Research Article

Open Access

Ranjan Mondal*, Moni Shankar Dey, and Bhabatosh Chanda

Image Restoration by Learning Morphological Opening-Closing Network

<https://doi.org/10.1515/mathm-2020-0103>

Received September 12, 2019; accepted August 6, 2020

Abstract: Mathematical morphology is a powerful tool for image processing tasks. The main difficulty in designing mathematical morphological algorithm is deciding the order of operators/filters and the corresponding structuring elements (SEs). In this work, we develop morphological network composed of alternate sequences of dilation and erosion layers, which depending on learned SEs, may form opening or closing layers. These layers in the right order along with linear combination (of their outputs) are useful in extracting image features and processing them. Structuring elements in the network are learned by back-propagation method guided by minimization of the loss function. Efficacy of the proposed network is established by applying it to two interesting image restoration problems, namely *de-raining* and *de-hazing*. Results are comparable to that of many state-of-the-art algorithms for most of the images. It is also worth mentioning that the number of network parameters to handle is much less than that of popular convolutional neural network for similar tasks. The source code can be found here <https://github.com/ranjanZ/Morphological-Opening-Closing-Net>

Keywords: Morphological Network, Opening-Closing Network, Dilation Erosion Neurons, Learning structuring element, De-Hazing, De-Raining

MSC: 68T07; 68T45; 68U10

1 Introduction

Mathematical Morphology was developed initially for the analysis of geometrical structures. But its strong theoretical foundation has facilitated its use in various domains, like digital image processing and graphs, to name a few. Although this is quite an old field, there has been a renewed interest in using morphological operators for various image processing tasks in recent years. This is mostly due to their effectiveness in solving various challenging problems [36, 41] involving shapes. Depending on the task at hand, the choice and size of the Structuring Element (SE) greatly affects the performance of the methods, apart from the choice of the morphological operator [9]. Also, a single operator seldom suffices to obtain the desired results. Usually, different operators need to be operated in a particular sequence. Choosing correct sequence of correct operators along with associated SEs can be extremely hard in practice. Recently Mondal *et al.* [30, 32] have shown that the linear combination of elementary morphological operations, i.e., dilations and erosions can approximate any smooth functions and the function to be approximated can be learned using back-propagation similar to neural networks. Inspired by these findings, in this paper we propose a technique to automatically learn the sequence of elementary operations along with associated SEs required to solve some real-life problems. This may be achieved while having access to pairs of input and output images. Morphological operations are based on min-max algebra resulting in the inherent non-linearity of the system, which eliminates the need for activation function. Dilation and erosion operations increase the possible number of decision boundaries,

*Corresponding Author: **Ranjan Mondal:** Indian Statistical Institute, Kolkata, India E-mail: ranjan15_r@isical.ac.in

Moni Shankar Dey: Indian Institute of Technology, Bombay, India E-mail: msdey.iitb@gmail.com

Bhabatosh Chanda: Indian Statistical Institute, Kolkata, India E-mail: chanda@isical.ac.in

due to which the network can learn effective representation using fewer parameters compared to CNN. Most of the computer vision models for tasks such as object detection, segmentation and classification are trained using data captured in well-controlled environments. Images captured in adverse environments degrade the image quality, and it becomes necessary to restore the data before using them in computer vision tasks.

Here we propose and implement morphological network replicating the effect of opening and closing (a combination of dilation and erosion) through automatic learning of SEs. To show the effectiveness of our approach, we have applied our proposed network on two image restoration problems: image de-raining and image de-hazing. We do not use any highly optimized or highly complex architecture for the tasks to demonstrate versatility of the proposed network. Rest of the paper is organized as follows. Section 2 describes some previous works related to our objective. The proposed method is presented and explained in Section 3. After that, we have presented experimental results with discussion on de-raining and de-hazing problems in Section 4. Finally, concluding remarks and future scope of work are given in Section 5.

2 Related Works

Mathematical morphology is based on min-max algebra which results in inherent non-linearity within the operation. It has a rich theoretical framework. Mathematical morphology is used in many image processing tasks. Morphology is very efficient in removing noise [42] and contrast enhancement [33]. Bovik [7] shown that it can be used for template matching. Morphological operations are also able to extract feature from images like edges [12] and peak/valley blob. It is highly suitable for many shape-oriented problems, including character recognition [27] and shape analysis [4]. Based on the combination of basic dilation & erosion operations, many complex image filters can be built. Convolutional Neural Network (CNN) are very popular in extracting image features and solving different problems in image processing like image denoising [49, 50], semantic segmentation [5, 34]), and object detection [16, 20, 37]. Image de-raining and de-hazing are ill-posed problems, considering that one needs to accurately determine the depth of (partially) obscured surface as well as restore color. There have been several approaches in dealing with hazy images. He *et al.* [18] proposed a simple method, called as ‘dark channel’ prior, where the statistic of the outdoor haze-free image was used to find the minimum value in any one of the colour channels. This information was used to determine the haze depth and transmittance and produced the de-hazed image. Ancuti *et al.* [3] used a fast method to estimate transmittance and airlight by identifying hazy region based on the difference in the hue of image and its inverse. Transmittance was also estimated by Meng *et al.* [28] from boundary constraints on a radiance cube. Fattal [10] used the concept of color line prior, based on the fact that colors of pixels over a small image patch have a linear structure in RGB space. This line is displaced from the origin for hazy images due to airlight, which can be used for estimating transmittance. Tang *et al.* [44] learned the mapping between the hand-crafted features and transmittance. Learning features from patches, and mapping between features and transmittance can also be achieved by CNN using regression [8]. Ren *et al.* [38] proposed CNN network at multiple scales to extract features and to establish their mapping to transmittance. Li *et al.* [23] used a modified haze equation where the parameters were unified to a single variable, which was estimated by a CNN.

For image de-raining, Luo *et al.* [26] assumed sparse rain streaks having similar orientations. They proposed a discriminative approach that approximated the clean background and removed rain streak components. Chen *et al.* used a low patch-rank prior to capture rain pattern. Patch-based GMM priors to model rain streak and background by decomposing input image was proposed by Li *et al.* [25]. Fu *et al.* [14] used CNN to learn the mapping between rainy image and clean image directly. Fu *et al.* [15] decomposed input image into background and rain in different layer, from which CNN is used to remove rain streaks. Recently Mondal *et al.* [31] proposed a network with dilation and erosion operations which learns the structuring elements to remove the raindrops in a grayscale image.

3 Proposed Method

Convolution operation and elementary morphological operations such as dilation and erosion are all neighbourhood operations defined in terms of the kernel. However, morphological operators are non-linear operators, while convolution operator is linear. Though both types of operators can be defined on n -dimensions, in this work we consider morphological operations only on two-dimensional (2D) images. In this section, we build layers using dilation and erosion operators. The networks built using the dilation and erosion layers are subsequently applied for de-raining and de-hazing the images.

In mathematical morphology, the objects or its parts in an image are considered to be sets, and the operators are defined in terms of set theoretic translation, union, intersection and complementation to examine/extract various properties of these objects, such as shape and texture, using small geometrical probe known as structuring element (SE). In binary images, the pixels belong to either foreground or background (i.e., complement of foreground) set. These sets are viewed as sets in 2D space. Similarly, a grayscale image X defined over domain $\mathcal{D} = \{(x, y)\}$ may be represented as a 3D set $U_X = \{(x, y, v) \mid v \leq X(x, y)\}$, called umbra of X . Structuring element W defined over $S = \{(x, y)\}$ may also be represented by a 3D set U_W . Dilation of U_X by U_W is defined as union of all U_X translated to every pixels of U_W . To get back dilated grayscale image top surface of the dilation result is extracted. We may get eroded grayscale image in similar manner. However, a concise and more straightforward formulation of grayscale dilation and erosion is proposed by Sternberg [43] as follows.

