Generalization of Double DIP: Unsupervised Image Decomposition

Deepanshu Parihar
Khoury college of Computer Sciences
Northeastern University
Boston, MA 02115
parihar.d@northeastern.edu

Mrinal Soni
Khoury college of Computer Sciences
Northeastern University
Boston, MA 02115
soni.mr@northeastern.edu

Ajeya Kempegowda
Khoury college of Computer Sciences
Northeastern University
Boston, MA 02115
kempegowda.a@northeastern.edu

Abstract

An image can be viewed as a mixture of "simpler" layers. Gandelsman et.al[1] proposed a unified framework for many seemingly unrelated computer vision tasks can be viewed as a special case of image decomposition into separate layers. For example, image segmentation; transparent layer separation; Image dehazing and more.Here we try to answer the question, Is the framework proposed capable of handling any input for all the image decomposition tasks that it claims to work on. We investigate the fidelity of the framework on different real-case applications - Dense haze images, Water mark removal on different variations of the watermark and explore how the choice of input noise, interpolation methods affects reconstruction accuracy and convergence to find if the framework proposed generalises well over different inputs.

1 Introduction

Deep convolutional neural networks (ConvNets) currently set the state-of-the-art in inverse image reconstruction problems such as denoising[3]. Generally, their excellent performance is imputed to their ability to learn realistic image priors from a large number of example images. Further, recent studies have shown that a generator network is sufficient to capture a great deal of low-level image statistics prior to any learning and show that a randomly-initialized neural network can be used as a handcrafted prior with excellent results in standard inverse problems such as denoising, superresolution, and inpainting[3, 4]. Further, generalization requires the structure of the network to "resonate" with the structure of the data. One such study[4] proposes a framework[1] that is capable of handling a variety of aforementioned image decomposition tasks and claims to achieve better results than some leading methods. Here we try to evaluate these claims to see if the framework generalises over a variety of images for specific image decomposition tasks and understand the limitations of the framework that shed light on areas of improvement.

Preprint. Under review.

2 Related Work

2.1 Dehazing

Haze represents one of the atmospheric phenomena most challenging for camera sensors and vision applications. The visibility of such hazy scene is highly degraded generating loss of contrast for the distant objects, selective attenuation of the light spectrum, and additional noise. For instance, the presence of haze has a great impact in the road traffic as it may severely reduce the visibility for drivers. As a result, restoring the contents in hazy images – process known as dehazing – is important for several outdoor image processing and computer vision applications such as visual surveillance and automatic driving assistance.

Typically, a hazy image *I*(x) is modeled[1][5] as :

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

where $A(x)$ is the Airlight map (A-map), $J(x)$ is the hazefree image, and $t(x)$ is the transmission (t-map), which exponentially decays with scene depth. The goal of image dehazing is to recover from a hazy image $I(x)$ its underlying haze-free image $J(x)$

2.2 Watermark Removal

Watermarking is the direct embedding of additional information into the original content or host signal.It is used mostly in digital media to hide proprietary information in digital media like photographs, digital music, or digital video. With the rise of the internet the sharing of information digitally is increasing and so is copyright infringements for the creators. In order to protect the interest of the content providers, these digital contents can be watermarked.A vast literature exists on digital watermarking [10,11]. This paper focuses on visible watermarks superimposed on images. There are various methods to attack these watermarks[1,12] and hence robustness is important. It can be kept in check by various types of inconsistencies or variations that can be introduced while embedding the watermark in each image. Randomly changing the position of the watermark, opacity or color,geometric variations like shadows,depth etc are some of them. The previous proposed method [12] is capable of handling most of these except for geometric variations. The framework proposed in [1] is able to handle such geometric variations.So the question that comes to mind is "Is the framework proposed in [1] sufficient to tackle any watermark?". We try to explore and find any fallacies that are present in the framework with respect to watermark removal and if not then what are the changes that need to take place so that watermarks are more robust.

2.3 Transparency separation

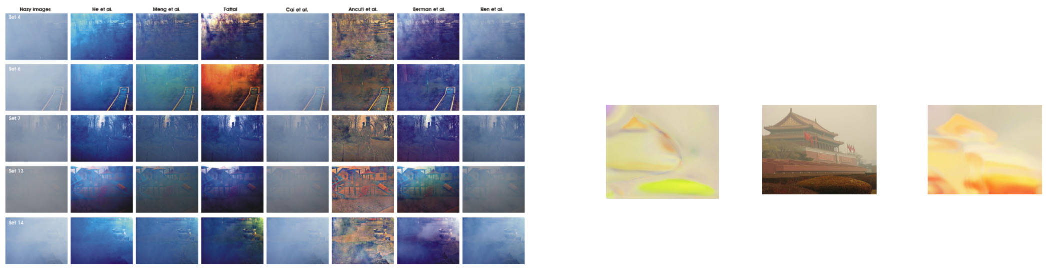
Transparency separation is a task in imaging that revolves around two or multiple images being superimposed on each other. In graphics, superimposition is the placement of an image or video on top of an already existing image or video, usually to add an overall image effect or to to conceal something (such as a different face being superimposed over the original image in a given photograph) or is the result of unwanted reflections (lights being reflected off a car's windshield either during bright sunny day or at night time).

