An In-Focus from a Single Input Image

1st Onyinyechukwu Onyemaobi
Computer Science Department (of GSU)
Immersive Media Computing REU (of GSU)
Atlanta, United States
oonyemaobi1@student.gsu.edu

2nd Saeid Belkasim
Computer Science Department (of GSU)
Immersive Media Computing REU (of GSU)
Atlanta, United States
sbelkasim@gsu.edu

***Abstract*—All in focus capture of an image is often limited by the quality of the camera. Usually, when images are captured with objects of different distances from the camera, all of the objects are not in focus. The object at the center of focus is clear, and objects that are behind or in front of it become less and less focused as the distance increases. My task this summer was to research the different ways to change a single image into an all in focus image. In this work, several different ways are applied to solve this out-of-focus problem including deep neural networks, convolution kernels, and histogram equalization. To the limited time concern, only the basic convolution kernels and histogram equalization methods are implemented, but we are able to extend implementation in the future.**

***Index Terms*—generative adversarial network, convolution filter, histogram equalization**

Fig. 1. This is the provided image for the project. The focal point is in the region with the strawberries. The other objects, fruits in this case, become less and less focused as they move further from the focal point—the strawberries.

Funding by National Science Foundation

I. Introduction

The focus of an image depends on the camera's aperture and light. Still, images taken with a wider aperture usually contain areas that are out-of-focus. Ideally, if an image have all pixels in focus, we will be able to generate what so called all-in-focus image. Post-capture control of the depth of field of an image would be very useful to photographers [1]. Even with modern technology, it is not exactly easy to capture all-in-focus images. Doing this requires certain pieces of information like the in-focus intensity and true point spread function of the focused point [1], the depth estimation map [3], or the focus measure [1]. According to Johnson, with an accurate depth map, the depth of an image can be manipulated with blurring and deblurring operations—moving objects further and closer, respectively [3]. Most deep learning algorithms use this kind of information to train their networks to generate in-focus images. My goal was to figure out a way to generate an in-focus image from an input image without this information. In this paper, I will discuss the various methods I researched this summer to convert the provided image (Fig. 1) into an all-in-focus image.

There are various amounts of research using deep learning neural networks to bring an image into focus. A deep neural network is a network with some level of complexity, and it has several hidden layers, which is why it is referred to as "deep" [2]. These hidden layers extract valuable information from the input images. There are several methods introduced to generate an in-focus image from an input image using deep neural networks. These networks are trained with the input image and other inputs like depth estimation, focus measure, true point spread function, and in-focus intensity. There is not a network that is trained to generate an in-focus image from an input image. This has to do with the lack of a data set with out of focus and in focus pairwise images.

There are also manual methods to focus an image. The visual result of focusing an image could be compared to sharpening an image or adding contrast to the image. An in-focus image has more luminance which means more edges can be detected. Sharpening and adding contrast to an image has these effects.

An image can be sharpened using a convolution filter or kernel. A filter or kernel is a matrix, usually 3x3, that is convoluted with an image by placing the filter on each pixel in the image and multiplying it and its neighbors with the corresponding pixel in the filter and adding these values together to create a new value for the pixel [4].

In another way, an image's contrast can be increased using histogram equalization. A histogram is the graphical representation of an image, so it records the number of pixels with a specific intensity value. Histogram equalization is a digital image processing technique where the histogram of an image is stretched to contain more intensity values. This is a global technique, and it results in the image gaining higher contrast.

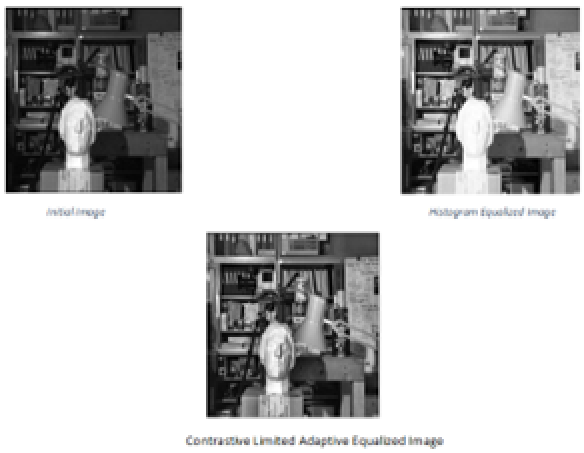

Fig. 2. This image is provided in Sudhakar's histogram equalization tutorial [7]. It shows the effects of histogram equalization. The contrast limited AHE image prevents the amplification of noise in the regions.