3.1 Morphological dilation and erosion layers

Let gray-scale image X is of size $M \times N$. The Dilation ($\oplus$) and Erosion ($\ominus$) operations on X are defined, respectively, as [43]

$$(X \oplus W_d)(x, y) = \max_{(l, m) \in S} (X(x - l, y - m) + W_d(l, m)) \quad (1)$$

$$(X \ominus W_e)(x, y) = \min_{(l, m) \in S} (X(x + l, y + m) - W_e(l, m)) \quad (2)$$

where $W_d(x, y) \in \mathbb{R}$ and $W_e(x, y) \in \mathbb{R}$ are the structuring elements of dilation and erosion operators respectively defined on domain $S = \{(l, m) \mid l \in \{1, 2, 3, \dots, a\}; m \in \{1, 2, 3, \dots, b\}\}$. Note that actual shape of the geometric probe (SE) may not always cover entire rectangular domain S , then

$$W(x, y) = \begin{cases} \geq 0 & (x, y) \in G \\ -\infty & \text{otherwise} \end{cases}$$

During implementation a large number is used in place of ∞ , and that number is at least 1 (one) plus the maximum possible pixel value in gray scale image. Both dilation and erosion are many-to-one mappings. That means

$$(X \oplus W_1)(x, y) = (X \oplus W_2)(x, y) \quad \text{for some } W_1 \text{ and } W_2 \quad (3)$$

may not imply $W_1(x, y) = W_2(x, y)$. For example, let C denotes a curve whose length is greater than the diameter of a disk SE D . Then it can be shown that

$$C \oplus D = C \oplus \delta D$$

where δD denotes the boundary of D . In discrete domain, connectivity of C and D should be chosen appropriately (e.g., if C is 8-connected, then D should be 4-connected and vice versa). Similarly,

$$(X \ominus W_1)(x, y) = (X \ominus W_2)(x, y) \quad \text{for some } W_1 \text{ and } W_2 \quad (4)$$

may not imply $W_1(x, y) = W_2(x, y)$. For example, let A be a simple blob (i.e., a connected component without any hole) and D is a disk, then it can be shown that

$$A \ominus D = A \ominus \delta D$$

The examples (as depicted in equations (3) and (4)) can also be verified by taking the SE as $\delta D \cup D'$, where $D' \subset D$. This notion is useful in learning SEs for compound operators, like opening and closing.

Now, similar to convolution layer, morphological layers can be formed using dilation and erosion operators. We call a layer with dilation (resp. erosion) operation as *dilation-layer* (resp. *erosion-layer*). We also define the output of dilation-layer as *dilation feature map* and that of erosion-layer as *erosion feature map*. Note that at each layer multiple dilation or erosion may be applied using different SEs. Applying k dilation operation will generate k dilation feature map in the next layer. We denote dilation applied on the input with k number of SE of size $a \times b$ by $D_{a \times b}^k$. Similarly, erosion operation by k number of structuring element is denoted by $E_{a \times b}^k$. Note that both dilation and erosion are increasing operations and also are dual to each other.

3.2 Morphological opening and closing layers

Elementary morphological operations, i.e., dilation and erosion, are applied in many image processing tasks, such as edge detection. However, opening and closing operations are far more useful and important as these are filters. Being filters these operations satisfy essential properties like increasing and idempotent. Moreover, opening is anti-extensive, while closing is extensive. Like dilation and erosion, opening and closing are also dual to each other. In fact, many operations that are increasing, idempotent and anti-extensive are termed as 'opening'. Some examples are area opening [45] and path opening [21] which follows the properties. Traditionally, opening and closing are defined as a compound operator by concatenating dilation and erosion operators. Thus, these operations are defined in terms of SEs as

$$\text{Opening : } (X \circ W_o)(x, y) = ((X \ominus W_o) \oplus W_o)(x, y) \quad (5)$$

$$\text{Closing : } (X \bullet W_c)(x, y) = ((X \oplus W_c) \ominus W_c)(x, y) \quad (6)$$

So, morphological opening or closing network can be constructed by cascading morphological dilation and erosion layers as defined in equations (5) and (6). Equations (3) and (4) suggest that $X \circ W_o$ may be resulted in by dilation and erosion with SEs different from W_o . In other words, for a class of problems

$$(X \circ W_o)(x, y) = ((X \ominus W_o') \oplus W_o'')(x, y) \quad (7)$$

where W_o , W_o' and W_o'' may not be equal to each other. Thus, we may call $((X \ominus W_1) \oplus W_2)(x, y)$ (for some W_1 and W_2) an opening operation if it is

1. Increasing: $X \subset Y$ implies $((X \ominus W_1) \oplus W_2)(x, y) \subseteq ((Y \ominus W_1) \oplus W_2)(x, y)$
2. Idempotent: $((((X \ominus W_1) \oplus W_2) \ominus W_1) \oplus W_2)(x, y) = ((X \ominus W_1) \oplus W_2)(x, y)$
3. Anti-extensive: $((X \ominus W_1) \oplus W_2)(x, y) \subseteq X$

Similar argument is true for closing operation also. Note that to develop opening (or closing) network we could have used and trained the same SEs for dilation and erosion layers following equation (5) (resp. equation (6)). In that case we need to specify the use of opening or closing networks and their order beforehand to solve the given problem. In this work we have trained the SEs in a sequence of dilation and erosion layers independently so that the opening/closing networks (consequently, alternate sequential filters) evolve along with appropriate SEs to solve the problem in hand. Moreover, multiple dilation/erosion operations with different SEs W_k ($k = 1, 2, 3, \dots$) may be applied on an image or on already computed feature map to produce multiple dilation or erosion feature maps. These multiple feature maps are expected to highlight different types of features in the image based on the profile of the SEs.

All the SEs are initialized randomly while building the network and are trained using back-propagation based on the training samples (images). It may be noted that the erosion and dilation operations use min and max operations, respectively. So, though not fully differentiable, these are at least piece-wise differentiable. Hence, there occurs no problem in back-propagating the gradient to update the SEs. Next we briefly present how back-propagation takes place in the morphological network to train the SEs.

3.3 Back-propagation in Morphological Network

Back-propagation algorithm is used to update the SEs in the morphological opening-closing network. We have already stated that our opening and closing networks are built using dilation and erosion layers. Second, dilation and erosion are dual operations, so describing the training of SEs for either of them is sufficient to describe the other. For simplicity, we consider here a single dilation layer. The propagation of gradient thorough the network is very similar to that of a neural network. To start with we recall the expression of the gradient. Suppose an input image X having a dimension of $M \times N \times c$ is passed through a dilation layer that produces an output feature map Y . The structuring element S in the dilation layer has size $A \times B \times c$ and, with appropriate padding, the size of Y is $M \times N$. The output at location (x, y) of Y can be expressed as :

$$Y(x, y) = \max_{l, m, n} \{X(x + l, y + m, n) + S(l, m, n)\} \quad (8)$$

where (l, m, n) denotes the index of the SE S . The SE is trained using back-propagation so that Y approaches desired dilation feature map $\tilde{Y}$. Let L be the loss or error between output of the dilation layer and the desired dilation feature map, i.e.,

$$L = L(\tilde{Y}, Y) \quad (9)$$

When SE S is properly trained, $Y \rightarrow \tilde{Y}$ which implies $L \rightarrow 0$. It is evident from equations (8) and (9) that Y and, consequently, L depend on SE S . Using chain rule due to partial derivatives we have

$$\begin{aligned} \frac{\partial L}{\partial S(x, y, z)} &= \sum_x \sum_y \frac{\partial Y(x, y)}{\partial S(x, y, z)} \frac{\partial L}{\partial Y(x, y)} \\ &= \sum_x \sum_y \nabla Y(x, y) \frac{\partial L}{\partial Y(x, y)} \end{aligned}$$

where

$$\nabla Y(x, y) = \begin{cases} 1 & \text{if } Y(x, y) = X(x + \delta x, y + \delta y, \delta z) + S(\delta x, \delta y, \delta z) \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

Thus the structuring element is updated as

$$S(x, y, z) = S(x, y, z) - \alpha \frac{\partial L}{\partial S(x, y, z)} \quad (11)$$

where α is the learning rate. If multiple SEs $S_k(x, y, z)$ ($k = 1, 2, 3, \dots$) are used, multiple dilation feature maps Y_k ($k = 1, 2, 3, \dots$) are obtained, but the back-propagation strategy would be same.

