

Single-Shot Retinal Image Enhancement Using Untrained and Pretrained Neural Networks Priors Integrated with Analytical Image Priors

Adnan Qayyum¹, Waqas Sultani¹, Fahad Shamshad¹, Rashid Tufail², and Junaid Qadir¹

¹ Information Technology University of the Punjab, Lahore, Pakistan

² Wajahat Surgical Hospital, Attock, Pakistan

Abstract—Retinal images acquired using fundus cameras are often visually blurred due to imperfect imaging conditions, refractive medium turbidity, and motion blur. In addition, ocular diseases such as the presence of cataract also result in blurred retinal images. The presence of blur in retinal fundus images reduces the effectiveness of the diagnosis process of an expert ophthalmologist or a computer-aided detection/diagnosis system. In this paper, we put forward a single-shot deep image prior (DIP)-based approach for retinal image enhancement. Unlike typical deep learning-based approaches, our method does not require any training data. Instead, our DIP-based method can learn the underlying image prior while using a single degraded image. To perform retinal image enhancement, we frame it as a layer decomposition problem and investigate the use of two well-known analytical priors, i.e., dark channel prior (DCP) and bright channel prior (BCP) for atmospheric light estimation. We show that both the untrained neural networks and the pretrained neural networks can be used to generate an enhanced image while using only a single degraded image. We evaluate our proposed framework quantitatively on five datasets using three widely used metrics and complement that with a subjective qualitative assessment of the enhancement by two expert ophthalmologists. We have compared our method with a recent state-of-the-art method cofe-Net using synthetically degraded retinal fundus images and show that our method outperforms the state-of-the-art method and provides a gain of 1.23 and 1.4 in average PSNR and SSIM respectively. Our method also outperforms other works proposed in the literature, which have evaluated their performance on non-public proprietary datasets, on the basis of the reported results.

Index Terms—Retinal image enhancement, Retinal image generation, Single image analysis

I. INTRODUCTION

In ophthalmic clinical practice, retinal images are routinely acquired using fundus photography, which is used in the diagnosis and treatment of different retinal diseases such as diabetic retinopathy, hypertensive retinopathy, and age-related muscular degeneration [1] [2]. In addition, these images have been used for developing different computer-aided detection/diagnosis systems for glaucoma detection [3], diabetic retinopathy classification [4], retinal arteries and veins classification [5], and blood vessel segmentation [6]. However, acquired retinal fundus images often contain blurriness due to different reasons such as dusty camera lenses, low-resolution camera, imperfect illumination, refractive medium turbidity, and incorrect focus [7]. In addition, the presence of cataract

(an eye disease due to which natural eye lens becomes foggy) also results in hazy and blurred retinal images [8], where the blur/haze intensity increases with the cataract severity [9]. An event of an eye blink and occlusion due to eyelashes can also result in a blurry retinal image. The presence of blur significantly affects the sensitivity of the diagnosis process (either performed by a human expert or computer-aided system), particularly, for progressive ophthalmological diseases as it is difficult to analyze vascular structure in visually blurred images [2]. A study focused on the automatic analysis of morphometric properties of vasculature in retinal images revealed that more than 25% images were not suitable for automatic analysis due to their bad quality [10].

In the literature, different approaches have been proposed for retinal fundus image enhancement ranging from traditional image processing methods like image transformation [2]; contrast adjustment [11] and normalization [12]; to deep learning (DL) based methods [13]. DL-based methods operate in a supervised learning fashion and are typically data-driven. These methods require a substantial amount of training data, which is scarcely available in practice, as obtaining high-quality representative medical data is very difficult, time-consuming, and expensive. In fact, it is documented that a major challenge in retinal fundus image enhancement research is the unavailability of an application-specific dataset in which the blur/hazy images have their corresponding reference as ground truth (that can be used for evaluating the performance of enhancement methods) [14].

Various databases have been used in the retinal fundus image enhancement literature such as DRIMDB [15], HRF [16], and DR2 [17]. However, these datasets suffer from different issues, e.g., the images in these databases are assigned binary labels, i.e., either accept or reject. In a recent study [18], the authors have attempted to address these challenges by manually annotating a sample of a large collection of retinal fundus images that have been acquired for diabetic retinopathy grading (i.e., EyePACS dataset). The authors considered three-class annotations, i.e., good, usable, and reject. Furthermore, they used the labeled images to train a DL model to evaluate the quality of the retinal fundus images. However, the aforementioned datasets do not contain paired images (i.e., the blurred image and its corresponding reference (clean) image). Furthermore, the acquisition of a good quality

reference image is another major challenge. The exact pixel-to-pixel tight reference image cannot be realistically acquired in clinical settings. This is because to acquire reference (clean) images, we have to take the retinal fundus image of the patients after cataract surgery. However, it is quite challenging due to factors such as varied refractive properties of the replaced lenses, possibly changed field of view, and the position of the eye. We note that typically the datasets used for fundus image enhancement are the ones that have been developed for diabetic retinopathy or retinal blood vessels segmentation. These datasets lack ground truth reference images, which hinder their use for full-reference performance evaluation. of the image enhancement.

In this paper, unlike data-driven DL-based methods, we present a new approach for the dehazing of retinal fundus images that does not require any training data. Our approach uses only a single degraded image for recovering a true estimate of a clean image without requiring a reference image. Our approach is inspired by the recent successes of untrained neural network priors (UNNP) [19], [20], which uses a single degraded image to solve different inverse imaging problems like denoising, deblurring, and inpainting [21]. We formulate the problem of retinal fundus image enhancement as a highly ill-posed inverse problem in which we aim to find a clean retinal image from a degraded one without having any prior knowledge of the clean image. In this paper, we build upon our previous work [22], in which we demonstrated that the structure of a convolutional neural network (CNN) can be used as a regularizer to solve such inverse problem. This approach was shown to be quite effective because it does not require paired data (hazy and clean image) to train the DL models. Also, since our approach does not utilize any training data, it does not suffer from the problem of distribution shifts and provides good generalization. With reference to our previous work [22], the following are the specific extensions made in this paper:

- 1) We present the use of analytical image priors for atmospheric light estimation integrated with coupled deep image priors (CDIPs) networks for retinal fundus image enhancement. In contrast to our previous work [22], which used three DIPs with a separate DIP employed for atmospheric light estimation, our current framework only utilized two DIP networks and is therefore, computationally less expensive.
- 2) We show the effectiveness of using pre-training for our single-shot DL approach. Also, we demonstrate the effect of pre-training for cross dataset analysis thorough ablation studies. Pretraining achieves comparable performance 3-5 times faster than using untrained CDIPs.
- 3) We have modified our experimental setup and have employed an early stopping strategy to avoid overfitting.
- 4) We conduct an extensive performance evaluation of the proposed framework on five different datasets (both quantitative and qualitative) and report promising results. Extending our previous work [22], we have also used a no-reference metric for quantitative analysis (i.e., BRISQUE). Moreover, to highlight the clinical significance, we have performed a subjective evaluation done

by two expert ophthalmologists.

- 5) We perform computational complexity analysis of our method and compare it with previous works.

Paper Organization: The rest of the paper is organized as follows. The related work is presented in Section II. The proposed methodology is presented in Section III. The experimental setup and results are described in Section IV. Limitations and promising directions for future work are identified in Section V. Finally, we conclude the paper in Section VI.

II. RELATED WORK

In recent years, different DL-based methods have been proposed for retinal fundus image enhancement. In [23], the authors proposed to integrate a shadow removal layer with a U-Net model for dehazing of retinal fundus images. The proposed framework is trained in two steps, i.e., U-Net is trained first, and then Inception-v3 model is fine-tuned using the learned parameters of U-Net. The U-Net model learns to estimate the transmission map which results in minimum classification error. The proposed framework was evaluated for the diabetic retinopathy classification task, however, no quantitative metrics were used for the performance evaluation of retinal image enhancement.

