Log-ShadowFormer: Exploring Shadow Removal in Log-RGB Space

Elizabeth Witten
Northeastern University
360 Huntington Ave, Boston, MA 02115
witten.e@northeastern.edu

Srijha Thammareddy
Northeastern University
360 Huntington Ave, Boston, MA 02115
thammareddy.s@northeastern.edu

Abstract

Data quality is a crucial element of machine learning performance. In computer vision, nearly all benchmark datasets have the same quality concern: compressed, nonlinear sRGB images. In this paper, we explore the impact of sRGB, linear-RGB, and log-RGB images in the domain of shadow removal. By working in the linear- and log-RGB domains, the shadow removal task is simpler. We present Log-ShadowFormer, a log-RGB variant of the state-of-the-art ShadowFormer, and demonstrate its robustness on high quality data. Further, we address nonlinearity issues in the existing ISTD dataset. In response, we introduce an early version of RAWShadowIO, a new dataset collected in RAW format, along with a clear processing pipeline from RAW to linear and sRGB images.

1. Introduction

In machine learning, model performance is highly dependent on data quality. In the computer vision space, nearly all prominent datasets use compressed, nonlinear data in the form of sRGB images [1]. This nonlinearity is part of the typical image processing pipeline, and it results in images that are more pleasing to the human eye.

In many practical applications, such as security systems, raw sensor data can be directly passed to a waiting application, rendering sRGB processing completely unnecessary. More importantly, processing images in this way breaks the laws of physics that govern the interaction of light and material. When networks are trained on such data, features such as color and intensity become less informative.

One computer vision task that is tied heavily to these laws of physics is shadow removal. Shadows often make vision tasks, including image segmentation, object detection and tracking [2, 3, 4], more difficult. As such, shadow removal, which aims to restore areas of an image where light has been obstructed, is often used as a preprocessing step before downstream tasks.

By framing shadow removal as a relighting problem, a restored image can be obtained by applying a scale factor to the shadow regions [5]. Given an RGB image as input, that factor would be a multiplicative constant in the linear domain. Given a log-RGB image, the scale factor would be an additive constant in the log domain.

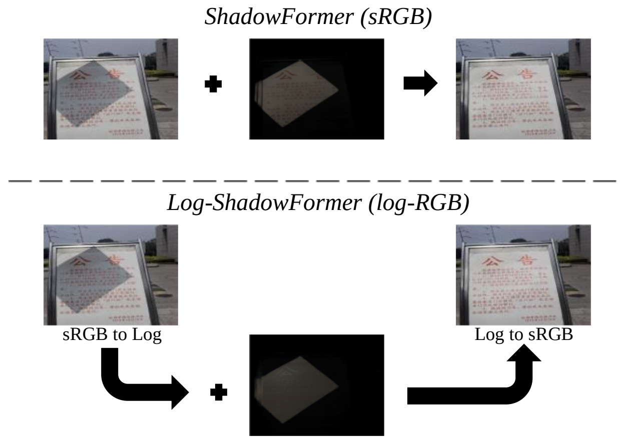


Figure 1. Comparison of relighting problem in sRGB and log-RGB space

Using this formulation, we propose that the use of linear-RGB and log-RGB images as input results in a simpler problem and solution. Whereas the use of sRGB images introduces mathematical complications, relighting shadows in log-RGB space is reduced to an addition problem, which can be modeled using a residual architecture. This process is illustrated in Figure 1.

In this paper, we investigate the use of log-RGB images with the ShadowFormer [6] architecture, a recent state-of-the-art transformer model that employs a global residual connection to relight an input image. We present a version of this model trained in log-RGB space, which we refer to as Log-ShadowFormer. We show that, compared to the original ShadowFormer, Log-ShadowFormer learns a simpler and more robust model for shadow removal.

Additionally, we perform an analysis of the benchmark dataset, ISTD [7]. We show that both the training and test sets contain images that have been greatly affected by nonlinear processing. First, we manually investigate the normalized illuminant spectral ratios, in pseudolinear-RGB space, for the best and worst performing test images. We find that these ratios are not always invariant across material, suggesting that the inverse sRGB operation was unable to return the images to a true linear-RGB space. Then, we compute the percentage of saturated pixels in all ISTD images. Saturated pixels are problematic because they have been clipped, in the case of 8-bit images, to 0 or 255. This lost information further obscures the illuminant spectral ratios.

To address the nonlinearity issues with the current shadow removal datasets, we incorporate an early version of RAWShadowIO, a new shadow removal dataset collected by Srijha Thammareddy as RAW images. In addition to bringing new indoor and outdoor scenes and lighting conditions, we also provide a pipeline to process the images from RAW to linear and sRGB, so that every transformation applied to the images is transparent.

The main contributions of this work are:

1. We present initial results showing that Log-ShadowFormer learns a robust model of shadow removal.
2. We analyze the benchmark ISTD [7] dataset for issues with saturation and nonlinearity.
3. We introduce an early version of RAWShadowIO and demonstrate its effectiveness with Log-ShadowFormer.

2. Related Work

2.1. Shadow Illumination Models

When considering shadow images, it is useful to be able to decompose the image into semantically meaningful elements or regions. Porter and Duff [8] introduced the concept of *matte* channels, which are alpha channels designating the shape of an element. In shadow removal, shadow mattes can be used to separate a shadow image into shadow and non-shadow regions:

$$I = I_m \odot I_s + (1 - I_m) \odot I_{ns} \tag{1}$$

Here, I_s and I_{ns} represent the shadow and non-shadow regions, respectively. I_m represents the shadow matte. Because the shadow matte can provide valuable information, estimating the shadow matte is often the first step in the shadow removal process.

Another way to break down an image is by considering the physical illumination model. According to Retinex Theory [9], the human vision system is able to normalize color perception, even under differing illumination conditions. Following [10], an image can be decomposed into illumination (L) and reflectance (R) components, using pixel-wise multiplication (denoted by $\odot$):

$$I = L \odot R \tag{2}$$

Illumination refers to the light that falls on a surface, while reflectance refers to the inherent properties of a material to reflect light. Importantly, reflectance is a property that is invariant to changes in illumination.

Arbel and Hel-Or [5], extend Equation (2) by noting that shadow regions are formed by a reduction in the illumination field, which can be represented by a multiplicative scalar:

$$I = L \odot R \odot C \tag{3}$$

Here, C contains the factors to correct the shadow pixels. In keeping with [8], the values of C within the shadow region can also be considered the shadow matte.

Further, Arbel and Hel-Or [5] move this equation into the log domain, where a shadow is now the result of an additive change in intensities:

$$\log I = \log R + \log L + \log C \tag{4}$$

These shadow and illumination models are foundational to the classical shadow removal techniques that follow, as well as several of the deep learning methods.

2.2. Classical Shadow Removal Methods

Automated shadow removal is a challenging task. A successful shadow removal technique should be able to preserve color and texture information, eliminate traces of the shadow edges, and produce results quickly to be used in preprocessing [11]. These criteria are especially difficult to achieve when given a single input image. Shadows come in many shapes, colors, and intensities. Even within a single shadow area, the shadow adjustment is often non-uniform.

Approaches across major classical techniques can be organized into a few categories [11]. Among these, the relighting methods most closely match the approach used in this paper. This general strategy aims to identify a scaling factor which can be applied to brighten shadow pixels. In one early approach, Fredembach and Finlayson [12] demonstrated a simple algorithm which minimizes variation in each color channel across a shadow edge to find a constant used relight the shadow region. Another approach, by Arbel and Hel-Or [5], uses the log domain to find an additive scale factor, while incorporating surface geometry into their scale factor computation.

Other classical methods include color transfer methods [13], which use statistical properties to transfer color information from non-shadow areas to shadow areas. Reintegration methods [14] work on the pixel level, using image gradients along shadow edges to reintegrate back to a shadow-free image, while patch-based methods [15], on the other hand, use correspondences between larger image patches.

2.3. Deep Learning Shadow Removal Methods

More recent work has shown excellent performance using deep learning. While the classical methods may struggle with the non-ideal conditions of real-world images, large-scale training data and deep architectures enable these models to better handle complex and non-uniform shadows.

DeshadowNet [16] is a fully convolutional deep neural network that follows the relighting approach to generate a shadow matte. DeshadowNet follows a similar formulation to [5], where the shadow image is the pixel-wise product of a shadow-free image and a shadow matte:

$$I_s = I_{ns} \odot C \tag{5}$$

During training, the shadow and shadow-free image pairs are converted to log space, so Equation (5) becomes:

$$\log I_s = \log I_{ns} + \log C \tag{6}$$

In [17], Le and Samaras incorporate the shadow matte into a two-network system based on a simple illumination model. The first network, based on ResNeXt [18], applies a linear transformation to relight the entire image, using a scaling factor and an additive constant. Then, the second network, based on U-Net [19], predicts and applies a shadow density matte to produce the shadow-free image. The use of the matting layer smooths out the boundary artifacts that may occur when using binary shadow masks.

ShadowFormer [6] also uses a classical illumination model as the foundation for its lightweight transformer architecture. ShadowFormer uses a combination of Equations (1) and (2) to model a shadow image:

$$I_s = I_m \cdot L_s \cdot R + (1 - I_m) \cdot L_{ns} \cdot R \tag{7}$$

Here, I_m is a mask denoting the shadow region of the image, L_s and L_{ns} are the illumination of the shadow and non-shadow regions, respectively, and R is the reflectance.

This model reflects two important qualities. First, the shadow image contains varying illumination conditions, while the shadow-free image ideally has spatially consistent illumination. Second, reflectance is a strong source of global contextual correlation between the shadow and non-shadow regions.

These properties are well-suited for the vision transformer [20, 21, 22], which is able to take advantage of long-range contextual information. Within its bottleneck stage, ShadowFormer employs a Shadow-Interaction Module with Shadow-Interaction Attention, which reweights the attention map to emphasize correlations between the shadow and non-shadow regions.

3. Log-ShadowFormer

3.1. Data Pipeline

As input, ShadowFormer expects a shadow image I_s and its shadow mask I_m. The shadow mask has positive pixels in the shadow region, and zero pixels elsewhere. For training and evaluation, we provide image triplets, as in the benchmark ISTD [7] dataset. Each triplet consists of the shadow image I_s, the shadow mask I_m, and the target shadow-free image I_{ns}.

Following the work done with the adjusted ISTD (ISTD+) dataset [17], we additionally apply target adjustment to each shadow-free image, in order to reduce illumination inconsistencies between the shadow and shadow-free images. This transformation is always performed in linear- or pseudo-linear space. Our implementation of target adjustment was provided by Heather Fryling.