The idea is straightaway extended to multi-layer network. Suppose an intermediate feature map denoted by X_i is passed through i -th dilation layer and produces an output feature map $X_{i+1,k}$ after dilating with SE $S_{i,k}$ in i -th layer. Then, keeping all other parameters same, equation (8) can be re-written as

$$X_{i+1,k}(x, y) = \max_{l, m, n} \{X_i(x + l, y + m, n) + S_{i,k}(l, m, n)\} \quad (12)$$

Let L is the final loss of the network by comparing desired output and the predicted output. As before we can calculate the gradient $\frac{\partial L}{\partial S_{i,k}(x, y, z)}$ with respect to the k -th SE by the following equation using chain rule:

$$\begin{aligned} \frac{\partial L}{\partial S_{i,k}(x, y, z)} &= \sum_x \sum_y \frac{\partial X_{i+1,k}(x, y)}{\partial S_{i,k}(x, y, z)} \frac{\partial L}{\partial X_{i+1,k}(x, y)} \\ &= \sum_x \sum_y \nabla X_{i+1,k}(x, y) \frac{\partial L}{\partial X_{i+1,k}(x, y)} \end{aligned}$$

where

$$\nabla X_{i+1,k}(x, y) = \begin{cases} 1, & \text{if } X_{i+1,k}(x, y) = X_i(x + \delta x, y + \delta y, \delta z) + S_{i,k}(\delta x, \delta y, \delta z) \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

Figure 1: (a) A simple opening network trained with 400 images of flicker data to verify properties, (b) trained structuring elements of erosion layer and dilation layer, respectively.

The term $\frac{\partial L}{\partial X_{i+1,k}(x, y)}$ can be obtained by computing recursively starting from the final layer. Finally, k -th SE $S_{i,k}(x, y, z)$ is updated using equation (11). Similarly, the gradient for the erosion layer can be derived. A worked-out example of gradient calculation for the erosion layer is shown in [13].

3.4 Experiments with Morphological Networks

In order to establish the efficacy of our concept of learning morphological opening and closing networks along with associated structuring elements (kernels), we have carried out some initial experiments before applying to real-life problems. However, we first need to show experimentally that the opening network realized by concatenating erosion and dilation layers with not necessarily the same SEs may satisfy desired properties of opening operation, i.e., increasing, idempotent and anti-extensive. In other words, a network made of an erosion layer followed a dilation layer having different SEs satisfies essential properties of opening and, thus, can work like opening network.

3.4.1 Verifying properties of opening operation

First, for training and testing the network to be developed, we have taken 400 natural images from flickr [35] and converted those to grayscale images of size 416×416 . Each image is opened with a disk SE of radius 7 to generate the groundtruth of opened image. Second, we construct an opening network by concatenating an erosion layer followed by a dilation layer (see figure 1(a)), where each layer has its own structuring element trained by back-propagation to minimize the loss. We take SEs of size 20×20 and initialize with random numbers from uniform distribution. Suppose image X is input to the network and at the output we obtain image $\gamma(X)$.

The experiment is carried out following a 10-fold cross-validation strategy. That means the total image set is divided into 10 groups, and images of 9 groups are used to train the network while the images of the remaining group are used as test images. It is repeated 10 times with each group as the test set and the result of all 400 images are accumulated while those are considered as test images. A sample pair of SEs trained in the network is shown in figures 1(b) and 1(c).

Verifying 'Increasing' property:

Given an image X we generate an intermediate image X' as

$$X'(x, y) = kX(x, y) + \mathcal{N}(0, 1)$$

where $\mathcal{N}$ is a noise image whose pixel values range from 0 to 1. Finally, X' is clipped between 0 and 1 inclusive to get the final noisy image $\hat{X}$ with bright noise. Thus $X < \hat{X}$, and increasing property ensures that $\gamma(X) \leq \gamma(\hat{X})$.

Figure 2: The distribution of (a) Increasing error, (b) Idempotent error, and (c) Anti-extensive(b).

So the opening network fails to satisfy this property if $\gamma(X)(x, y) - \gamma(\hat{X})(x, y) > 0$ for any (x, y) . Hence, for u -th image X_u we calculate error $\mathcal{E}_{Iu}$ for increasing property as

$$\mathcal{E}_{Iu} = \frac{1}{|X_u|} \sum_{x,y} \gamma(X_u)(x, y) - \gamma(\hat{X}_u)(x, y) \quad \text{if } \gamma(X_u)(x, y) - \gamma(\hat{X}_u)(x, y) > 0 \quad (14)$$

where $|X_u|$ denotes the size of image X_u . Histogram of this error $\{\mathcal{E}_{Iu} | u = 1, 2, \dots, 400\}$ is shown in figure 2(a). The mean and standard deviation of $\mathcal{E}_{Iu}$ are 0.000 and 0.000, respectively, which implies that the network satisfies the increasing property.

Verifying ‘Idempotent’ property: As before an Image X feed to trained opening network produces $\gamma(X)$. Now if we feed $\gamma(X)$ again to the network, we get $\gamma(\gamma(X))$. To hold the idempotent property, $\gamma(X)$ should be equal to $\gamma(\gamma(X))$. In other words, the network fails to satisfy idempotent property if $|\gamma(X)(x, y) - \gamma(\gamma(X))(x, y)| \neq 0$, for any (x, y) . Thus, for each u -th image X_u we calculate error $\mathcal{E}_{du}$ as

$$\mathcal{E}_{du} = \frac{1}{|X_u|} \sum_{x,y} |\gamma(X_u)(x, y) - \gamma(\gamma(X_u))(x, y)| \quad (15)$$

Histogram of this error $\{\mathcal{E}_{du} | u = 1, 2, \dots, 400\}$ is shown in figure 2(b). Mean and standard deviation of $\mathcal{E}_{du}$ are 0.019 and 0.004, respectively. Since these values are not equal to zero (0) as before, we can at most say that the network very closely satisfies the idempotent property and expect it to satisfy fully if the network is trained better with a larger number of images.

Verifying ‘Anti-extensive’ property: This property may be verified in exactly similar way for increasing property because anti-extensive property states that $\gamma(X_u)(x, y) \leq X_u(x, y)$. Hence, corresponding error $\mathcal{E}_{au}$ may be computed as

$$\mathcal{E}_{au} = \frac{1}{|X_u|} \sum_{x,y} \gamma(X_u)(x, y) - X_u(x, y) \quad | \gamma(X_u)(x, y) - X_u(x, y) > 0 \quad (16)$$

Histogram of this error $\{\mathcal{E}_{au} | u = 1, 2, \dots, 400\}$ is shown in figure 2(c). The mean and standard deviation of $\mathcal{E}_{au}$ are 0.000 and 0.000, respectively, which implies that the network satisfies the anti-extensive property.

Note that we could have trained the same SE for both erosion and dilation layers to realize opening (and similarly closing) network, but in that case, we need to define the sequence of operations to solve a given problem. On the other hand, training the SEs independently for each layer allows us to develop the required network to solve the problem at hand. Next, we show how different networks evolve from the elementary morphological layers based on the trained SEs.