Zhao et al. [13] proposed a generative adversarial network (GAN) based retinal image enhancement method that works in a weakly supervised fashion, i.e., it uses unpaired clean and blurry retinal images. Two GANs were trained for deblurring of retinal fundus images using a training set of 949 images with $4\times$ data augmentation, i.e., a total of 3796 images were used for the training. In a similar study, You et al. [24] proposed to integrate a convolutional block attention module (CBAM) with a CycleGAN for enhancement of retinal fundus images. The method proposed in [24] was trained using unpaired clean and blurry retinal images, as it is quite challenging to acquire strictly clean and blurry paired retinal images for training a GAN. The proposed method was evaluated qualitatively and quantitatively in terms of average peak signal to noise ratio (PSNR) and structural similarity (SSIM). In [18], the authors first annotated a sample of a large collection of retinal images for image quality assessment purposes (they consider three classes, i.e., Good, Usable, and Reject). The data was originally annotated for the diabetic retinopathy classification task and then they employed a data-driven DL model to automatically assess the quality of retinal fundus images. In [25], authors first developed a synthetic model for introducing visual artifacts in retinal fundus images and then proposed a DL model to surpass these artifacts. In our previous work [22], we present a single-shot unsupervised three DUP-based framework and incorporated DCP loss into the overall loss. A comprehensive survey focused on the retinal fundus image quality enhancement is presented in [14].

III. METHODOLOGY

The key idea of the proposed method is image decomposition, where we decomposed the input degraded image into individual components and then we use the image formation

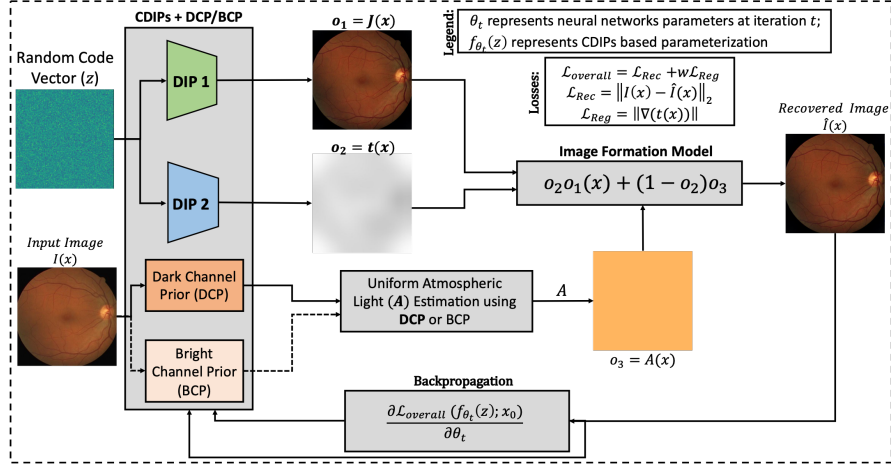


Fig. 1. Untrained coupled deep image priors (CDIPs) integrated with analytical priors. The input image $I(x)$ is decomposed into three layers $o_1 = J(x)$ (the enhanced image by DIP_1), $o_2 = t(x)$ (the transmission map by DIP_2), and $o_3 = A$ the uniform atmospheric light (estimated using either DCP or BCP). In our previous work [22], we employed a separate DIP network for estimating non uniform $A(x)$. To reconstruct the recovered image all individual components are mixed using image information model. The figure depicts the actual results at last iteration.

model to get the enhanced image. In our proposed framework we employ multiple DIP networks to decompose the input image into its basic components. Unlike conventional DL, our method does not require any training data (paired or unpaired) and works by using only a single degraded image. This makes it quite useful for realistic medical applications, where paired data is usually not available and acquisition and annotation are costly. The proposed method is described below.

A. Image Decomposition using CDIPs

CDIPs leverages the well-known concept of image decomposition in computer vision, where the goal is to decompose an image into its basic layers. For instance, in image segmentation, the objective is to decompose the image into foreground and background layers; and in image dehazing, we are interested in decomposing the hazy image into a clear image and a haze map.

1) *Illumination Compensation via CDIPs*: In the literature, the illumination compensation problem is widely viewed as an haze removal problem. The model to describe a hazy image $I(x)$ is given by [26] as the following:

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

where $t(x)$, $J(x)$ and A represents transmission map (t-map), restored clean image (haze free), and atmospheric light, respectively. As described above, we formulated our problem as layer decomposition problem in which we aim to decompose a blurred/hazy image $I(x)$ into its aforementioned three layers, i.e., clean image ($o_1(x) = J(x)$), transmission map ($o_2 = t(x)$), and atmospheric light ($o_3 = A$), as shown in Fig. 1. Following our formulation, the Eq. 1 can be expressed as

$$\hat{I}(x) = o_2 o_1(x) + (1 - o_2) o_3, \quad (2)$$

The pipeline of the proposed CDIPs framework integrated with conventional priors is depicted in Fig. 1. As shown in the figure, DIP_1 and DIP_2 networks take randomly sampled

uniform noise vector z , respectively and attempts to generate different layers $o_1(x) = J(x)$ and $o_2 = t(x)$ using the input image $I(x)$, respectively. As depicted in the Fig. 1, the atmospheric light is either estimated using dark channel prior (DCP) or bright channel prior (BCP), which have been widely used in haze removal problem (we will discuss the respective details in the coming section). The outputs of two DIP networks are mixed with the output of analytical prior using Eq. 2. The loss function optimized by the proposed CDIPs framework is given as:

$$\mathcal{L}_{overall} = \mathcal{L}_{Rec} + w\mathcal{L}_{Reg}, \quad (3)$$

where \mathcal{L}_{Rec} denotes the reconstruction loss i.e., $\|I - \hat{I}\|_2$, \mathcal{L}_{Reg} is regularization loss of DIP_2 that is defined as the norm of Laplacian which is minimized to enforce the estimated mask (o_2) to be smooth and finally, $\mathcal{L}_{overall}$ represents the overall loss of CDIPs architecture.

B. Atmospheric Light Estimation

Atmospheric light is another component of an image that needs to be estimated along with the transmission map $t(x)$ and it is used in a hazy image formation model to get the recovered/enhanced image (Eq. 1). In our previous work, we have estimated non-uniform atmospheric light as a separate layer using a third DIP network, while in this paper, we show that uniformly estimated atmospheric light also provides comparable results, despite being computationally less expensive that reduces time and space utilization. We assume that there is a uniform scattering of light that can be easily estimated using analytical priors such as dark channel prior (DCP) and bright channel prior (BCP). In the literature, most single image dehazing methods are based on this assumption [26]–[29]. In [23], authors assume $A = 1$ for retinal fundus image enhancement.

1) *Atmospheric Light Estimation Using DCP*: In contrast to our previous work [22], where we incorporated DCP loss into the overall loss of the CDIPs framework, in this paper, we

use DCP to estimate the atmospheric light. In the literature, DCP has been widely used in single-image dehazing of natural images [26], [30]. It is based on the observation that in most of the local patches of haze-free color images, pixels have very low intensity in one of the color channels, which is called the dark channel. The dark channel prior for an image $I(x)$ can be computed as:

$$I^{Dark}(x) = \min_{C \in \{R, G, B\}} \left(\min_{y \in \Omega(x)} (I^C(y)) \right), \quad (4)$$

where $I^C(y)$ is the color channel of image $I(x)$ and $\Omega(x)$ is local patch centered at x . In our previous work [22], we incorporated DCP loss into optimization loss of the CDIPs framework which is given as $\mathcal{L}_{DC} = \sum_{j=1}^{W \times H} |dark(\hat{I}(x))_i|$. The computation of DCP loss is computationally very expensive [31], [32], as it involves the computation of patch-based DCP for the generated image at each iteration. To overcome this issue, we have not incorporated DCP loss into the overall loss of the CDIPs framework (as we did in our previous work), instead, we have used DCP directly for estimating the uniform atmospheric light. Uniform atmospheric light is estimated by selecting the top 0.1% darkest pixels from DCP as candidate pixels. The intensity of dark pixels is mainly contributed by atmospheric light in the haze images, therefore, these candidate pixels can directly provide an accurate estimate of haze transmission. Unlike, natural outdoor images that have sky regions, retinal fundus images do not have correlated patches as that of the estimated transmission map, which makes it feasible to recover haze-free retinal images.