Let $I_k(x, y)$ represent the intensity value of color channel $k \in \{R, G, B\}$ at pixel (x, y) in image I. Then, for each k, we compute scaling factor γ_k as the ratio between the noisy and clean average non-shadow pixel values:

$$\gamma_k = \frac{\sum_{(x,y)\in I_m} I_{s_k}(x, y)}{\sum_{(x,y)\in I_m} I_{ns_k}(x, y)} \tag{8}$$

Then, the adjusted target image I'_{ns} is:

$$I'_{ns_k} = \gamma_k \cdot I_{ns_k} \tag{9}$$

During training, we follow the original ShadowFormer paper [6] in applying basic geometric data augmentation: random crop to 320 by 320, random rotation, and random flip.

We also experiment with applying intensity and color augmentation after the target adjustment and geometric transformations. We generate three static test sets from ISTD with: (1) 50% intensity augmentation; (2) 50% color balance augmentation; and (3) 25% intensity augmentation, 25% color balance augmentation, 25% both augmentations.

For sRGB images, we convert to pseudolinear-RGB by applying an inverse sRGB transformation. We convert to pseudolog- or log-RGB by taking the natural logarithm of a pseudolinear- or linear-RGB image. We store all images using 32-bit floating point values.

3.2. ShadowFormer Architecture

In this section, we provide an overview of the ShadowFormer architecture. Full details may be found in the original ShadowFormer paper [6].

Given input (I_s, I_m), ShadowFormer applies a linear projection to obtain a low-level feature embedding X_0. The embedding X_0 is fed into the transformer-based encoder and decoder.

In the bottleneck stage, the Shadow-Interaction Module (SIM) is used to calculate the global contextual correlation across both spatial and channel dimensions. Within the SIM, shadow-interaction attention uses the shadow mask to emphasize the similarities between shadow and non-shadow regions in the image.

Following the decoder layers, ShadowFormer applies a

Table 1. Mean evaluation results on each dataset for ISTD-trained models

Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
ISTD									
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	**31.576**	**0.957**	**4.002**	**28.163**	**0.990**	**6.392**	33.030	**0.969**	3.556
Linear-ShadowFormer	31.474	**0.957**	4.093	27.891	**0.990**	6.679	32.978	**0.969**	3.622
Log-ShadowFormer	31.500	**0.957**	4.052	27.811	**0.990**	7.205	**33.319**	**0.969**	**3.504**
RAWShadowIO (linear TIFF)									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.361	0.857	9.985	19.110	**0.974**	**19.198**	27.539	0.887	8.309
Linear-ShadowFormer	23.193	0.844	12.044	18.756	0.973	21.319	25.308	0.874	10.499
Log-ShadowFormer	**24.620**	**0.860**	**9.737**	**19.262**	**0.974**	19.884	**27.918**	**0.889**	**7.914**
RAWShadowIO (sRGB JPG)									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.080	0.855	7.583	19.396	**0.974**	**17.759**	28.547	0.884	5.896
Linear-ShadowFormer	25.038	0.856	8.095	19.476	**0.974**	18.723	28.379	0.885	6.178
Log-ShadowFormer	**25.248**	**0.859**	**7.570**	**19.661**	0.974	**18.307**	**28.757**	**0.888**	**5.647**

linear projection to obtain the residual image I_r. The final output is computed using the global residual connection as: $\hat{I} = I_s + I_r$.

ShadowFormer uses an L1 loss function:

$$\mathcal{L}(I_{gt} - \hat{I}) = \|I_{gt} - \hat{I}\| \tag{10}$$

Where $\hat{I}$ is the output image, and I_{gt} is the ground truth shadow-free image.

3.3. Modifications to ShadowFormer

During training, we do not apply mix-up augmentation for linear- or log-RGB images. Mix-up is a data augmentation technique where pairs of inputs within a batch are blended using a random mixing coefficient sampled from a beta distribution. While mix-up augmentation improves generalization for the sRGB model, it further breaks the physical relationships that Log-ShadowFormer aims to learn. In our experiments, we see that Log-ShadowFormer achieves good generalization without this augmentation.

We make no changes to the ShadowFormer architecture itself.

4. Supplementary Dataset: RAWShadowIO

RAWShadowIO is a new unpublished dataset, collected by Srijha Thammareddy as RAW images. It consists of 323 high-quality raw images: 228 in the training set, and 95 in the test set. There are 38 scenes in total, with a variety of lighting conditions: 19 indoor scenes, and 19 outdoor scenes.

The value in the RAWShadowIO dataset comes from it being collected directly as RAW images, so that later transformations of the data are well-defined for the whole dataset. In addition to the dataset, we provide a RAW image pipeline, which outputs the images in a format ready for use[1].

The first step in the pipeline is to use the Python library rawpy to process the RAW images into linear-RGB images. For our post-processing parameters, we use a gamma of (1,1), an autobright threshold of 0.001, and a bit depth of 8.

Then, the linear target-shadow pairs are aligned using an affine transformation. This step is used to mitigate any small changes in camera angle between the collection of the two images. After alignment, we apply a border crop to mitigate noise, and then resize to 640 by 480, the input dimensions for ShadowFormer. The final images are each saved in three formats: linear TIFF files, sRGB PNGs, and sRGB JPGs.

Using the final linear target-shadow pairs, we generate shadow masks using the process described in [7]. We threshold the target-shadow difference image to create an initial binary mask. We then clean up the mask using morphological operations. By manual inspection, we identify any masks that were improperly generated and redraw them by hand.

4.1. Combined Dataset

To better investigate the impact of the highly saturated ISTD images, we created a combined dataset with the RAWShadowIO dataset and ISTD, which we will refer to as ISTD-RAWShadowIO. We maintain the existing train and set splits for each dataset.

[1] The RAW image pipeline is available on GitHub: https://github.com/wittene/raw-shadow-pipeline.

Table 2. Mean evaluation results on each dataset for ISTD-RAWShadowIO-trained models

Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
ISTD									
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	**31.486**	**0.957**	**4.072**	**27.716**	**0.990**	**6.974**	**33.147**	**0.969**	3.542
Log-ShadowFormer	31.238	0.956	4.072	27.229	0.989	7.150	33.140	**0.969**	**3.513**
RAWShadowIO (linear TIFF)									
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	27.012	**0.875**	8.299	22.795	**0.978**	14.322	**29.204**	**0.900**	7.252
Log-ShadowFormer	**27.059**	0.872	**8.284**	**22.934**	0.977	**14.028**	29.154	0.899	**7.221**
RAWShadowIO (sRGB JPG)									
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	27.795	**0.874**	**6.037**	23.445	**0.978**	**12.157**	**30.157**	**0.899**	**4.964**
Log-ShadowFormer	**27.845**	0.872	6.123	**23.448**	0.977	12.162	30.126	**0.899**	4.978

From RAWShadowIO, we use the sRGB PNG images. From ISTD, we include only those images with less than 0.5% saturation—we call this subset ISTD-Unsaturated.

The resulting ISTD-RAWShadowIO dataset has 1507 training images and 612 test images, for a total of 2119 images. In the training set, 1279 images are from ISTD, and 228 images are from the supplementary dataset. In the test set, 517 images are from ISTD, and 95 images are from the supplementary dataset.

5. Experiments

5.1. Implementation

Networks. We use a modified version of the official ShadowFormer [6] implementation in PyTorch [2]. Following the original paper, we train each model using an AdamW optimizer with a momentum of $(0.9, 0.999)$ and a weight decay of 0.02. We use a cosine annealing learning rate scheduler, starting with an initial learning rate of 2×10^{-4} and decreasing to 1×10^{-6}. We also experiment with other hyperparameter settings, but we find that the published hyperparameters perform the best in our experiments.

Of the two proposed ShadowFormer variants, we use the large model in our experiments. This model has a four-scale encoder-decoder structure, and the first embedding dimension has size 32.

For our Log-ShadowFormer variations, we set the maximum log-RGB pixel value to 256.

Experimental Conditions. We train three experimental models using the same training data: ShadowFormer (using sRGB images), Linear-ShadowFormer (using pseudolinear-RGB images), and Log-ShadowFormer (using pseudolog-RGB images).

For each image type, we train several variations. Our base model for each image type applies target adjustment in linear space, plus the mix-up augmentation for the sRGB model. We also try variations where we apply target adjustment in sRGB space, before the inverse sRGB transformation. Finally, we experiment with intensity and color balance data augmentation for ShadowFormer and Log-ShadowFormer.

In Section 5.2, we compare the best-performing variations. For detailed results for all experiments run, see Appendix A through Appendix D.

Datasets. The primary dataset we use for both training and testing is ISTD [7], which consists of 1330 training triplets and 5404 test triplets. We also work with RAWShadowIO, as described in Section 4.1.

Evaluation Metrics. Following previous works [6, 7, 16, 17], we evaluate the root mean square error (RMSE) between the shadow and target images in LAB color space. In sRGB color space, we compute the peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM). For RMSE, lower values represent better results. For PSNR and SSIM, higher values represent better results.

Additionally, we compute these three metrics for the overall image, for the shadow region only, and for the non-shadow region.

Prior work presents evaluation over the mean of the test set. We also present evaluation results over the following: mean of the best-performing 25% of results, mean of the worst-performing 25% of results, median, and trimean.

5.2. Results

Models Trained on ISTD. In Table 1, we present the best-performing variations for each image type based on the ISTD test set, referred to simply as ShadowFormer, Linear-ShadowFormer, and Log-ShadowFormer.

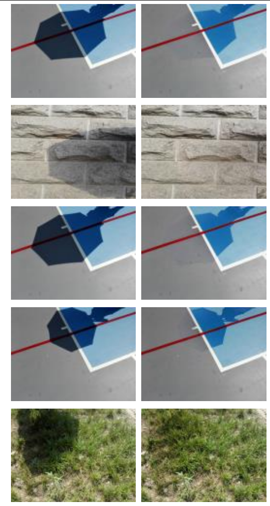
Of the ShadowFormer variations, the model trained with

[2] Log-ShadowFormer is available on GitHub: https://github.com/wittene/log-shadowformer.

Table 3. Best performing ISTD test inputs
(Top: Highest SPSNR, Bottom: Lower SPSNR)

ShadowFormer	Log-ShadowFormer
	

Table 4. Worst performing ISTD test inputs
(Top: Lowest SPSNR, Bottom: Higher SPSNR)

ShadowFormer	Log-ShadowFormer
	

intensity augmentation performed the best. For Linear-ShadowFormer and Log-ShadowFormer, the best variations applied target adjustment prior to the inverse sRGB transformation, with no additional data augmentation. The detailed results, including the best- and worst-25% mean, the median, and the trimean metrics, for all variations can be found in Appendix A through Appendix D.

