Pre-Processing Methods for Diabetic Retinopathy

Alex Gaudio
Electrical Computer Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
agaudio@andrew.cmu.edu

Ankit Shrivastava
Civil and Environmental Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
ashriva2@andrew.cmu.edu

Wenbo Zhao
Mechanical Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
wenbozha@andrew.cmu.edu

Yuanyuan Ji
Civil and Environmental Engineering
Carnegie Mellon University
Pittsburgh, PA 15213
yuanyuaj@andrew.cmu.edu

Abstract

Diabetic Retinopathy (DR) is the leading cause of preventable blindness among working age adults, yet detection of the disease typically requires challenging and time-consuming analysis of retinal fundus images. While many DR detection and lesion segmentation algorithms have been proposed, there is less emphasis on image enhancement pre-processing and poor understanding of why one pre-processing technique might be more useful than another. To explain how pre-processing image enhancement methods facilitate DR lesion segmentation, we implement consistency and separability scores that are task specific and conditionally independent of the downstream image processing model given the pre-processing method. We then use this evaluation to guide development of a novel pre-processing method useful to DR lesion segmentation. We reinterpret a well-known theory for dehazing natural images as a generalized framework for image enhancement, showing the theory performs pixel color amplification. We provide a theoretical toolset for creating a family of pre-processing methods under the framework, and we provide examples, both novel and existing, of image enhancement algorithms derived via this toolset. Our best performing enhancement method significantly increases separability of DR lesions, requires only a single image, does not use labels, and does not require training or a dataset.

1 Introduction

Image enhancement is a process of adding or removing noise from images in order to improve performance on a future image processing task. We consider image-to-image pre-processing methods that are intended to facilitate a downstream image processing task such as Diabetic Retinopathy lesion segmentation, where the goal is to identify which pixels in an image of a human retina are pathological. In this setting, image enhancement does not in itself perform segmentation, but it ideally improves observability of relevant features. Example image enhancement techniques are color illumination, image dehazing, deblurring and edge detection. Fig. (1) shows our best performing enhancement method.

Evaluation of image enhancement pre-processing methods is challenging. A typical black box strategy to justify using a given pre-processing method is to incorporate it into the image processing model and evaluate the model's change in performance via cross validation or ablative analysis. These black box evaluation approaches are specific to the task and model, and they ask whether a particular

Preprint. Under review.

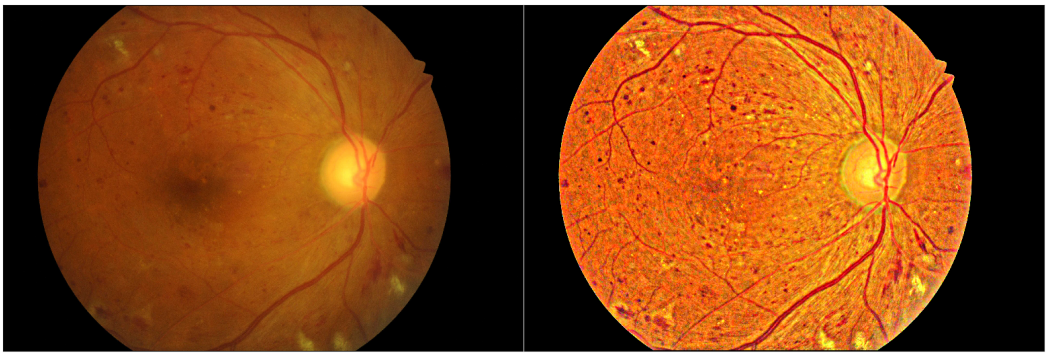

Figure 1: Comparing unmodified image (left) to our Illuminate-Sharpen enhancement (right).

downstream image processing model, such as a disease segmentation model, performs better or worse with image pre-processing.

