RetinexDL: A Deep Learning Low-Light Image Enhancer using Retinex Theory

John Asencio-Giron
June 12, 2024

Project Theme: How many photons does it take to see an object?

1. Introduction:

Although we have made many advancements in cameras and image processing, a big problem still presents itself, taking images in low light environments. Low-light environments often result in limited photon counts resulting in images with poor visibility, high noise levels, and reduced contrast. This limits their utility in practical applications. Enhancing such images to improve visibility and detail retrieval has become a critical area of research in computer vision and image processing.

The application of such a task spans across many fields. In astronomy, capturing images of space in low-light conditions is essential for the study of distant galaxies, stars, and other celestial bodies. Similarly, in surveillance, low-light image enhancement plays a crucial role in improving the effectiveness of surveillance cameras, enabling better identification of objects and people in dark environments. In medical imaging, being able to obtain clear images in low-light conditions is vital for diagnostic purposes, such as in endoscopy or radiography. This is why it is an important task to research.

In this paper I plan on discussing and experimenting with different techniques that have been used to enhance low-light Images. There are many manual ways to do this. The simplest one involving no algorithmic work is tinkering with individual pixel values via exposure, contrast and RGB values, via software like Adobe Lightroom or Photoshop. What are other ways we can tackle this problem? In my preliminary research, I found utilizing deep learning and neural networks was a promising approach to our problem.

1.1 Retinex Theory:

I plan on creating a simple model showing the power of such methods, inspired by retinex theory. According to the principles of the Retinex theory, a low-light image I of dimensions $H \times W \times 3$ can be separated into two components, the reflectance image R (of the same dimensions) and an illumination map L of dimensions $H \times W$. This separation is represented mathematically as:

$$I = R \times L$$

It's important to note that in the Retinex model, we assume that the observed image I is free from any corruptions. However, this assumption contradicts the reality of under-exposed scenes where problems are common.

2. Related Work:

2.1 Self-Supervised Deep Learning Methods:

Self-supervised approaches for low-light image enhancement have gained significant attention due to their ability to learn from unpaired or low-light-only data. Zhang et al. [3] proposed a self-supervised method inspired by Retinex theory and information entropy theory. Their network, trained exclusively on low-light images, quickly and effectively enhances image quality, making it a promising solution for low-light photo enhancement. Additionally, I found another method which is guided but not all the way supervised [4]. This approach enhances images by utilizing a multi-branch neural network guided by attention maps. By training on a synthetic dataset for increased diversity, their method also provides ideal results. There are also semi-supervised approaches which provide ideal results. [6]

2.2 Retinex Theory-Based Methods:

Retinex theory, based on human visual perception, has provided a basis for the development of low-light image enhancement techniques. Building upon this theory, the neural network KinD (KinDling the Darkness), decomposes low-light images into illumination and reflectance components [2]. KinD offers light adjustment capabilities while maintaining efficiency, reflecting its suitability for practical applications. Similarly, Retinexformer, a one-stage Retinex-based transformer processes images in a similar manner [5]. By integrating a Retinex-based model with an Illumination Guided Transformer, Retinexformer effectively uncovers hidden details in low-light conditions. This fusion of theory-driven design and transformer architecture showcases the potential of combining traditional insights with modern deep learning techniques. These two methods are very important to my research as I want to build my model drawing upon the methods used in these techniques.

3. Methods:

In my preliminary research I found that a lot of the methods used to enhance low-light images use Retinex theory coupled with deep learning. For example, Wang et al. created a one-stage Retinex based CNN, called DeepUPE, to directly predict the illumination map.