2) *Atmospheric Light Estimation using BCP*: Unlike DCP, BCP is based on the observation that in most of the local patches of haze-free color images, pixels have very high intensity in one of the color channels, which is called the bright channel. BCP has been widely used in single image dehazing literature [33]. The bright channel for an image $I(x)$ is computed as:

$$I^{Bright}(x) = \max_{y \in \Omega(x)} \left(\max_{C \in \{R, G, B\}} (I^C(y)) \right), \quad (5)$$

where $I^C(y)$ is the color channel of image $I(x)$ and $\Omega(x)$ is local patch centered at x . Similar to DCP, uniform atmospheric light is estimated by selecting top 0.1% brightest pixels from BCP as candidate pixels. After that, the maximum intensity value from candidate pixels in the hazy image is given as atmospheric light [30]. BCP have been shown quite effective for atmospheric light estimation [34] for the application of retinal image enhancement because the color intensities of retinal structures (such as vessels, optic disc, etc.) are inherently different from estimated atmospheric light (therefore, BCP prominently highlights such regions).

C. Untrained and Pretrained CDIPs

To capture useful information about the images, the parameters of DL model are tuned/learned from the training data. However, in the literature [35], it has been also shown that pretrained neural networks contain a significant amount of information about the task at hand, e.g., dehazing in our case. In this paper, we leverage the idea of pretraining and

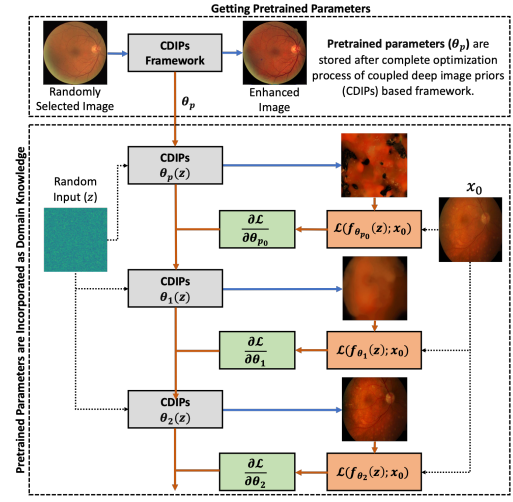


Fig. 2. Proposed CDIPs framework **with pretraining**. Legend: z : random code vector; θ_{p0} : pretrained weights; $f_{\theta_{p0}}(z)$: CDIPs based parameterization; \mathcal{L} : overall loss; x_0 : input degraded image;

investigate its effect in CDIPs application for retinal image enhancement. We show that the knowledge learned from one retinal fundus image can be used to effectively model other retinal fundus images in comparatively less amount of time while getting a comparable performance.

In Fig. 2, we demonstrate that how pretraining works for our problem using CDIPs. Starting from the pretrained parameters θ_{p0} loaded into CDIPs framework and randomly initialized input code vector z , we attempt to model probability distribution $p(x|x_0)$ given that x is unknown clean image and x_0 is the degraded (hazy) image. The output at the first iteration contains randomization due to random code vector z and the optimizer iteratively optimizes neural network's parameters θ for the given input image using backpropagation. It has been shown in the literature [19], [36] that the choice of neural network architectures has a direct impact on the performance, e.g., we can design/handcraft a particular neural network architecture for modeling a specific image [21]. This serves as a solution space when modeling images using untrained neural networks priors. We found that the optimization process tends to destabilize for few images (evident from the high standard deviation (SD) in iterations reported in Table II that indicates that the optimization was sometimes stopped for a few images at very fewer iterations. This phenomenon highlights that the pretrained neural networks are not best for those particular images, otherwise pretraining works pretty well for most of the images. To overcome this issue, we rely on the early stopping.

IV. EXPERIMENTS AND RESULTS

A. Setup and Dataset Description

We perform extensive experiments on the five publicly available fundus image datasets that have different number of images and dimensions—namely, (i) DRIVE [37], (ii) STARE [38], (iii) Messidor [39], and (iv and v) DIARET DB calibration level 0 and 1 [40]. Unlike our previous work, we applied some pre-processing steps before feeding the input to the proposed framework. Firstly, all images from each dataset

TABLE I
PERFORMANCE EVALUATION FOR **CDIPs FRAMEWORK WITHOUT PRETRAINING** INTEGRATED WITH DCP AND BCP IN TERMS OF THE AVERAGE PSNR, SSIM, AND BRISQUE (NO REFERENCE). HIGHER PSNR AND SSIM ARE DESIRABLE WHILE LOWER BRISQUE IS DESIRABLE.

| Dataset | #Images | CDIPs + DCP | | | | | | | | CDIPs + BCP | | | | | | | |
|------------|---------|-------------|-----|----------------|------|---------------|-------|----------------|-------|-----------------|-----|----------------|------|---------------|-------|----------------|-------|
| | | #Iterations | | PSNR | | SSIM | | BRISQUE | | Avg. Iterations | | PSNR | | SSIM | | BRISQUE | |
| | | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Messidor | 1200 | 2976 | 149 | 42.6261 | 2.27 | 0.9617 | 0.01 | 39.7356 | 12.74 | 3250 | 507 | 39.6052 | 1.85 | 0.978 | 0.01 | 42.5132 | 14.38 |
| STARE | 397 | 3469 | 352 | 38.9792 | 3.01 | 0.9708 | 0.02 | 35.3496 | 18.37 | 3443 | 99 | 39.9813 | 1.94 | 0.9792 | 0.01 | 34.6286 | 17.02 |
| DRIVE | 40 | 3447 | 48 | 38.6594 | 0.59 | 0.9757 | 0.002 | 19.2102 | 7.52 | 3451 | 42 | 38.2847 | 1.04 | 0.9750 | 0.003 | 23.7266 | 6.95 |
| DIARET DB0 | 130 | 3452 | 41 | 41.0517 | 1.31 | 0.9803 | 0.01 | 37.4146 | 13.9 | 3459 | 38 | 41.3724 | 1.37 | 0.9816 | 0.004 | 37.3445 | 13.27 |
| DIARET DB1 | 89 | 3464 | 47 | 40.734 | 1.28 | 0.9709 | 0.02 | 39.2502 | 14.44 | 3453 | 42 | 41.8525 | 1.44 | 0.9829 | 0.004 | 40.6751 | 13.79 |

are centered cropped such that the field-of-view (FoV) is preserved. Secondly, each image was resized to a standard size of 512×512 . We note here that DRIVE and STARE are the benchmark databases for *retinal blood vessel segmentation* while Messidor and DIARET DB (0 and 1) is the benchmark databases for *diabetic retinopathy classification*. We quantitatively evaluate the proposed framework using three widely known metrics, i.e., PSNR, SSIM, and Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE). PSNR is a widely used metric and it measures the similarity between the two images. Higher PSNR represents better quality of the image. SSIM is a perception-based image quality metric that is used to measure the perceived change in structural information between two images. SSIM exploits inter-dependencies among spatially close pixels and a higher SSIM value is desirable. BRISQUE is a no-reference image quality metric that scores the perceptual quality of an input image based on its naturalness, where the lower value of BRISQUE is desirable. Finally, to highlight the clinical significance, we also perform the subjective qualitative assessment of the proposed method through two expert ophthalmologists.