This evaluation strategy is fundamentally limited. First, the ablative analysis, or the systematic evaluation of all combinations of a candidate set of pre-processing methods, is computationally inefficient. Comparing even a small set of pre-processing methods and their respective hyperparameters may be intractable or expensive. Secondly, this black box evaluation sheds little intuition to whether a pre-processing method may be useful to other models or tasks. For example, high capacity models, such as deep networks, often benefit by image augmentation techniques [9] or regularization constraints that make the task more difficult. On the other hand, low capacity models such as a simple thresholding function, benefit from pre-processing methods that remove noise or make the data easier to analyze. The model, as a black box, does not explain how well the pre-processing works, what it does, or in many cases, whether the observed performance change generalizes to other models. Third, in the case of image enhancement for medical image analysis, the image processing model is often a human physician. Obtaining willing physicians is difficult. Evaluating change in diagnostic performance under different pre-processing algorithms cannot risk patient health or waste time. In lab settings, obtaining human feedback is time consuming. To facilitate the process of choosing and implementing a pre-processing method, we argue that pre-processing methods should be primarily evaluated independently of the image processing model.

To evaluate pre-processing methods for DR segmentation independently of the model, we utilize consistency and separability measures. Section 5 describes these measures. We use the results of those measures to make informed decisions on the development of a novel pre-processing method for DR segmentation.

In the following sections, we develop a theoretical framework and example algorithms for creating a variety of image enhancement pre-processing methods. We show how a distortion model traditionally used for dehazing also performs general image enhancement by amplifying differences between the input image and a baseline "atmosphere" term. Section 2 interprets dehazing theory in context of image distortion. Section 3 generalizes the theory for other forms of image enhancement besides dehazing. Section 5 presents our evaluation methods. Section 5.2 describes our results, where we show evaluation scores on a variety of pre-processing methods and describe how the evaluation guided our development of a novel pre-processing method for DR lesion segmentation.

2 Related Work

Natural images are distorted by refraction of light as it travels through the transmission medium (such as air), causing modified pixel intensities in the color channels of the image. A widely used physical model for this distortion has traditionally been used for single image dehazing [2, 5, 8, 14]:

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

where each pixel location, $\mathbf{x}$, in the distorted RGB image, $\mathbf{I}$, can be constructed as a function of the distortion-free radiance image $\mathbf{J}$, a grayscale transmission map image $\mathbf{t}$ quantifying the relative portion of the light ray coming from the observed surface in $\mathbf{I}(\mathbf{x})$ that was not scattered (and where

Dehazing (Dark Channel Prior)		Illumination Correction (Inverted Dark Channel Prior)	
$\mathbf{A} = (r, g, b).$	(2)	$\mathbf{A} = (1, 1, 1).$	(6)
$\mathbf{t} = \text{guidedFilter}(1 - \min_c \min_{y \in \Omega_{I(\mathbf{x})}} \frac{I^c(\mathbf{y})}{A^c}, \mathbf{I}).$	(3)	$\mathbf{J} = 1 - f_{\text{DCP}}(1 - \mathbf{I}, \mathbf{A})$	(7)
$\mathbf{J}(\mathbf{x}) = \frac{\mathbf{I}(\mathbf{x}) - \mathbf{A}}{\max(t(\mathbf{x}), \epsilon)} + \mathbf{A}.$	(4)		
$\implies \mathbf{J} = f_{\text{DCP}}(\mathbf{I}, \mathbf{A})$	(5)		

Figure 2: **Left:** Dark Channel Prior (DCP) method for dehazing. **Right:** Inverted DCP for color illumination. Math follows from simple algebraic manipulation of Eq. (1) and the DCP assumption. Given an (inverted) image $\mathbf{I}$, obtain a dark channel using the min operator, compute and refine a transmission map $\mathbf{t}$, and last, obtain the undistorted (and un-inverted) image $\mathbf{J}$.

$t(\mathbf{x}) \in [0, 1] \; \forall \; \mathbf{x}$), and an atmosphere term, $\mathbf{A}$, which is typically a RGB vector that approximates the color of the uniform scattering of light. This image degredation model, Eq. (1), expresses the physical assumption that scattering of light conserves energy, and therefore, distortion due to light scattering is a t-weighted average of the noise-free RGB pixel with a constant color $\mathbf{A}$. We refer to [2] for a deeper treatment of the physics behind the model in a dehazing context. Obtaining a distortion free image $\mathbf{J}$ via this model is typically a three step process. Given $\mathbf{I}$, define an atmosphere term $\mathbf{A}$, solve for the transmission map $\mathbf{t}$, and then solve for $\mathbf{J}$.