II. Related Works

There are several methods introduced centered around creating an all-in-focus image using deep learning neural networks. Sakurikar et. al's introduce the RefocusGAN by training two generative adversarial networks (GAN) using a data set of light field images with depth information [1]. It takes a single input image and outputs a fully focused image before sending it as input into the refocus network to change the focal point [1]. Xu et. al proposes a blind deconvolution method that removes defocus blurring from an image by suppressing segmentation error outputs and providing more accurate kernel estimation [8]. Tung and Hwang develop quantization and joint deblurring process to correct the effects created from depth estimation errors [3]. Kupyn et. al presented DeblurGAN, a, end-to-end learning model for motion deblurring based on the conditional GAN and content loss [16], and improves upon this method with DeblurGAN v2, which used a double discriminator network and introduces the Feature Pyramid Network into deblurring [18]. Jiang et. al proposes introduce EnlightenGAN, which does not use a pairwise image dataset, but information from the input itself to train the GAN to enhance the image. The method is said to be easily adaptable to enhancing various kinds of real-world images [17].

There are various convolution filters for sharpening. They are matrices that generally all have a positive number equal to

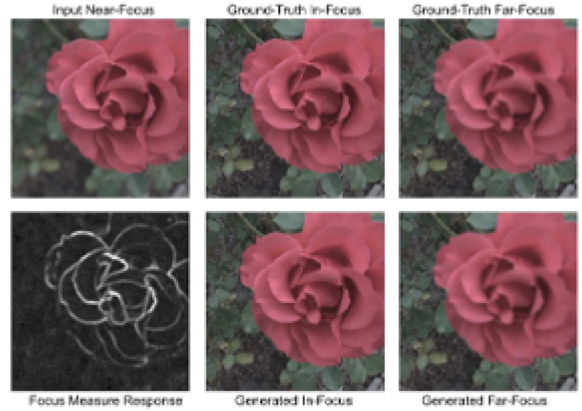

Fig. 3. This image shows the results obtained from the RefocusGAN implementation [1].

or greater than 5 in the middle element and are surrounded by -1's [5] [10]. The filter matrices can either be 3x3, 5x5, or 7x7 [10]. This can be adjusted or chosen by the user, depending on the input images.

Histogram equalization is a global technique, and it results in the higher contrast of an image. However, for out-of-focus issues like the one presented, a local technique would be more effective. Adaptive histogram equalization (AHE) is another equalization technique improves local contrast and edge definition in different regions of the image. However, this technique is known to amplify noise [7]. Contrast limited adaptive histogram equalization (CLAHE) was created to prevent the amplification of noise. The contrast is limited by using a clip limit, typically between 2 and 4 [6] [9], and it's used to clip the histogram before calculating the continuous distribution function [6]. Yadav et. al uses CLAHE for fog and noise removal to improve video quality in real time system [12].

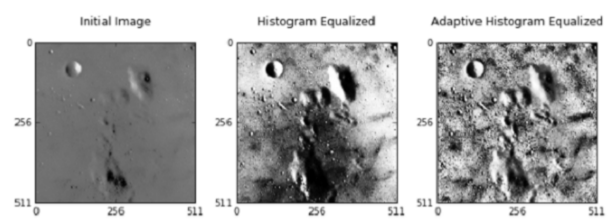

Fig. 4. This image was provided from Sudhakar's histogram equalization tutorial [7]. It shows the results of adaptive histogram equalization.

III. Research Activity

A. Deep Neural Networks

In the beginning of my research, my assignment was to create a neural network that could generate an in-focus image from a single image input. First, I familiarized myself with digital image processing and deep learning. I used the materials given to me from my mentor to do this. I also went through several Matlab tutorials and documentation—also, Matlab is the primary programming language for this project.