B. Implementation Details

We use the U-Net as the generator network for generating the enhanced retinal fundus images [41]. U-Net is a deep CNN-based generative model that was specifically developed for biomedical image segmentation [41]. It has an hourglass-like architecture and skip-connections between encoder layers and decoder layers. The same model architecture was used in our previous work [22]. Each DIP uses one U-Net and overall framework is optimized using ADAM optimizer [42] with a learning rate of 0.004 and default values of $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 1e^{-8}$. We perform different ablation studies using untrained U-Net and pretrained U-Net. In untrained experiments, all images from each dataset are processed for 3500 iterations and in pretraining experiments, every image is allowed to be processed for maximum 2500 iterations. Furthermore, we employ an early stopping strategy to avoid overfitting for all experiments. The number of trainable parameters in the proposed framework are approximately 1.15M.

C. Using Untrained CDIPs with DCP and BCP

As described in the earlier sections, we used analytical priors for estimating uniform atmospheric light in the retinal fundus image. In this section, we present the results of using DCP and BCP for atmospheric light estimation. We also report the standard deviation in addition to average values

to provide an exact idea about the statistical distribution of achieved performance across different datasets which contain a different number of images. The comparative results of using CDIPs (i.e., one for getting recovered image and the other for estimating transmission map) and using DCP and BCP for atmospheric light estimation are presented in Table I. It can be observed from the table that the average performance of using BCP with CDIPs is comparatively higher than using DCP with CDIPs. This indicates that most of the fundus images have brighter pixels that are efficiently leveraged by the BCP. This trend can be seen for almost every dataset in terms of all metrics used except for the DRIVE dataset where DCP performance is higher, probably due to low intensity in images of DRIVE. Similarly, it can be seen that the performance of CDIPs coupled with analytical priors (DCP and BCP) is relatively low for blood vessel segmentation databases (DRIVE and STARE) as compared to diabetic retinopathy classification datasets (Messidor, DIARET DB 0 and 1). This is due to the fact that the image quality of segmentation databases is relatively higher (as they are purposely developed for retinal blood vessels segmentation) than diabetic retinopathy classification datasets.

D. Using Pretrained CDIPs with DCP and BCP

In this section, we highlight the benefit of pretraining for our proposed CDIPs framework. Pretraining works in two steps: (i) we randomly choose one image from each dataset and fit with our proposed CDIPs framework and upon its completion, we store the learned/optimized parameters of both DIP networks as pretrained parameters (this process is performed once for each dataset); and then (ii) we load these saved parameters when fitting other images. The results for pretrained CDIPs coupled with DCP and BCP are presented in Table II, which highlights that pretraining provides comparable performance for all datasets when modeled with the BCP-based CDIPs framework while requiring less number of iterations than untrained neural network-based approach. This is because in pretraining, domain knowledge is incorporated into the network that facilitates the optimization process. Whereas, when DCP is used with CDIPs, there is a certain difference between the performance in terms of different metrics, i.e., PSNR, SSIM, and BRISQUE. The key noticeable thing is the significantly reduced iterations required for optimizing the neural network parameters for a given input image. It is evident from Table II that pretraining provides comparable performance as compared with random parameters initialization and converges in fewer iterations.

TABLE II

PERFORMANCE EVALUATION FOR **CDIPs FRAMEWORK WITH PRETRAINING** INTEGRATED WITH DCP AND BCP THAT ARE USED FOR ATMOSPHERIC LIGHT ESTIMATION. THE RESULTS ARE REPORTED IN TERMS OF THE AVERAGE PSNR, SSIM, AND BRISQUE.

| Dataset | #Images | CDIPs + DCP | | | | | | | | CDIPs + BCP | | | | | | | |
|------------|---------|-------------|-----|----------------|------|---------------|------|----------------|-------|-------------|-----|----------------|------|---------------|-------|----------------|-------|
| | | #Iterations | | PSNR | | SSIM | | BRISQUE | | #Iterations | | PSNR | | SSIM | | BRISQUE | |
| | | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Messidor | 1200 | 1859 | 804 | 37.8598 | 2.69 | 0.9696 | 0.02 | 49.4879 | 18.2 | 2001 | 730 | 38.1572 | 2.8 | 0.9722 | 0.01 | 47.4018 | 17.12 |
| STARE | 397 | 1793 | 966 | 35.5338 | 3.89 | 0.9506 | 0.04 | 50.8993 | 22.51 | 2133 | 762 | 37.2896 | 3.14 | 0.9652 | 0.02 | 43.3222 | 20.63 |
| DRIVE | 40 | 2379 | 285 | 37.6970 | 0.83 | 0.9707 | 0.01 | 23.4503 | 8.49 | 2463 | 30 | 37.0994 | 1.67 | 0.9705 | 0.001 | 28.9008 | 8.99 |
| DIARET DB0 | 130 | 2209 | 713 | 39.6497 | 2.84 | 0.9742 | 0.02 | 45.2281 | 19.65 | 2383 | 400 | 40.7116 | 2.33 | 0.9794 | 0.01 | 39.0192 | 15.27 |
| DIARET DB1 | 89 | 2235 | 678 | 39.9232 | 3.13 | 0.9732 | 0.03 | 50.4220 | 16.66 | 2444 | 171 | 41.4633 | 1.62 | 0.9824 | 0.01 | 43.6905 | 13.7 |

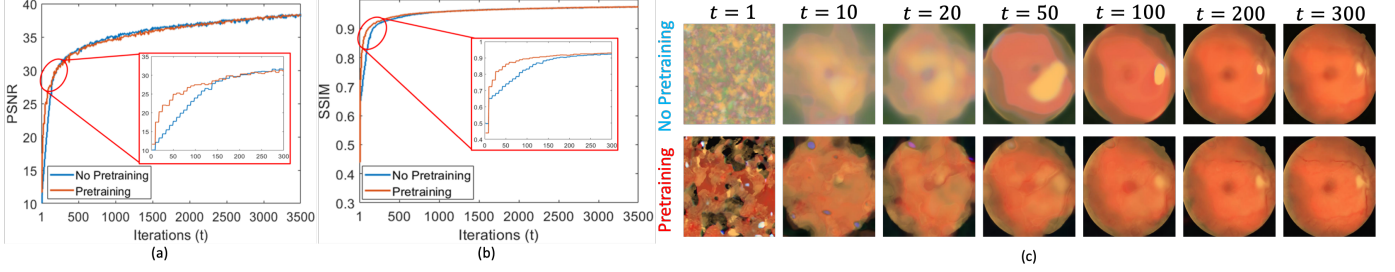


Fig. 3. Improvement in PSNR (a) and SSIM (b) over iterations t for a given input image using **randomly initialized network (no pretraining)** vs. using **pretrained network** (initialized with parameters of a network that generated a fundus image of the same dataset). It is clear from the zoom-in version of the plots that pretraining allows the network to capture the unknown image statistics at very few iterations. Intermediate results generated by both methods at iteration $t = 1, 10, 20, 50, 100, 200$, and 300 are shown in (c). It can be seen that the pre-trained network was able to learn retinal blood vessels after 200 iterations while untrained network was not able to capture such details.

The difference in the performance of pretrained CDIPs is expected and it indicates that the pretrained neural network parameters being used are not optimal for certain images. Due to this reason, the optimization process for a few images stop after very fewer iterations that results in low-quality images. This observation is supported by the standard deviation values, e.g., the high standard deviation in average iterations used for optimization for certain datasets (e.g., STARE) indicates that a few images were early stopped at a relatively less number of iterations as compared to others. This situation can be avoided by employing an optimal pretraining strategy. We further note that relatively high average SSIM values (Table I and II)—indicates that our method is able to reconstruct retinal images with high similarity between the structure of local patterns of the original image and enhanced image, which implies good quality preservation of the vascular structure.