When tested on ISTD, ShadowFormer performed best in the shadow regions and in the overall image, while Log-ShadowFormer had better performance in the non-shadow regions. Linear-ShadowFormer was generally a little bit worse than ShadowFormer.

However, when tested on images from RAWShadowIO, Log-ShadowFormer generally outperformed ShadowFormer and Linear-ShadowFormer, for both the linear images and the compressed sRGB images. These results indicate that Log-ShadowFormer generalizes to unseen scenes better than ShadowFormer.

Models Trained on ISTD-RAWShadowIO. We train the basic ShadowFormer and Log-ShadowFormer variations using the ISTD-RAWShadowIO combined dataset. The mean evaluation results are shown in Table 2.

As with the ISTD-trained models, ShadowFormer performs best on the ISTD test set, while Log-ShadowFormer performs slightly better on the RAWShadowIO test sets. These are very promising results, as they indicate that Log-ShadowFormer may outperform ShadowFormer when given higher-quality data.

Models Trained on ISTD-Unsaturated. We also train the same basic ShadowFormer and Log-ShadowFormer variations using ISTD-Unsaturated. The detailed results can be found in Appendix A and Appendix C. We find that, as with the ISTD-trained models, ShadowFormer outperforms Log-ShadowFormer. We also find that the models trained on the full ISTD dataset have better performance, on average, over the unsaturated subset. This supports our hypothesis that highly saturated images are more difficult, possibly due to lost relationships within the data.

6. Analysis

In the following sections, we analyze the ISTD-trained ShadowFormer and Log-ShadowFormer models.

6.1. Error Analysis

Ranked Shadow Region PSNR. To better understand the strengths and weaknesses of each model, we rank each test image based on shadow region PSNR (SPSNR).

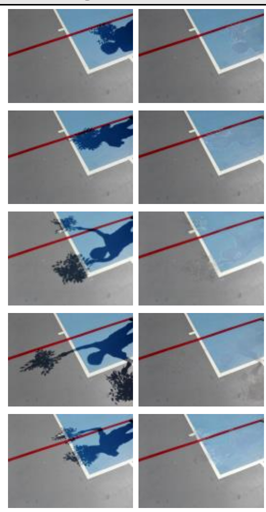
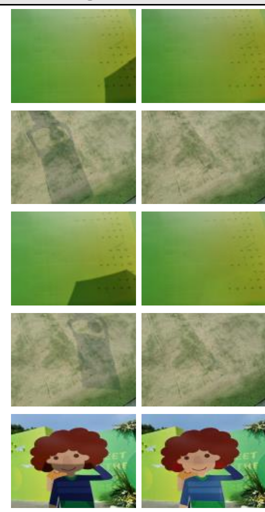
In Table 3, we show the best performing ISTD test images. For both models, we see that the images that achieve the highest SPSNR are generally gray with a smooth surface texture.

In Table 4, we show the worst performing ISTD test images. In contrast with the best performing inputs, these images tend to be brightly colored, or have a highly textured surface.

Comparing ShadowFormer and Log-ShadowFormer. Because some scenes are difficult for both ShadowFormer and Log-ShadowFormer, we also look at their relative performances. To do this, we look at the difference between ShadowFormer and Log-ShadowFormer SPSNR for each test image. These comparisons are shown in Table 5.

We see that, relative to ShadowFormer, Log-ShadowFormer had the greatest issue with one particular tennis court scene. In fact, of the worst-25 images by

Table 5. Relative performance with ISTD test inputs
(Top: Largest SPSNR difference,
Bottom: Lower SPSNR difference)

ShadowFormer with higher SPSNR	Log-ShadowFormer with higher SPSNR
	

SPSNR difference, 23 of them belong to this same scene.

It does not appear that color by itself is a weakness of Log-ShadowFormer, as the images with the greatest SPSNR differences in Log-ShadowFormer's favor are also brightly colored.

In the next section, we look at the data in more detail to determine whether there are any quality issues in the ISTD dataset that may explain these performance differences.

6.2. Data Analysis

Illuminant Spectral Ratio Analysis. To guarantee that the pseudolinear- and pseudolog-RGB images are high-quality, the original sRGB image must not have any nonlinear processing, aside from the sRGB transformation. After applying the inverse sRGB transformation, we check the pseudolinear-RGB images for linearity using illuminant spectral ratio analysis.

We compute the illuminant spectral ratio S for a material x using a lit pixel x_{ns} and a shadowed pixel x_s on the same material:

$$S = \frac{x_s}{x_{ns} - x_s} \tag{11}$$

To compare illuminant spectral ratios across materials, we use the normalized illuminant spectral ratios. Because this ratio is dependent only on the direct and ambient light, the illuminant spectral ratio should be invariant to material in a linear image.

Images used in this analysis are detailed in Appendix E. In the pseudolinear-RGB ISTD test set, we found that the worst-performing images had inconsistent normalized

Table 6. Residual comparison

Input	ShadowFormer	Log-ShadowFormer
		

spectral ratios. This indicates that there is still nonlinearity in the pseudolinear-RGB images, and that there are illumination relationships that are unrecoverable with the inverse sRGB transform.

Saturation Analysis. Another thing that impacts data quality is pixel saturation. If pixels are saturated, then their values are clipped, and illumination relationships are lost. Since we are working with 8-bit images, we define a saturated pixel as having a 0 or 255 value in any color channel. We define a highly saturated image as having at least 0.5% pixels saturated.

We find that 51 out of 1330 ISTD training images, and 23 out of 540 ISTD test images, are highly saturated.

6.3. Residual Comparison

Table 6 compares the ShadowFormer and Log-ShadowFormer residuals for a selection of ISTD and RAWShadowIO test images. The top three examples belong to ISTD, while the bottom three examples belong to RAWShadowIO.

We see that the ShadowFormer residuals are very detailed, with colors that match the input image. Meanwhile, the Log-ShadowFormer residuals are typically more uniform.

However, it should be noted that there is still more variation in the log residuals than expected, particularly for the ISTD images. Given more data, as RAWShadowIO continues to be collected, it will be interesting to see if the log residuals become more uniform.

7. Conclusion

In this work, we explore the log-RGB domain for shadow removal through in-depth experimentation with ShadowFormer and Log-ShadowFormer. We present initial findings on the robustness of training in the log-RGB domain. We also demonstrate the importance of using a high-quality data pipeline by uncovering saturation and nonlinearity issues in the ISTD dataset, and we introduce RAWShadowIO as a high-quality supplement.

There is plenty of room in this project for future work. Right now, the RAWShadowIO dataset is in its early stages. As the dataset grows, more experiments can be done where ShadowFormer and Log-ShadowFormer models are trained entirely on RAWShadowIO. These experiments will give more information on the benefits of using true linear-RGB images, run through a well-defined image pipeline, rather than pseudolinear-RGB images with unknown processing applied to them.

There is also room to experiment with altering the Log-ShadowFormer architecture. For example, one concept that we explored was altering Log-ShadowFormer to output a six-channel residual. In this model, three channels capture a smooth body reflection residual, and three channels collect any remaining corrections, such as specular reflection. From our initial exploration of this, additional hyperparameter tuning is required for the model to learn.

As prior work has shown success using the log-RGB domain with CNN architectures, it may also be interesting to replace Log-ShadowFormer's first linear projection with a small CNN.

8. Acknowledgements

I'd like to thank Professor Bruce Maxwell for being a wonderful advisor and for guiding me along this project. I'd also like to thank Heather Fryling—for sharing her initial Log-ShadowFormer implementation and helping me to kick off my own work, and for being a great resource as I worked with the ISTD dataset—and Srijha Thammareddy—for collecting and sharing the images for the RAWShadowIO dataset.

9. References

[1] H. Fryling, "Rethinking image formats for computer vision," Northeastern University, 2024.

[2] W. Zhang, X. Zhao, J.-M. Morvan and L. Chen, "Improving shadow suppression for illumination robust face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 41, no. 3, pp. 611-624, 2018.

[3] R. Cucchiara, C. Grana, M. Piccardi, A. Prati and S. Sirotti, "Improving shadow suppression in moving object detection with HSV color information," in *IEEE Intelligent Transportation Systems (ITSC)*, 2001.

[4] I. Mikic, P. C. Cosman, G. T. Kogut and M. M. Trivedi, "Moving shadow and object detection in traffic scenes," in *International Conference on Pattern Recognition (ICPR)*, 2000.

[5] E. Arbel and H. Hel-Or, "Texture-preserving shadow removal in color images containing curved surfaces," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2007.

[6] L. Guo, S. Huang, D. Liu, H. Cheng and B. Wen, "ShadowFormer: global context helps shadow removal," in *AAAI Conference on Artificial Intelligence*, 2023.

[7] J. Wang, X. Li and J. Yang, "Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018.

[8] T. Porter and T. Duff, "Compositing digital images," in *Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, 1984.

[9] E. H. Land, "The retinex theory of color vision," *Scientific American,* vol. 237, no. 6, pp. 108-129, 1977.

[10] H. Barrow, J. Tenenbaum, A. Hanson and E. Riseman, "Recovering intrinsic scene characteristics," *Computer Vision Systems,* vol. 2, pp. 3-26, 1978.

[11] S. Murali, V. Govindan and S. Kalady, "A survey on shadow removal techniques for single image," *International Journal of Image, Graphics and Signal Processing,* vol. 8, no. 12, 2016.

[12] C. Fredembach and G. Finlayson, "Simple shadow removal," in *International Conference on Pattern Recognition*, 2006.

[13] E. Reinhard, M. Adhikhmin, B. Gooch and P. Shirley, "Color transfer between images," *IEEE Computer Graphics and Applications,* vol. 21, no. 5, pp. 34-41, 2001.

[14] G. Finlayson and C. Fredembach, "Fast re-integration of shadow free images," in *Color and Imaging Conference*, 2004.

[15] L. Zhang, Q. Zhang and C. Xiao, "Shadow Remover: Image shadow removal based on illumination recovering optimization," *IEEE Transactions on Image Processing,* vol. 24, no. 11, pp. 4623-4636, 2015.

[16] L. Qu, J. Tian, S. He, Y. Tang and R. W. H. Lau, "DeshadowNet: A multi-context embedding deep network for shadow removal," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[17] H. Le and D. Samaras, "Shadow removal via shadow image decomposition," in *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019.

[18] S. Xie, R. Girshick, P. Dollár, Z. Tu and K. He, "Aggregated residual transformations for deep neural networks," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

[19] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, 2015.

