Coupled Deep Image Prior Systems

Nathaniel M. Haddad*
Khoury College of Computer Sciences
Northeastern University 02115
Boston, MA
haddad.na@northeastern.edu

Abstract

It is well known that deep image priors are capable of learning low-level statistics of images and videos and can be used as a handcrafted prior for standard inverse problems. Using these statistics under the assumption that images are mixtures of multiple layers of sub-images, deep image priors are able to successfully decompose images into separate layers. When coupled together (Double-DIP), deep image priors become even more versatile and can be used for tasks such as image dehazing and image segmentation. In this paper, I show that coupled deep image priors can be linked together into systems that are able to produce iterative solutions to problems. I introduce two tasks to demonstrate that it is possible to create more complex systems of deep image priors via iterative methods and also by increasing the number of deep image priors in an existing Double-DIP model. I show that together, these systems can be used solve problems in which the resulting layers of image decomposition are many. Furthermore, this paper reinforces claims made about the inductive bias of deep image priors and their ability to generalize well to new problems.

1 Introduction

In computer vision, image layer decomposition refers to the problem of breaking a given image into separate layers. This assumes that a given image is a mixture of layers. In unsupervised image decomposition, a single image is used to train a generative network to output the separated layers from the input mixture. These architectures are a state-of-the-art alternative to deep convolutional neural networks that use large datasets for training in a number of tasks, including image super resolution, image in-painting, image denoising, image dehazing, image segmentation, and others.

It has been shown that deep image priors devised by Ulyanoc et al. are capable of learning low-level statistics about an image using the structure of the untrained encoder-decoder network (U-net) [1]. This allows deep image priors to successfully reconstruct an image as a mixture of layers as image generation problem using the structure of a randomly initialized deep convolutional neural network. The weights of the network are optimized to fit measurements of a given image.

Gandelsman et al. proposed coupling deep image priors together under a unified framework of layer-decomposition (the Double-DIP method) in which each separate layer has a stronger *internal self-similarity* than the original image [2]. They call this the problem of decomposing images into many separate layers *multi-task layer decomposition*, which can be used with Double-DIP to accomplish a number of computer vision problems (image segmentation, image dehazing, transparency separation, and watermark removal) in images and video [2].

I show that it is possible to solve multi-task layer decomposition problems using systems of deep image priors. The Double-DIP solutions for these problems may be expanded to larger systems of

*Code available at https://github.com/nmhaddad/python-deep-learning-deep-image-prior-systems

(a) original. (b) two-layer segmentation. (c) three-layer segmentation.

Figure 1: Comparison of image segmentation methods. a) original image; b) two-layer segmentation result using Double-DIP method; c) three-layer segmentation result using Double-DIP system.

deep image priors that together solve a problem simultaneously. Under the assumption that images are mixtures of many layers and that deep image priors may be coupled together to perform multi-task layer decomposition, I show that coupled deep image priors can be stacked on top of each other to create iterative solutions. This proves useful in tasks described in Ulyanoc et al. and Gandelsman et al. including image segmentation and image dehazing. I show that three-layer image segmentation is possible using Double-DIP and introduce an image dehazing method called *iterative image dehazing* which approaches the problem of recovering a haze-free image from a hazy/foggy image via an iterative solution derived from a system of layers recovered from an image.

2 Image segmentation systems

In previous work with coupled deep image priors (Double-DIP), it was shown that coupled deep image priors are able to successfully separate foreground and background layers from an input image [2]. Given a *hint*, each deep image prior is able to generate a texture from the image, guided by the provided hint corresponding to a specific part of the image (e.g. foreground hint, background hint). These layers are then multiplied by a binary mask and then combined to produce a learned image. In this section, I show that it is possible to extract a third layer from an input image using three Double-DIP models and I propose adding a third deep image prior net to the Double-DIP model to generate a third layer for future work. For this experiment, I chose to identify a *middle-ground* texture, or the part of an image not included in the foreground or the background layers.