E. Ablation Studies

1) *Effect of Using Pretraining*: The effect of using pretraining in our proposed CDIPs framework is depicted in Fig. 3. The figure demonstrates the intermediate results generated by our method at different iterations along with corresponding PSNR and SSIM, i.e., $t = 1, 10, 20, 50, 100, 200$, and 300 for both cases: (i) when the input image is modeled using randomly initialized network (no pretraining); and (ii) when the input image is modeled using pretrained network (initialized with parameters of a network that generated a fundus image of the same dataset). The difference in the performance of both cases is easily noticeable in zoom-in versions of the plots, as depicted in Fig. 3. Moreover, it can be seen that pretrained networks were able to capture/learn the low-level structural details of retinal blood vessels after 200 iterations while the untrained network was not able to capture such

details. Randomization seen in the output of the first iteration is due to the use of random code vector z (please see Fig. 2). Note that only parameters of the encoder network in the U-Net model are fitted using a single degraded image, i.e., without data-driven training.

2) *Pretraining for Cross Dataset Analysis*: We have also investigated the effect of pretraining for cross dataset-based analysis, i.e., when pretrained parameters of an image from one dataset are used to fit the images of another dataset. The quantitative results for cross dataset analysis are summarized in Table III. We used three datasets for this analysis in which one is developed for retinal blood vessels segmentation (i.e., DRIVE) and the other two are diabetic retinopathy databases. It is evident from the table that cross dataset based pretraining works pretty well and it provides better results when the images are of similar nature, for instance, the performance on the DB1 dataset when pretrained with DB0 image is significant as compared to when images from DRIVE are fitted using pretrained parameters of DB0 image. This is because of the fact that both DB0 and DB1 are of the same category, i.e., diabetic retinopathy datasets. Pretraining using an image from DRIVE database does not provide good performance in terms of average SSIM on images of DB1 database, which is expected as both databases are different. Similar to other pretraining experiments, in cross dataset pretraining, each image is processed for 2500 iterations and we used early stopping to avoid overfitting. Average iterations and iterations SD is also reported in Table III. It can be seen from the table that as compared to pretraining with DCP, pretraining with BCP was more prone to overfitting (evident from higher SD in the number of iterations). Because the domain knowledge (incorporated as pretrained parameters) increases the overall pigmentation of the fundus images enhanced using CDIPs integrated with BCP. Fig. 4 presents the results for an image

TABLE III
QUANTITATIVE RESULTS FOR CROSS DATASET PRETRAINING IN TERMS OF AVERAGE PSNR, SSIM, AND BRISQUE.

| Dataset | Pretrained On | Method | Iterations | | PSNR | | SSIM | | BRISQUE | |
|-------------|----------------------|-------------|------------|---------|----------------|--------|---------------|---------|---------|-------|
| | | | Average | SD | Average | SD | Average | SD | Average | SD |
| DB0 | One image from DRIVE | CDIPs + DCP | 2456 | 37 | 38.9114 | 1.28 | 0.9713 | 0.01 | 46.4396 | 13.32 |
| | | CDIPs + BCP | 2437 | 217 | 38.8235 | 2.36 | 0.9688 | 0.02 | 45.6849 | 13.45 |
| CDIPs + DCP | | 2388 | 274 | 38.7113 | 1.76 | 0.9476 | 0.01 | 46.4661 | 14.15 | |
| CDIPs + BCP | | 2364 | 484 | 39.0811 | 1.69 | 0.9378 | 0.01 | 46.8970 | 17.14 | |
| DRIVE | One image from DB0 | CDIPs + DCP | 2453 | 39 | 37.5131 | 0.78 | 0.9689 | 0.01 | 30.7329 | 8.9 |
| | | CDIPs + BCP | 2459 | 40 | 37.2651 | 2.06 | 0.9708 | 0.01 | 25.8046 | 10.93 |
| | | CDIPs + DCP | 2459 | 34 | 41.1129 | 1.23 | 0.9818 | 0.01 | 43.7756 | 14.02 |
| | | CDIPs + BCP | 2444 | 117 | 37.2340 | 2.18 | 0.9710 | 0.01 | 24.8941 | 8.05 |

from the STARE database that is fitted using pretrained parameters of the DIARET DB1 database. It is clear from the figure that the pretraining significantly boost the performance of our proposed framework in recovering faithful estimates of clean image. This is due to the fact that the untrained neural network is provided with domain knowledge (i.e., how a fundus image looks like in this case).

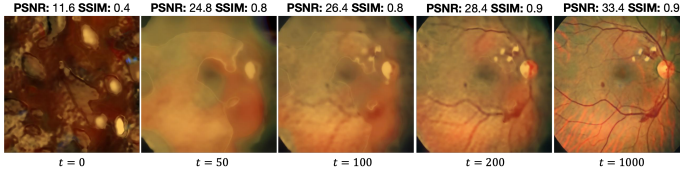


Fig. 4. Improvement in PSNR and SSIM when an image from STARE database is fitted using pretrained parameters of an image from DIARET DB1 database. The proposed framework starts learning low-level image statistics, i.e., retinal blood vessels at fewer iterations.

F. Subjective Evaluation of Enhancement

The subjective assessment of enhancement quality was performed by two expert ophthalmologists (one of them is the co-author of the paper who has been actively involved in the project since the beginning). For this purpose, we carefully selected five images from each dataset to ensure that the selected images are a good representative of our best, good, and average results to ensure a fair evaluation. Note that we select only five images per dataset in the interest of the expert's time and efforts. We then asked the expert ophthalmologists to subjectively evaluate the quality of the images by keeping in mind the following important features of retinal images: (i) visibility of optic disc (F1); (ii) visibility of fovea (F2); (iii) visibility of retinal blood vessels including venules, arterioles, and capillaries including their branching and termination (F3); (iv) overall general fundus (F4); and (v) overall assessment on the quality of the enhanced image as compared to original (hazy) image (F5). Subjective evaluation scores were provided based on the comparative analysis of the original (without enhancement) and enhanced images, as we show both images to the expert ophthalmologists for grading images generated by our method (to ensure a fair comparison and evaluation).

To get the subjective assessment scores, we used a grading based on five scales: 5: *Excellent* Results; 4: *Very Good* Results; 3: *Good* Results; 2: *Average* Results; and 1: *Bad* Results. The average values of subjective evaluation metrics along with respective standard deviation scores are presented in Table IV. From the table, it is evident that the expert

ophthalmologists rate the overall quality of the enhanced images to nearly excellent quality (nearly 5) for each dataset. Promising subjective assessment scores highlights the efficacy and clinical significance of our proposed method.

Below we describe the overall assessment of the expert ophthalmologists on the efficacy of our method, the ophthalmologists have the following specific observations about our proposed method.

- 1) It enhances the overall image clarity by $\sim 40\%$ by enhancing contrast and defining the outlines of various fundal structures like the optic disc, vessels, and fovea.
- 2) In the enhanced images, it is very easy to diagnose and analyze the retinal abnormalities that include fibrosis, lesions, and hard exudates.
- 3) Blood vessel termination and branching points become clearly visible in enhanced images as compared to the original images, which makes the detection of abnormal vessels and leakage points very easy.
- 4) In pigmented fundus images, the solution clarifies various lesions like degenerations, fibrosis, thinning, etc.
- 5) In light pigmented fundus images, it improves the visibility of various dark lesions like hemorrhage, melanomas, and naevi, etc.
- 6) The only noted limitation of the proposed method is slight obscuration of the outline of fundal structures with less than normal pigmentation such as seen in high myopes or Albinism.