3.4.2 Implementing various morphological network using common framework

Here we again take the same set of grayscale images and try to simulate various morphological networks using a sequence of dilation and erosion layers. To generate the ground truth, first, we apply a few morphological

operations in a conventional way on the set of images. Then we build a network architecture as a sequence of dilation and erosion layers. We train the morphological network, especially, the structuring elements using the back-propagation algorithm and mean squared error loss between the output by conventional method and the output produced by the network.

In Table 1, we have shown the SEs that are learned based on assigned task and simulated input and output pair. The first column indicates the morphological operations used to generate the output image (Y) from the input image (X). The second column shows the network architecture we have employed to achieve the operation stated in the first column. After training the network (for both the required operation and the associated SE), it is expected to generate the desired output Y given the input X . The learned SEs are shown in the third column. Note that in case of redundant or unnecessary layer, corresponding SE would be an impulse, which realizes an identity transform. For example, in the first row, we try to simulate 'dilation' operation, while the network architecture consists of an erosion layer followed by a dilation layer. After convergence, the trained SE for erosion is found to be approximately an impulse; while SE for the dilation layer is a disk of desired dimension. Similar network is used to simulate erosion and opening as shown in the second and third rows, respectively. The trained SEs justifies their objectives. In the fourth and fifth rows, networks with four layers are used to simulate opening and closing. In the fifth row, it can be seen that the SEs of the first layer (erosion) and the fourth layer (dilation) are approximately impulses as these layers are redundant and unnecessary, and so are expected to implement identity transform; whereas in the second (dilation) and the third (erosion) layers the learned SEs are similar to disks. However, in the fourth row, combination of first and second layers or combination of third and fourth layers can simulate opening. Secondly, opening is an idempotent operation. So we see trained SEs are more or less similar and are of significant size in all four layers.

In this arrangement We obtain a single dilated or eroded feature map after each layer. However, by using multiple SEs at each layer, multiple feature maps can be generated for performing complex tasks like de-raining and de-hazing.

Though in this simulation experiment, the input and output of the network are both graylevel images, this may be extended to colour image where both input to our network and corresponding output are in (R,G,B) format. We do not apply any strategic treatment to handle colour information, rather let the network to learn the three-dimensional SE of size $A \times B \times 3$ so that it can produce desired output. Second, we have used multiple SEs in each layer. It may help to generate and process features from the colour image and produce desired output. Moreover, in this experiment we use a single SE at each layer, so single feature map is obtained after each layer. By using multiple SEs at each layer, multiple feature maps can be generated for performing complex tasks like de-raining and de-hazing as described in the next sections.

3.5 Image De-raining

The degradation of rainy images depends on several factors such as raindrop size, track of raindrops, rain density and lighting condition. There may be other types of noises present simultaneously in the input image. Morphological filters such as opening and closing are capable of removing noise from an image while preserving the edges. Removing raindrops from an image can be considered as removing mostly bright noise of particular shape and size from the image. This suggests that alternate sequence of dilation and erosion layers, forming opening/closing operations, in the morphological network should be able to realize the de-raining operations by learning the appropriate SEs. However, it is challenging to know the size and shape of the raindrops and other factors beforehand.

Hence, for the purpose, we propose, a morphological network architecture consisting of a sequence of multiple pairs of dilation and erosion layers, which may result in alternating sequential filters (ASF) in parallel paths. Schematic diagram of the network is shown in Fig. 3. The output feature maps from the two paths are then linearly combined to get the output maps. This step is essential for recovering undesired removal of some features by opening or closing filters. Finally, sigmoid activation function is applied to it to produce the final output (image). In Table 2 we have shown three architectures, which we have used for our experiments.

Table 1: Illustrates the efficacy of proposed morphological network. Network architecture given in 2nd column tries to simulate the operation suggested in 1st column. Corresponding learned SEs are shown in 3rd column. (For detail see the text.)

Morphological Operation to be Achieved	Network Architecture	Learned Structuring elements from the first layer
Dilation	$X \rightarrow E_{10 \times 10}^1 \rightarrow D_{10 \times 10}^1 \rightarrow Y$
Erosion	$X \rightarrow E_{10 \times 10}^1 \rightarrow D_{10 \times 10}^1 \rightarrow Y$
Opening	$X \rightarrow E_{10 \times 10}^1 \rightarrow D_{10 \times 10}^1 \rightarrow Y$
Opening	$X \rightarrow E_{10 \times 10}^1 \rightarrow D_{10 \times 10}^1 \rightarrow E_{10 \times 10}^1 \rightarrow D_{10 \times 10}^1 \rightarrow Y$
Closing	$X \rightarrow E_{10 \times 10}^1 \rightarrow D_{10 \times 10}^1 \rightarrow E_{10 \times 10}^1 \rightarrow D_{10 \times 10}^1 \rightarrow Y$

In [31], the authors have shown that opening network, as expected, is more effective in removing bright noise. Usually closing network removes dark noise, which does not have much to do in this application. However, to make our network more general, we take a linear combination of the output of opening path and closing path and make a single generic architecture. Since the output of the de-rained image should be of the same size as the input image, appropriate zero padding is used in each layer. For the color image processing, a number of feature maps after each dilation layer and erosion layer using as many SEs. Multiple feature maps help in propagating the color information to the output.

During training, we initialize all the SEs and the weights of the linear combination layers with random numbers. To train the network, we define a loss using the structural similarity index measure (SSIM) between predicted output P and the ground truth T of the data defined as

$$\text{SSIM}(P, T) = \frac{(2\mu_P\mu_T + c_1)(2\sigma_{PT} + c_2)}{(\mu_P^2 + \mu_T^2 + c_1)(\sigma_P^2 + \sigma_T^2 + c_2)} \quad (17)$$

Here μ_P and μ_T are the mean for image P and T , respectively; while σ_P and σ_T are their standard deviations. The term σ_{PT} is the co-variance between the images, and c_1 and c_2 are constants, whose values are set to 0.0001 and 0.0009, respectively to avoid divide by zero. We use the structural dissimilarity (DSSIM) as the loss function considering all small patches of the output and those of corresponding ground truth for training the network. So the loss function is defined as

$$\text{Loss}_{total} = \text{DSSIM}(I_{out}, I_{gt}) = \frac{1}{M} \sum_i \frac{1 - \text{SSIM}(P_{out}^i, P_{gt}^i)}{2} \quad (18)$$

where I_{out} and I_{gt} are the output image and the ground truth image, and P_{out}^i and P_{gt}^i are the i^{th} patches, respectively. There are a total of M patches. In our experiment, we have taken the patch size as 10×10 . We minimize the loss function Loss_{total} by back-propagation algorithm and learn the parameters of the network.

Figure 3: The network for image de-raining. It consists of 2 parallel paths containing a complementary sequence of Erosion and Dilation layers. The structuring element for each operation is 8×8 . The outputs from each of the paths are taken together for linear combination to produce three feature maps, which pass through *sigmoid* activation to get the final output.

Table 2: The architectures of paths shown in Fig. 3. $D_{8 \times 8}^8$ denotes a layer with 8 dilation operators with different trainable SEs of size 8×8 . $E_{8 \times 8}^8$ is defined similarly. In the last opening layer of closing network and last closing layer of opening network, we have taken c number of structuring elements where c is 1 for grayscale image and 3 for color image processing task.