We started the literature review with a few keywords: sharpen, refocus, shallow depth of field. This led to finding many papers, and we decided to re-implement the method described in Sakurikar et. al's paper introducing the RefocusGAN [1]. A generative adversarial network has two neural networks called a generator and a discriminator. The generator takes random inputs and generates images, and the discriminator is supposed to classify the output as real or fake [14]. The generator's goal is to fool the discriminator into thinking the generated image is a real image. The first GAN in their model generates an in-focus image from an input image and its' corresponding focus measure. However, their dataset was not exactly available in a format I could use, and the description was slightly too vague for me to re-implement.

I decided to follow a GAN tutorial [15], and tweak that to take a single input image and generate it. This tutorial was in Python, and I found that there were countless in-depth tutorials for neural networks using the Python Language instead of Matlab. Python was also the language used in majority of the papers I read. After a long search, I came to find that a dataset with in-focus and out-of-focus pairwise images did not exist, and time would not allow me to make my own. Still, my research within the deep learning field allowed me to create a plan to tackle the problem.

B. Convolution Filter

During my research, I also wanted to explore manual ways of focusing an image. I met with a PhD research assistant, with a background in Image Processing, and she introduced pre-processing algorithms for images using convolution filters. I followed an article in implementing several convolution filters—including sharpening filters [5]. Even though the images were not fully restored, one can see in Fig. 5 and Fig. 6, the images improved moderately. The filters are global technique's so there will always be a difference between the sharpness of the focused area and the out-of-focus area. If we could regionalize these application of the convolution filters to preserve local sharpness like contrast limited AHE is regionalized to preserve local contrast, then we would possibly see better results.

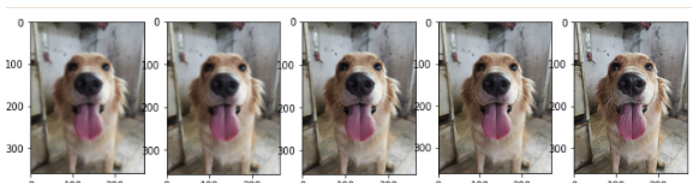

Fig. 5. This image was provided from Manansala's tutorial [5]. I have applied a box blur filter, then an unsharp mask filter, then three sharpening filters. These are the results from the consecutive application of the filters. First, the box blur; second, the unsharp mask on the box blur image, and so on.

C. Histogram Equalization

After a summer of research in image processing, I realized that the histograms of the images had a lot to do with its luminance. With this in mind, I found out about histogram equalization. Because this method increased the contrast of an image, it also increased the edges in an image, so I thought I could combine this method with the convolution filter. In Fig. 7, one can see that the contrast was enhanced however the color of the image has changed. If we could adjust the CLAHE method to add contrast to detected edges and less of the actual object then maybe the colors of the image would not change as much.

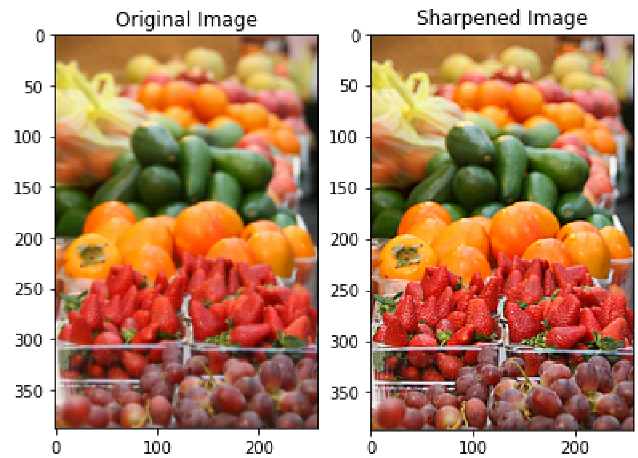

Fig. 6. This image shows a side by side comparison of the provided image and the image after applying the convolution sharpen filter.

Fig. 7. This image shows the provided image after CLAHE and the convolution sharpen filter.