G. Visual Evaluation of Enhancement

The visual comparison of generated (enhanced) images using different quality of blurry images and different methods proposed in this paper is shown in Fig. 5. These images have been selected from the subset images that we used for getting expert evaluation (we choose one image from each dataset) and we ensure that these images are representative of our best, good, and average results. It can be clearly seen from Fig. 5 that the proposed method can recover good quality images in which retinal blood vessels are prominently highlighted. Note that the clean image is recovered without any prior knowledge of the enhanced image. From Fig. 5, it can be seen that DCP favors some images and similarly, BCP also provides good performance in some images (e.g., the choroidal tissue defects near optic disc in the image of DIARET DB0 are prominently highlighted by BCP). This highlights that DCP leverages images with a darker background and BCP favors images having a bright background, i.e., DCP effectively handles the images having pixels with low intensities, which

TABLE IV
SUBJECTIVE EVALUATION OF IMAGES ENHANCED USING PROPOSED FRAMEWORK. LEGEND: F1–VISIBILITY OF OPTIC DISC; F2–VISIBILITY OF FOVEA; F3–VISIBILITY OF RETINAL BLOOD VESSELS INCLUDING VENULES, ARTERIOLES AND CAPILLARIES; F4–OVERALL GENERAL FUNDUS; F5–OVERALL ASSESSMENT ON THE QUALITY OF ENHANCED IMAGE AS COMPARED TO ORIGINAL (HAZY) IMAGE;

| Dataset | F1 | | F2 | | F3 | | F4 | | F5 | |
|-----------|------|------|------|------|------|------|------|------|------|------|
| | Mean | SD | Mean | SD | Mean | SD | Mean | SD | Mean | SD |
| Messidor | 3.8 | 0.4 | 3.6 | 0.49 | 4.2 | 0.4 | 3.8 | 0.75 | 5 | 0 |
| STARE | 3.2 | 0.75 | 3.2 | 0.75 | 3.8 | 0.75 | 3.8 | 0.75 | 4.6 | 0.49 |
| DRIVE | 3.6 | 0.8 | 3.6 | 0.48 | 3.4 | 0.49 | 4.8 | 0.4 | 5 | 0 |
| DIART DB0 | 3.4 | 0.49 | 3.4 | 0.8 | 3.2 | 0.4 | 4.2 | 0.75 | 4.8 | 0.4 |
| DIART DB1 | 3.4 | 0.49 | 3.6 | 0.48 | 3.2 | 0.4 | 3.6 | 0.49 | 4.8 | 0.4 |

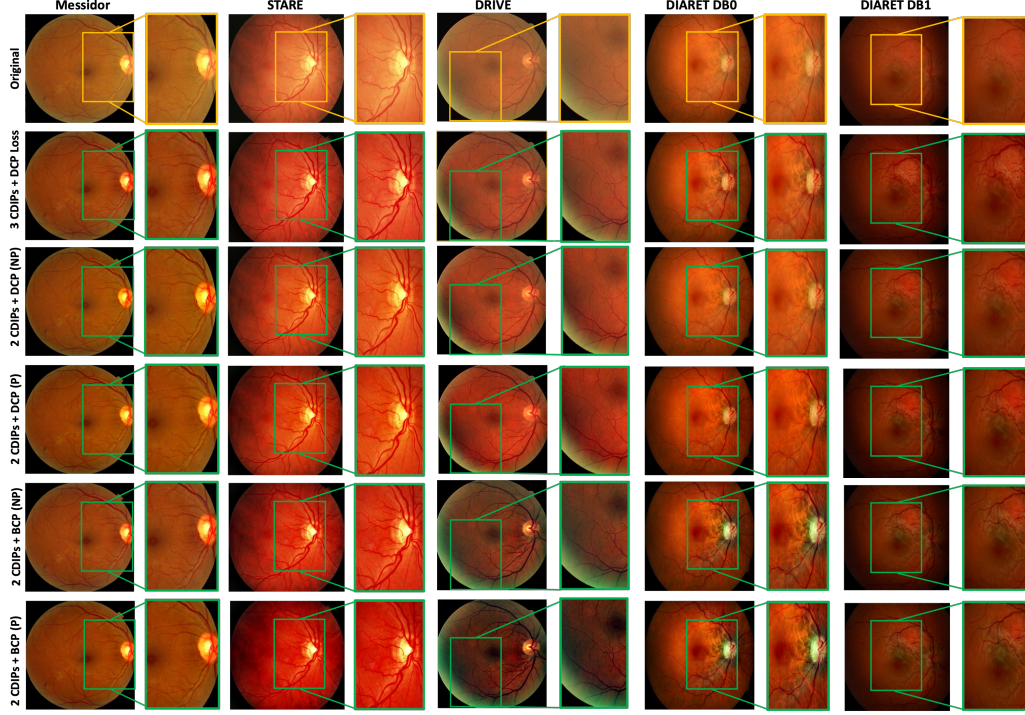


Fig. 5. Visual comparison of reconstructed (enhanced) images for different quality of blurry images. Blurry input images are shown in the first row and the rest of the rows contains results for a different method that has been proposed in this study (except three CDIPs + DCP loss [22]). Legend: CDIPs: Coupled Deep Image Priors; DCP: Dark Channel Prior; BCP: Bright Channel Prior; NP: No Pretraining Used; P: Pretraining Used.

makes it suitable for pigmented retinal fundus images and BCP effectively handles those images that have high intensities pixels, which makes it suitable for light pigmented retinal fundus images (i.e., that have sufficient brightness).

H. Evaluation on Synthetic Data

The acquisition of reference (ground truth clean) images in retinal fundus image enhancement problem is very challenging due to the difficulties in obtaining exact pixel-to-pixel tight image before and after cataract surgery. To fill in this gap, Shen et al. [25] have presented a fundus image degradation model to synthetically introduce visual artifacts and haze/blurriness to retinal fundus image. In Table V, we perform a comparative analysis of our method in terms of average PSNR and SSIM on synthetically degraded DRIVE dataset ([25]). To ensure a fair comparison, we have used a similar experimental setup as used in [25], where the synthetic retinal artifacts were introduced in the images of the DRIVE dataset. The original images were regarded as ground truth (reference images), and the synthetically degraded images are enhanced using our proposed CDIPs based framework. It is evident from Table

V that our method significantly outperforms existing methods despite being single-shot and unsupervised. Note that we did not use any pretraining for our analysis on the synthetic dataset. In addition, we used the method of Shen et al. [25] for synthetically generating different degraded images by introducing different artifacts in retinal fundus images that include uneven illumination (haze), blur, and artifacts due to the dusty lens of the fundus camera. The performance of the proposed framework in terms of average PSNR and average SSIM on the DRIVE dataset using different degradation methods is summarized in Table VI.

I. Complexity Analysis

As described above, in our previous work [22], we incorporated DCP loss into the overall loss of three CDIPs network. The DCP loss involves the computation of small patches from the generated image at every iteration that inherently slows the overall framework. In this paper, we present an alternative approach, where we used only two DIP networks that are integrated with conventional image priors, i.e., DCP and BCP. These priors are used for estimating uniform atmospheric light

TABLE V
COMPARISON WITH THE STATE-OF-THE-ART METHODS ON **SYNTHETICALLY DEGRADED IMAGES** FROM DRIVE DATASET AND A PROPRIETARY DATASET (*PUBLICLY NOT AVAILABLE, AUTHORS REPORTED RESULTS ARE USED FOR COMPARISON.)