The well known Dark Channel Prior (DCP) method [5, 7] addresses the dehazing task for Eq. (1) by imposing a prior assumption on RGB images. The assumption differentiates the noisy (hazy) image, $\mathbf{I}$, from its noise free (dehazed) image, $\mathbf{J}$. That is, in any haze-free three-channel region of a RGB image, at least one pixel has zero intensity in at least one channel ($\{(0, g, b), (r, 0, b), (r, g, 0)\}$), while a hazy region will have no pixels with zero intensity ($(r > 0, g > 0, b > 0)$). Indeed, Eq. (1) shows distortion is simply an additive airlight term $\mathbf{A}(1 - t(\mathbf{x}))$. To quantify haze distortion in an image, the assumption justifies creating a fourth channel, known as the dark channel, by applying a min operator convolutionally to each region of the images $\mathbf{I}$ and $\mathbf{J}$. For example, $\tilde{I}^{\text{dark}}(\mathbf{x}) = \min_c \min_{\mathbf{y} \in \Omega_{\mathbf{I}}(\mathbf{x})} \frac{I^{(c)}(\mathbf{y})}{A^c}$, where c denotes the color channel (red, green or blue) and $\Omega_{\mathbf{I}}(\mathbf{x})$ is a set of pixels in $\mathbf{I}$ neighboring pixel $\mathbf{x}$. While the min operator causes $\tilde{I}^{\text{dark}}$ to lose fine detail, an edge-preserving filter known as the guided filter [6] restores detail $\mathbf{I}^{\text{dark}} = g(\tilde{\mathbf{I}}^{\text{dark}}, \mathbf{I})$. By observing that the distortion free image $\mathbf{J}^{\text{dark}}$ is entirely zero while $\mathbf{I}^{\text{dark}}$ is not entirely zero, solving Eq. (1) for $\mathbf{t}$ leads to Eq. (3) and then Eq. (4) in Fig. 2. In practice, the denominator of (4) is $\max(t(\mathbf{x}), \epsilon)$ to avoid numerical instability or division by zero; this amounts to preserving a small amount of distortion in heavily distorted pixels. Fig. 2 summarizes the mathematics.

It was recently shown [12, 3, 13] that the distortion model in Eq. (1) and associated dehazing methods can be used for illumination correction. The central idea is to invert the image, apply the dehazing method, and then invert the dehazed result. We demonstrate the mathematics of this inverted DCP method in Figure 2. Prior work requires the important assumption that $\mathbf{A} = (1, 1, 1)$, meaning the image is white-balanced. In the dehazing context, this assumption would mean the distorted pixels are too bright, but in the color illumination context, distorted pixels are too dark. In the Methods section, we expand on this concept.

3 Methods

While Eq. (1) has traditionally been utilized for dehazing natural images, the model generalizes to any setting where a distorted signal $\mathbf{I}$ is a $\mathbf{t}$-weighted average of the latent noise-free signal $\mathbf{J}$ with a vector $\mathbf{A}$. This insight implies that the model can be generalized a large variety of linear image manipulation, including modifications that are not necessarily physically valid, but that still perform useful image enhancement. We emphasize three primary components when developing image enhancement methods under the theory: choose $\mathbf{A}$, choose $\mathbf{t}$, and decide on a transform for $\mathbf{I}$.

3.1 The distortion model amplifies pixel intensities

We assume that $\mathbf{A}$, $\mathbf{I}$ and $\mathbf{J}$ share the same space of pixel intensities, so that in any given channel, c, the intensities A^c, $I^c(\mathbf{x})$ and $J^c(\mathbf{x})$ can all have the same maximum or minimum value. We can derive the simple equation $t(\mathbf{x}) = \frac{I^{(c)}(\mathbf{x})-A^{(c)}}{J^{(c)}(\mathbf{x})-A^{(c)}} \in [0,1]$ from Eq. (1) by noting that the distortion model is linear system containing three channels. The range of $\mathbf{t}$ implies the numerator and denominator must have the same sign. For example, if $\mathbf{A}^{(c)} = 1$, then the numerator and denominator are negative and $J^{(c)}(\mathbf{x}) \leq I^{(c)}(\mathbf{x}) \leq A^{(c)}$. This has important implications that dehazing methods using $A = (1,1,1)$ will always make the recovered image $\mathbf{J}$ darker, and by extension, inverted dehazing methods make the recovered image brighter. Likewise, if $\mathbf{A}^{(c)} = 0$, then $J^{(c)}(\mathbf{x}) \geq I^{(c)}(\mathbf{x}) \geq A^{(c)}$, and solving for $\mathbf{J}$ becomes trivial $\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})t(\mathbf{x}) + 0$. With the DCP method, inverted dehazing when $\mathbf{A} = (1,1,1)$ has the same effect as dehazing with $\mathbf{A}^{(c)} = 0$, except that the latter case suffers division by zero as shown in Eq. (3). Thus, the purpose of inverted dehazing is to avoid division by zero. The key insight is that the choice of $\mathbf{A}$ determines how the color of each pixel in the recovered image $\mathbf{J}$ changes. Models that recover $\mathbf{J}$ using Eq. (1) will simply amplify color values for each pixel $\mathbf{x}$ in the direction $\mathbf{I}(\mathbf{x}) - \mathbf{A}$.