Description of the MorphoN	
Closing Network	$X \rightarrow D_{8 \times 8}^8 \rightarrow E_{8 \times 8}^8 \rightarrow D_{8 \times 8}^8 \rightarrow E_{8 \times 8}^8 \rightarrow D_{8 \times 8}^8 \rightarrow E_{8 \times 8}^c \rightarrow Y$
Opening Network	$X \rightarrow E_{8 \times 8}^8 \rightarrow D_{8 \times 8}^8 \rightarrow E_{8 \times 8}^8 \rightarrow D_{8 \times 8}^8 \rightarrow E_{8 \times 8}^8 \rightarrow D_{8 \times 8}^c \rightarrow Y$
Opening-Closing Network	$ \begin{array}{c} \begin{array}{c} \text{X} \end{array} \begin{array}{l} \xrightarrow{\quad} E_{8 \times 8}^8 \rightarrow D_{8 \times 8}^8 \rightarrow E_{8 \times 8}^8 \rightarrow D_{8 \times 8}^8 \xrightarrow{w_c^c} \\ \xrightarrow{\quad} D_{8 \times 8}^8 \rightarrow E_{8 \times 8}^8 \rightarrow D_{8 \times 8}^8 \rightarrow E_{8 \times 8}^8 \xrightarrow{w_s^c} \end{array} \begin{array}{c} \text{Sigmoid} \\ \otimes \end{array} \rightarrow Y \end{array} $

3.6 Image De-Hazing

When light rays travel through a turbid medium, it undergoes a phenomenon known as scattering, wherein the light is scattered in different directions due to its interactions with particles, such as dust and aerosols floating in it. Haze occurs when the concentration of these particulate matters exceed a certain threshold. Any image taken in hazy conditions suffers from visibility degradation such as reduced contrast, saturation attenuation and color shifting. Besides, scattered environmental light appears like a veil over the scene. These days, image de-hazing has become one of the trending problems. Here we intend to restore such hazy images using the opening-closing network. The observed hazy image can be physically modelled using the following equation [22]

$$I(x) = t(x)J(x) + (1 - t(x))A, \quad (19)$$

where $I(x)$ is the observed or recorded intensity of the hazy image at location x and $J(x)$ is the intensity of corresponding non-hazy (ideal) image at location x . A is the airlight, which characterizes the constant environmental illumination. $t(x)$ is the transmittance coefficient that determines the amount of light reaching the observer (camera) from the objects after travelling through the medium. Transmittance $t(x)$ intuitively measures the amount of haze present in a particular location x and, in general, depends upon the depth of the scene. We modify equation 19 as

$$I(x) = t(x)J(x) + K(x), \quad (20)$$

where the bias like term $K(x)$ is the space-variant airlight representing $(1 - t(x))A(x)$. Also note that, since $0 \leq t(x) \leq 1$, contrast of $J(x)$ is reduced to $J(x)t(x)$. It is known that morphological opening can estimate local bias if flat SE of appropriate size is used. Second, Opening (resp. Closing) operator can detect local bright

Figure 4: The network for image de-hazing. It consists of 2 parallel paths containing a complementary sequence of Erosion and Dilation layers. The structuring element for each operation is 4×4 . The output from each of the paths is concatenated and passed through three 1×1 convolution layers. The airlight and transmittance map is generated by sending the output to 2 parallel sub-networks. The output from each of the sub-network is linearly combined to get the final output.

(resp. dark) features whose scaling affects the contrast. The morphological network is employed to estimate the airlight map as well as the transmittance map in order to recover the haze-free image. As shown in Fig. 4, we have taken two separate paths: one for opening and the other closing. The outputs of the paths undergo linear combination. The network produces the transmittance map $t(x)$ and the airlight map $K(x)$ as output. Since $0 \leq t(x) \leq 1$ and $0 \leq K(x) \leq 1$, in the last layer of transmittance path and also of airlight path we have employed sigmoid activation function to limit the estimated values of $t(x)$ and $K(x)$ within the said range. Once $t(x)$ and $K(x)$ are estimated, we can determine de-hazed image as

$$J_{out}(x) = \min \left(\frac{I(x) - K(x)}{t(x)}, 1 \right) \quad (21)$$

During training of the network, with given $t(x)$ and $K(x)$, hazy image is generated as

$$I(x) = J(x)t(x) + K(x); \quad (22)$$

Now, given a pair of hazy and haze-free clear images, the network learns the SEs and also the weights of the linear combination layer. Network gives $\hat{t}(x)$ and $\hat{K}(x)$ at every iteration of the training process. We reconstruct the hazy image $\hat{I}_{out}$ and estimated de-hazed image $\hat{J}_{out}$ based on these $\hat{t}(x)$ and $\hat{K}(x)$ using the equations 21 and 22. We define the loss function L , which is very similar to bi-directional consistency loss [29] given by

$$L = L_1 + L_2, \quad (23)$$

where,

$$L_1(x) = \text{DSSIM}(I, \hat{I}_{out}), \quad (24)$$

$$L_2(x) = \text{DSSIM}(J, \hat{J}_{out}); \quad (25)$$

DSSIM is calculated using equation 18. We minimize the loss L and learn the network parameters. In the next section, we present experimental results to justify our claim.

4 Experimental results

In this section, we evaluate performance of the proposed network both quantitatively and qualitatively on de-raining and de-hazing tasks.

Table 3: Test results of de-rain network, quantified by SSIM and PSNR, on both grayscale and color images of Rain dataset. **Closing Net** and **Opening Net** show the results of the two parallel paths when trained individually. The result of de-rain network (**opening-closing net**) is compared with **CNN**

	Metric	Base Line	Closing Net	Opening Net	Opening-Closing Net	CNN	MorphoN[31]
Gray	SSIM	0.85	0.90	0.90	0.91	0.93	0.92
	PSNR	24.41	26.00	25.99	27.29	29.24	28.03
Color	SSIM	0.84	0.88	0.87	0.89	0.91	-
	PSNR	24.06	24.81	24.28	25.01	27.33	-

4.1 Experimental Setup and Data Set

We have implemented the network in Python using Keras with TensorFlow library at the backend. We carried out our experiments on Intel machine with a GPU of memory 12GB. For all the experiments, we have initialized the structuring elements randomly using standard glorot uniform initializer [17]. To minimize the loss, we have used RMSProp optimizer for all the networks.

Training of image de-raining network is done with benchmark rain dataset [14]. This rain dataset has 1,000 clean images. For each clean image, there are 14 different synthesized rainy images with different streak orientations and sizes. So, a total of 14,000 sample images are available. We have considered 80% of the data for the training and 10% of the data for validation, and the rest 10% is used as test data. Images in the dataset are of different sizes. However, as the network takes fixed-size input, all the images are resized to 416×416 through bi-linear interpolation.

For de-hazing, we have used O-HAZE [1] dataset for training. This dataset was first proposed in O-HAZE [3] challenge. The dataset has 35 images for the training and five images for validation and five images for testing. In this set, images are of different sizes, and the sizes are of the order of 5500×3500 pixels. So we have resized each image to 1024×1024 . However, while testing, we have taken the full resolution of the images. Since the O-HAZE dataset is very small, we have also used NYU dataset of D-Hazy dataset [2] for the training. For quantitative evaluation of our Opening/Closing network, we have used validation dataset of the O-HAZE as test set because their ground truths are available. We also test our network in Fattal and Middlebury dataset of D-Hazy [2].

We have trained both the de-rain and de-haze network until the loss is converged. Note that, the loss function involving max/min (due to dilation/erosion) operation is only piece-wise differentiable. In practice, we can find the sub-gradient of each morphological operation and back-propagate through the network. In the next two subsections, we have evaluated our Morphological network developed for De-hazing and De-raining network.

4.2 Results of Image De-Raining

Here we present the quantitative and qualitative evaluation of morphological network for image de-raining.