| Paper | Method | Degradation | Dataset | Images | PSNR | SSIM | BRISQUE |
|------------------------|------------------------------------------------------------|-------------|--------------|--------|--------------|--------------|--------------|
| Li et al. [43] | Histogram distribution prior. | Synthetic | DRIVE | N/A | 9.51 | 0.543 | N/A |
| Fu et al. [44] | Retinex-based single image enhancing. | | | | 10.19 | 0.580 | |
| Guo et al. [45] | Illumination map estimation. | | | | 14.10 | 0.703 | |
| Cheng et al. [46] | Guided retinal image filtering. | | | | 14.97 | 0.648 | |
| Tian et al. [47] | Global and local contrast adaptive enhancement. | | | | 15.42 | 0.721 | |
| Fu et al. [48] | Weighted variational model. | | | | 15.56 | 0.722 | |
| He et al. [26] | Dark channel prior (DCP). | | | | 15.78 | 0.559 | |
| Zuiderveld et al. [49] | Contrast limited adaptive histogram equalization | | | | 15.93 | 0.740 | |
| Eilertsen et al. [50] | Deep convolutional neural network (CNN). | | | | 19.01 | 0.755 | |
| Shen et al. [25] | Multiple GANs are used in the proposed Cofe-Net framework. | | | | 21.24 | 0.758 | |
| Zhao et al. [13] | Two GANs are trained in supervised learning. | Real | Proprietary* | 10 | 19.24 | 0.89 | 40.62 |
| Zhou et al. [11] | Luminosity and contrast adjustment. | | | | 17.73 | 0.73 | 46.13 |
| Mitra et al. [8] | Histogram and intensity equalization. | | | | 16.38 | 0.78 | 45.16 |
| Xiong et al. [9] | Used image formation model. | | | | 17.26 | 0.87 | 43.61 |
| <i>This Paper</i> | Coupled deep image priors (CDIPs) integrated with DCP. | Synthetic | DRIVE | 40 | 22.47 | 0.897 | 43.02 |

TABLE VI
PERFORMANCE OF THE PROPOSED CDIPs FRAMEWORK USING **SYNTHETICALLY DEGRADED IMAGES ON DRIVE**.

| Degradation | Method | Images | PSNR | SSIM |
|---------------|-------------|--------|-------|-------|
| Haze | CDIPs + DCP | 40 | 22.47 | 0.897 |
| | CDIPs + BCP | | 21.92 | 0.864 |
| Blur | CDIPs + DCP | | 21.49 | 0.893 |
| | CDIPs + BCP | | 21.23 | 0.848 |
| Haze and Blur | CDIPs + DCP | | 22.21 | 0.895 |
| | CDIPs + BCP | | 20.7 | 0.846 |
| Haze and Dust | CDIPs + DCP | | 21.23 | 0.868 |
| | CDIPs + BCP | | 21.08 | 0.844 |

TABLE VII
COMPLEXITY ANALYSIS IN TERMS OF PERCENTAGE IMPROVEMENT IN AVERAGE TIME AND AVERAGE MEMORY UTILIZED PER IMAGE.

| Method | Avg. Time/Image | Avg. Memory/Image | Parameters |
|--------------------------------|-----------------|-------------------|---------------|
| Three CDIPs with DCP Loss [22] | 35 mins | 4.4 GB | 1.72M |
| CDIPs + DCP/BCP | 20 mins | 2.9 GB | 1.15M |
| Percentage Improvement | 27.27% | 20.54% | 19.86% |

in contrast to our prior work [22], where we used a separate DIP network for non-uniform atmospheric light estimation. In this section, we perform a complexity analysis of both approaches in terms of the percentage improvement in average time per image, average memory utilized per image, and the number of parameters saved. Table VII highlights that we have 27.27%, 20.54%, and 19.86% percent improvement in time, memory, and parameters respectively. These scores are based on an image size of 512×512 and when 3500 iterations were used to recover the input images. The image size and number of iterations have a direct impact on time and memory usage. Note that these scores are for the overall end-to-end framework that involves all steps from image loading to saving, etc. We further note that the sole optimization of CDIPs takes only 0.1 sec per iteration for our current approach.

V. LIMITATIONS AND FUTURE WORK

The key limitation of the CDIPs-based approach is the time required for reconstructing a recovered image. We see that increasing the number of iterations results in a gradual increase in performance, e.g., PSNR and SSIM. To address this time utilization issue, we have also investigated the use of pretraining, which has provided promising results for the majority of images. However, in some cases, pretraining

results in overfitting that indicate the pretrained parameters being used were not optimal for specific images. To overcome this issue, we aim to develop an optimal pretraining strategy to fully uncover the potential of incorporating domain knowledge (i.e., through pretraining) into untrained neural networks in our future work. Also, the development of optimal early stopping criteria could be another possible solution towards reducing the time taken and avoiding overfitting. Another limitation of our approach is the slight obscuration of the outline of fundal structures that have been noted by the expert ophthalmologists while performing the subjective evaluation.

VI. CONCLUSIONS

In this paper, we have presented a unified single-shot deep learning (DL) framework for the enhancement of retinal fundus images. For this purpose, we have employed coupled untrained neural networks known as deep image priors (DIP) which are integrated with a conventional image prior, i.e., dark channel prior (DCP) and bright channel prior (BCP). The proposed work reconstructs the enhanced image using a single degraded image without the requirement of end-to-end data-driven training. We quantitatively evaluate the proposed approach on five different retinal fundus image datasets in terms of average peak signal to noise ratio (PSNR), structural similarity index (SSIM), and Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) score. We also evaluate our method using a synthetic dataset (in which different image degradation were synthetically introduced in the original images prior to enhancement). We perform performance evaluation of our proposed approach with existing similar methods and our method outperforms them. To highlight the clinical significance, we incorporate the subjective assessment of the enhancement that is performed by two expert ophthalmologists. In addition to using untrained neural networks, we have also investigated pretraining that have provided promising results while reducing the computational cost and time.

ACKNOWLEDGMENT

We are very thankful to Dr. Kanwal Zareen Abbasi Associate Professor of Eye at HBS Medical and Dental College, Islamabad, Pakistan, for subjective qualitative assessment.

