Automated Portrait Mode for Single Camera Images

Boris Sobolev
UC Berkeley
sobolev@berkeley.edu

Abstract

Phone cameras have small focal lengths and small apertures, which result in all objects in a picture being in focus. Some phones feature solutions to recreate a shallow depth-of-field effect in cameras while taking a picture known as a "Portrait Mode". We propose a solution for automated post-production software that transforms single-camera images taken without this mode into imitations of professionally taken portraits. We use semantic segmentation and depth estimation pre-trained models to separate a person from its' background. Along with applying an out-of-focus effect, our software performs color correction, white balance correction, automated cropping, and color intensity adjustments to enhance the visual appeal of the picture. Furthermore, unlike existing solutions, it features custom aperture shapes to improve the quality of bokeh.

1. Introduction

Depth-of-field is an important aesthetic quality that can enrich visual appeal of the picture. It works especially well for portraits. Current solutions exist as online pipelines used in modern phones tuned to work with specific hardware. However, there are a lot of pictures taken before this technology became available. All these pictures could be automatically improved.

1.1. Related Work

[4] describes the process of creating portraits for single camera images. [5] explains how the Portrait Mode can be improved to produce SLR-like bokeh along with depth estimation improvements.

We used both results in our implementation to come up with our approach to solve this problem. Mostly high level ideas were analysed as our main goal was to test our approach on regular images.

Figure 1. Automatically Cropped Portrait with Background Objects removed

2. Portrait Effect

The main method used to produce the shallow depth-of-field effect is to find regions of the picture which are in focus, and the ones that are not.

There are different techniques to produce those regions and in this project we tried instance segmentation, depth field estimation and proposed how to use the combination of both.

2.1. Binary Background Blur

Given an original image and one region - in focus one, we cut out the object and create a background mask. We use log gamma correction to fix colour in both regions. Then we blur the image without the object as well as the object mask. We add blurred image to the cutout object and get picture with artifacts on the border. To remove those we add the multiplication of the blurred image and object mask to the

Figure 2. Semantic Segmentation Masks

result picture.

2.2. Gamma Correction

Apart from improving visual appeal, using gamma correction twice helps us better model the light. Optical cameras detects prioritize brighter regions over the darker ones due to its construction; however, when we manipulate pixels all the values are equal to convolutions. Gamma correction allows us to restore the more natural effect.

We use log gamma correction to improve color intensity of the image. We do it after we separate two regions and also after we get a final result.

2.3. Improving Bokeh quality

To model the aperture we blur the image by performing convolution of the Disk Kernel like in [5] and [4]. We also added two Heart Shaped kernels of set size to use on smaller images.

2.4. Gaussian Blur

Gaussian blur removes noise along with the very important information we can use to model light in the image. If we subtract from picture blurred using Gaussian kernel, picture that was blurred using aperture imitation convolution, we can see that Gaussian blur removed a lot of pixel which contained light.

2.5. Saturation

One of the artistic techniques implemented in our software is saturation adjustment for different regions. We convert image to HSV format and multiply S channel by a constant.

Grey-scale background effect is achieved by making saturation close to zero.

3. Instance Segmentation

We use notion of different objects in the picture for out computations. We used pre-trained models from [6] model zoo for all tasks which involved instance segmentation.

3.1. Person Mask

A person (or a group of people) play central role in the portrait. First step to synthesize a portrait is to separate the person from the background. We go through all instances classified as person in the picture and the most important ones based on the relative area of the mask. Post-processed picture can feature multiple people on for imitated SLR camera to focus.

3.2. Imitating Depth Field

All large objects that are close to the person might be in focus since those share the similar distance from the camera. Thus we can include the ones below the masks of the person (people) in focus region. Current implementation doesn't fill the gaps between close objects, but this can be done using simple dynamic programming algorithm in future.

3.3. Small Background Objects Removal

To improve the visual appeal of the portrait we can remove small objects from the background. This procedure reduces background noise, preserves color and flow of the image. Since the background is blurred imperfections of used algorithms do not produce disturbing artifacts.

One of the supported methods is Seam Carving [2]. We used it with standard energy function. Furthermore, we applied protective mask to important objects in the picture, while decreasing the we decided to delete as proposed in [2]. Since this method does not preserve size of the image, we have an alternative.

