LOW-LIGHT IMAGE ENHANCEMENT USING GENERATIVE ADVERSARIAL NETWORKS

Daryl Fernandes[1], Joel Dsouza[2], Lloyd Dsouza[3], Kunal Meher[4]

[1-4]Department of Computer Engineering, Xavier Institute of Engineering, Mahim(w), Mumbai 400016, India

---***---

Abstract - *Owing to environmental and/or technological limitations, many images are captured under poor lighting conditions. These involve images captured at night and/or unbalanced ambient lighting conditions, or if the object is located in front of a light source, image is under-exposed during the capture of photographs. Such low-light images, have increased noise levels and minimum visibility, hence suffering from compromised quality and detail. Our proposed approach takes a low-light image as an input and creates a well-enhanced image at its output. To increase the efficiency of translating low light or night time images to daytime images, we use a Generative Adversarial Network (GAN) implementation.*

Key Words: Low Light, GAN, Image Enhancement, Computer Vision

1. INTRODUCTION

Many photos are often captured under low ambient lighting conditions due to environmental and/or technical constraints. These include low-light and/or unbalanced lighting conditions in the environment, placement of object in front of a light source, and under-exposure during image capturing. Such low-light photos suffer from compromised quality and information, along with increased noise levels. This quality affects viewers experience while loss of information leads to wrong message being communicated, such as inaccurate object/face recognition. In particular, our proposed method will take a low-light image as an input and produce a well enhanced image as its output. We use a modified version of the original Cycle Generative Adversarial Network (CycleGAN) [3] to enhance the performance of the conversion of low light or night time images to daytime images. Furthermore, we have also considered using attention modules from EnlightenGAN [4] to enhance the background details of the image. Using our modified CycleGAN and attention modules from EnlightenGAN we generate an output image with a much better quality, with more information and lower noise levels.

1.1 SCOPE

Our proposed system will enhance the low-light image to a more informative/detailed image, similar to the ones captured in well-lit conditions. The user will be able to access our system through a user-friendly website. The current proposed system works with images of resolution 256x256. Since images with a higher resolution take longer to process and train. We plan to further improve this in the future by processing higher resolution images.

2. RELATED WORK

In today's fast-paced world users desire fast and simple technologies which make their lives easier and are more reliable. Our method allows an average smartphone user to enjoy the simple click experience while capturing a decent picture without the hassle of all the related technical jargon. Under good lighting conditions, there is no issue as the camera and software are intelligent enough to adjust those parameters to the user's liking but in low light conditions, even the best cameras tend to give a poor result.

2.1 RETINEX-NET

Retinex-Net [8] has carefully designed constraints and parameters for this highly ill-posed decomposition, which may be limited by model capacity. It assumes that observed images can be decomposed into reflectance and illumination. The drawback of this method is that the parameters of the kernels tend to depend on artificial settings, which leads to reduced accuracy and flexibility in some cases. Some lightness enhancement is conducted on illumination by a network called Enhance-Net.

2.2 LIME (LOW-LIGHT IMAGE ENHANCEMENT)

LIME [10] uses Illumination Map Estimation to illuminate each pixel is first estimated individually by finding the maximum value in R, G, B channel, based on Retinex modelling. Then it refines the initial illumination map by imposing a structure before it, as the final illumination map. This method has a hard time with over-exposed artifacts and suffers severe noise amplification of the image.

2.3 ZERO-DCE (ZERO REFERENCE DEEP CURVE ESTIMATION)

Zero-DCE [2] formulates light enhancement as a task of image-specific curve estimation with a deep network. There are three objectives in the design of such a curve. Each pixel value of the enhanced image should be in the normalized range to avoid information. The form of this curve should be as simple as possible and differentiable in the process of gradient back propagation.

2.4 PIX2PIX

Pix2Pix [7] titled "Image-to-Image Translation with Conditional Adversarial Networks" It is an extension of the Conditional Adversarial Networks algorithm. The networks not only learn the mapping from the input image to the output image but also the loss function to train this mapping, allowing the same generic approach to problems that traditionally would require very different loss formulations.