In case of image reflection each pixel value in image I(x) is a convex combination of a pixel from the transmission layer y_1 (x) and the corresponding pixel in the reflection layer y_2(x).

3 Experiments

3.1 Dehazing

We've made use of the *code* provided by the authors of the double DIP paper. In this experiment we've applied dehazing technique on *DenseHaze dataset*[2]. The dataset provides dense haze images and haze free(ground truth) images that can be used for benchmarking the results across various dehazing techniques. The authors of DenseHaze[2] ran an experiment that compares the performance of dehazing techniques developed by He et al. [5], Meng et al. [6], Fattal [7], Cai et al. [8], Ancuti et



(a) Comparative results of representative single image dehazing techniques.

(b) Recovered airlight maps when the network was fed with Uniform noise(left) and Gaussian noise(right). The image in the center represent the input provided to dehaze

Figure 1: **Qualitative evaluation on Dehazing**

al. [9], Berman et al. [13], Ren et al. [14]. Further, we've compared these results to the technique developed by Gandelman et.al[1].

Finally, after the qualitative results obtained in Figure 1, we replaced the bilinear interpolation by bicubic in order to achieve better reconstructions. Also, we explored the choice of noise input the network by changing it from Uniform to Normal distribution.

3.2 Watermark Removal

There are two major categories of input that the framework needs to be tested on *Natural Images* and *Document images* as these are the ones that are mostly watermarked to prevent copyright infringement. Here we use the multi image solution of the framework that does not require the bounding box of the watermark but needs multiple images containing the same watermark.

Natural Images: The proposed framework is capable of handling geometric variations[1]. So we test our method for color and opacity. We take the images proposed in the framework and watermark it with an inclined transparent object. We superimpose the same original image with a more opaque inclined object. Then we run these as input over multiple values of iteration to check for the removal of watermark and quality of image recovered.

Document Images: We superimpose the images of documents with a different color of watermark object than that proposed in the original framework[1]. Then we run these as input over multiple values of iteration to check for the removal of watermark and quality of image recovered. The quality of image is a very important criteria for documents.

3.3 Transparency Separation

In this section we evaluate DoubleDIP framework on *Black and White Natural Images* as the authors do not address separation of monochromatic images. This requires a significant changes in the original code base to accommodate the color schemes of the input.

4 Results

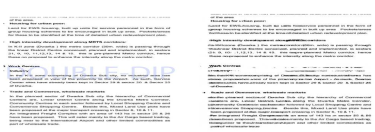
4.1 Dehazing

The first row in Table 1 shows the hazy images and the last row shows the ground truth. The middle rows, from left to right, present the results of He et al. [5], Meng et al. [6], Fattal [7], Cai et al. [8], Ancuti et al. [9], Berman et al. [13], Ren et al. [14] and Gandelman et.al[1] images. Therfore, 1 we infer that the double DIP framework produces better quantitative results by achieving higher PSNR values. However, the image quality is still severely degraded as shown in Figure 1.

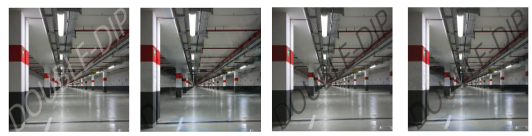
The choice of input noise provided evidence of unwanted artifacts in the recovered images, in particular Gaussian noise. The recovered airlight map can be visualised in Figure 2.

Table 1: **Quantitative evaluation**. In this table are presented 10 randomly picked up sets from our Dense-Haze dataset (the hazy images, ground truth and the results are shown in Fig.1). Using the haze-free (ground-truth) images we can compute the PSNR and values for the dehazed images produced by the evaluated techniques

Image	He et al.	Meng et al.	Fattal	Cai et al.	Ancuti et al.	Berman et al.	Ren et al.	Gandelman et.al
Set 4	14.12	14.57	11.10	10.08	14.39	12.55	11.39	40.869
Set 4	15.99	15.47	12.32	10.97	15.74	14.79	12.46	39.19
Set 4	15.65	16.97	14.59	13.95	18.48	15.29	16.45	38.68
Set 13	13.43	15.03	11.64	14.54	17.57	12.98	15.34	43.319
Set 14	15.06	14.67	13.11	9.58	11.18	12.62	11.78	41.892