4.2.1 Quantitative Evaluation

In order to evaluate quantitatively the opening-closing network applied on rainy images, we have used two objective measures such as SSIM [46] and Peak signal-to-noise ratio (PSNR). In [31], image de-raining is done only on grayscale images. We have extended that work for de-raining of color images. The extension is not straightforward and a minor modification of network is needed to preserve hue in the image. We trained the network on the rainy image dataset until the convergence of training loss is achieved. Opening Net, Closing Net, and the opening-closing net are trained separately to study their relative performances. The architec-

Figure 5: Result of the De-rain network over grayscale images from Rain dataset. The output from Closing Net, Opening Net, De-rain network and CNN, along with input image and ground truth are shown for qualitative comparison.

tures of these networks are illustrated in table 2. The estimated de-rained image from each network is quantitatively evaluated against the rain-free image available as ground truth in the dataset. The average values of SSIM and PSNR for grayscale and color test images, are reported in Table 3. In [31], we have shown that opening path (network) performs better than the closing path (network) as the opening filter removes bright noise. However, as shown in Table 3 closing network performs similarly well as opening network. We believe this is because we have taken multiple dilation/erosion layers and problem dependent learning of SEs has created effectively an opening path emphasizing bright noise removal. We have also compared our results with a standard convolutional neural network (CNN) of U-net architecture [40]. This Table also reveals that the opening-closing network gives results comparable to that of CNN with U-net architecture. However, we observe that CNN has more generalization capability than the opening-closing network, but needs to train a huge number of parameters, which incurs a very high cost. It can also be seen that in the PSNR metric, we do not get much improvement for both grayscale and color images. It may be because of training the network with SSIM loss only. Incorporating mean squared error in the loss may result in better PSNR.

4.2.2 Qualitative Evaluation

For qualitative evaluation, we have shown the results of different networks applied on a large number of images to nine evaluators (research fellows in our lab) and asked them to assign ranks to the output. Order of average rank is similar to quantitative evaluation. That is CNN comes out to be the best closely followed by opening-closing network. Results of opening network and closing network are comparable. Here we display results for only 3 grayscale images in Fig. 3 and 3 color images in Fig. 6. In both figures, we compare the output of each network with ground truth.

In grayscale images (Fig. 3) it may be noted that though the opening-closing network did remove raindrops successfully, blobs of uniform intensity are present. This effect is a bit more severe in case of opening network and closing network (especially in the top row images). In the case of color images, as seen in Fig. 6, the opening-closing network can effectively de-rain and also reproduces the vivid colors with desired contrast. The output of the closing network has less contrast than the output of the opening network, whereas

the opening-closing Net is able to address the issue. Secondly, the blob effect, as seen in grayscale images, are not visible here.

The results shown above are on images from benchmark rain dataset [14], where rain images are synthesized from clean images; hence groundtruth for images are available. We also show few results on real rain images in Fig. 7 for which the groundtruth are not available. The opening-closing network is able to clear the rain perfectly, whereas the output of CNN(U-Net) has surprisingly few small rain streaks in some images (see bottom row of the figure). This may be due to presence of large number of parameters in CNN (U-net) leading to over-fitting.

Table 4: Quantitative evaluation on validation dataset of O-HAZE images in terms of PSNR and SSIM metrics.

Method	36.png PSNR/SSIM	37.png PSNR/SSIM	38.png PSNR/SSIM	39.png PSNR/SSIM	40.png PSNR/SSIM	Average PSNR/SSIM
DCP[18]	18.18/0.45	16.09/0.49	14.12/0.08	12.88/0.36	14.21/0.39	15.10/0.35
CAP[51]	17.47/0.50	16.17/0.45	15.14/0.18	14.80/0.41	16.37/0.57	15.99/0.42
MS-CNN[38]	16.59/0.49	15.76/0.43	13.25/0.19	12.79/0.40	16.53/0.56	14.99/0.41
DehazeNet[8]	16.92/0.43	14.98/0.48	15.54/0.34	17.65/0.48	17.04/0.54	16.43/0.45
AOD-Net[23]	17.10/0.45	16.47/0.39	16.12/0.12	15.04/0.34	15.95/0.50	16.13/0.36
DCPDN[47]	17.14/0.44	15.29/0.42	14.66/0.11	15.24/0.36	17.78/0.52	16.02/0.36
Zhang et al.[48]	24.67/0.73	22.41/0.66	23.75/0.72	21.91/0.63	22.29/0.68	23.00/0.68
Ours	20.22/0.75	21.11/0.75	19.45/0.75	19.63/0.76	14.57/0.64	19.00/0.73

4.3 Results of Image De-Hazing

In this subsection, we evaluate the performance of the opening-closing network quantitatively and qualitatively with respect to the de-hazing problem. De-hazing problem is relatively more widely attempted. So

Figure 6: Result of De-rain network over color images from Rain dataset. The output from Closing Net, Opening Net, De-rain network and CNN, along with input image and ground truth are shown for qualitative comparison.

Figure 7: Results from Closing Net, Opening Net, De-rain network and CNN on real rainy images

Table 5: Quantitative evaluation of opening-closing network with state of art algorithms on 12 images of Fattal dataset in terms of SSIM and PSNR metrics.

Image	He et al. [19] PSNR/SSIM	Fattal [11] PSNR/SSIM	Ren et al. [39] PSNR/SSIM	Berman et al. [6] PSNR/SSIM	Li et al. [24] PSNR/SSIM	Opening-Closing Net PSNR/SSIM
church	11.16/0.78	21.43/0.96	14.17/0.87	15.68/0.89	9.44/0.62	12.32/0.82
couch	18.4/0.86	20.8/ 0.9	17.99/ 0.88	17.24/0.87	16.77/0.83	16.04/0.81
dolls	19.73/ 0.85	21.29/0.77	16.93/ 0.86	15.69/0.83	17.21/0.85	17.76/0.81
flower1	14.1/ 0.88	30.01/0.98	9.08/0.43	12.15/0.72	12.22/0.79	14.71/0.88
flower2	14.37/ 0.86	31.94/0.99	10.81/0.6	11.86/0.68	13.13/0.79	15.16/0.85
lawn1	13.84/0.8	24.49/0.97	14.37/0.83	14.78/0.86	11.32/0.69	11.63/0.77
lawn2	11.2/0.74	24.94/0.97	13.29/0.77	15.32/0.87	10.98/0.68	11.55/0.77
mansion	17.45/0.87	26.96/0.97	17.69/0.89	17.33/ 0.87	14.24/0.7	15.73/0.86
moebius	12.66/0.78	19.01/ 0.9	16.36/ 0.9	14.58/0.85	13.22/0.77	13.82/0.86
raindeer	18.12/0.83	26.22/0.94	16.82/0.81	16.59/0.82	16.53/0.8	16.41/0.85
road1	12.95/0.8	25.74/0.96	14.11/0.84	16.3/0.88	11.75/0.66	12.05/0.78
road2	15.84/0.84	23.6/0.96	16.45/0.88	18.22/0.9	11.96/0.62	12.44/0.78
Average	14.98/0.82	24.7/0.94	14.84/0.8	15.48/0.84	13.23/0.73	14.13/0.82

there exist quite a few benchmark datasets and many good algorithms. We have compared our method with multiple classical methods like Dark channel Prior(DCP) [18],color attenuation prior(CAP) [51] and multiple deep learning methods like Multiscale CNN(MS-CNN) [38], DehazeNet [8] AOD-Net [23] and DCPDN [47]. DCP method is based on the observation that in hazy images intensity of color channels is uniformly contributed by the air-light. In CAP method, they have approximated the haze by the difference of brightness and saturation. Recent deep learning based methods like MS-CNN and DehazeNet predict transmittance map using the network. AOD-Net reformulates the haze equation using a single parameter and estimates it using CNN based network. DCPDN algorithm incorporates the concept of image pyramid inside the network.

