LAB_ShadowFormer

By Dhanush Adithya & Harshil Bhojwani
CS 7180 - Advanced Perception

Overview

Our project focuses on **shadow removal** of images using a Transformer based model, specifically the Shadow Former architecture.

Intention was to fine tune the model by introducing few tweaks to enhance the performance of the model.

We work with the ISTD dataset, starting with fixing the dataset by reducing the illumination inconsistencies between the shadowed images and the ground truth shadow free image.

DataSet

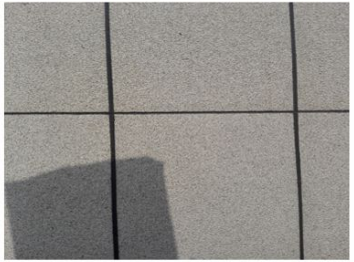
ISTD Dataset includes 1300 training and 540 testing triplets (shadow images, masks, and shadow free images).

ISTD data set suffers from a illumination inconsistencies where The shadow/no-shadow pairs in ISTD have differen values **outside** the shadow area.

And images were suffering from having more than 2% of the pixels being saturated. This can distort color perception and affects the model that usually requires a balanced intensity distribution.

Shadow Image

Shadow Free



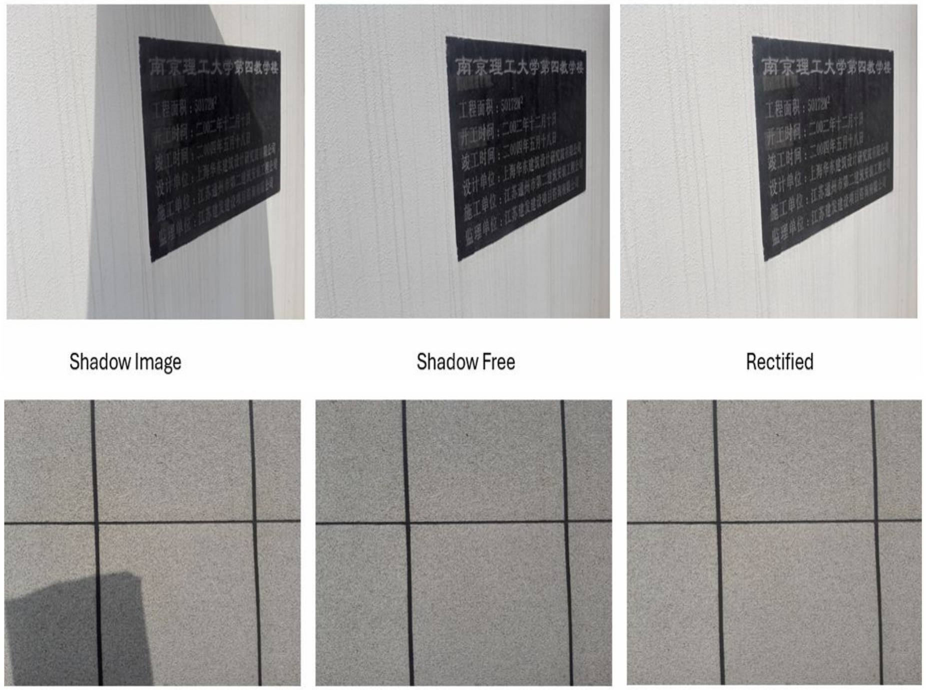
DataSet

First task was to Fix the Dataset by using the pixels outside the shadow mask and computing the difference between Shadow and Shadow Free image in that area.

Fitting a plane in between the shadow and shadow free image either as combination of 3 channels or doing it per channels.

Projecting the Shadow Free Image onto that to best match the Shadow Image, using it as Ground Truth

Once Fixed removing all the images having more than 2% saturated pixels

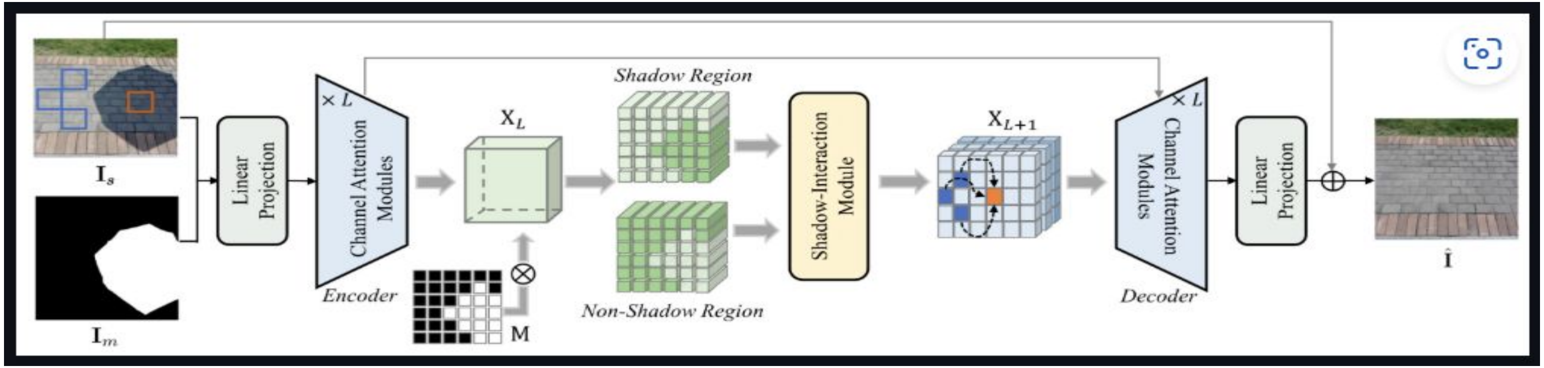


ShadowFormer

Shadow Former is a Retinex-based shadow model, to exploit non-shadow regions to help shadow region restoration.

