Mid-Term Report: Low-light Image Enhancement using Deep Learning

Ping-Han Chuang (pchuang4), Sunny Shen (sshen98)

University of Wisconsin-Madison

1 Introduction

In this project, the goal is to explore and research how we could use deep learning to enhance low light images. Low-light image enhancement has wild applications in our daily life and in different scientific research fields such as night surveillance, automated driving, fluorescence microscopy, etc. This technique can improve the usefulness of an image to satisfy human viewing. Also, it can make low-light images applicable for applications in autonomous driving, scientific data capture, and general visual enhancement. For example, self-driving car can see humans in the dark environment to prevent accidents from happening with this technology. Therefore, low-light image enhancement is very significant and is worth studying.

Regardless of the technological advancements, there is still a long way to improve this task. Low-light images typically suffer from two problems which are low visibility and high noise. Low visibility shows that the image has small pixel values due to few amount of photon, and high noise would dominate and disrupt the image content, so the overall signal-to-noise ratio is low. We are trying to survey as well as implement methods based on the Awesome Low Light Image Enhancement GitHub repository [1] which provides a list of resources related to low light image enhancement, including datasets, methods, papers, and metrics to see the performance. Overall, we'd like to re-implement a method and then try to add some modifications upon it to see if we could improve the performance.

2 Related Work

Many applications aim to enhance brightness, contrast, and reduce noise from the images in an on-board real-time manner. Lore et al. [2] propose a deep autoencoder-based approach LLNET to identify signal features from low-light images and adaptively brighten images without over-amplifying/saturating the lighter parts in images with a high dynamic range. LLNET shows that a variant of the stacked-sparse denoising autoencoder (SSDA), which are a sparsity-inducing variant of deep autoencoders that ensures learning the invariant features embedded in the proper dimensional space of the dataset in an unsupervised manner, can learn from underlying signal characteristics/synthetically darkened and noise-added training examples to adaptively enhance images taken from natural low-light environment.

Since it is hard to collect training data with low-light and normal-light images, Jiang et al. [3] use GAN (Generative Adversarial Network) to do unsupervised learning. They use information extracted from input to regularize training. In this model, there are one generator and two discriminators. For the generator, it adopt U-Net [4] with attention mechanism to get multi-level features with rich texture information so that generator can generates real enough images to fake discriminators. As for discriminators, one is local and the other is global. Global discriminator is responsible for discriminating real or fake on image-level, while local discriminator is for local patch level. With this structure, the network can prevent over- or under-enhancement in local region.

3 Solution

We plan to implement an algorithm called KinD proposed by Zhang et al. [5]. KinD is a deep learning-based solution built on top of the Retinex theory. By decomposing images into illumination layer and reflectance layer, adjusting light on illumination layer won't affect details and possible degradations on reflectance layer and degradation removal can be done on reflectance layer. There are three processes in this solution: layer decomposition on input image, illumination adjustment on illumination layer, and degradation removal on reflectance layer. Each of them correspond to a neural network, so there are three sub-networks in this solution. Three sub-networks use either l1 norm or l2 norm as their loss functions.

Since there is no ground truth for layer decomposition, layer decomposition layer is trained by paired images. To be more specific, for a specific scene, there are short-exposure image and long-exposure image. With paired image, loss function is designed by comparing illumination layers and reflectance layers of paired images. For the illumination adjustment net and reflectance restoration net, there is also no ground truth, so similar concept is used to design their loss functions. For reflectance restoration net, it tries to make processed reflectance layer of low-exposure image approach that of high-exposure image. Similarly, illumination adjustment net make adjusted illumination layer of low-exposure image close to that of high-exposure image. During testing, only low-exposure image is needed.

Zhang et al. [5] used LOL dataset [6] to train the model. We plan to use LOL dataset as baseline and then maybe use MIT-Adobe FiveK [7] to see whether there is performance difference.