[20] H. Chen, Y. Wang, T. Guo, C. Xu, Y. Deng, Z. Liu, S. Ma, C. Xu, C. Xu and W. Gao, "Pre-trained image processing transformer," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[21] Z. Zhang, Y. Jiang, J. Jiang, X. Wang, P. Luo and J. Gu, "STAR: A structure-aware lightweight transformer for real-time image enhancement," in *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021.

[22] Z. Wang, X. Cun, J. Bao, W. Zhou, J. Liu and H. Li, "Uformer: A general u-shaped transformer for image restoration," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022.

1. Appendix A

Detailed Experimental Results on ISTD

1.1. Models Trained on ISTD

Table 7. ISTD-trained, with target adjustment in sRGB space; experimental results on ISTD

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.520	0.957	4.077	28.015	0.990	6.616	33.041	0.969	3.613
Linear-ShadowFormer	31.474	0.957	4.093	27.891	0.990	6.679	32.978	0.969	3.622
Log-ShadowFormer	31.500	0.957	4.052	27.811	0.990	7.205	33.319	0.969	3.504
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.968	0.989	2.085	33.368	0.998	3.205	38.868	0.994	1.883
Linear-ShadowFormer	36.852	0.989	2.050	33.325	0.998	2.940	38.715	0.994	1.879
Log-ShadowFormer	37.064	0.989	2.027	33.262	0.998	3.093	39.205	0.994	1.820
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.325	0.897	7.586	21.707	0.974	12.927	26.588	0.921	6.741
Linear-ShadowFormer	25.447	0.900	7.676	22.151	0.974	13.448	26.550	0.923	6.801
Log-ShadowFormer	24.997	0.898	7.622	21.343	0.973	15.389	26.862	0.924	6.533
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.823	0.970	3.360	28.549	0.994	5.287	33.337	0.979	2.983
Linear-ShadowFormer	31.752	0.962	3.357	27.964	0.993	5.222	33.239	0.979	2.952
Log-ShadowFormer	31.793	0.970	3.283	28.225	0.994	5.251	33.646	0.980	2.877
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.315	0.972	3.228	28.870	0.994	5.034	33.548	0.980	2.872
Linear-ShadowFormer	32.002	0.970	3.279	27.991	0.994	5.181	33.432	0.980	2.855
Log-ShadowFormer	32.124	0.971	3.175	28.274	0.994	5.172	33.594	0.981	2.767

Table 8. ISTD-trained, with target adjustment in linear space; experimental results on ISTD

Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
Mean									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.478	0.956	4.056	27.783	0.990	6.891	33.043	0.969	3.542
Log-ShadowFormer	31.217	0.957	4.043	27.232	0.990	7.184	33.135	0.969	3.481
Mean – Best 25%									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.951	0.989	2.063	33.166	0.998	3.249	38.780	0.994	1.852
Log-ShadowFormer	36.737	0.989	2.078	32.654	0.998	3.253	38.876	0.994	1.851
Mean – Worst 25%									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.296	0.897	7.504	21.417	0.973	13.747	26.645	0.922	6.551
Log-ShadowFormer	25.138	0.897	7.392	21.183	0.973	14.197	26.932	0.923	6.364
Trimean									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.795	0.970	3.342	28.303	0.993	5.394	33.303	0.980	2.941
Log-ShadowFormer	31.390	0.970	3.378	27.573	0.993	5.783	33.413	0.979	2.891
Median									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.070	0.971	3.215	28.541	0.994	5.179	33.400	0.981	2.811
Log-ShadowFormer	31.710	0.972	3.227	27.496	0.994	5.549	33.344	0.980	2.749

Table 9. ISTD-trained, with target adjustment and intensity augmentation in linear space; experimental results on ISTD

Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
Mean									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.576	0.957	4.002	28.163	0.990	6.392	33.030	0.969	3.556
Log-ShadowFormer	31.245	0.956	4.101	27.437	0.990	7.290	33.022	0.968	3.534
Mean – Best 25%									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.743	0.989	2.094	33.328	0.998	3.182	38.678	0.994	1.887
Log-ShadowFormer	36.576	0.989	2.147	32.692	0.998	3.414	38.610	0.994	1.890
Mean – Worst 25%									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.765	0.897	7.329	22.679	0.974	12.018	26.720	0.921	6.546
Log-ShadowFormer	25.200	0.896	7.479	21.497	0.974	14.509	26.823	0.921	6.486
Trimean									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.865	0.970	3.323	28.314	0.994	5.261	33.325	0.979	2.969
Log-ShadowFormer	31.535	0.969	3.419	27.759	0.994	5.714	33.316	0.979	2.945
Median									
Method	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.103	0.972	3.157	28.416	0.994	5.156	33.482	0.980	2.839
Log-ShadowFormer	31.793	0.971	3.325	27.829	0.994	5.516	33.261	0.980	2.810

Table 10. ISTD-trained, with target adjustment, intensity augmentation, and color balance in linear space; experimental results on ISTD

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.534	0.956	4.134	28.110	0.990	6.979	32.990	0.968	3.607
Log-ShadowFormer	31.305	0.955	4.198	27.494	0.990	7.779	33.045	0.967	3.553
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.756	0.989	2.183	33.298	0.998	3.428	38.466	0.993	1.958
Log-ShadowFormer	36.723	0.989	2.169	32.746	0.998	3.690	38.609	0.994	1.896
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.521	0.895	7.547	22.077	0.973	13.648	26.719	0.920	6.549
Log-ShadowFormer	25.193	0.892	7.633	21.479	0.974	15.105	26.785	0.917	6.482
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.899	0.970	3.404	28.407	0.994	5.550	33.399	0.979	3.026
Log-ShadowFormer	31.632	0.969	3.577	27.888	0.993	6.303	33.411	0.979	2.968
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.080	0.971	3.263	28.580	0.994	5.320	33.517	0.980	2.877
Log-ShadowFormer	31.799	0.970	3.404	28.031	0.994	5.959	33.444	0.979	2.837

1.2. Models Trained on ISTD-RAWShadowIO

Table 11. ISTD-RAWShadowIO-trained, with target adjustment in linear space; experimental results on ISTD

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.339	0.957	4.086	27.463	0.990	7.140	33.039	0.969	3.538
Log-ShadowFormer	31.076	0.956	4.166	27.066	0.989	7.519	32.906	0.969	3.581
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.751	0.989	2.099	32.942	0.998	3.329	38.679	0.994	1.867
Log-ShadowFormer	36.409	0.988	2.133	32.486	0.998	3.372	38.319	0.993	1.902
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.188	0.898	7.538	21.142	0.973	14.241	26.702	0.923	6.476
Log-ShadowFormer	25.093	0.897	7.551	20.884	0.973	14.984	26.834	0.922	6.478
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.651	0.970	3.364	27.885	0.993	5.584	33.275	0.980	2.965
Log-ShadowFormer	31.323	0.969	3.511	27.454	0.993	5.963	33.255	0.979	3.023
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.888	0.971	3.217	28.017	0.994	5.365	33.455	0.980	2.828
Log-ShadowFormer	31.515	0.970	3.437	27.718	0.994	5.707	33.379	0.980	2.933

1.3. Models Trained on ISTD-Unsaturated

Table 12. ISTD-Unsaturated-trained, with target adjustment in linear space; experimental results on ISTD

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.534	0.957	4.093	27.782	0.990	6.883	33.110	0.969	3.589
Log-ShadowFormer	31.215	0.957	4.141	26.966	0.990	7.570	33.149	0.969	3.542
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.913	0.989	2.044	32.881	0.998	3.264	38.910	0.994	1.834
Log-ShadowFormer	36.714	0.989	2.134	32.269	0.998	3.610	39.014	0.994	1.855
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.431	0.898	7.730	21.890	0.973	13.701	26.666	0.923	6.785
Log-ShadowFormer	25.370	0.898	7.564	21.175	0.973	14.689	26.945	0.923	6.497
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.912	0.970	3.322	28.127	0.994	5.427	33.389	0.980	2.925
Log-ShadowFormer	31.348	0.970	3.445	27.164	0.993	6.076	33.405	0.980	2.974
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.174	0.971	3.196	28.279	0.994	5.272	33.582	0.981	2.797
Log-ShadowFormer	31.630	0.972	3.342	27.047	0.994	5.842	33.346	0.981	2.855

2. Appendix B

Detailed Experimental Results with Intensity and Color Augmentation
All models trained on variations of ISTD

2.1. Experimental Results on ISTD

Table 13. ISTD-trained, with target adjustment in linear space; experimental results on ISTD

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.478	0.956	4.056	27.783	0.990	6.891	33.043	0.969	3.542
Log-ShadowFormer	31.217	0.957	4.043	27.232	0.990	7.184	33.135	0.969	3.481
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.951	0.989	2.063	33.166	0.998	3.249	38.780	0.994	1.852
Log-ShadowFormer	36.737	0.989	2.078	32.654	0.998	3.253	38.876	0.994	1.851
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.296	0.897	7.504	21.417	0.973	13.747	26.645	0.922	6.551
Log-ShadowFormer	25.138	0.897	7.392	21.183	0.973	14.197	26.932	0.923	6.364
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.795	0.970	3.342	28.303	0.993	5.394	33.303	0.980	2.941
Log-ShadowFormer	31.390	0.970	3.378	27.573	0.993	5.783	33.413	0.979	2.891
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.070	0.971	3.215	28.541	0.994	5.179	33.400	0.981	2.811
Log-ShadowFormer	31.710	0.972	3.227	27.496	0.994	5.549	33.344	0.980	2.749

Table 14. ISTD-trained, with target adjustment and intensity augmentation in linear space; experimental results on ISTD

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.576	0.957	4.002	28.163	0.990	6.392	33.030	0.969	3.556
Log-ShadowFormer	31.245	0.956	4.101	27.437	0.990	7.290	33.022	0.968	3.534
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.743	0.989	2.094	33.328	0.998	3.182	38.678	0.994	1.887
Log-ShadowFormer	36.576	0.989	2.147	32.692	0.998	3.414	38.610	0.994	1.890
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.765	0.897	7.329	22.679	0.974	12.018	26.720	0.921	6.546
Log-ShadowFormer	25.201	0.896	7.479	21.497	0.974	14.509	26.823	0.921	6.486
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.865	0.970	3.323	28.314	0.994	5.261	33.325	0.979	2.969
Log-ShadowFormer	31.535	0.969	3.419	27.759	0.994	5.714	33.316	0.979	2.945
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.103	0.972	3.157	28.416	0.994	5.156	33.482	0.980	2.839
Log-ShadowFormer	31.793	0.971	3.325	27.829	0.994	5.516	33.261	0.980	2.810

Table 15. ISTD-trained, with target adjustment, intensity augmentation, and color balance augmentation in linear space; experimental results on ISTD