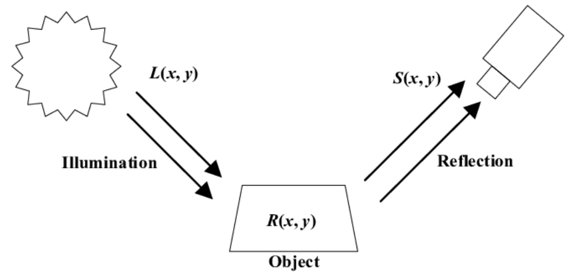

Figure 1: Diagram illustrating Retinex Theory
https://www.researchgate.net/figure/Illustration-of-Retinex-theory_fig1_348964295

When making my model, I wanted to draw upon that theory and split the images into the reflectance image and the illumination map. From figure one we see that from the object we must split into the reflection and illumination, the model must predict the illumination, and can then calculate the reflection based off of that. I propose a convolutional neural network model that can be trained only on low-light images and does not need any ground truth images, or light images. Therefore, it is self supervised. This can be done by having different layers in the model. Now, the problem Retinex has is it assumes there are no corruptions. Corruptions could be things like aberrations, but most importantly noise. Noise is very common in low-light images.

3.1 Process and Preliminary Results:

I do not have too much experience with Deep Learning, so the first thing I wanted to do was set up my model and my testing environment. The model as it stands currently consists of 3 layers. This is not including my plans for using retinex theory. I would like to research my process for that a little bit longer, but I needed to see how I could do it. My data is from the index of images from kodak. I load 15 of those images in, preprocess them, then feed them into the model. The model's performance seems to learn more with each epoch, which is what we want.

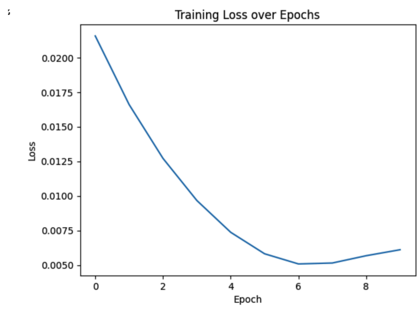

Figure 2: Graph output of average loss across epochs

It is to note, that the loss is expected to be low as the model is self supervised, and I do not want to use the original normal images. The model is being built like this to simulate real learning. I used a mean-squared loss to guide the learning process of the model, comparing the pixel values of the original low-light image, and the enhanced image. The resulting loss value serves as a measure of how well the model's predictions align with the original low-light images. Using techniques like gradient descent, the model learns to minimize loss, while generating enhanced images that appear better than the original. However, currently without any of the proposed Retinex techniques, the model seems to output an image that is less noisy as a byproduct of the current state.

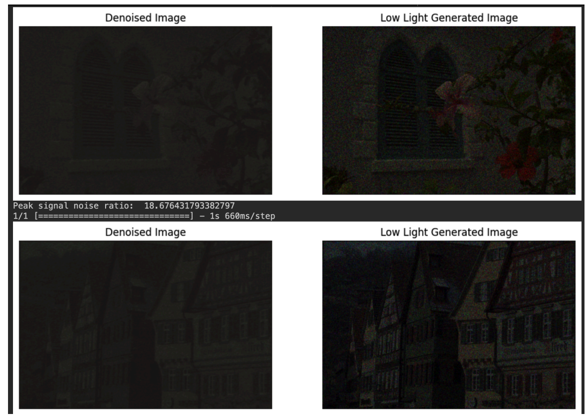

Figure 3: Model Outputs without Retinex techniques. The preprocessed low-light image is on the right, the enhanced image generated by the model is on the left.

In the outputs, the model also seems to dim the image and desaturate it a bit to induce a less noisy image it deems as better. I also took the Peak Signal Noise Ratio for each of the 15 images and took the average of all 15. The average Peak Signal Noise Ratio for every image was ~19.40476. This number is not the highest, but it does indicate moderate enhancement between the original and enhanced image.

After doing research on implementations [5, 3, 2], I saw that one of the negatives of using Retinex techniques is indeed the image may be lighter, but due to the assumption that there are no corruptions the image produced may still be noisy. My assumption was that after implementing the Retinex techniques, I will have to apply another technique in order to denoise the image. If this current state of the model translates over then there will be no need for that. The next steps will be implementing the Retinex Theory techniques to the model which will involve adding more layers.