Table 6: Quantitative results of opening-closing network on Middlebury portion of D-Hazy dataset with PSNR and SSIM metrics. The opening-closing network has the second highest average PSNR and SSIM values over the other methods

Image	He et al. [19] PSNR/SSIM	Ren et al. [39] PSNR/SSIM	Berman et al. [6] PSNR/SSIM	Li et al. [24] PSNR/SSIM	Opening-Closing Net PSNR/SSIM
Adirondack	16.02/0.82	14.39/ 0.89	16.74/0.88	14.18/ 0.89	16.54/0.87
Backpack	14.4/0.85	16.21/0.87	12.24/0.82	16.1/0.91	14.22/0.85
Bicycle1	12.39/0.81	20.66/0.93	12.61/0.82	23.21/0.96	14.2/0.86
Cable	12.95/0.7	7.65/0.64	9.93/0.63	6.95/0.64	12.54/0.68
Classroom1	20.17/0.87	10.91/0.74	20.95/0.89	10.02/0.72	12.34/0.76
Couch	18.68/0.81	10.13/0.61	13.76/0.7	10.56/0.63	12.01/0.64
Flowers	17.73/0.89	10.47/0.78	17.45/0.9	9.25/0.76	14.85/0.83
Jadeplant	13.48/0.69	7.78/0.6	7.06/ 0.65	7.65/0.59	13.1/0.64
Mask	15.88/ 0.89	14.15/0.85	14.18/0.84	14.3/ 0.91	15.88/0.85
Motorcycle	13.81/0.79	13.2/ 0.81	11.6/0.62	12.25/ 0.82	14.49/0.79
Piano	18.66/0.86	12.4/0.71	15.08/0.78	13.89/0.75	16.8/0.79
Pipes	15.52/0.79	10.9/0.68	13.81/0.74	10.34/0.69	15.42/0.78
Playroom	17.7/0.85	13.42/0.77	17.64/ 0.83	13.24/0.78	15.47/0.79
Playtable	18.58/0.9	15.09/0.86	16.63/0.88	14.73/0.86	15.99/0.86
Recycle	12.5/0.82	18.3/0.95	13.43/0.88	16.62/0.9	13.03/0.87
Shelves	15.47/0.83	20.43/0.94	16.9/0.88	16.52 ?/ 0.92	16.88/0.87
Shopvac	13.87/0.8	7.62/0.66	11.58/0.78	6.89/0.64	11.15/0.73
Sticks	16.96/0.9	20.5/0.96	20.41/0.93	19.13/ 0.96	19.28/0.93
Storage	17.38/0.88	11.23/0.82	16.36/ 0.88	10.24/0.79	17.71/0.87
Sword1	15.06/0.87	15.48/0.91	12.57/0.83	14.29/ 0.91	14.37/0.86
Sword2	15.66/0.89	12.89/0.88	14.89/0.88	12.8/ 0.9	15.91/0.89
Umbrella	10.4/0.8	14.92/0.9	9.63/0.72	14.58/0.91	12.11/0.84
Vintage	14.63/0.86	19.27/0.96	14.09/0.83	16.82/0.94	13.76/0.85
Average	15.56/0.83	13.82/0.81	14.33/0.81	13.35/ 0.82	14.70/0.82

4.3.1 Quantitative Evaluation

In De-hazing, we quantitatively compare the performance of our opening-closing network with other state-of-the-art methods. We have trained the opening-closing network using the O-HAZE training dataset and tested the network on O-HAZE validation dataset. The result is reported in Table 4. The PSNR and SSIM values of the results of different algorithms for each of the 5 images of validation dataset are shown in the table. It is seen that the output of the opening-closing network gives the best results in terms of SSIM in most of the images. However, in terms of PSNR, it ranked 2nd best in the table next to that of [11].

We have also trained the network using NYU portion of D-Hazy dataset and tested it on Fattal [11] and Middlebury portion of D-Hazy dataset. We report these results in Table 5 and Table 6, respectively. From Table 5 it is revealed that the performance of the opening-closing network on Fattal dataset is not as good as that on O-HAZE dataset (Table 4). Probably it happens because NYU portion of D-Hazy dataset has only white airlight, while in Fattal dataset the airlight may not be white. However, in Table 6, we see that the performance of our morphological network is comparable to that of state-of-the-art algorithms on Middlebury portion of D-Hazy dataset.

4.3.2 Qualitative Evaluation

For qualitative evaluation of the de-hazed image, we have used the same strategy as has been used for de-rained images. However, comparing the quality of de-hazed images is so difficult that grading is often inconsistent. We show the output of the opening-closing network on only three challenging images from O-HAZE

Figure 8: Results of opening-closing de-haze network over O-HAZE dataset. The transmittance and airlight map along with ground truth is shown for qualitative evaluation.

Figure 9: Results of opening-closing De-haze network on real outdoor images. Transmittance and airlight map along with ground truth is shown for qualitative evaluation.

dataset along with groundtruth. The result of our network (fourth column) along with input image (first column) and the groundtruth (fifth column) are shown in Fig. 8. The second and third columns show estimated transmittance map and airlight, respectively. The average subjective quality over results of our method on this dataset given by the evaluators is 'reasonably good' and tends to be of 'low brightness'. It should be noted

that most of the state-of-the-art methods assume one of the two parameters (i.e., transmittance and airlight) is either known or can be determined heuristically, and the other (commonly transmittance) is estimated. On the other hand, the opening-closing network estimates both the parameters simultaneously from the input image and de-haze the same by applying the appropriate transformations.

Apart from the known benchmark datasets, we have also applied our morphological network on some real-world hazy images popularly used by the research community. Results on three images are shown in Fig. 9. We observe that our network is able to effectively de-haze the images and manages to reproduce the colors faithfully. However, in some instances, it was unable to reproduce lighter shades of some color. This can be attributed to the fact that we haven't used a highly optimized network for de-hazing images. In the real scene, the amount of haze is increased when depth is increased. The depth effect can be clearly seen in the transmittance map (second column (b)). Although the opening-closing network is able to de-haze the image, sometimes it fails and cannot obtain the haze-free image correctly. In Fig. 10 we have reported a few failure cases of opening-closing network. In the first row and last row in Fig. 10, it can be seen that the output image is darker. This may be because of the overestimation of the transmittance map. In the 2nd row, the estimated haze-free image has its table color green from white. From the estimated airlight and transmittance map, we can see that it considers the white table as haze.

Figure 10: In this figure we have shown few failure cases of opening-closing de-haze network. The predicted transmittance (b) and airlight (c) maps are shown along with output (d). The network is able to remove haze considerably but fails to preserve the color in the image(2nd row). Whereas in 1st and 3rd it removes the white haze but at the same time it makes the image darker.

5 Conclusions

In this paper, we have renewed the concept of learning morphological structuring elements (SEs) and shown that the SE of dilation and erosion can be learned. We have built a network that creates the effect of opening and closing with dilation and erosion layers placed sequentially in two different parallel paths and then by taking a linear combination of their outputs. To establish the efficacy of the proposed network, we have exploited it on two classes of image restoration problems, namely de-raining and de-hazing. Like CNN training methods, the proposed network learns SEs and various weights as parameters that lead to desired morphological network in terms of both operations and SEs. Once learned, the network then can estimate degrading parameters from the test input image and, subsequently, restore it. Although we get comparable results with state-of-the-art methods, we believe improving loss function can give better results. Second, to minimize the loss function, the structuring elements are learned by back-propagation algorithm, which is done by exploiting subgradient method as dilation/erosion operators are defined in terms of piece-wise differentiable max/min operators. Employing soft max/min may lead to fully differentiable loss function and may thereby simplify the back-propagation.