The choice of atmosphere term $\mathbf{A}$ is traditionally a single color value, that is a $A = (r,g,b)$, a vector with three scalar values. Without loss of generality, it is physically valid to also consider $\mathbf{A}$ as a three channel image since the atmospheric light may shift color across the image, for instance due to a change in light source. Considering $\mathbf{A}$ as an image significantly expands the capabilities of Eq. (1) as a distortion model, as we can control the direction of color amplification for each individual pixel, and indeed, of each color channel of any given pixel. In this regard, the distortion model can be utilized to recover an image $\mathbf{J}$ where the pixel-wise difference between the images $\mathbf{I}$ and $\mathbf{A}$ are amplified.

3.1.1 Image Sharpening

One way to utilize the pixel amplification property is to amplify fine detail. In this section, show a connection between dehazing and *unsharp mask*, a deblurring method and standard image sharpening technique [10]. Consider what might happen if we apply a physically invalid non-linear blurring operator to define $\mathbf{A}$ as a three channel image, $\mathbf{A} = \text{blurry}(\mathbf{I})$. Solving Eq. (1) for $\mathbf{J}$ gives $\mathbf{J} = \frac{1}{\mathbf{t}}\mathbf{I} - \frac{(1-\mathbf{t})}{\mathbf{t}}\mathbf{A}$. Since each scalar value $t(\mathbf{x})$ is in $[0,1]$, we can represent the fraction $t(\mathbf{x}) = \frac{1}{u(x)}$. Substituting, we have the simplified matrix form $\mathbf{J} = \mathbf{u} \circ \mathbf{I} - (\mathbf{u} - 1) \circ \text{blurry}(\mathbf{I})$ where the $\circ$ operator denotes element-wise multiplication with broadcasting across channels. This form is precisely *unsharp mask*, where $\mathbf{u}$ is either a constant, or $\mathbf{u}$ is a 1-channel image matrix determining how much to sharpen each pixel. The latter case is known as locally adaptive unsharp masking. Thus, we show the distortion model in Eq. (1) is equivalent to image sharpening by choosing $\mathbf{A}$ to be a blurred version of the original input image.

3.2 Selectively amplifying pixel intensities

While the choice of $\mathbf{A}$ determines the direction of pixel amplification, a choice for the transmission map, t or $\mathbf{t}$, determines the amount or rate of pixel amplification by averaging. For instance, when $t(\mathbf{x}) = 1$, the distorted and undistorted pixel are the same, since $\mathbf{I}(\mathbf{x}) = \mathbf{J}(\mathbf{x})1 + 0$, and as $t(\mathbf{x})$ approaches zero, the distortion increases.

Methods that find non-trivial transmission maps encode prior domain knowledge. Methods based on DCP, for example, make the assumption that the dark channel is zero in a dehazed image. Rather than replace these methods with our own, we propose to augment the capabilities of existing dehazing operators that rely on assumed differences between distorted and undistorted images.

3.2.1 Amplifying in better color spaces

The purpose of DCP is to find a $\mathbf{t}$ that works well for dehazing natural images. It is sub-optimal for dehazing when one color channel in the distorted image is very dark, since that channel may contribute too many zeros to the transmission map. Likewise, the operator does not work well for color illumination if one color channel is saturated. Thus, the utility of DCP depends on the color

space of the input images. In this section, we show how an invertible matrix transform changes the color space and manipulate the distortion model Eq. (1).