4. Applying Retinex Theory:

Applying Retinex Theory to my deep learning model was fairly straightforward, but it did have a lot of roadblocks. First I started by obtaining an illumination map from my image. After getting an illumination map, we can get the reflection. At first I was getting only images with light that were very noisy. I knew this was going to be one of the problems I would encounter because of the assumption of Retinex theory. My hypothesis was since the model at first was producing albeit dark images, they were a lot smoother than the original low-light images and this would transfer over to the image enhancement. This assumption was wrong and my model would produce an image that looked like this:

Figure 4: Model Outputs without Retinex techniques. The image on the left is the output of my model and the image passed in was the one on the right.

As seen in Figure 4, the image is lighter and you can make something out of it, but it is very noisy. The model needed to produce a completely light image without any noise. To counter this roadblock, I introduced new denoising layers to our illumination map, that in theory would have helped us produce a non noisy image. In Yu Zhang's model they introduced smoothness functions in their model to combat this. I took a similar approach, using smoothness operations to determine a loss, and aggregating that to my total loss. While this did add positive effects to my output it was not to say it was a success. Since time was dwindling, efforts were shifted to post processing techniques to enhance my image. I decided to use some of the denoising techniques we used in class for my images. The denoise_wavelet was the first denoise function experimented with.

Figure 5: Denoise Wavelet Post Processing output. Sharpened image using Lappachian filters is on the right.

As you can see, the image looks too blurry and the sharpening makes it a little sharper at the cost of sacrificing more noise. Experimenting with other denoising techniques did not yield significantly better results, although the denoise_tv_bregman method was my personal best performing method.

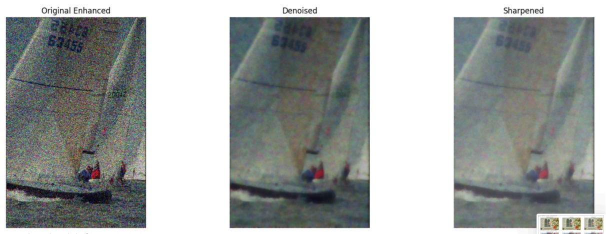

Figure 6: Final RetinexDL outputs with post processing techniques.

5. Possible Reasons for Noisy Images:

5.1 Loss Function

Once I added the extra layers the loss for each epoch was much much greater. Yes the trend was the same, it would decline rapidly after each epoch but it was much higher. This could be attributed to two things. One being I added a smoothing factor to the total loss which would cause it to rise. However even before I added that it was still very high. I believe adding in new layers was the culprit. This could be a reason that the model could not produce non noisy images. The model was not able to learn efficiently. Now, I did try experimenting with different ranges of epochs, but I stayed at a solid 10 for a majority of the time.

5.2 Reflection Image and Illumination Map Calculation

My original attempt at trying to produce a non noisy image was adding more layers to the illumination map. However this did not work. Could it be that I did not even produce an accurate illumination map? If I did not this could have thrown the calculation off for the reflection image, throwing off the enhanced image.

5.3 Time Constraints

The timeline was a factor in the development of RetinexDL. Yes, the project was a term-long project, but the true time used to develop the model was about half of that, on and off. Most of the time was spent researching and finding ways to use deep learning to enhance low light images. Once the model was built, it would take approximately 3 minutes each training run, which for testing purposes is not the best. In the future, I look to continue my work on this project, and apply it to other useful scenarios.

6. Use Cases

A very interesting use case for this project is in low-light video recording. In theory, the model could continuously learn live from select frames and enhance video image quality. This could be via an app on a phone using the phone's normal camera to receive the input images. This use case is the most interesting to me because it is the most practical. Newer phones are getting good at capturing images in low-light situations, so enhancing only an image might not yield better results than the existing phone camera.