REFERENCES

- [1] B. Chen, Y. Chen, Z. Shao, T. Tong, and L. Luo, "Blood vessel enhancement via multi-dictionary and sparse coding: Application to retinal vessel enhancing," *Neurocomputing*, vol. 200, pp. 110–117, 2016.
- [2] D. Li, L. Zhang, C. Sun, T. Yin, C. Liu, and J. Yang, "Robust retinal image enhancement via dual-tree complex wavelet transform and morphology-based method," *IEEE Access*, vol. 7, pp. 47 303–47 316, 2019.
- [3] U. Raghavendra, H. Fujita, S. V. Bhandary, A. Gudigar, J. H. Tan, and U. R. Acharya, "Deep convolution neural network for accurate diagnosis of glaucoma using digital fundus images," *Information Sciences*, vol. 441, pp. 41–49, 2018.
- [4] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017.
- [5] W. Ma, S. Yu, K. Ma, J. Wang, X. Ding, and Y. Zheng, "Multi-task neural networks with spatial activation for retinal vessel segmentation and artery/vein classification," in *MICCAI*. Springer, 2019, pp. 769–778.
- [6] J. I. Orlando, E. Prokofyeva, and M. B. Blaschko, "A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images," *IEEE transactions on Biomedical Engineering*, vol. 64, no. 1, pp. 16–27, 2016.
- [7] L. Cao, H. Li, and Y. Zhang, "Retinal image enhancement using low-pass filtering and α -rooting," *Signal Processing*, vol. 170, p. 107445, 2020.
- [8] A. Mitra, S. Roy, S. Roy, and S. K. Setua, "Enhancement and restoration of non-uniform illuminated fundus image of retina obtained through thin layer of cataract," *Computer methods and programs in biomedicine*, vol. 156, pp. 169–178, 2018.
- [9] L. Xiong, H. Li, and L. Xu, "An enhancement method for color retinal images based on image formation model," *Computer methods and programs in biomedicine*, vol. 143, pp. 137–150, 2017.
- [10] T. J. MacGillivray, J. R. Cameron, Q. Zhang, A. El-Medany, C. Mulholland, Z. Sheng, B. Dhillon, F. N. Doubal, P. J. Foster, E. Trucco *et al.*, "Suitability of uk biobank retinal images for automatic analysis of morphometric properties of the vasculature," *PLoS One*, vol. 10, 2015.
- [11] M. Zhou, K. Jin, S. Wang, J. Ye, and D. Qian, "Color retinal image enhancement based on luminosity and contrast adjustment," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 3, pp. 521–527, 2017.
- [12] M. Foracchia, E. Grisan, and A. Ruggeri, "Luminosity and contrast normalization in retinal images," *Medical image analysis*, vol. 9, no. 3, pp. 179–190, 2005.
- [13] H. Zhao, B. Yang, L. Cao, and H. Li, "Data-driven enhancement of blurry retinal images via generative adversarial networks," in *MICCAI*. Springer, 2019, pp. 75–83.
- [14] A. Raj, A. K. Tiwari, and M. G. Martini, "Fundus image quality assessment: survey, challenges, and future scope," *IET Image Processing*, vol. 13, no. 8, pp. 1211–1224, 2019.
- [15] U. Sevik, C. Kose, T. Berber, and H. Erdol, "Identification of suitable fundus images using automated quality assessment methods," *Journal of biomedical optics*, vol. 19, no. 4, p. 046006, 2014.
- [16] T. Köhler, A. Budai, M. F. Kraus, J. Odstrčilik, G. Michelson, and J. Hornegger, "Automatic no-reference quality assessment for retinal fundus images using vessel segmentation," in *IEEE international symposium on computer-based medical systems*, 2013, pp. 95–100.
- [17] R. Pires, H. F. Jelinek, J. Wainer, and A. Rocha, "Retinal image quality analysis for automatic diabetic retinopathy detection," in *SIBGRAPI Conference on Graphics, Patterns and Images*, 2012, pp. 229–236.
- [18] H. Fu, B. Wang, J. Shen, S. Cui, Y. Xu, J. Liu, and L. Shao, "Evaluation of retinal image quality assessment networks in different color-spaces," in *MICCAI*. Springer, 2019, pp. 48–56.
- [19] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *IEEE CVPR*, 2018, pp. 9446–9454.
- [20] Y. Gandelsman, A. Shocher, and M. Irani, "Double-DIP: Unsupervised image decomposition via coupled deep-image-priors," in *IEEE CVPR*, vol. 6, 2019, p. 2.
- [21] A. Qayyum, I. Ilahi, F. Shamshad, F. Boussaid, M. Bennamoun, and J. Qadir, "Untrained neural network priors for inverse imaging problems: A survey," 2021.
- [22] A. Qayyum, W. Sultani, F. Shamshad, J. Qadir, and R. Tufail, "Single-shot retinal image enhancement using deep image priors," in *MICCAI*. Springer International Publishing, 2020, pp. 636–646.
- [23] A. Smailagic, A. Sharan, P. Costa, A. Galdran, A. Gaudio, and A. Campilho, "Learned pre-processing for automatic diabetic retinopathy detection on eye fundus images," in *International Conference on Image Analysis and Recognition*. Springer, 2019, pp. 362–368.
- [24] Q. You, C. Wan, J. Sun, J. Shen, H. Ye, and Q. Yu, "Fundus image enhancement method based on CycleGAN," in *IEEE EMBC*. IEEE, 2019, pp. 4500–4503.
- [25] Z. Shen, H. Fu, J. Shen, and L. Shao, "Modeling and enhancing low-quality retinal fundus images," *IEEE TMI*, 2020.
- [26] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE PAMI*, vol. 33, no. 12, pp. 2341–2353, 2010.
- [27] Y. Bahat and M. Irani, "Blind dehazing using internal patch recurrence," in *2016 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2016, pp. 1–9.
- [28] D. Berman, S. Avidan *et al.*, "Non-local image dehazing," in *IEEE CVPR*, 2016, pp. 1674–1682.
- [29] G. Meng, Y. Wang, J. Duan, S. Xiang, and C. Pan, "Efficient image dehazing with boundary constraint and contextual regularization," in *IEEE international conference on computer vision*, 2013, pp. 617–624.
- [30] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, and M.-H. Yang, "Single image dehazing via multi-scale convolutional neural networks," in *European conference on computer vision*. Springer, 2016, pp. 154–169.
- [31] K. B. Gibson, D. T. Vo, and T. Q. Nguyen, "An investigation of dehazing effects on image and video coding," *IEEE TIP*, vol. 21, no. 2, pp. 662–673, 2011.
- [32] J.-P. Tarel and N. Hautiere, "Fast visibility restoration from a single color or gray level image," in *IEEE 12th International Conference on Computer Vision*. IEEE, 2009, pp. 2201–2208.
- [33] L. Tao, C. Zhu, J. Song, T. Lu, H. Jia, and X. Xie, "Low-light image enhancement using CNN and bright channel prior," in *IEEE ICIP*. IEEE, 2017, pp. 3215–3219.
- [34] D. Singh and V. Kumar, "Single image haze removal using integrated dark and bright channel prior," *Modern Physics Letters B*, vol. 32, no. 04, p. 1850051, 2018.
- [35] S. F. Asim, Muhammad and A. Ahmad, "Blind image deconvolution using pretrained generative priors," in *BMVC*, 2019.
- [36] Y.-C. Chen, C. Gao, E. Robb, and J.-B. Huang, "NAS-DIP: Learning deep image prior with neural architecture search," *arXiv preprint arXiv:2008.11713*, 2020.
- [37] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE TMI*, vol. 23, no. 4, pp. 501–509, 2004.
- [38] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE TMI*, vol. 19, no. 3, pp. 203–210, 2000.
- [39] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay *et al.*, "Feedback on a publicly distributed image database: the messidor database," *Image Analysis & Stereology*, vol. 33, no. 3, pp. 231–234, 2014.
- [40] R. Kälviäinen and H. Uusitalo, "DIARETDB1 diabetic retinopathy database and evaluation protocol," in *Medical Image Understanding and Analysis*, vol. 2007. Citeseer, 2007, p. 61.
- [41] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015, pp. 234–241.
- [42] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *ICLR*, 2015.
- [43] C.-Y. Li, J.-C. Guo, R.-M. Cong, Y.-W. Pang, and B. Wang, "Underwater image enhancement by dehazing with minimum information loss and histogram distribution prior," *IEEE TIP*, vol. 25, 2016.
- [44] X. Fu, P. Zhuang, Y. Huang, Y. Liao, X.-P. Zhang, and X. Ding, "A retinex-based enhancing approach for single underwater image," in *IEEE ICIP*. IEEE, 2014, pp. 4572–4576.
- [45] X. Guo, Y. Li, and H. Ling, "Lime: Low-light image enhancement via illumination map estimation," *IEEE TIP*, vol. 26, 2017.
- [46] J. Cheng, Z. Li, Z. Gu, H. Fu, D. W. K. Wong, and J. Liu, "Structure-preserving guided retinal image filtering and its application for optic disk analysis," *IEEE TMI*, vol. 37, no. 11, pp. 2536–2546, 2018.
- [47] Q.-C. Tian and L. D. Cohen, "Global and local contrast adaptive enhancement for non-uniform illumination color images," in *IEEE International Conference on Computer Vision Workshops*, 2017, pp. 3023–3030.
- [48] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in *IEEE CVPR*, 2016, pp. 2782–2790.
- [49] K. Zuiderveld, "Contrast limited adaptive histogram equalization," *Graphics gems*, pp. 474–485, 1994.
- [50] G. Eilertsen, J. Kronander, G. Denes, R. K. Mantiuk, and J. Unger, "Hdr image reconstruction from a single exposure using deep CNNs," *ACM transactions on graphics (TOG)*, vol. 36, no. 6, pp. 1–15, 2017.