Consider a simple example where the red, green and blue (RGB) channels are converted to cyan, magenta and yellow (CMY). We can represent this system of linear equations as a matrix transform $\mathbf{I}M$, where M is a 3×3 mixing matrix and the columns define different linear combinations of the red, green and blue channels. We assume M is invertible, square, and its columns sum to 1.

$$\mathbf{I}_{CMY} = \mathbf{I}_{RGB}M = \mathbf{I}_{RGB}\begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \end{bmatrix} \tag{8}$$

Any matrix M satisfying the above assumptions will not violate the physics of the model if used to transform $\mathbf{I}$ and recover $\mathbf{J}$ according to Eq. 1. As proof, we show conditions when the matrix transform has no effect on the recovered $\mathbf{J}$. Consider the setting where we multiply both sides of Eq. (1) by M. We assume that $\mathbf{t}$ is either scalar or a grayscale image, while $\mathbf{A}$ is either an (r, g, b) vector or the same shape as $\mathbf{I}$.

$$\mathbf{I}M = (\mathbf{J} \circ \mathbf{t})M + (\mathbf{A} \circ (1 - \mathbf{t}))M = \mathbf{J}M \circ \mathbf{t} + \mathbf{A}M \circ (1 - \mathbf{t}) \tag{9}$$

$$\implies \mathbf{J}M = \frac{(\mathbf{I} - \mathbf{A})M}{\mathbf{t}} + \mathbf{A}M. \tag{10}$$

Eq. (10) implies that if $\mathbf{I}$ and $\mathbf{A}$ are transformed, then we can restore $\mathbf{J}$ to the original color space by multiplying it with the inverted matrix: $(\mathbf{J}M)M^{-1}$.

The motivation for an invertible matrix transform is in the ways it interacts with non-linear operators. Non-linearities behave in interesting ways with color space transforms. In the case of image sharpening, for instance, if we incorporate the blurring step such that $\mathbf{A}M \triangleq \text{blur}(\mathbf{I}M)$, then the restored $(\mathbf{J}M)M^{-1}$ after sharpening may have surprising results because $\mathbf{A}MM^{-1} \neq \text{blur}(\mathbf{I}M)M^{-1}$. On the other hand, if we performed sharpening so that $AM \triangleq \text{blur}(\mathbf{I})M$, then the matrix transform passes through without affecting the model. Interestingly, while Eq. (9) suggests $\mathbf{t}$ is not affected by the matrix transform, this is not always the case. The DCP method solves for $\mathbf{t}$ as a function of $\mathbf{I}$ and $\mathbf{A}$, as shown in Eq. (3). Therefore, transforming either $\mathbf{I}$ or $\mathbf{A}$ before solving for $\mathbf{t}$ may affect the rate of amplification in unexpected ways.

Last, we note that any invertible transform, including invertible non-linear functions may work with the distortion model in Eq. (1). For instance, a deep network may be trained to generate a transform for any given image, in which case $M = f(\mathbf{I})$.

3.2.2 The ideal matrix transform

Suppose there exists a optimal M^* that maximizes performance of pre-processing method. Then given a dataset of images and an objective function, numerical optimization can find a local optimum for M. Assuming the existence of M^* implies that each distinct image domain and processing objective has a distinct transform M^*.

We discuss our approach to estimating M^* using Covariance Matrix Adaptation [4] in the results section.

4 Competing Methods

We evaluate our pre-processing methods against several existing methods.

4.1 Contrast Stretching

The image is rescaled to include all intensities that fall within the 2nd and 98th percentiles(by default).

4.2 Histogram Equalization

Histogram Equalization spreads out the most frequent intensity values throughout range of pixel values. After equalization, the image has a roughly linear cumulative distribution function.

4.3 Adaptive Equalization

Adaptive equalization is used for local contrast enhancement. Local details can therefore be enhanced even in regions that are darker or lighter than most of the image.

4.4 Multi-Scale Retinex with Color Restoration (MSRCR)

Retinex methods were shown to have a duality with Eq. (1) [3], based on the property that pixel intensities are always increasing $\mathbf{J}^c(\mathbf{x}) \geq \mathbf{I}^c(\mathbf{x}) \geq \mathbf{A}^c$. The general effect of retinex processing on images is a "graying out" of the image due to brightening, known as a gray-world violation [1]. Color restoration by clipping the min and max intensities in each channel reduces gray-world violations.