Double-DIP image segmentation as proposed by Gandelsman et al. describes decomposing an image I into a foreground layer, a background layer, and a binary mask $m(x)$ such that input image

$$I(x) = m(x)y_1(x) + (1 - m(x))y_2(x).$$

By adding a third layer (such as middle-ground), we add $y_3(x)$ to our equation. Doing so changes the binomial distribution of the binary mask to a multinomial distribution. Therefore

$$I(x) = \sum_{i=1}^{n} y_i(x)m_i(x).$$

where n is the number of layers to be segmented from the image.

Hand engineering Double-DIP showed that while it is possible to segment images into foreground and background, a heavy amount of hand engineering is required [2]. In Gandelsman et al., Double-DIP models for image segmentation used a foreground and background hint: these hints are masks generated by manipulating the results of image histogram equalizations. In order to create a three-layer segmentation, I show that a third hint may generated using this same technique to estimate what the third layer is expected to be in the input image.

Three-Layer Segmentation There are two techniques using this third hint that become possible for three-layer segmentation. The first technique I present is the simplest: a system of Double-DIPs that uses the results of the segmentation process described in Gandelsman et al. [2]. The Double-DIP process is run three times consecutively: *1) with background and foreground hints; 2) with*

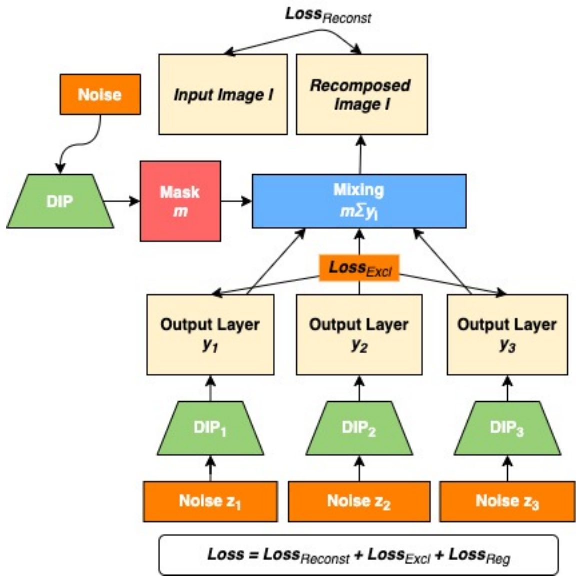

Figure 2: **Three-layer-segmentation system.** Using the Double-DIP framework, three deep image priors decompose an input image into separate layers y_1, y_2, y_3. Mixing reconstructs image $I(x)$ using a learned mask and multinomial distribution of mixture.

background and middle-ground hints; 3) with foreground middle-ground hints. The resulting images can then be used identify the true middle-ground layer by

$$I(x) = y_1(x)m(x) + y_2(x)m(x) + y_3(x)m(x).$$

This follows the proposed multinomial distribution described above.

In the second technique for three-layer image segmentation, I propose expanding the existing Double-DIP model to estimate foreground, background and middle-ground layers for the purposes of image segmentation. I introduce a third deep image prior for the third-layer so that each layer of an image (foreground, middle-ground, background) can be generated. I then propose changing the mask from a binomial distribution to a multinomial distribution. While I was able to show that it is possible to generate a three-layer segmentation problem using Double-DIP systems, I was not able to produce great results with this method. I feel that with a team and adequate resources and time, I would be able to show good results using this method.

3 Iterative image dehazing

Multi-task layer decomposition via Double-DIP has been shown to be successful in image dehazing tasks. Image dehazing is a problem that seeks to remove particle scattering from degraded images, reducing or removing elements such as fog, pollution, or underwater particulate and lighting phenomena [3, 5]. An image dehazing task can be viewed as an image decomposition problem where a hazy/foggy image $I(x)$ is a mixture of separate layers of images, where image

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

and $A(x)$ is the Airlight map (A-map), $J(x)$ is the haze-free image, and $t(x)$ is the transmission (t-map), which decays with scene depth, as images grow more degraded with depth [2]. The goal of a image dehazing problem is to recover from a degraded image $I(x)$ a haze-free image,

$$y_1(x) = J(x).$$

Using the image dehazing method described above, I will show that it is possible to use an iterative technique using deep image priors to achieve a significantly dehazed image $J(x)$. Double-DIP image

(a) original. (b) single dehaze. (c) double dehaze.