4 Evaluation

We will evaluate the results by objective and subjective methods. For the objective evaluations, we will also compute PSNR and SSIM. They are commonly used metrics for image quality. PSNR is the peak-signal-to-noise ratio. Higher PSNR indicates that result images are closer to the ground truths. SSIM is structural similarity index measure, which shows the similarity between structures of two images. Higher SSIM is better. Although these two objective metrics are popular, they can't reflect the visual perception of humans. Therefore, we will also compare input images, ground truths, and output images by simply using our eyes. By placing those images side-by-side, we can tell whether there are any artifacts, like color inconsistency or over-enhancement on output images.

As for the testing data, we plan to use LOL [6] and MIT-Adobe FiveK [7]. Besides, we will test the method on photos captured by our cell phones.

5 Progress

Since submitting the proposal, We have researched the KinD algorithm, taken deep dive into the architecture of three layers network, collected the dataset from LOL [6], created a project website using Github.io [8] as well as Docusaurus [9] which saved the implementation time and thus helped us to focus on the page content, and started to implement the decomposition, illumination adjustment, and reflectance restoration nets using PyTorch [10].

There are two reasons why we use PyTorch to re-implement:

- PyTorch is the most popular framework to design deep learning algorithms in recent years.
- The source code provided by the authors was implemented in an older version of TensorFlow which is hard to make modifications.

Our goal is to achieve the result presented in Kindling the Darkness: A Practical Low-light Image Enhancer from Zhang et al. [5]. For example, given the inputs (Figure 1, 2, and 3), the corresponding results are Figure 4, 5, and 6.

6 Timeline

- September 25th - Topic brainstorm
- September 30th - Survey
- October 6th - Project proposal complete
- Early-October - Deep dive into KinD (paper study)
- Mid-October - Create a webpage, collect datasets, and start experiments to create baseline result
- End-October - Research to implement KinD with PyTorch [10]

Figure 1: Example 1

Figure 2: Example 2

- November 20th - Finish implementation and start training
- November 27th - Evaluate on LOL dataset and measure quantitative metric
- December 6th - Final presentation
- December 15th - Project webpage complete

Figure 3: Example 3

Figure 4: Example 1 after KinD

References

[1] "Awesome Low Light Image Enhancement". In: URL: `https://github.com/dawnlh/awesome-low-light-image-enhancement#readme`.

[2] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. "LLNet: A deep autoencoder approach to natural low-light image enhancement". In: *Pattern Recognition* 61 (2017), pp. 650–662.

Figure 5: Example 2 after KinD

Figure 6: Example 3 after KinD

[3] Yifan Jiang et al. *EnlightenGAN: Deep Light Enhancement without Paired Supervision*. 2019. DOI: `10.48550/ARXIV.1906.06972`. URL: `https://arxiv.org/abs/1906.06972`.

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 2015. DOI: `10.48550/ARXIV.1505.04597`. URL: `https://arxiv.org/abs/1505.04597`.

[5] Yonghua Zhang, Jiawan Zhang, and Xiaojie Guo. *Kindling the Darkness: A Practical Low-light Image Enhancer*. 2019. DOI: `10.48550/ARXIV.1905.04161`. URL: `https://arxiv.org/abs/1905.04161`.

[6] Chen Wei et al. "Deep Retinex Decomposition for Low-Light Enhancement". In: *British Machine Vision Conference*. 2018.

[7] Vladimir Bychkovsky et al. "Learning Photographic Global Tonal Adjustment with a Database of Input / Output Image Pairs". In: *The Twenty-Fourth IEEE Conference on Computer Vision and Pattern Recognition*. 2011.

[8] *Github.io*. Github, 2022. URL: `https://pages.github.com/?(null)` (visited on 11/10/2022).

[9] *Docusaurus*. Meta Platforms, 2022. URL: `https://docusaurus.io/` (visited on 11/10/2022).

[10] *PyTorch*. The Linux Foundation, 2022. URL: `https://pytorch.org/` (visited on 11/09/2022).