MSRCR operates on a $[1, 256]$ normalized image $\mathbf{I}$ by averaging the log of $\frac{\mathbf{I}}{\text{blur}(\mathbf{I},\sigma)}$ for multiple scales of σ, multiplying by a color restoration term $\log(\alpha \mathbf{I} / \sum_{\mathbf{x}} \mathbf{I}(\mathbf{x}))$ for scalar α hyperparameter, and renormalizing. Division is element-wise.

$$\hat{\mathbf{J}}^c = \left(\frac{1}{n} \sum_{\sigma} \log \left(\frac{\mathbf{I}^c}{\text{blur}(\mathbf{I}^c, \sigma)} \right) \right) \log \left(\frac{\mathbf{I}^c \alpha}{\sum_c \mathbf{I}^c} \right) \tag{11}$$

$$\mathbf{J}^c = \text{renormalize}(\hat{\mathbf{J}}^c) \tag{12}$$

5 Evaluation

5.1 Indian Diabetic Retinopathy Dataset (IDRiD)

The Indian Diabetic Retinopathy Dataset [11] contains 54 retinal fundus images. Each image contains pixel level annotations segmenting healthy from diseased tissue for up to four lesion types: Microaneurysms (MA), Hemorrhages (HE), Hard Exudates (EX) and Soft Exudates (SE). We use the images to develop pre-processing methods and the labels to evaluate how well a pre-processing method separates healthy and diseased pixels.

5.1.1 Measuring separability with the Two Sample Kolmogorov-Smirnov Test

For the purpose of Diabetic Retinopathy detection, an image is enhanced if the distribution of pixel intensities for health pixels is significantly different from the distribution of diseased pixel intensities. To quantify how well a given retinal fundus image from the IDRiD dataset separates diseased from healthy pixels, we implement the Two-Sample Kolmogorov-Smirnov Test (KS) under the null hypothesis that two independent samples are drawn from the same continuous distribution. A low KS score (close to zero) indicates the healthy and diseased pixels are hard to separate, while a KS score close to one indicates the healthy and diseased pixels have distinct and non-overlapping generating distributions. Since the KS test evaluates each max absolute difference between the cumulative distributions of the two samples, we can evaluate each image channel independently and return the max KS score across channels.

Our KS statistic for separability is defined as

$$S_{\text{KS}}(\mathbf{I}^c) = \sup_{x \in \mathbf{I}^c} |H(x) - D(x)|, \tag{13}$$

Where $H(x)$ and $D(x)$ are the cumulative distributions of healthy and diseased pixels in the c^{th} channel of image $\mathbf{I}$.

5.1.2 Measuring consistency of pixel ranges

It is also highly desirable to ensure that all enhanced images in the dataset map the distribution of diseased pixels into the same range. For each enhanced image in the dataset, we have a distribution of

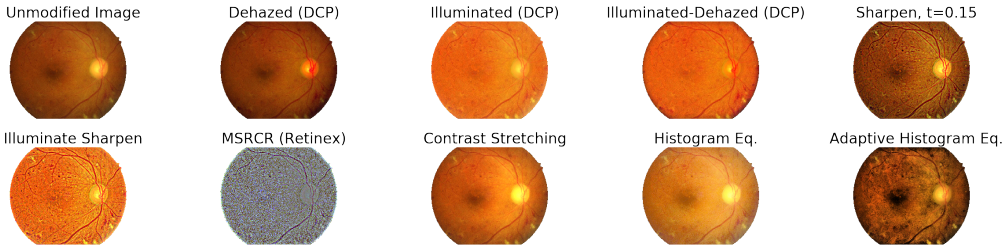

Figure 3: Caption

diseased pixel intensities $p(r, g, b|\text{diseased})$. According to the Central Limit Theorem, the distribution obtained by averaging these diseased pixel distributions over the dataset of images is approximately Gaussian. Therefore, the standard deviation of this Gaussian distribution estimates how much the range of diseased pixel values varies. We assign a consistency score to each pre-processing method based on standard deviation of the distribution of diseased pixels. Pre-processing methods with smaller consistency scores are desirable as they map diseased pixels into the same range across the dataset.

Our consistency statistic is defined as the standard deviation of the average distribution of pixel intensities across the dataset for each image channel.

$$\text{std}\left\langle \frac{1}{|D|}\sum_{i\in D}^{|D|} p(\text{channel}|\text{diseased}, \mathbf{I}, D)\right\rangle, \tag{14}$$

where D is the 54 image IDRiD dataset and $channel$ is either red, green or blue.