A multi-scale channel attention framework is employed to hierarchically capture the global information.

It Introduces Shadow-Interaction Module (SIM) with Shadow-Interaction Attention (SIA) in the bottleneck region to effectively model the context correlation between shadow and non-shadow regions



Modification- Changing the Input from sRGB to LAB

Shadow removal can be seen as a sub problem of Intrinsic Imaging, since it requires identifying the illumination component (shadows) and modifying it, thus we decided to use LAB color space as our model input. Lumininance separation in LAB allows for more precise manipulation of brightness and shadows without directly affecting color information. LAB has a wider color Gamut that may aid with better preservation of color nuances during shadow removal

We first converted sRGB to XYZ:

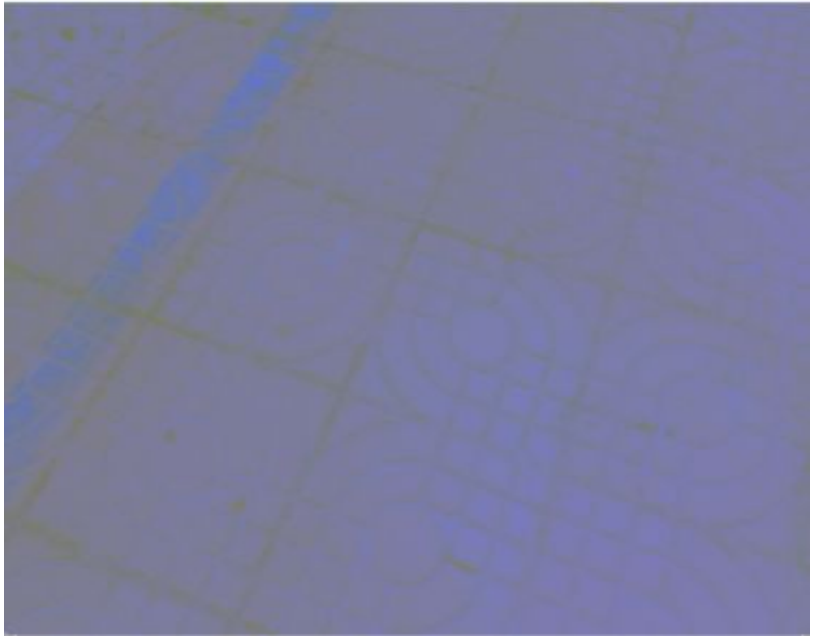
```
x = (r * 0.4124 + g * 0.3576 + b * 0.1805) / 0.95047

y = (r * 0.2126 + g * 0.7152 + b * 0.0722) / 1.00000

z = (r * 0.0193 + g * 0.1192 + b * 0.9505) / 1.08883
```

then XYZ to LAB

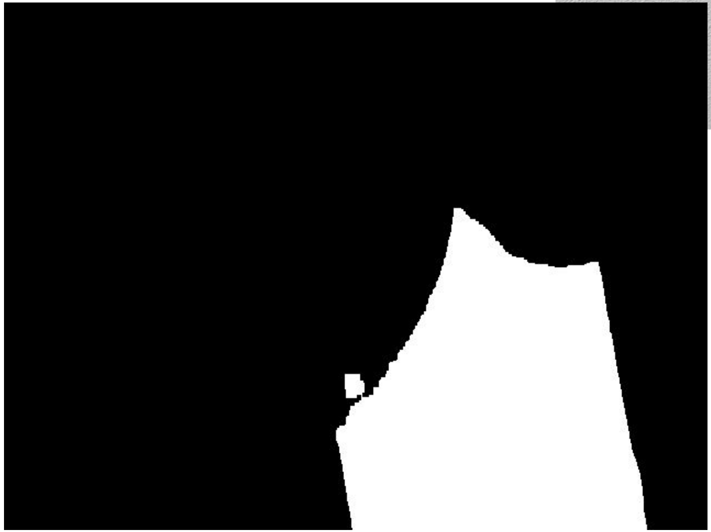
```
x = x ** (1 / 3) if x > 0.008856 else (7.787 * x) + (16 / 116)

y = y ** (1 / 3) if y > 0.008856 else (7.787 * y) + (16 / 116)

z = z ** (1 / 3) if z > 0.008856 else (7.787 * z) + (16 / 116)
```

```
L = (116 * y) - 16

A = 500 * (x - y)

B = 200 * (y - z)
```

Modification- Changing the Input from sRGB to LAB



Modification- Applying a Gaussian Blur to Shadow Mask



Modification- Changing the Loss to Cosine Distance (Image Space)

Cosine similarity in Lab* color space provides superior color vector alignment by capturing directional relationships between color vectors independent of absolute intensity. Its illumination-invariant nature allows for more robust color matching across different lighting conditions, respecting the inherent perceptually uniform characteristics of the Lab* color space.