Alternative method preserves size. We used the approach described in [3] to fill in the gaps produced by deleting small objects from the picture. For example, this method successfully removes people from the background.

3.4. Automated Cropping

Automated Cropping uses semantic segmentation and key points estimation to crop the original picture to make

Figure 3. From Left to Right. Original Image; Generated Portrait Image With Depth Field Imitation Using Semantic Segmentation, Saturation Adjustments and Disk Blur; Generated Portrait Image Using Grey-Scale Background and Disk Blur.

Figure 4. Keypoints Used to Crop the Image

Figure 5. Original Image

it look more like a portrait. It places in focus objects either in center or more to the side by using thirds rules.

Cropping could be further improved by using dense pose estimation.

4. Single Picture Depth Estimation

Depth estimation could be use for a more precise synthesis of SLR-like pictures since objects could be divided in multiple groups based on the predicted distance from the camera. Different strength of blur could be applied to those groups to enhance the bokeh.

We used a pre-trained model by [1] to produce depth estimates; however, we went with instance segmentation as our main method due to the low resolution of the output images.

References

[1] I. Alhashim and P. Wonka. High quality monocular depth estimation via transfer learning. *arXiv e-prints*, abs/1812.11941, 2018.

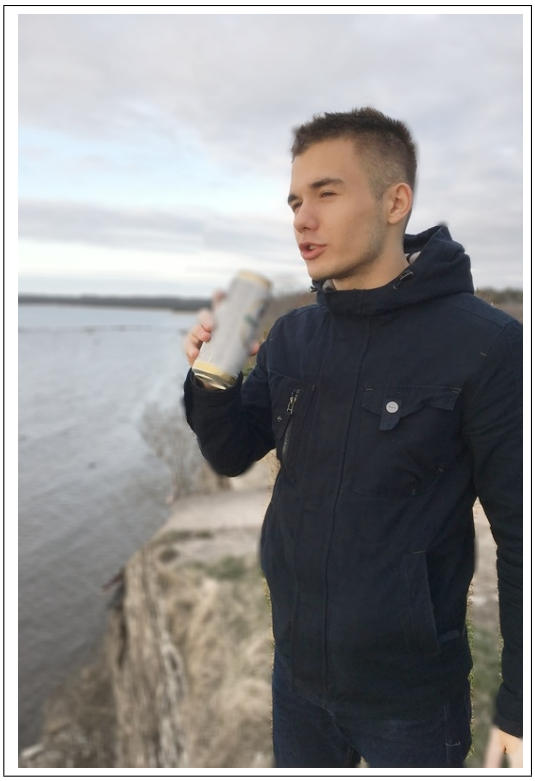


Figure 6. Portrait Effect with Saturation Decrease on the Background, Small Objects removed with Seam Carving Producing Visible Distortion

[2] S. Avidan and A. Shamir. Seam carving for content-aware image resizing. In *ACM SIGGRAPH 2007 papers*, pages 10–es. 2007.

[3] A. Telea. An image inpainting technique based on the fast marching method. *Journal of graphics tools*, 9(1):23–34, 2004.

[4] N. Wadhwa, R. Garg, D. E. Jacobs, B. E. Feldman, N. Kanazawa, R. Carroll, Y. Movshovitz-Attias, J. T. Barron, Y. Pritch, and M. Levoy. Synthetic depth-of-field with a single-camera mobile phone. *ACM Transactions on Graphics (TOG)*, 37(4):1–13, 2018.

[5] N. Wadhwa and Y. ZhangAggarwal. Improvements to portrait mode on the google pixel 4 and pixel 4 xl. `https://ai.googleblog.com/2019/12/improvements-to-portrait-mode-on-google.html`, 2019.