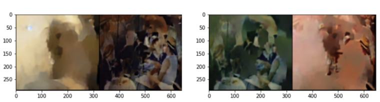


(a) Document watermarked images

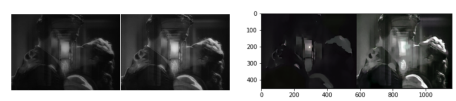


(b) Opaque and non opaque watermarks

Figure 2: **Qualitative evaluation on watermark removal**



(a) Image separation at 1000th iteration with Normal noise (left) and Uniform noise (right) as input



(b) Black and White images

Figure 3: **Qualitative evaluation on Transparency separation**

Further, changing the interpolation method to bicubic took a toll on convergence, where the execution time increased by five fold.

4.2 Watermark Removal

The framework seems unable to handle watermarks over a certain level of opacity even after optimizing the framework to run at multiple different hyperparameter iterations as seen in Figure 3. Further, the quality of image obtained is not very high with many parts of the image not being recovered properly. The image recovered after watermark removal needs to be processed further at least in the case of documents to prevent information loss.

4.3 Transparency separation

The framework doesn't seem to handle *Black and White Images* despite making the desired changes in the architecture (changing color channel from RGB to Binary), even after 12000 iterations which is a lot more than what the framework ideally requires for coloured images. While trying to separate *Black and White Images* the framework results in separating the lesser prominent image from the mixture of two images. But, when it comes to the darker (more prominent image) it produces a completely dark image which is nowhere near the actual image in the mixture. Finally, we infer that the left image (a) in Figure 4 has a better initial learning at 999th iteration produced by supplying Gaussian noise as opposed to uniform noise, inferring that noise type and noise intensity could be a factor in improving initial learning of the framework. Combining these with a changes in the noise statistics(variance) we believe it could lead to better convergence.

5 Conclusion

Our works provides evidence that the Double DIP framework doesn't readily generalize to handle all the variations that are currently proposed in the world of watermark embedding and dehazing. As revealed, by Dense-Haze dataset, the existing image dehazing techniques are not prepared to deal with dense hazy scenes and leaves significant room for improvement both qualitatively and quantitatively. Finally, the framework fails while separating superimposed monochromatic images and needs further improvement for non RGB images.n. We intend to investigate and improve more on this unified framework in the future.

Our work has been published in the GitHub repo.

References

[1] Y. Gandelsman, A. Shocher, and M. Irani. (2018) "Double-DIP" : Unsupervised Image Decomposition via Coupled Deep-Image-Priors. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* arXiv preprint arXiv:1812.00467, 2018

[2] C. O. Ancuti, C. Ancuti, M. Sbert and R. Timofte(2019), Dense-Haze: A Benchmark for Image Dehazing with Dense-Haze and Haze-Free Images, 2019. *IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019*, pp. 1014-1018.

[3] H. C. Burger, C. J. Schuler, and S. Harmeling. Image denoising: Can plain neural networks compete with bm3d? In Proc. CVPR, pages 2392–2399, 2012.

[4] Dmitry Ulyanov, Andrea Vedaldi, Victor Lempitsky;Deep image prior in the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 9446-9454

[5] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. IEEE Trans. Pattern Anal. Mach. Intell., 33(12):2341–2353, Dec. 2011

[6] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," In IEEE Int. Conf. on Computer Vision, 2013.

[7]Raanan Fattal, "Dehazing using color-lines," ACM Trans. on Graph., 2014.

[8] 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, 2016

[9]C. Ancuti, C. O. Ancuti, A.C. Bovik, and Christophe De Vleeschouwer, "Night time dehazing by fusion," IEEE ICIP, 2016.

[10] C. I. Podilchuk and E. J. Delp. Digital watermarking: algorithms and applications. Signal Processing Magazine, IEEE, 18(4):33–46, 2001

[11] V. M. Potdar, S. Han, and E. Chang. A survey of digital image watermarking techniques. In Industrial Informatics,2005. INDIN'05. 2005 3rd IEEE International Conference on, pages 709–716. IEEE, 2005.

[12] T. Dekel, M. Rubinstein, C. Liu, and W. T. Freeman. On the effectiveness of visible watermarks. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[13] D. Berman, T. Treibitz, and S. Avidan, "Non-local image dehazing," IEEE Intl. Conf. Comp. Vision, and Pattern Recog, 2016

[14] W. Ren, S. Liu, H. Zhang, X. Cao J. Pan, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," Proc. European Conf. Computer Vision, 2016.