2.5 CYCLE GAN

CycleGAN is a technique that involves the automatic training of image-to-image translation models without paired examples. The transfiguration makes sense given the images are similar in size and structure, except for their colouring. This simple yet powerful technique achieves visually impressive results in translating the photographs.

2.6 ENLIGHTEN GAN

EnlightenGAN trains on unpaired images, thus making it flexible to adapt to new images that the model was not trained on. It enhances and adjusts various aspects of the image such as the contrast and white balance in a way that doesn't overexpose the image which most other architectures tend to do.

3. PROPOSED METHOD

In our proposed system we have used a similar architecture to that of CycleGAN and EnlightenGAN. These models are an extension of GAN networks, the main difference is their generator Neural Network. This network comprises of Encoder and Decoder modules, the encoder extracts the main features from the image while the decoder utilizes the extracted features to generate the output.

3.1 PREPARING DATASET

For training we have used LOw Light Paired (LOL) [9], Synthetic Image Pairs from Raw Images [9], Sony and Fuji low-light images [1], Single Image Contrast Enhancement (SICE) [6] datasets. Since there is no single dataset containing enough images to train the model, we decided to club together various datasets as per our requirements. We have used the Google Scraped Images Dataset [5] and created a noisy and underexposed version of each image to create a paired training dataset for our use.

4. METHODOLOGY

We propose a Generative Adversarial Network (GAN) based approach for low-light image enhancement. GAN's learn to generate an image, thus we can also get images that the training dataset has never seen before. Although we have used paired dataset for training our model, GAN's are capable of learning from un-paired dataset.

In CycleGAN model, the shape of artifacts in the output image may be different from the ground truth as this model recreates an image based on the features extracted from the input image. In EnlightenGAN model, it gives better results in face, building, and natural scenes, but works poorly on other objects. Our model contains two major components Generator and Discriminator.

4.1 GENERATOR

The main objective of the generator is to produce a meaningful output from noisy input. Our Generator comprises of Encoder and Decoder network.

4.1.1 ENCODER

The Encoder Network receives a 4-dimensional array as the input which has 2 parts the first being the number of images and the second is an RGB image of resolution 256 x 256. Encoder Network is made up of 21 convolutional layers. Where each layer has a Convolution2D, BatchNormalization, LeakyReLU. Convolution2D layer creates a convolution kernel that is convoluted with the layer input to produce a tensor of outputs. Batch normalization applies a transformation that maintains the mean output close to 0 and the output standard deviation close to 1. LeakyReLU is a leaky version of Rectified Linear Unit (ReLU), it allows a small gradient when the unit is not active. The input 4-dimensional array is processed and passed on to the Decoder.

4.1.2 DECODER

The Decoder Network receives a 4-dimensional array from the Encoder as the input. Our Decoder Network is made up of 10 convolutional layers. Where each layer has a UpSampling2D, LeakyReLU, Convolution2D, BatchNormalization. UpSampling2D increases the dimensions by the specified size, there are two methods of interpolation Nearest Neighbour, Bilinear transformation. We decided to use UpSampling2D in place of Convolution2DTranspose, since it is cheap in terms of computational power. Convolution2D layer creates a convolution kernel that is convolved with the layer input to produce a tensor of outputs. Batch normalization applies a transformation that maintains the mean output close to 0 and the output standard deviation close to 1. LeakyReLU is a leaky version of Rectified Linear Unit (ReLU), it allows a small gradient when the unit is not active. The input 4-dimensional array is processed and passed on to the Discriminator.

4.2 DISCRIMINATOR

Discriminator is a Convolutional Neural Network (CNN) that discriminates between the generated image and the input given to the generator. Discriminator Network consists of 6 Layers which comprises of Convolution2D, BatchNormalization, LeakyReLU.

4.3 LOSS METRICS

We have used Mean Absolute Error (MAE) as our loss function for Generator Network while Binary Cross Entropy for the Discriminator Network.