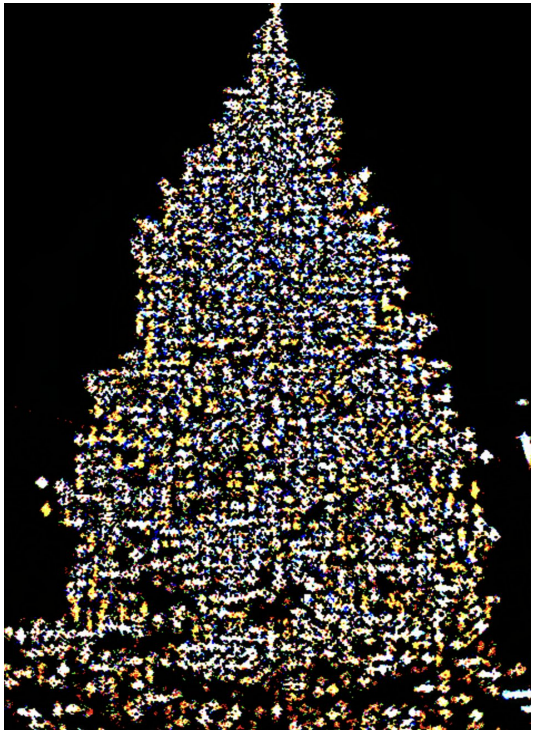
[6] Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick. Detectron2. `https://github.com/facebookresearch/detectron2`, 2019.

Magic Portraits

Boris Sobolev
CS194-26 Final Project

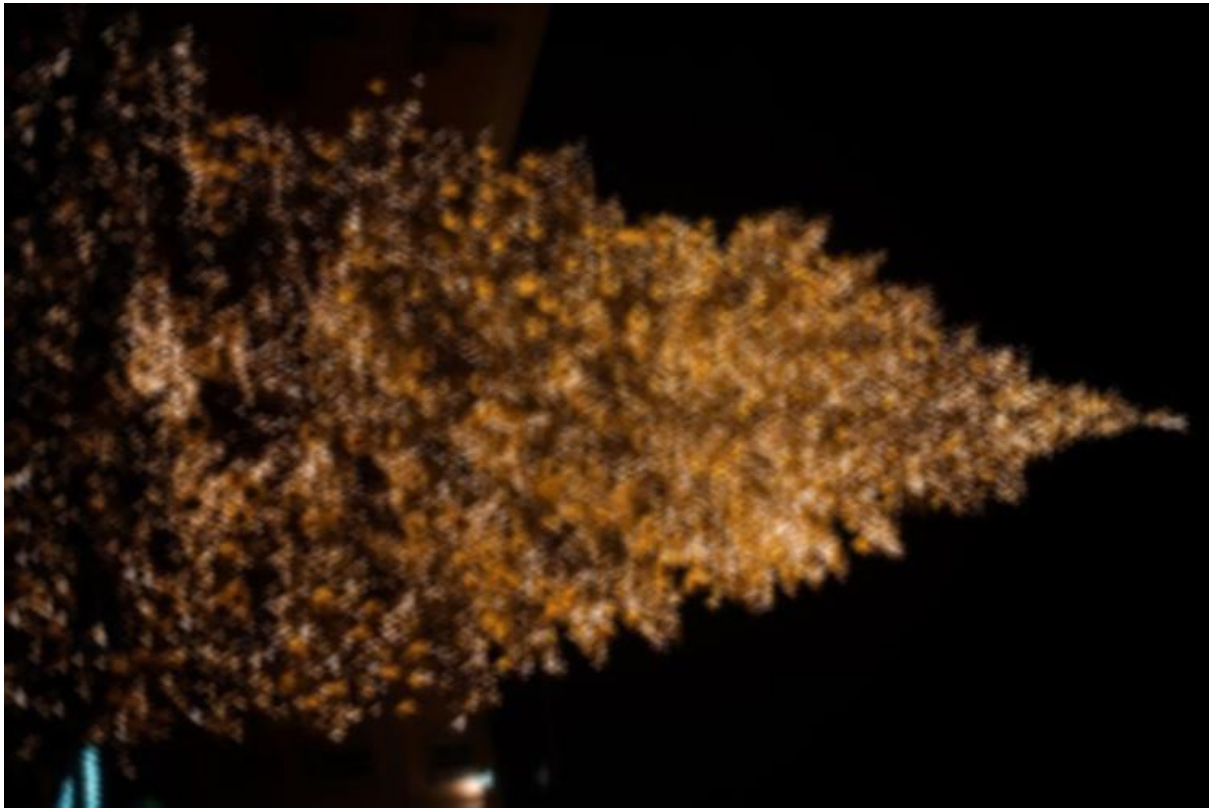
Turning single camera pictures into real camera portraits.