IV. Conclusion

A method that can create an in-focus image from a single input image is not an easy method to come across. After extensive research, I have found several potential ways to carry out this task. There are deep neural networks that can be trained to generate images. There are convolution kernels that can sharpen images. Histogram equalization can increase the contrast of an image by stretching the histogram to contain all intensity values. Next, we would like to create a dataset with out of focus and corresponding in focus images to train a GAN to generate in focus images without any other information. We would also like to test the use of the convolution sharpen filter on regions of the image.

ACKNOWLEDGMENT

I would like to acknowledge and thank Annie Ji, a PhD student at Georgia State University, for helping me edit and revise this report.

This research was fully given assistance to realize this project by the Department of Creative Media Industries Institute at Georgia State University. I thankfully acknowledge the National Science Foundation and Georgia State University for funding this research and providing and giving access to the needed resources.

REFERENCES

[1] Sakurikar, P., Mehta, I., Balasubramanian, V. N., and Narayanan, P. J. "Refocusgan: Scene refocusing using a single image." Proceedings of the European Conference on Computer Vision (ECCV). (pp. 497-512). 2018

[2] Johnson, J. "What's a Deep Neural Network? Deep Nets Explained." BMCBlogs. 2020. https://www.bmc.com/blogs/deep-neural-network/

[3] Tung, S., and Hwang, W. "Multiple depth layers and all-in-focus image generations by blurring and deblurring operations." Pattern Recognition. 69: 184-198. 2017.

[4] Ormesher, I. "Convolution Filters." Medium. 2020. https://medium.com/@ianormy/convolution-filters-4971820e851f

[5] Manansala, J. "Image Processing with Python: Image Effects using Convolutional Filters and Kernels." Medium. 2021. https://medium.com/swlh/image-processing-with-python-convolutional-filters-and-kernels-b9884d91a8fd

[6] Htoon, K. S. "A Tutorial to Histogram Equalization." Medium. 2020. https://medium.com/@kyawsawhtoon/a-tutorial-to-histogram-equalization-497600f270e2

[7] Sudhakar, S. "Histogram Equalization." Towards Data Science. 2017. https://towardsdatascience.com/histogram-equalization-5d1013626e64

[8] Xu, G., Liu C., and Ji, H. "Removing out-of-focus blur from a single image." arXiv preprint arXiv:1808.09166. 2018.

[9] "Adaptive Histogram Equalization." ImageMagick. https://imagemagick.org/script/clahe.php

[10] Liao, Y. "Practical electron microscopy and database." An Online Book. 2006. https://www.globalsino.com/EM/page1371.html

[11] Ndandala, S. "Contrast Limiting Adaptive Histogram Equalization." Programming Notes. 2018. https://programmingnotes.net/code-snippets/computer-vision/contrast-limiting-adaptive-histogram-equalization/

[12] Yadav, G., Maheshwari, Y., and Agarwal, A. "Contrast limited adaptive histogram equalization based enhancement for real time video system." 2014 international conference on advances in computing, communications and informatics (ICACCI). IEEE, 2014.

[13] Srinivasan, P. P., Wang, T., Sreelal, A., Ramamoorthi, R., and Ng, R. "Learning to synthesize a 4D RGBD light field from a single image." In Proceedings of the IEEE International Conference on Computer Vision. (pp. 2243-2251). 2017.

[14] Brownlee, J. "A Gentle Introduction to Generative Adversarial Networks (GANs)." Machine Learning Mystery. 2019. https://machinelearningmastery.com/what-are-generative-adversarial-networks-gans/

[15] Jauregui, A. F. "How to Code a Generative Adversarial Network (GAN) in Python." https://anderfernandez.com/en/blog/how-to-code-gan-in-python/

[16] Kupyn, O., Budzan, V., Mykhailych, M., Mishkin, D., and Matas, J. "Deblurgan: Blind motion deblurring using conditional adversarial networks." In Proceedings of the IEEE conference on computer vision and pattern recognition. (pp. 8183-8192). 2018.

[17] Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P. and Wang, Z. "Enlightengan: Deep light enhancement without paired supervision." IEEE Transactions on Image Processing, 30, pp.2340-2349. 2021.

[18] Kupyn, O., Martyniuk, T., Wu, J. and Wang, Z. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 8878-8887). 2019.