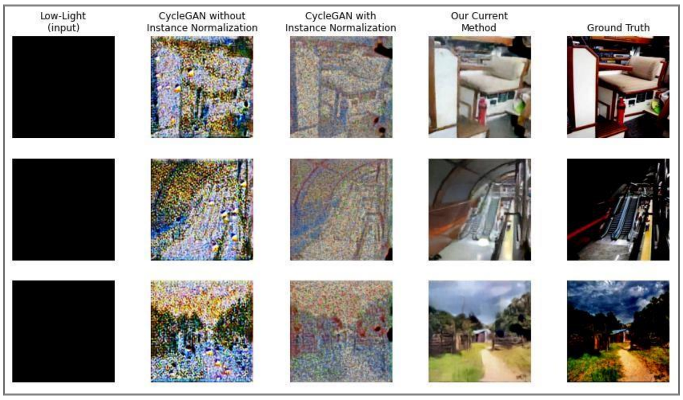

Fig -1: Comparison between our approaches

5. EXPERIMENTS

We have tried multiple different approaches by tuning the parameters and model network. We started by training using Cycle GAN approach. In Cycle Gan, we trained on two different model networks; one with Instance Normalization and another without while changing the convolutions in the Generator network. In this approach we train the model to translate the image from low-light domain to enhanced domain and vice versa. By doing so we trained the Discriminator further so it could discriminate between real and fake image in both domains.

Our Current model consists of skip connections between the Encoder and Decoder Networks. These connections extract details of an image and pass on to the decode network, these details help in generating the edges of the objects. There is another Transformation network between the Encoder and Decoder, that also processes the details from the Encoder and convolved image.

The output comparison of the various models that we have implemented are shown in Fig -1. As we can see the Cycle GAN without Instance does not give reproduce proper colours. Cycle GAN outputs also do not have accurate edges, while our model gives much better edge detection.

We trained our model using Google Colab, 12GB RAM, Nvidia Tesla T4 16GB GPU / Nvidia Tesla P100 16GB GPU. Due to limited RAM memory, training was done on batches of size 12 and images were resized to 256x256.

6. CONCLUSION

In this paper, we have proposed a method to tackle the challenging task of enhancing low-light images. The proposed method utilizes GAN architecture. We have utilized paired dataset of low-light and bright images, which comprises of five datasets merged together and a custom dataset. As seen from experimentation we are able to enhance images with low lighting scenarios. There is, of course, room for improvement in the colour accuracy department. The current model works on images of resolution 256 x 256 which can further be upgraded with the availability of better hardware system.

References

[1] C. a. C. Q. a. X. J. a. K. V. Chen, "Learning to See in the Dark," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2108.

[2] C. G. a. C. L. a. J. G. a. C. C. L. a. J. H. a. S. K. a. R. Cong, Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement, vol. 2001.06826, 2020.

[3] N. R. B. J. H. K. M. A. a. K. R. P. S. W. Cho, "Semantic Segmentation With Low Light Images by Modified CycleGAN-Based Image Enhancement," IEEE Access, vol. 8, pp. 93561-93585, 2020.

[4] Y. a. G. X. a. L. D. a. C. Y. a. F. C. a. S. X. a. Y. J. a. Z. P. a. W. Z. Jiang, "Enlightengan: Deep light enhancement without paired supervision," CoRR, vol. abs/1906.06972, no. 1906.06972v1, 2019.

[5] V. Basu, "Low-light Image Enhancement using CNN,"

2018. [Online]. Available:

https://www.kaggle.com/basu369victor/low-lightimage-enhancement-with-cnn/data.

[6] J. a. G. S. a. Z. L. Cai, "Learning a Deep Single Image Contrast Enhancer from Multi-Exposure Images," IEEE Transactions on Image Processing, vol. 27, pp. 20492062, 2018.

[7] P. I. a. J.-Y. Z. a. T. Z. a. A. A. Efros, Image-to-Image Translation with Conditional Adversarial Networks, vol. 1611.07004, 2018.

[8] C. W. a. W. W. a. W. Y. a. J. Liu, Deep Retinex Decomposition for Low-Light Enhancement, vol. 1808.04560, 2018.

[9] J. L. Chen Wei and W. Wang and W. Yang, "Deep Retinex Decomposition for Low-Light Enhancement," in BMVC, 2018.

[10] X. G. a. Y. L. a. H. Ling, "LIME: Low-Light Image Enhancement via Illumination Map Estimation," IEEE Transactions on Image Processing, vol. 26, pp. 982-993, 2017.