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.534	0.956	4.134	28.110	0.990	6.979	32.990	0.968	3.607
Log-ShadowFormer	31.305	0.955	4.198	27.494	0.990	7.779	33.045	0.957	3.553
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.756	0.989	2.182	33.298	0.998	3.428	38.466	0.993	1.958
Log-ShadowFormer	36.723	0.989	2.169	32.746	0.998	3.690	38.609	0.994	1.896
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.521	0.895	7.547	22.077	0.973	13.648	26.719	0.920	6.545
Log-ShadowFormer	25.192	0.892	7.633	21.479	0.974	15.105	26.785	0.917	6.482
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.899	0.970	3.404	28.407	0.994	5.550	33.399	0.979	3.026
Log-ShadowFormer	31.632	0.969	3.577	27.888	0.993	6.303	33.411	0.979	2.968
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.080	0.971	3.263	28.580	0.994	5.320	33.517	0.980	2.877
Log-ShadowFormer	31.799	0.970	3.404	28.031	0.994	5.959	33.444	0.979	2.837

2.2. Experimental Results on ISTD with Intensity Augmentation

Table 16. ISTD-trained, with target adjustment in linear space; experimental results on ISTD with intensity augmentation

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.093	0.958	3.881	28.284	0.990	6.741	33.661	0.970	3.381
Log-ShadowFormer	31.559	0.958	3.998	27.265	0.990	7.392	33.620	0.970	3.410
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	37.512	0.990	2.001	33.601	0.998	3.157	39.226	0.994	1.792
Log-ShadowFormer	36.709	0.989	2.076	32.470	0.998	3.294	39.030	0.994	1.826
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	26.046	0.901	7.080	21.993	0.974	13.332	27.444	0.926	6.132
Log-ShadowFormer	25.757	0.899	7.306	21.426	0.974	14.633	27.536	0.923	6.203
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.444	0.971	3.240	28.710	0.994	5.322	33.926	0.980	2.836
Log-ShadowFormer	31.776	0.971	3.384	27.526	0.993	5.915	33.886	0.980	2.846
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.793	0.972	3.106	28.852	0.972	3.106	28.852	0.994	5.141
Log-ShadowFormer	32.063	0.973	3.203	27.502	0.994	5.541	34.089	0.981	2.713

Table 17. ISTD-trained, with target adjustment and intensity augmentation in linear space; experimental results on ISTD with intensity augmentation

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.505	0.960	3.720	29.071	0.991	5.937	33.954	0.971	3.308
Log-ShadowFormer	32.188	0.960	3.085	28.382	0.990	6.727	33.951	0.971	3.281
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	37.789	0.990	1.958	34.325	0.999	2.990	39.591	0.994	1.759
Log-ShadowFormer	37.579	0.990	1.985	33.699	0.998	3.137	39.509	0.994	1.744
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	26.603	0.904	6.803	23.335	0.976	11.251	27.634	0.926	6.033
Log-ShadowFormer	26.122	0.903	6.952	22.354	0.975	13.409	27.774	0.927	5.983
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.783	0.973	3.109	29.319	0.994	4.918	34.146	0.981	2.752
Log-ShadowFormer	32.523	0.972	3.188	28.702	0.994	5.338	34.166	0.981	2.734
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	33.034	0.974	2.978	29.552	0.995	4.751	34.250	0.982	2.630
Log-ShadowFormer	32.793	0.973	3.061	28.724	0.994	5.123	34.126	0.982	2.637

Table 18. ISTD-trained, with target adjustment, intensity augmentation, and color balance augmentation in linear space; experimental results on ISTD with intensity augmentation

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.494	0.959	3.830	29.047	0.990	6.444	33.954	0.970	3.350
Log-ShadowFormer	32.257	0.959	3.898	28.449	0.990	7.213	33.983	0.970	3.298
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	37.815	0.990	2.026	34.208	0.999	3.200	39.542	0.994	1.805
Log-ShadowFormer	37.745	0.990	2.019	33.821	0.998	3.380	39.580	0.994	1.756
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	26.412	0.901	6.988	22.974	0.975	12.604	27.621	0.925	6.057
Log-ShadowFormer	26.081	0.901	7.118	22.292	0.975	14.137	27.706	0.924	5.984
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.889	0.973	3.195	29.469	0.994	5.109	34.200	0.981	2.801
Log-ShadowFormer	32.544	0.972	3.316	28.793	0.994	5.802	34.216	0.981	2.730
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	33.116	0.974	3.080	29.744	0.995	4.798	34.319	0.981	2.690
Log-ShadowFormer	32.685	0.973	3.167	29.073	0.994	5.479	34.215	0.981	2.629

2.3. Experimental Results on ISTD with Color Balance Augmentation

Table 19. ISTD-trained, with target adjustment in linear space; experimental results on ISTD with color balance augmentation

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.475	0.958	4.661	27.176	0.990	9.966	33.431	0.970	3.661
Log-ShadowFormer	31.200	0.957	4.605	26.546	0.990	10.216	33.521	0.970	3.612
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.606	0.989	2.267	32.533	0.998	3.672	38.732	0.994	1.974
Log-ShadowFormer	36.326	0.989	2.282	31.882	0.998	3.738	38.740	0.994	1.960
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.835	0.901	8.470	21.497	0.973	19.654	27.491	0.926	6.458
Log-ShadowFormer	25.609	0.899	8.331	21.043	0.974	19.295	27.621	0.924	6.435
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.719	0.971	4.025	27.351	0.993	8.431	33.679	0.980	3.153
Log-ShadowFormer	31.327	0.971	3.988	26.638	0.993	9.091	33.762	0.980	3.055
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.964	0.972	3.911	27.212	0.994	7.880	33.935	0.981	2.989
Log-ShadowFormer	31.587	0.973	3.756	26.683	0.994	8.676	34.048	0.981	2.885

Table 20. ISTD-trained, with target adjustment and intensity augmentation in linear space; experimental results on ISTD with color balance augmentation

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.848	0.960	4.477	27.794	0.990	9.196	33.702	0.971	3.568
Log-ShadowFormer	31.504	0.959	4.568	27.137	0.990	9.736	33.666	0.971	3.597
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.740	0.989	2.247	32.866	0.998	3.571	38.882	0.994	1.957
Log-ShadowFormer	36.493	0.989	2.312	32.156	0.998	3.873	38.787	0.994	1.985
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	26.358	0.905	8.175	22.580	0.975	17.964	27.740	0.928	6.334
Log-ShadowFormer	25.897	0.903	8.150	21.649	0.974	18.592	27.815	0.927	6.310
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.084	0.972	3.790	27.800	0.994	7.717	34.038	0.981	3.046
Log-ShadowFormer	31.766	0.972	3.974	27.361	0.994	8.321	33.990	0.981	3.086
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.259	0.974	3.678	27.761	0.994	7.185	34.091	0.982	2.916
Log-ShadowFormer	31.962	0.973	3.885	27.324	0.994	8.009	34.068	0.982	2.952

Table 21. ISTD-trained, with target adjustment, intensity augmentation, and color balance augmentation in linear space; experimental results on ISTD with color balance augmentation

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.461	0.959	3.963	28.988	0.990	6.729	33.928	0.971	3.448
Log-ShadowFormer	32.222	0.959	4.013	28.402	0.990	7.399	33.956	0.970	3.398
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	37.588	0.990	2.128	34.093	0.999	3.382	39.303	0.994	1.902
Log-ShadowFormer	37.559	0.990	2.097	33.750	0.998	3.476	39.377	0.994	1.843
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	26.522	0.901	0.718	23.048	0.975	12.906	27.755	0.925	6.217
Log-ShadowFormer	26.192	0.900	7.283	22.320	0.975	14.401	27.816	0.924	6.162
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.863	0.973	3.284	29.394	0.994	5.414	34.311	0.981	2.866
Log-ShadowFormer	32.539	0.972	3.382	28.721	0.994	5.996	34.274	0.981	2.854
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.989	0.974	3.150	29.627	0.995	5.197	34.452	0.982	2.731
Log-ShadowFormer	32.784	0.974	3.219	28.844	0.994	5.713	34.278	0.982	2.726

2.4. Experimental Results on ISTD with Intensity Augmentation and Color Balance Augmentation

Table 22. ISTD-trained, with target adjustment in linear space; experimental results on ISTD with intensity augmentation and color balance augmentation

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.640	0.959	4.534	27.386	0.990	9.649	33.526	0.971	3.579
Log-ShadowFormer	31.228	0.958	4.579	26.558	0.990	10.100	33.526	0.970	3.570
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	36.694	0.989	2.256	32.570	0.998	3.654	38.723	0.993	1.949
Log-ShadowFormer	36.140	0.989	2.301	31.665	0.998	3.791	38.568	0.994	1.966
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	26.014	0.903	8.295	21.691	0.974	19.341	27.652	0.928	6.305
Log-ShadowFormer	25.782	0.901	0.8332	21.209	0.974	19.505	27.674	0.926	6.344
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.861	0.971	3.869	27.604	0.993	7.955	33.794	0.980	3.076
Log-ShadowFormer	31.438	0.971	3.906	26.651	0.993	8.760	33.857	0.980	3.021
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.092	0.972	3.731	27.553	0.994	7.338	34.103	0.981	2.937
Log-ShadowFormer	31.657	0.973	3.761	26.685	0.994	8.380	34.114	0.982	2.883

Table 23. ISTD-trained, with target adjustment and intensity augmentation in linear space; experimental results on ISTD with intensity augmentation and color balance augmentation

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.070	0.960	4.340	28.081	0.991	8.826	33.898	0.972	3.469
Log-ShadowFormer	31.791	0.960	4.401	27.541	0.990	9.303	33.893	0.972	3.470
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	37.102	0.990	2.192	33.225	0.998	3.418	39.183	0.994	1.913
Log-ShadowFormer	36.876	0.990	2.257	32.671	0.998	3.647	39.144	0.994	1.920
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	26.488	0.906	8.008	22.806	0.975	17.645	27.942	0.929	6.134
Log-ShadowFormer	26.206	0.906	7.978	21.988	0.975	18.358	28.086	0.930	6.096
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.277	0.972	3.642	28.151	0.994	7.238	34.203	0.981	2.980
Log-ShadowFormer	32.061	0.972	3.768	27.767	0.994	7.770	34.219	0.981	2.991
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.530	0.974	3.495	28.076	0.995	6.669	34.264	0.982	2.845
Log-ShadowFormer	32.280	0.974	3.631	27.784	0.995	7.217	34.217	0.982	2.875