5.2 Results

5.2.1 Qualitative Analysis

We show results of various image enhancement methods in Fig. 3. The unmodified image is characteristic of IDRiD data: a good quality image with some blurring. Dehazing via DCP performs poorly on retinal fundus images because the blue channel in distorted fundus images is mostly zeros. Color Illumination yields characteristically brighter images, sometimes too bright. Illuminated-Dehazed applies illumination and then dehazing, and is visually appealing. Sharpening adds detail but suffers from salt and pepper noise on dark images. Illuminate-Sharpen is our best method, which has bright, sharp images with a lot of detail. MSRCR produces interesting results similar to Illuminate-Sharpen, though in a different color space and noisier. Contrast stretching produces pleasing results, but the fundus is still unevenly illuminated. Histogram equalization also produces visually appealing results, but the image appears too washed out. Adaptive histogram equalization makes the macula and several areas too dark; since the macula is the most important for vision and DR grading, any enhancement method should not make this area harder to see.

5.2.2 Quantitative Analysis

The top row of Fig. 4 shows results of our separability scores across all lesion types available in IDRiD, all pre-processing methods and all color channels. The most relevant evaluated lesions for DR are Microaneurysms (MA) and Hemorrhages (HE). Optic Disc does not benefit by pre-processing, as evidenced by its large separability in the red channel, and is largely irrelevant to DR. The green channel has the highest separability scores and our results show it is the most useful for retinal fundus analysis. The Illuminate-Sharpen method and Sharpen method have the highest and second highest separability score for MA and HE, which matches intuition that sharpening amplifies fine detail such as MA lesions and smaller HE lesions. Soft Exudates (SE) were best separated by standard color illumination, as shown in the green channel, while Hard Exudates (HE) were best separated by the Illuminate-Dehaze method. SE and EX are larger features, and we suspect that sharpening with a larger blur radius may improve the our two sharpen-based methods on these features.

The bottom row of Fig. 4 shows our consistency scores across lesion types, pre-processing methods and color channels for diseased pixels. The Illuminated-Sharpen method has the smallest (best) value

on all lesions and channels except the blue channel. The MSRCR and Sharpen methods also show a smaller standard deviation.

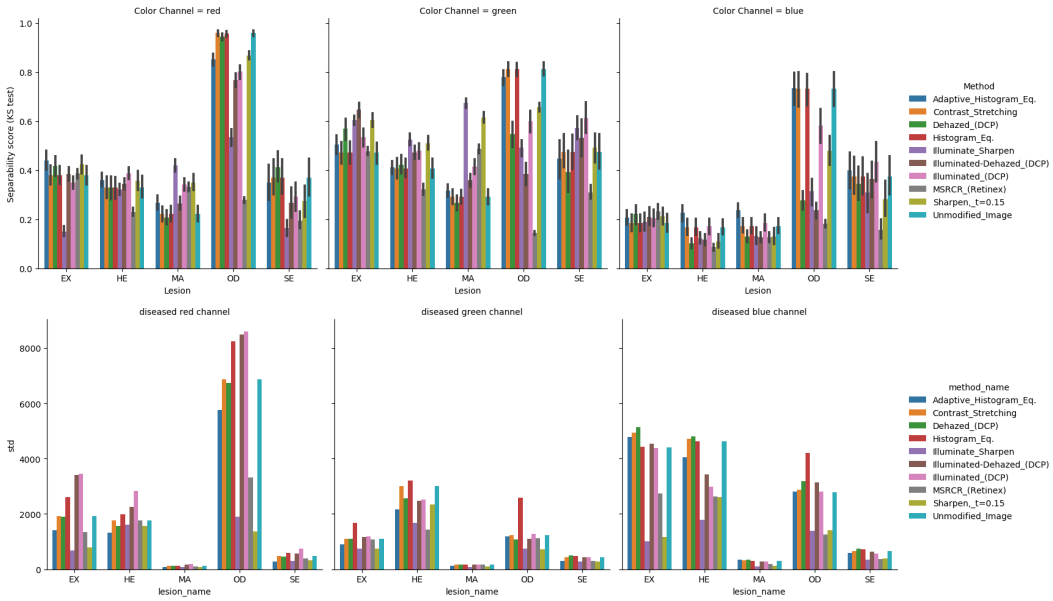

Figure 4: We evaluate different lesion types and methods. Microaneurysms (MA) and Hemorrhages (HE) are the most relevant lesions to DR. **Top:** Separability scores. Separability closer to 1 is better. **Bottom:** Consistency scores. Consistency closer to 0 is better.