Figure 3: Comparison of image dehazing methods. a) original image; b) dehazed image using Double-DIP method and 4000 iterations; c) dehazed image using iterative image dehazing with 1000 and 3000 iterations.

dehazing assumes that a degraded image is a mixture composed of multiple layers where output $J(x)$ is a haze-free image. *iterative image dehazing* assumes that output $J(x)$ itself is a mixture of multiple layers of separate images. Therefore, it can be assumed that the haze-free image $J(x)$ can be decomposed into separate layers. This is an iterative process and requires that multiple problems be solved simultaneously such that

$$I(x_i) = t(x_i)J(x_i) + (1 - t(x_i))A(x_i),$$

where i is the *ith* image dehazing iteration. Together, this system of coupled deep image priors produces image dehazing solutions that result in haze-free image

$$y_1(x_N) = J(x_N)$$

where N is the maximum number of iterations. This allows iterations of image decomposition to occur consecutively as part of the solution for a single task, such as an extremely degraded image where a single image dehazing operation is insufficient for removing scattering phenomena. This is significant as it shows that it is possible to dehaze an image multiple times and achieve a better resulting haze-free image with each iteration.

Compared to the Double-DIP image dehazing solution, my proposed iterative image dehazing solution is capable of generating stronger haze-free results. However, while scattering is reduced, the resulting output image becomes more and more distorted with each iteration due to a loss of the original signal with each iteration.

Adjusting the number of iterations Double-DIP image dehazing assumes a uniform number of iterations for convergence of each deep image prior across multiple image dehazing iterations. However, I found that with each image dehazing operation, the number of iterations needed to produce an ideal output image increases by some factor. This is a result of the output image increasing in complexity, possibly due to the upsampling process. Further work is needed to show what the ideal iteration scheduling should be and why these images are growing. This proves troublesome as finding such a factor becomes computationally expensive as the factor and number of iterations grow in size.

4 Conclusion

In this paper, I showed that it is possible to increase the number of deep image priors used for a solution to a problem. This highlights the inductive bias captured by the structure of these unsupervised networks and their ability to generalize to new tasks. I presented two techniques for achieving three-layer image segmentation using a system of Double-DIPs and by increasing the number of deep image priors in the Double-DIP model. This is interesting work as it shows that with hints, unsupervised image decomposition can be used to separate specific parts of images. This problem is useful for segmenting images for which no trained model exists or not enough data exists to segment an image. Further work may address the possibility of increasing the number of segmented layers, improving the quality of the implementation, removing or automating hints from the process, and research in semantic segmentation using deep image priors.

I also show that Double-DIP image dehazing can be used to iteratively remove scattering mediums from degraded images through a proposed technique called iterative image dehazing. I also would like to highlight preliminary research I conducted that shows that both Double-DIP and iterative image dehazing can be used to remove underwater scattering in images and videos where objects are occluded by particulate matter, lighting, and other phenomena. Future work should consider implementing image dehazing techniques designed specifically to remove layers specific to underwater scattering problems. Future work should also consider the use of Double-DIP image dehazing in video tasks as well. Collections of frames can take advantage of blind dehazing, a technique used to recover unknown hazing parameters from an image [3]. Gandelsman et al. utilized this technique to estimate the haze Airlight color for image dehazing tasks. In iterative image dehazing, this estimation can be passed along with each iteration as a prior. Since video frames are strongly correlated, this value should improve the output of video dehazing problems.

Acknowledgments

I would like to acknowledge the authors of *Deep Image Prior* and *"Double-DIP": Unsupervised Image Decomposition via Coupled Deep-Image-Priors*. The code submitted with this project was built on top of code provided by the authors of the Double-DIP paper. I would also like to thank Paul Hand for providing direction and feedback on this project.

References

[1] Ulyanov, D. & Vedaldi, A. & Lempitsky, V. (2018) Deep image prior. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9446–9454.

[2] Gandelsman, Y. & Schocher, A. & Irani, M. (2018) "Double-DIP": Unsupervised Image Decomposition via Coupled Deep-Image-Priors. *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11026–11035.

[3] Bahat, M. & Irani, M. (2016) Blind dehazing using internal patch recurrence. *2016 IEEE International Conference on Computational Photography (ICCP)*, pp. 1-9.