Table 24. ISTD-trained, with target adjustment, intensity augmentation, and color balance augmentation in linear space; experimental results on ISTD with intensity augmentation and color balance augmentation

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.656	0.961	3.832	29.194	0.991	6.501	34.113	0.971	3.339
Log-ShadowFormer	32.418	0.960	3.883	28.603	0.991	7.166	34.145	0.971	3.288
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	37.863	0.990	2.082	34.267	0.999	3.310	39.498	0.994	1.859
Log-ShadowFormer	37.804	0.990	2.053	33.838	0.999	3.447	39.582	0.994	1.801
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	26.808	0.904	6.856	23.348	0.975	12.441	28.008	0.928	5.916
Log-ShadowFormer	26.492	0.903	6.950	23.676	0.976	13.864	28.098	0.926	5.867
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	33.035	0.973	3.245	29.442	0.994	5.228	34.450	0.981	2.831
Log-ShadowFormer	32.558	0.973	3.320	28.946	0.994	5.808	34.377	0.981	2.790
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	33.225	0.974	3.111	29.645	0.995	4.989	34.434	0.982	2.705
Log-ShadowFormer	32.698	0.974	3.165	29.209	0.994	5.524	34.276	0.982	2.680

3. Appendix C

Detailed Experimental Results on ISTD-Unsaturated

3.1. Models Trained on ISTD

Table 25. ISTD-trained, with target adjustment in linear space; experimental results on ISTD-Unsaturated

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.756	0.957	3.849	28.115	0.990	6.418	33.305	0.969	3.374
Log-ShadowFormer	31.460	0.957	3.868	27.427	0.990	6.694	33.369	0.969	3.341
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	37.039	0.989	2.047	33.241	0.998	3.225	38.879	0.994	1.838
Log-ShadowFormer	36.823	0.989	2.063	32.726	0.998	3.225	38.963	0.994	1.840
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.797	0.896	6.837	22.128	0.974	12.156	27.150	0.921	5.994
Log-ShadowFormer	25.431	0.896	6.859	21.522	0.974	12.603	27.357	0.921	5.906
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.132	0.970	3.272	28.612	0.994	5.200	33.576	0.980	2.861
Log-ShadowFormer	31.653	0.970	3.284	27.779	0.993	5.587	33.645	0.980	2.828
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.404	0.973	3.143	28.889	0.994	5.058	33.679	0.982	2.768
Log-ShadowFormer	31.867	0.972	3.160	27.800	0.994	5.414	33.470	0.981	2.728

3.2. Models Trained on ISTD-Unsaturated

Table 26. ISTD-Unsaturated-trained, with target adjustment in linear space; experimental results on ISTD-Unsaturated

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.823	0.957	3.828	28.077	0.990	6.367	33.396	0.969	3.366
Log-ShadowFormer	31.452	0.957	3.928	27.215	0.990	6.967	33.389	0.969	3.373
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	37.003	0.989	2.027	32.947	0.998	3.244	39.019	0.994	1.820
Log-ShadowFormer	36.814	0.989	2.117	32.337	0.998	3.585	39.104	0.994	1.842
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	26.009	0.897	6.842	22.469	0.974	11.944	27.270	0.922	6.020
Log-ShadowFormer	25.706	0.896	6.864	21.579	0.973	12.639	27.383	0.922	5.933
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.084	0.971	3.245	28.449	0.994	5.202	33.611	0.980	2.857
Log-ShadowFormer	31.621	0.971	3.387	27.385	0.993	5.914	33.640	0.980	2.904
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.304	0.973	3.115	28.564	0.994	5.072	33.708	0.982	2.771
Log-ShadowFormer	31.846	0.972	3.278	27.258	0.994	5.747	33.466	0.982	2.791

4. Appendix D

Detailed Experimental Results on RAWShadowIO

4.1. Experimental Results on Linear Images

Table 27. ISTD-trained, with target adjustment in sRGB space; experimental results on RAWShadowIO linear images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.391	0.856	10.053	18.965	0.974	19.841	27.561	0.885	8.426
Linear-ShadowFormer	23.193	0.844	12.044	18.756	0.973	21.319	25.308	0.874	10.499
Log-ShadowFormer	24.620	0.860	9.737	19.262	0.974	19.884	27.918	0.889	7.914
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	29.018	0.952	5.636	24.948	0.995	9.295	31.328	0.972	4.861
Linear-ShadowFormer	27.091	0.948	6.333	24.963	0.994	10.363	29.271	0.971	5.426
Log-ShadowFormer	28.917	0.955	5.515	24.929	0.995	9.637	31.994	0.974	4.203
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	18.879	0.708	16.391	12.400	0.938	33.499	23.053	0.748	13.460
Linear-ShadowFormer	18.786	0.693	19.434	13.058	0.936	34.704	20.614	0.735	17.227
Log-ShadowFormer	19.326	0.716	16.131	13.219	0.938	34.233	23.328	0.755	12.698
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.787	0.882	9.165	19.209	0.981	18.497	27.917	0.913	7.857
Linear-ShadowFormer	23.418	0.866	11.536	18.526	0.981	19.924	25.627	0.894	9.752
Log-ShadowFormer	25.155	0.883	8.678	19.515	0.982	18.067	28.069	0.916	7.390
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.861	0.886	8.977	19.881	0.982	18.150	27.957	0.930	7.709
Linear-ShadowFormer	23.650	0.867	11.033	18.719	0.982	18.828	25.654	0.895	9.523
Log-ShadowFormer	25.439	0.886	8.537	20.113	0.983	16.999	38.197	0.932	7.243

Table 28. ISTD-trained, with target adjustment in linear space; experimental results on RAWShadowIO linear images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.379	0.856	9.839	18.619	0.974	20.147	27.855	0.886	8.074
Log-ShadowFormer	23.979	0.858	9.893	18.235	0.973	21.027	27.811	0.888	7.870
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	29.017	0.955	5.469	24.339	0.995	9.756	31.720	0.973	4.500
Log-ShadowFormer	28.420	0.952	5.466	24.003	0.995	10.387	31.807	0.974	4.187
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	19.099	0.708	16.026	12.809	0.937	31.793	23.202	0.748	13.090
Log-ShadowFormer	18.675	0.712	16.257	12.575	0.937	34.659	23.142	0.752	12.811
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.793	0.880	8.980	18.298	0.981	19.338	28.215	0.912	7.473
Log-ShadowFormer	24.505	0.882	8.891	18.139	0.981	19.973	28.068	0.914	7.342
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.866	0.885	8.872	18.361	0.982	19.582	28.343	0.929	7.391
Log-ShadowFormer	24.806	0.889	8.611	18.075	0.982	18.923	28.355	0.930	7.408

Table 29. ISTD-trained, with target adjustment and intensity augmentation in linear space; experimental results on RAWShadowIO linear images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.361	0.857	9.985	19.110	0.974	19.198	27.539	0.887	8.309
Log-ShadowFormer	24.504	0.860	9.722	19.128	0.973	19.239	27.876	0.890	7.941
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	28.493	0.955	5.570	24.572	0.995	9.820	31.259	0.974	4.517
Log-ShadowFormer	28.607	0.955	5.447	24.645	0.995	9.660	31.894	0.975	4.469
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	18.970	0.710	16.400	12.767	0.935	31.805	23.017	0.750	13.363
Log-ShadowFormer	19.175	0.718	16.153	13.308	0.936	31.512	23.285	0.759	12.808
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.878	0.882	8.945	19.425	0.981	17.825	27.809	0.915	7.750
Log-ShadowFormer	24.957	0.884	8.610	19.205	0.982	17.932	28.078	0.917	7.308
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.051	0.886	8.885	19.788	0.982	17.815	28.145	0.932	7.860
Log-ShadowFormer	25.154	0.887	8.553	19.501	0.983	17.923	28.267	0.933	7.193

Table 30. ISTD-RAWShadowIO-trained, with target adjustment in linear space; experimental results on RAWShadowIO linear images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	27.012	0.875	8.299	22.795	0.978	14.322	29.204	0.900	7.252
Log-ShadowFormer	27.059	0.872	8.284	22.934	0.977	14.028	29.154	0.899	7.221
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.043	0.959	4.521	28.957	0.996	6.394	33.618	0.974	3.528
Log-ShadowFormer	31.100	0.956	4.598	29.579	0.996	6.557	33.447	0.972	3.690
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	21.119	0.732	14.570	15.506	0.943	26.285	23.735	0.770	12.457
Log-ShadowFormer	21.024	0.725	14.722	16.024	0.941	25.907	23.635	0.766	12.342
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	27.872	0.900	7.126	23.221	0.985	12.277	29.591	0.929	6.632
Log-ShadowFormer	27.925	0.897	7.015	23.241	0.985	11.940	29.659	0.927	6.575
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	28.079	0.898	6.955	23.863	0.987	12.147	30.091	0.940	6.428
Log-ShadowFormer	28.089	0.900	6.771	23.290	0.986	11.885	30.202	0.939	6.353

4.2. Experimental Results on sRGB PNG Images

Table 31. ISTD-trained, with target adjustment in sRGB space; experimental results on RAWShadowIO sRGB PNG images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.767	0.861	9.696	19.301	0.975	19.203	27.951	0.890	8.114
Linear-ShadowFormer	24.732	0.862	10.103	19.408	0.974	19.791	27.785	0.890	8.330
Log-ShadowFormer	24.956	0.865	9.431	19.578	0.975	19.370	28.275	0.893	7.633

Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	29.340	0.953	5.471	25.370	0.995	9.004	31.626	0.973	4.676
Linear-ShadowFormer	29.253	0.955	5.326	25.597	0.995	8.322	31.299	0.974	4.775
Log-ShadowFormer	24.956	0.865	9.431	19.578	0.975	19.370	28.275	0.893	7.633

Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	19.237	0.718	15.852	12.751	0.939	32.296	23.444	0.757	13.016
Linear-ShadowFormer	19.302	0.718	17.035	13.404	0.939	32.939	23.307	0.757	13.282
Log-ShadowFormer	19.624	0.725	15.675	13.545	0.940	33.238	23.672	0.763	12.316

Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.204	0.886	8.694	19.477	0.982	17.968	28.353	0.916	7.468
Linear-ShadowFormer	25.233	0.885	9.157	19.360	0.983	8.895	28.210	0.917	7.732
Log-ShadowFormer	25.343	0.887	8.428	19.699	0.983	17.641	28.347	0.919	7.180

Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.247	0.891	8.520	19.989	0.983	17.737	28.460	0.932	7.270
Linear-ShadowFormer	25.453	0.887	8.974	19.701	0.983	18.834	28.498	0.933	7.654
Log-ShadowFormer	25.651	0.890	8.343	20.097	0.984	16.713	28.498	0.935	7.045

