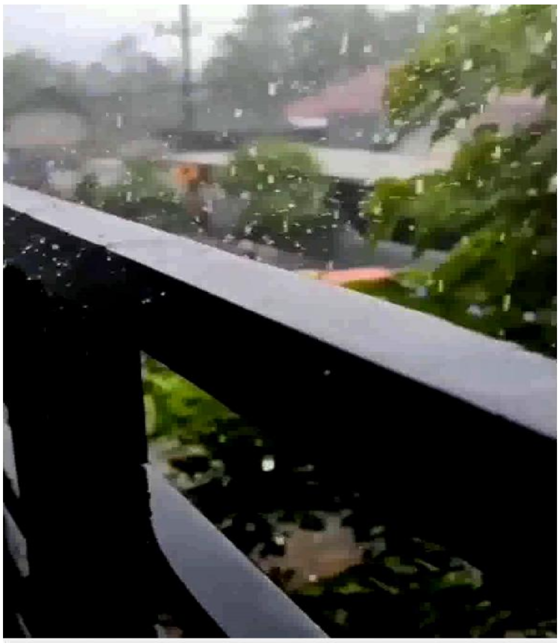

Figure 6: Output Dehazed Image

Pre requirements for implementing this project is having Core i5 Windows operating system consists of x86 or 64-bit processor with the frequency of 1.5GHz.

Image Quality Metrics:

The quality of output dehazed image can be measured with the help of dehazed image and input hazed image. Quality measures of images are Peak Signal to Noise Ratio (PSNR), Mean Square Error (MSE), and Structural Similarity Index Measurement (SSIM).

Mean Square Error is the difference between the original image and dehazed image. It is a non-negative parameter, and our desired value of MSE is nearer to the zero. It is measured by averaging the squared intensity of input hazy image and the resultant image pixels

$$\boldsymbol{MSE = \frac{1}{NM}\sum_{m=0}^{M=1}\sum_{n=0}^{N=1} e(m,n)^2}$$

Where, e(m, n) are the difference between the input and resultant images. Lower the MSE value, lower the error. The above output image MSE is 50.7296.

Peak Signal to Noise Ratio (PSNR) computes the quality measurement between input haze image and resultant image. It is a peak error measure. The higher the value of PSNR, better quality of dehazed image. It is defined as

$$\boldsymbol{PSNR = 10\ \log\frac{s^2}{MSE}}$$

Where, s = 255 for an 8- bit image. To above input image the PSNR value is 43.8380.

Structural Similarity Index Measure is a perceptual metric that qualifies image quality degradation caused by processing sech as data compression or by losses in data transmission. SSIM measures the perceptual difference between input haze and resultant dehazed image similarities. The Structural Similarity Index Metrics (SSIM) parameter is based on three important comparative measurements between the samples of X and Y are*luminance, contrast* and *structure* as in

$$S(x,y) = f\big(l(x,y)\big), c(x,y), s(x,y))$$

Resulting index value must be between 0 and 1 .The resultant SSIM for above output image is 0.6789.

Conclusion:

Hence In this work, we presented a modern deep learning approach for dehazing an image through encoding – decoding process. The transmission map can be estimated automatically by trained end to end system. We exploit the information of image through segmenting and estimated transmission map with better accuracy and decreased mean square error as compared to previous methods.

References:

[1] A. S. Parihar and O. P. Verma, "Contrast enhancement using entropy based dynamic sub-histogram equalization," IET Image Processing, vol. 10, pp. 799–808(9), November 2016.

[2] A. S. Parihar, "Histogram modification and dct based contrast enhancement," in 2018 2nd International Conference on Inventive Systems and Control (ICISC), 2018, pp. 614–618.

[3] A. S. Parihar, "Fuzzy adaptive gamma correction for contrast enhancement," in 2017 International Conference on Intelligent Sustainable Systems (ICISS), 2017, pp. 625–629.

[4] A. S. Parihar, O. P. Verma, and C. Khanna, "Fuzzy-contextual contrast enhancement," IEEE Transactions on Image Processing, vol. 26, no. 4, pp. 1810–1819, 2017.

[5] A. S. Parihar, O. P. Verma, and D. Yadav, "Image contrast enhancement using differential evolution," in Advances in Communication, Devices and Networking.Singapore: Springer Singapore, 2018, pp. 517–526.

[6] E. J. McCartney, "Optics of the atmosphere: scattering by molecules and particles," New York, John Wiley and Sons, Inc., 1976. 421 p., 1976.

[7] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," IEEE Transactions on pattern analysis and machine intelligence, vol. 33, no. 12, pp. 2341–2353, 2010.

[8] Q. Zhu, J. Mai, and L. Shao, "A fast single image haze removal algorithm using color attenuation prior," IEEE Transactions on image processing, vol. 24, no. 11, pp. 3522–3533, 2015.

[9] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in European conference on CV. Springer, 2016, pp. 154–169.

[10] B. Cai, X. Xu, K. Jia, C. Qing, and D. Tao, "Dehazenet: An end-to-end system for single image haze removal," IEEE Transactions on Image Processing, vol. 25, no. 11, pp. 5187–5198, 2016.

[11] Y. Song, J. Li, X. Wang, and X. Chen, "Single image dehazing using ranking convolutional neural network," IEEE Transactions on Multimedia, vol. 20, no. 6, pp. 1548–1560, 2017.

[12] B. Li, X. Peng, Z. Wang, J. Xu, and D. Feng, "Aod-net: All-in-one dehazing network," in IEEE Conference on Computer Vision, 2017, pp. 4770–4778.

[13] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in IEEE Conference on CVPR, 2018, pp. 3194–3203.