5.2.3 Optimal Matrix Transform Prior

We used an evolutionary optimization method, Covariance Matrix Adaptation, to search for a transform matrix maximizing the average Kolmogorov-Smirnov score of a 3D histogram of the transformed image $\mathbf{I}M$ across all lesion types. We found a matrix that increased the score of the unmodified images from 0.4 to 0.72, but we were not able to obtain results incorporating this matrix with our Illuminate-Sharpen method, and we did not evaluate it against our other pre-processing methods.

5.2.4 How the evaluation led to development of Illuminate-Sharpen

The Illuminate-Sharpen method applies Illuminate and Sharpen methods sequentially. This was motivated by the fact Sharpen was the best performing method, except on Soft Exudates, where Illuminate outperformed it. We also noticed in qualitative plots that the Sharpen method performed less well on poorly illuminated images. Thus, the combination of our results significantly improved performance.

6 Conclusion

In this paper, we provide a way to explain and develop image enhancement pre-processing methods for Diabetic Retinopathy lesion segmentation. We provide consistency and separability scores to evaluate our methods. We present dehazing theory as a general framework for image enhancement that amplifies differences in color between corresponding pixels. We use the theory to derive the unsharp mask algorithm, find an ideal color space for retinal fundus images, and develop a novel single-image enhancement method which we call Illuminate-Sharpen that increases separability, is more consistent than competing methods, and does not use image labels or require a dataset.

References

[1] Qiang Chen, Xin Xu, Quansen Sun, and Deshen Xia. A solution to the deficiencies of image enhancement. *Signal Processing*, 90(1):44 – 56, 2010.

[2] Raanan Fattal. Single image dehazing. *ACM Trans. Graph.*, 27(3):72:1–72:9, August 2008.

[3] Adrian Galdran, Alessandro Bria, Aitor Alvarez-Gila, Javier Vazquez-Corral, and Marcelo Bertalmio. On the duality between retinex and image dehazing. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun 2018.

[4] Nikolaus Hansen. The CMA evolution strategy: A tutorial. *CoRR*, abs/1604.00772, 2016.

[5] K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(12):2341–2353, Dec 2011.

[6] K. He, J. Sun, and X. Tang. Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(6):1397–1409, June 2013.

[7] Sungmin Lee, Seokmin Yun, Ju-Hun Nam, Chee Sun Won, and Seung-Won Jung. A review on dark channel prior based image dehazing algorithms. *EURASIP Journal on Image and Video Processing*, 2016(1):4, Jan 2016.

[8] Srinivasa G. Narasimhan and Shree K Nayar. Vision and the atmosphere. *IJCV*, 48(3):233–254, January 2002.

[9] Luis Perez and Jason Wang. The effectiveness of data augmentation in image classification using deep learning, 2017.

[10] Maria Petrou and Costas Petrou. *Image Processing: The Fundamentals*, pages 357–360. John Wiley & Sons, Ltd, Chichester, UK, 2011-01-27.

[11] Prasanna Porwal; Samiksha Pachade; Ravi Kamble; Manesh Kokare; Girish Deshmukh; Vivek Sahasrabuddhe and Fabrice Meriaudeau. Indian diabetic retinopathy image dataset (idrid), 2018.

[12] B. Savelli, A. Bria, A. Galdran, C. Marrocco, M. Molinara, A. Campilho, and F. Tortorella. Illumination correction by dehazing for retinal vessel segmentation. In *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, pages 219–224, June 2017.

[13] Asim Smailagic, Anupma Sharan, Pedro Costa, Adrian Galdran, Alex Gaudio, and Aurélio Campilho. Learned pre-processing for automatic diabetic retinopathy detection on eye fundus images. In Fakhri Karray, Aurélio Campilho, and Alfred Yu, editors, *Image Analysis and Recognition*, pages 362–368, Cham, 2019. Springer International Publishing.

[14] R. T. Tan. Visibility in bad weather from a single image. In *2008 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, June 2008.