Table 32. ISTD-trained, with target adjustment in linear space; experimental results on RAWShadowIO sRGB PNG images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.791	0.862	9.462	19.055	0.975	19.356	28.255	0.890	7.775
Log-ShadowFormer	24.432	0.863	9.527	18.682	0.974	20.245	28.209	0.892	7.574
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	29.397	0.957	5.305	24.646	0.995	9.439	32.152	0.974	4.352
Log-ShadowFormer	28.865	0.954	5.331	24.285	0.995	10.102	32.209	0.975	3.969
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	19.519	0.718	15.444	13.198	0.939	30.525	23.597	0.757	12.624
Log-ShadowFormer	19.134	0.721	15.717	13.057	0.939	33.441	23.527	0.760	12.395
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.172	0.885	8.602	19.033	0.982	18.482	28.646	0.916	7.129
Log-ShadowFormer	25.024	0.888	8.526	18.594	0.981	19.100	28.458	0.918	7.051
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.242	0.890	8.523	19.429	0.982	18.503	28.663	0.932	6.931
Log-ShadowFormer	25.269	0.893	8.316	18.541	0.982	18.034	28.682	0.933	7.017

Table 33. ISTD-trained, with target adjustment and intensity augmentation in linear space; experimental results on RAWShadowIO sRGB PNG images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.685	0.862	9.725	19.457	0.974	18.630	27.814	0.890	8.110
Log-ShadowFormer	24.751	0.864	9.503	19.368	0.974	18.822	28.136	0.894	7.760
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	28.824	0.956	5.448	24.967	0.995	9.539	31.505	0.974	4.418
Log-ShadowFormer	28.862	0.956	5.344	24.934	0.995	9.401	32.116	0.975	4.363
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	19.240	0.718	16.006	13.028	0.936	31.072	23.311	0.758	13.035
Log-ShadowFormer	19.395	0.729	15.800	13.516	0.937	30.877	23.556	0.765	12.517
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.214	0.886	8.666	19.844	0.982	17.226	28.112	0.918	7.587
Log-ShadowFormer	25.209	0.888	8.423	19.486	0.982	17.545	28.357	0.920	7.150
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.426	0.889	8.633	20.285	0.984	17.264	28.460	0.934	7.682
Log-ShadowFormer	25.390	0.890	8.361	19.833	0.983	17.498	28.547	0.935	7.048

Table 34. ISTD-RAWShadowIO-trained, with target adjustment in linear space; experimental results on RAWShadowIO sRGB PNG images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	27.344	0.879	8.074	23.223	0.979	13.808	29.447	0.903	7.065
Log-ShadowFormer	27.348	0.877	8.116	23.162	0.977	13.843	29.445	0.902	7.032
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	31.272	0.959	4.465	29.615	0.996	6.082	33.684	0.974	3.500
Log-ShadowFormer	31.368	0.956	4.560	29.725	0.996	6.537	33.625	0.972	3.640
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	21.523	0.742	14.037	15.940	0.945	24.918	24.116	0.778	12.047
Log-ShadowFormer	21.345	0.737	14.347	16.350	0.942	25.258	24.055	0.777	11.924
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	28.214	0.904	6.967	23.470	0.986	12.022	29.817	0.932	6.479
Log-ShadowFormer	28.137	0.902	6.846	23.498	0.985	11.730	29.913	0.929	6.391
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	28.465	0.900	6.798	24.108	0.988	11.958	30.311	0.943	6.313
Log-ShadowFormer	28.328	0.905	6.638	23.568	0.987	11.734	30.479	0.939	6.221

4.3. Experimental Results on sRGB JPG Images

Table 35. ISTD-trained, with target adjustment in sRGB space; experimental results on RAWShadowIO sRGB JPG images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.080	0.855	7.583	19.396	0.974	17.759	28.547	0.884	5.896
Linear-ShadowFormer	25.038	0.856	8.095	19.476	0.974	18.723	28.379	0.885	6.178
Log-ShadowFormer	25.248	0.859	7.570	19.661	0.974	18.307	28.757	0.888	5.647
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	29.891	0.947	4.216	25.442	0.995	7.933	32.572	0.968	3.507
Linear-ShadowFormer	29.724	0.949	4.099	25.646	0.994	7.632	32.412	0.969	3.386
Log-ShadowFormer	29.753	0.951	4.345	25.312	0.995	8.559	32.992	0.971	3.487
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	19.331	0.711	13.246	12.7912	0.937	30.218	23.751	0.751	9.983
Linear-ShadowFormer	19.369	0.712	15.213	13.438	0.937	31.772	23.557	0.752	11.142
Log-ShadowFormer	19.734	0.718	13.304	13.611	0.938	31.683	23.960	0.758	9.433
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.486	0.879	6.586	19.602	0.981	16.608	28.927	0.912	5.098
Linear-ShadowFormer	25.401	0.878	6.677	19.490	0.982	17.770	28.781	0.913	5.130
Log-ShadowFormer	25.664	0.880	6.432	19.838	0.982	16.431	28.961	0.913	4.807
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.550	0.881	6.363	20.214	0.982	16.213	29.180	0.926	4.939
Linear-ShadowFormer	25.747	0.878	6.213	19.833	0.984	17.968	28.998	0.929	4.953
Log-ShadowFormer	26.097	0.882	6.189	20.307	0.983	15.543	29.033	0.928	4.722

Table 36. ISTD-trained, with target adjustment in linear space; experimental results on RAWShadowIO sRGB JPG images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.072	0.856	7.500	19.115	0.974	17.954	28.787	0.885	5.719
Log-ShadowFormer	24.658	0.857	7.676	18.756	0.973	19.086	28.653	0.887	5.607
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	29.890	0.951	4.118	24.794	0.994	8.231	32.863	0.970	3.426
Log-ShadowFormer	29.232	0.949	4.203	24.408	0.995	9.316	32.748	0.970	3.336
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	19.587	0.711	12.987	13.228	0.938	28.603	23.892	0.751	9.703
Log-ShadowFormer	19.205	0.714	13.522	13.083	0.937	32.047	23.800	0.755	9.519
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.511	0.877	6.494	19.105	0.981	17.645	29.090	0.911	4.954
Log-ShadowFormer	25.165	0.881	6.600	18.686	0.980	17.726	29.085	0.913	4.756
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.576	0.882	6.337	19.400	0.983	18.017	29.232	0.927	4.820
Log-ShadowFormer	25.314	0.884	6.283	18.644	0.981	17.025	29.361	0.938	4.616

Table 37. ISTD-trained, with target adjustment and intensity augmentation in linear space; experimental results on RAWShadowIO sRGB JPG images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	24.878	0.855	7.729	19.543	0.974	17.200	28.235	0.885	6.017
Log-ShadowFormer	24.987	0.858	7.656	19.447	0.973	17.595	28.572	0.888	5.818
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	29.092	0.950	4.459	25.120	0.995	8.548	31.981	0.970	3.632
Log-ShadowFormer	29.218	0.950	4.211	25.003	0.995	8.634	32.760	0.971	3.409
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	19.329	0.710	13.268	13.032	0.934	28.979	23.591	0.751	10.144
Log-ShadowFormer	19.459	0.718	13.494	13.553	0.935	29.191	23.818	0.760	9.811
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.479	0.878	6.606	19.891	0.982	15.911	28.659	0.911	5.141
Log-ShadowFormer	25.535	0.881	6.537	19.535	0.981	15.963	28.894	0.913	5.100
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	25.724	0.878	6.485	20.237	0.983	15.896	28.969	0.920	5.012
Log-ShadowFormer	25.749	0.881	6.361	19.898	0.982	15.793	28.882	0.926	5.062

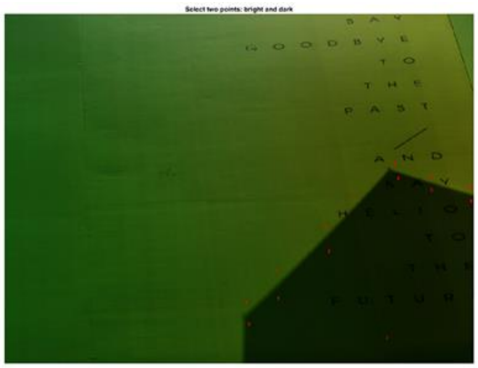
Table 38. ISTD-RAWShadowIO-trained, with target adjustment in linear space; experimental results on RAWShadowIO sRGB JPG images

Mean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	27.795	0.874	6.037	23.445	0.978	12.157	30.157	0.899	4.964
Log-ShadowFormer	27.845	0.872	6.123	23.448	0.977	12.162	30.126	0.899	4.978
Mean – Best 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	32.088	0.956	3.336	30.168	0.996	4.530	34.797	0.972	2.815
Log-ShadowFormer	32.288	0.953	3.355	30.339	0.996	4.827	34.623	0.970	2.861
Mean – Worst 25%									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	21.657	0.735	11.143	15.975	0.943	23.061	24.386	0.773	9.011
Log-ShadowFormer	21.490	0.730	11.497	16.452	0.940	23.369	24.301	0.771	8.970
Trimean									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	28.577	0.898	4.904	23.560	0.985	10.106	30.774	0.926	4.063
Log-ShadowFormer	28.684	0.899	4.792	23.700	0.985	10.055	30.730	0.926	4.169
Median									
Method	Overall Image			Shadow Region (S)			Non-Shadow Region (NS)		
	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓	PSNR ↑	SSIM ↑	RMSE ↓
ShadowFormer	28.720	0.898	4.793	24.202	0.988	10.049	31.070	0.935	3.988
Log-ShadowFormer	28.696	0.905	4.655	23.828	0.986	10.054	31.254	0.937	4.132

5. Appendix E

Detailed Illuminant Spectral Ratio Analysis

Table 39. ISTD test images with normalized spectral ratios and image annotated with sampled points

Annotated Linear Shadow Image	Normalized Illuminant Spectral Ratios (RGB)		
	0.376	0.806	0.457
	0.386	0.822	0.422
	0.510	0.774	0.375
	0.522	0.752	0.403
	0.323	0.766	0.556
	0.558	0.747	0.363
	0.403	0.735	0.546
	0.338	0.425	0.840
	0.405	0.507	0.761
	0.487	0.659	0.573
	0.307	0.542	0.782
	0.626	0.527	0.575
	0.577	0.577	0.577
	0.760	0.531	0.375
	0.619	0.563	0.551
	0.506	0.621	0.599
	0.501	0.585	0.638
	0.335	0.517	0.788