References

- [1] Codruta O Ancuti, 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*, pages 754–762, 2018.
- [2] Cosmin Ancuti, Codruta O Ancuti, and Christophe De Vleeschouwer. D-hazy: A dataset to evaluate quantitatively dehazing algorithms. In *2016 IEEE International Conference on Image Processing (ICIP)*, pages 2226–2230. IEEE, 2016.
- [3] Cosmin Ancuti, Codruta O Ancuti, Radu Timofte, Luc Van Gool, Lei Zhang, Ming-Hsuan Yang, Vishal M Patel, He Zhang, Vishwanath A Sindagi, Ruhao Zhao, et al. Ntire 2018 challenge on image dehazing: Methods and results. In *CVPR Workshops*, volume 1, 2018.
- [4] Jesus Angulo. A mathematical morphology approach to cell shape analysis. In *Progress in Industrial Mathematics at ECMI 2006*, pages 543–547. Springer, 2008.
- [5] Amy Bearman, Olga Russakovsky, Vittorio Ferrari, and Li Fei-Fei. What’s the point: Semantic segmentation with point supervision. In *European conference on computer vision*, pages 549–565. Springer, 2016.
- [6] D. Berman, T. Treibitz, and S. Avidan. Non-local Image Dehazing. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1674–1682, June 2016. 10.1109/CVPR.2016.185.
- [7] Alan C Bovik. Handbook of image and video processing. 2005.
- [8] Bolun Cai, Xiangmin Xu, Kui Jia, Chunmei Qing, and Dacheng Tao. Dehazenet: An end-to-end system for single image haze removal. *IEEE Transactions on Image Processing*, 25(11):5187–5198, 2016.
- [9] Nick Efford. *Digital image processing: a practical introduction using java (with CD-ROM)*. Addison-Wesley Longman Publishing Co., Inc., 2000.
- [10] Raanan Fattal. Dehazing using color-lines. *ACM transactions on graphics (TOG)*, 34(1):13, 2014.
- [11] Raanan Fattal. Dehazing Using Color-Lines. *ACM Trans. Graph.*, 34(1):13:1–13:14, December 2014. ISSN 0730-0301. 10.1145/2651362. URL <http://doi.acm.org/10.1145/2651362>.
- [12] Richard J Feehs and Gonzalo R Arce. Multidimensional morphological edge detection. In *Visual Communications and Image Processing II*, volume 845, pages 285–292. International Society for Optics and Photonics, 1987.
- [13] Gianni Franchi, Amin Fehri, and Angela Yao. Deep morphological networks. *Pattern Recognition*, page 107246, 2020.
- [14] Xueyang Fu, Jiabin Huang, Xinghao Ding, Yinghao Liao, and John Paisley. Clearing the skies: A deep network architecture for single-image rain removal. *IEEE TIP*, 26(6):2944–2956, 2017.
- [15] Xueyang Fu, Jiabin Huang, Delu Zeng, Yue Huang, Xinghao Ding, and John Paisley. Removing rain from single images via a deep detail network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3855–3863, 2017.
- [16] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 1440–1448, 2015.
- [17] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256, 2010.
- [18] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2010.

- [19] Kaiming He, Jian Sun, and Xiaoou Tang. Single Image Haze Removal Using Dark Channel Prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2341–2353, December 2011. ISSN 0162-8828. 10.1109/TPAMI.2010.168.
- [20] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017.
- [21] HJAM Heijmans, Michael Buckley, and Hugues Talbot. Path-based morphological openings. In *2004 International Conference on Image Processing, 2004. ICIP'04.*, volume 5, pages 3085–3088. IEEE, 2004.
- [22] Harald Koschmieder. Theorie der horizontalen sichtweite. *Beitrage zur Physik der freien Atmosphäre*, pages 33–53, 1924.
- [23] Boyi Li, 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*, pages 4770–4778, 2017.
- [24] Boyi Li, 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*, pages 4770–4778, 2017.
- [25] Yu Li, Robby T Tan, Xiaojie Guo, Jiangbo Lu, and Michael S Brown. Rain streak removal using layer priors. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2736–2744, 2016.
- [26] Yu Luo, Yong Xu, and Hui Ji. Removing rain from a single image via discriminative sparse coding. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3397–3405, 2015.
- [27] Gili Mendel. Optical character recognition using morphological attributes. 1993.
- [28] Gaofeng Meng, Ying Wang, Jiangyong Duan, Shiming Xiang, and Chunhong Pan. Efficient image dehazing with boundary constraint and contextual regularization. In *Proceedings of the IEEE international conference on computer vision*, pages 617–624, 2013.
- [29] Ranjan Mondal, Sanchayan Santra, and Bhabatosh Chanda. Image dehazing by joint estimation of transmittance and airlight using bi-directional consistency loss minimized fcn. In *CVPR Workshops*, pages 920–928, 2018.
- [30] Ranjan Mondal, Deepayan Chakraborty, and Bhabatosh Chanda. Learning 2d morphological network for old document image binarization. In *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pages 65–70. IEEE, 2019.
- [31] Ranjan Mondal, Pulak Purkait, Sanchayan Santra, and Bhabatosh Chanda. Morphological networks for image de-raining. In *International Conference on Discrete Geometry for Computer Imagery*, pages 262–275. Springer, 2019.
- [32] Ranjan Mondal, Sanchayan Santra, and Bhabatosh Chanda. Dense Morphological Network: An Universal Function Approximator. *arXiv e-prints*, art. arXiv:1901.00109, January 2019.
- [33] Susanta Mukhopadhyay and Bhabatosh Chanda. A multiscale morphological approach to local contrast enhancement. *Signal Processing*, 80(4):685–696, 2000.
- [34] Hyeonwoo Noh, Seunghoon Hong, and Bohyung Han. Learning deconvolution network for semantic segmentation. In *Proceedings of the IEEE international conference on computer vision*, pages 1520–1528, 2015.
- [35] Bryan A Plummer, Liwei Wang, Chris M Cervantes, Juan C Caicedo, Julia Hockenmaier, and Svetlana Lazebnik. Flickr30k entities: Collecting region-to-phrase correspondences for richer image-to-sentence models. In *Proceedings of the IEEE international conference on computer vision*, pages 2641–2649, 2015.
- [36] Kendall Preston Jr and Michael JB Duff. *Modern cellular automata: theory and applications*. Springer Science & Business Media, 2013.
- [37] Joseph Redmon and Ali Farhadi. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*, 2018.
- [38] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang. Single image dehazing via multi-scale convolutional neural networks. In *European conference on computer vision*, pages 154–169. Springer, 2016.
- [39] Wenqi Ren, Si Liu, Hua Zhang, Jinshan Pan, Xiaochun Cao, and Ming-Hsuan Yang. Single Image Dehazing via Multi-scale Convolutional Neural Networks. In *Computer Vision - ECCV 2016, Lecture Notes in Computer Science*, pages 154–169. Springer, Cham, October 2016. 10.1007/978-3-319-46475-6_10.
- [40] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015.
- [41] A Rosenfeld. Ac kak digital picture processing vol. 2, 1982.
- [42] Dan Schonfeld and John Goutsias. Optimal morphological pattern restoration from noisy binary images. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (1):14–29, 1991.
- [43] Stanley R Sternberg. Grayscale morphology. *Computer vision, graphics, and image processing*, 35(3):333–355, 1986.
- [44] Ketan Tang, Jianchao Yang, and Jue Wang. Investigating haze-relevant features in a learning framework for image dehazing. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2995–3000, 2014.
- [45] Luc Vincent. Morphological area openings and closings for grey-scale images. In *Shape in Picture*, pages 197–208. Springer, 1994.
- [46] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE TIP*, 13(4):600–612, 2004.
- [47] He Zhang and Vishal M Patel. Densely connected pyramid dehazing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3194–3203, 2018.
- [48] He Zhang, Vishwanath Sindagi, and Vishal M Patel. Multi-scale single image dehazing using perceptual pyramid deep network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 902–911, 2018.

- [49] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.
- [50] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE Transactions on Image Processing*, 27(9):4608–4622, 2018.
- [51] Qingsong Zhu, Jiaming Mai, and Ling Shao. A fast single image haze removal algorithm using color attenuation prior. *IEEE transactions on image processing*, 24(11):3522–3533, 2015.