We used a pretrained vgg16 model for feature extraction and then calculated cosine similarity on the extracted features.

```
cosine_sim=torch.mul(enhanced_features,ground_truth_features).sum(1)/(torch.mul(torch.pow((torch.pow(enhanced_features,2)).sum(1),0.5),torch.pow((torch.pow(ground_truth_features,2)).sum(1),0.5))+1e-8)
```

Continuation - cosine loss

We experienced numerical instability and gradient explosion with the initial implementation of cosine loss

We then introduced
Cosine Similarity Clamping : Preventing numerical instability by constraining cosine similarity values to a safe range, avoiding potential division-by-zero or overflow errors during backpropagation.
Gradient Norm Clipping: Mitigating exploding gradients by limiting the magnitude of parameter updates, ensuring stable training and preventing potential model divergence.

Modification- Changing the Loss to Cosine Distance (Combined Loss)

We combined cosine similarity and Charbonnier loss to create a comprehensive loss function that captures both directional color similarities and robust pixel-wise error measurements.

The cosine similarity component ensures color vector alignment and perceptual color integrity by focusing on directional relationships, while the Charbonnier loss provides resilience to outliers and smooth gradient behavior across different image regions.

Architecture tweaking

Made the model deeper by adding an encoder block and decoder block

The transition from sRGB to Lab* color space needs a deeper model, as the perceptually uniform and luminance-separated nature of Lab* enables more sophisticated feature extraction that can be progressively leveraged through deeper architectural layers. By introducing additional encoder and decoder blocks, the model gains enhanced ability to disentangle and reconstruct color information, particularly exploiting the Lab* space's unique channel separation of luminance from chrominance, thereby it improved color-aware feature learning and transformation capabilities.

Results

Shadow image

Ground truth

Final model

RGB

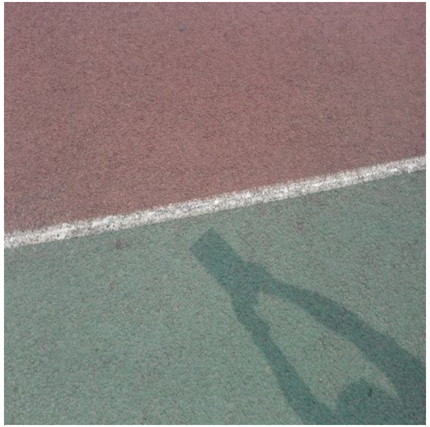
LAB

LAB with Cosine

grad-clip loss-clamp

Combined loss

Results - Continued

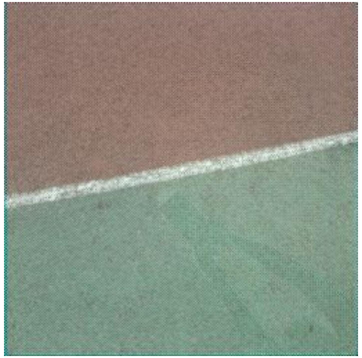

Shadow image

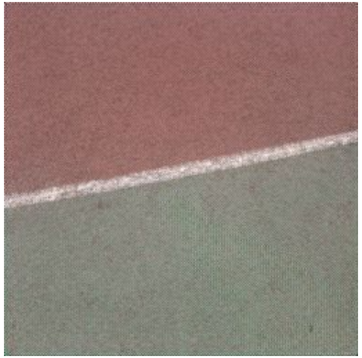

Ground truth


Final model


RGB

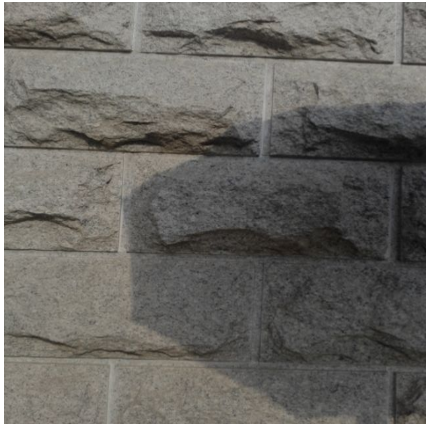

LAB


LAB with Cosine


grad-clip loss-clamp


Combined loss

Results - Continued



Shadow image

Ground truth

Final model

RGB

LAB

LAB with Cosine

grad-clip loss-clamp

Combined loss

Quantitative Results

Model	PSNR	SSIM
RGB	30.421639	0.965128
LAB	30.555857	0.981886
LAB with cosine	25.431281	0.949069
-With grad clip	28.973819	0.953715
- Combined loss	30.358224	0.983740
Deeper model	31.777911	0.985228

Failures

Discussion

This time our whole presentation is open for discussion instead of a single question