0.186	0.409	0.894
0.227	0.380	0.897
0.255	0.373	0.892

0.413	0.551	0.725
0.407	0.564	0.718
0.481	0.621	0.618

0.378	0.439	0.815
0.292	0.410	0.864
0.285	0.404	0.869

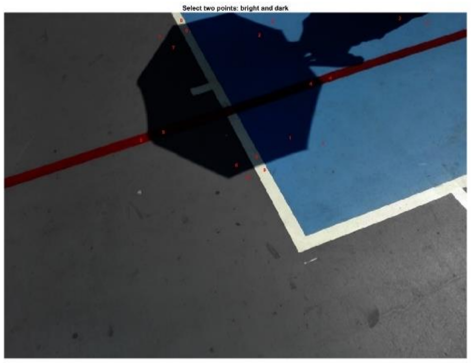
0.372	0.489	0.789
0.400	0.489	0.775
0.370	0.536	0.759
0.352	0.512	0.784

0.497	0.525	0.691
0.464	0.554	0.692
0.280	0.328	0.902
0.391	0.552	0.736

0.427	0.524	0.737
0.393	0.496	0.774
0.499	0.558	0.663

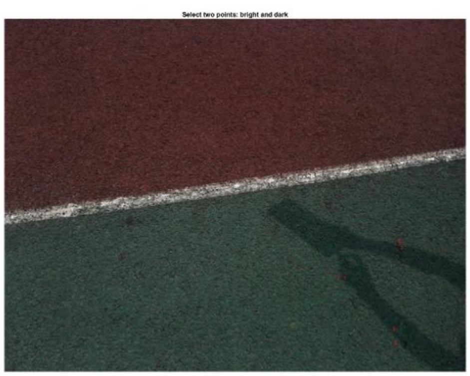
0.519	0.566	0.641
0.203	0.532	0.822
0.328	0.582	0.744
0.386	0.490	0.781
0.289	0.765	0.576
0.322	0.681	0.657
0.526	0.589	0.614
0.321	0.440	0.839

0.388	0.508	0.769
0.684	0.554	0.475
0.410	0.506	0.759
0.379	0.554	0.741
0.237	0.490	0.839

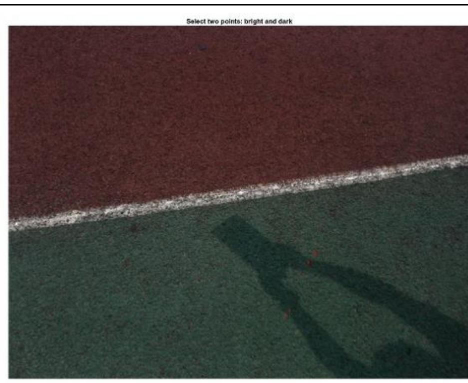


0.029	0.231	0.973
0.023	0.254	0.967
0.000	0.230	0.973
0.415	0.000	0.908
0.115	0.420	0.900
0.217	0.436	0.873
0.229	0.423	0.877
0.374	0.540	0.754
0.371	0.553	0.746

0.186	0.295	0.937
0.393	0.920	0.000
0.431	0.558	0.709
0.354	0.536	0.767
0.042	0.278	0.960
0.009	0.239	0.971
0.288	0.514	0.808
0.308	0.421	0.853



0.555	0.522	0.648
0.691	0.424	0.586
0.296	0.663	0.688

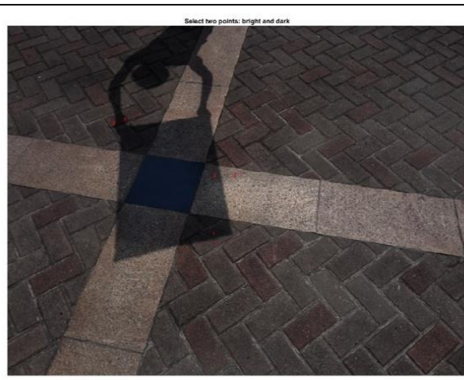


0.452	0.557	0.697
0.294	0.716	0.633

0.268	0.283	0.921
0.281	0.364	0.888
0.473	0.532	0.703

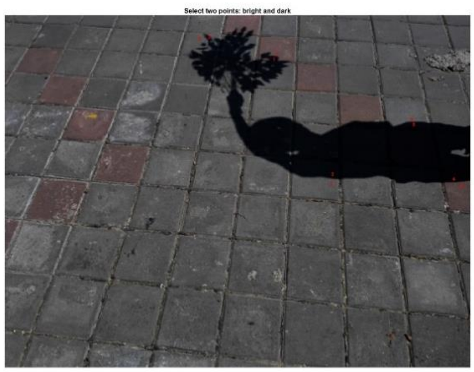
0.310	0.328	0.892
0.239	0.338	0.910
0.452	0.519	0.726

0.489	0.553	0.675
0.461	0.519	0.720
0.406	0.515	0.755

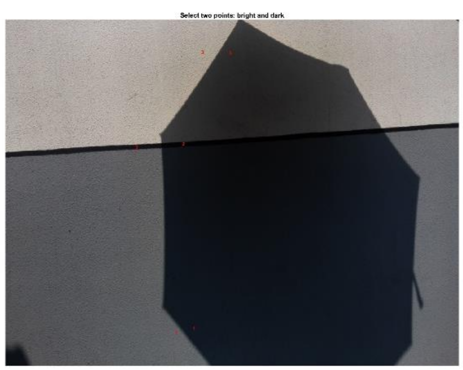


0.445	0.532	0.721
0.490	0.559	0.669
0.403	0.582	0.706

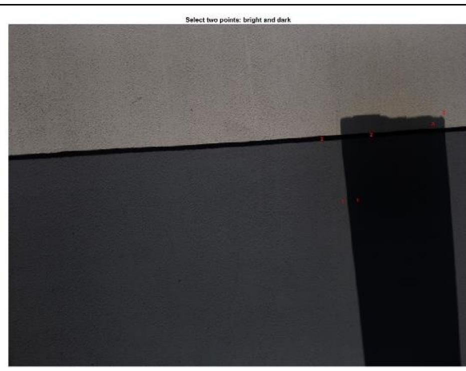
0.306	0.503	0.809
0.389	0.475	0.790
0.293	0.456	0.840
0.400	0.442	0.803



0.437	0.511	0.740
0.302	0.454	0.839
0.247	0.479	0.843
0.297	0.423	0.856
0.298	0.454	0.840

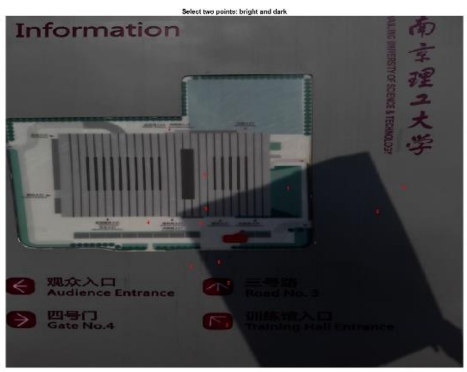


0.312	0.512	0.801
0.567	0.551	0.612
0.530	0.568	0.630



0.427	0.526	0.736
0.575	0.553	0.603
0.503	0.549	0.667

0.459	0.531	0.712
0.244	0.882	0.404
0.382	0.508	0.772
0.358	0.541	0.761
0.491	0.555	0.672
0.475	0.539	0.696
0.511	0.525	0.681

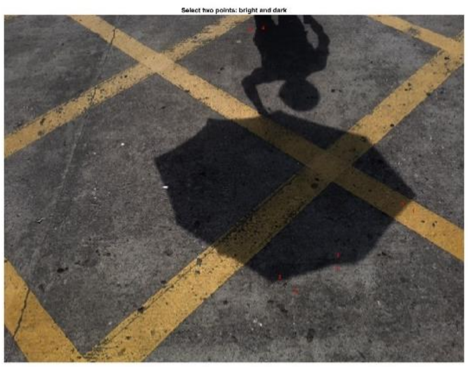


0.324	0.526	0.786
0.401	0.497	0.770
0.228	0.737	0.636
0.499	0.542	0.676
0.499	0.534	0.683
0.428	0.498	0.754

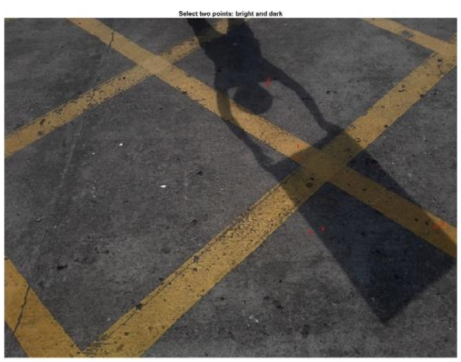
0.336	0.554	0.762
0.345	0.470	0.812

0.471	0.383	0.795
0.560	0.501	0.660
0.340	0.479	0.809
0.328	0.463	0.824

0.356	0.477	0.803
0.392	0.549	0.739
0.453	0.513	0.729

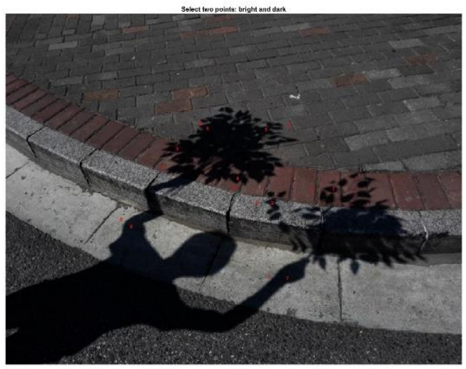


0.489	0.532	0.691
0.500	0.543	0.675
0.479	0.543	0.689
0.488	0.524	0.699



0.507	0.577	0.641
0.466	0.540	0.701
0.491	0.518	0.700
0.445	0.533	0.719

0.306	0.477	0.824
0.368	0.400	0.819
0.317	0.485	0.816
0.324	0.519	0.791
0.393	0.492	0.777
0.287	0.426	0.858



0.379	0.537	0.754
0.308	0.441	0.843
0.390	0.498	0.774
0.419	0.519	0.745
0.333	0.490	0.806
0.370	0.491	0.789
0.292	0.481	0.826
0.331	0.503	0.798

0.294	0.494	0.818
0.383	0.522	0.762
0.279	0.474	0.836

0.440	0.543	0.715
0.430	0.511	0.744
0.380	0.490	0.785

0.524	0.494	0.694
0.519	0.564	0.643
0.468	0.501	0.728

0.489	0.437	0.755
0.414	0.503	0.759
0.466	0.546	0.696

0.196	0.387	0.901
0.181	0.405	0.896
0.226	0.316	0.921
0.239	0.374	0.896

0.242	0.452	0.859
0.383	0.474	0.793
0.248	0.448	0.859
0.309	0.477	0.823