7. Conclusion

There are many different ways to enhance low light images. One way to do it is by using deep learning models. You can use models to train with light and dark images, and some have even been developed to train only on dark images. The model cannot do it on its own though, it needs an algorithm behind it to help it produce light images. Mine used Retinex theory hence the name RetinexDL. Retinex theory splits the image into two components, a reflection image and an illumination map, which can be used to enhance the light in the image. One of the downsides of Retinex theory is that it does not account for corruptions inside of the image i.e noise. While RetinexDL did produce a lighter image, the quality is not as crisp as intended, and needed to be post processed to help the smoothness of the image. The techniques used to enhance the image inside of the model did not seem to change the output too much. This could have been attributed to a poorly optimized loss function. Another factor was time constraints. While my RetinexDL was able to produce a lighter image, expectations were higher. In future iterations I hope to enhance it and apply it to the use case mentioned above.

References:

[1] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, Runmin Cong. *Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement*. https://paperswithcode.com/paper/zero-reference-deep-curve-estimation-for-low, The paper presents Zero-Reference Deep Curve Estimation (Zero-DCE), a deep learning method for enhancing image brightness. Zero-DCE trains a lightweight network, DCE-Net, to estimate dynamic range adjustment curves without needing paired data. Through specially designed loss functions, it achieves efficiency and effectiveness across different lighting conditions, outperforming existing methods.

[2] Yonghua Zhang, Jiawan Zhang, Xiaojie Guo. *Kindling the Darkness: A Practical Low-light Image Enhancer,* https://paperswithcode.com/paper/190504161, This work introduces KinD (KinDling the Darkness), a network aimed at enhancing low-light images by decomposing them into illumination and reflectance components. It is inspired by Retinex theory, which suggests that human vision relies on comparing light from different parts of a scene. The model is trained with paired images taken under different exposure conditions, eliminating the need for illumination data. Their experiments show KinD's superior performance, offering user-friendly light adjustment capabilities, all while maintaining efficiency for practical use.

[3] Yu Zhang, Xiaoguang Di, Bin Zhang, Chunhui Wang. *Self-supervised Image Enhancement Network: Training with Low Light Images Only.* https://paperswithcode.com/paper/self-supervised-image-enhancement-network, This paper presents a self-supervised method for improving low-light images using deep learning. Also inspired by the Retinex theory and by the information entropy theory, they propose a simple network that separates illumination and reflectance. Trained only on low-light images, their method achieves state-of-the-art results quickly and effectively, making it a promising solution for enhancing low-light photos. The model is even capable of being trained on just one photo.

[4] Feifan Lv, Yu Li, Feng Lu. *Attention Guided Low-light Image Enhancement with a Large Scale Low-light Simulation Dataset.* https://paperswithcode.com/paper/attention-guided-low-light-image-enhancement, The paper presents a new method for improving low-light images, using a multi-branch neural network guided by attention maps. They create a synthetic dataset for training, which offers more diversity than existing datasets. Their approach outperforms current methods in enhancing brightness, reducing noise, and improving overall image quality.

[5] Yuanhao Cai, Hao Bian, Jing Lin. *Retinexformer: One-stage Retinex-based Transformer for Low-light Image Enhancement.* https://paperswithcode.com/paper/retinexformer-one-stage-retinex-based, This paper also introduces another method to enhance images based on Retinex theory. They call their model Retinexformer. Retinexformer enhances low-light images by employing a one-stage Retinex-based framework combined with an Illumination-Guided Transformer (IGT). This approach effectively estimates illumination information and restores corruptions hidden in the dark or introduced during the light-up process.

[6] Wenhan Yang, Shiqi Wang, Yuming Fang, Yue Wang, and Jiaying Liu. *Band Representation-Based Semi-Supervised Low-Light Image Enhancement: Bridging the Gap Between Signal Fidelity and Perceptual Quality.* http://39.96.165.147/Pub%20Files/2021/ywh_tip21_2.pdf