Why Gaussian Blur Doesn’t Work

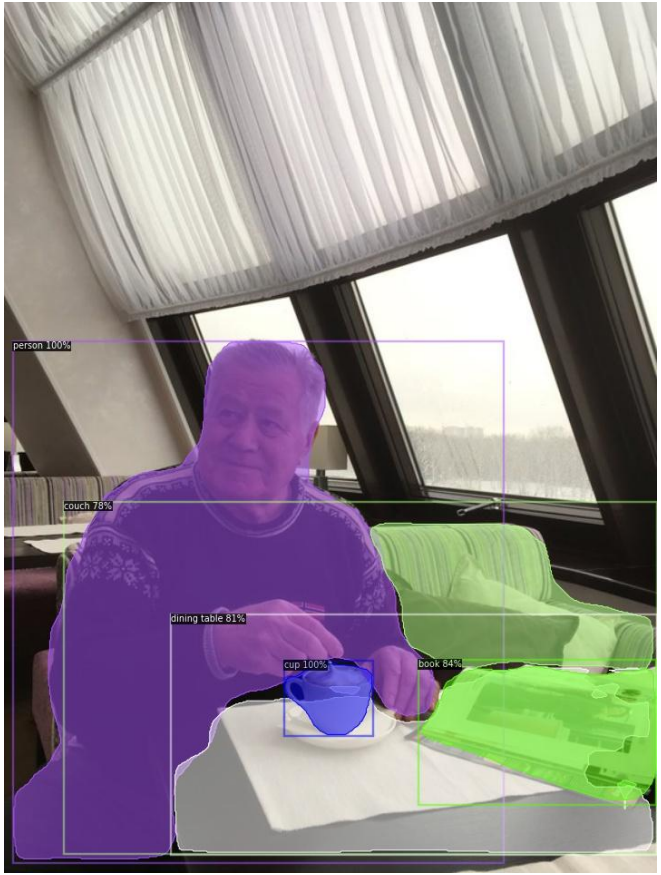


Custom Aperture

Disk, Heart-Shaped



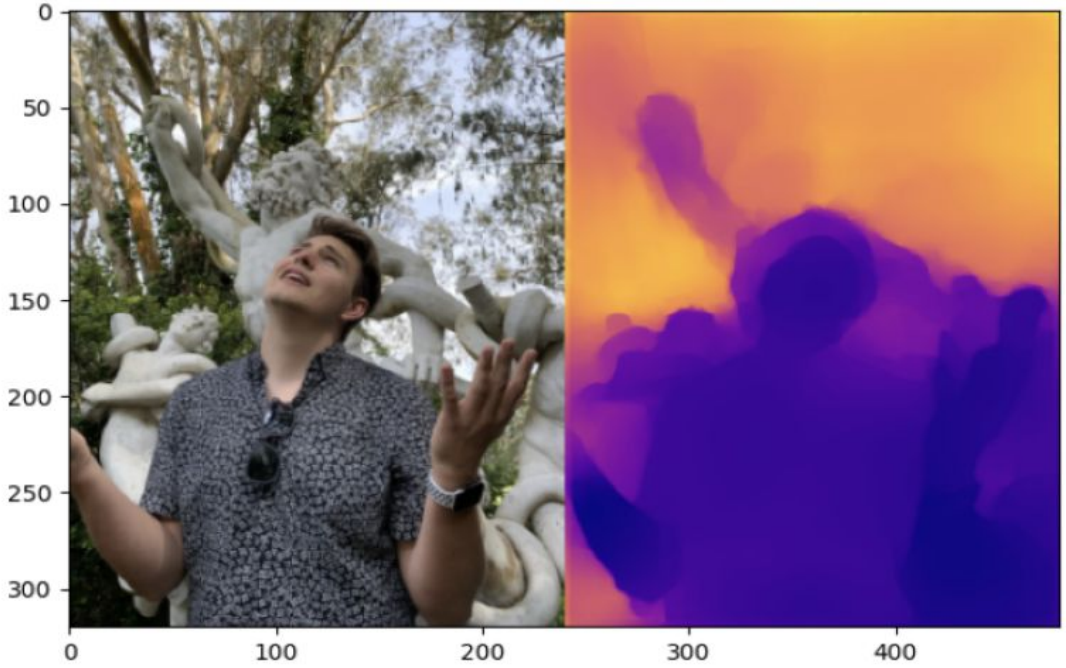
Depth imitation



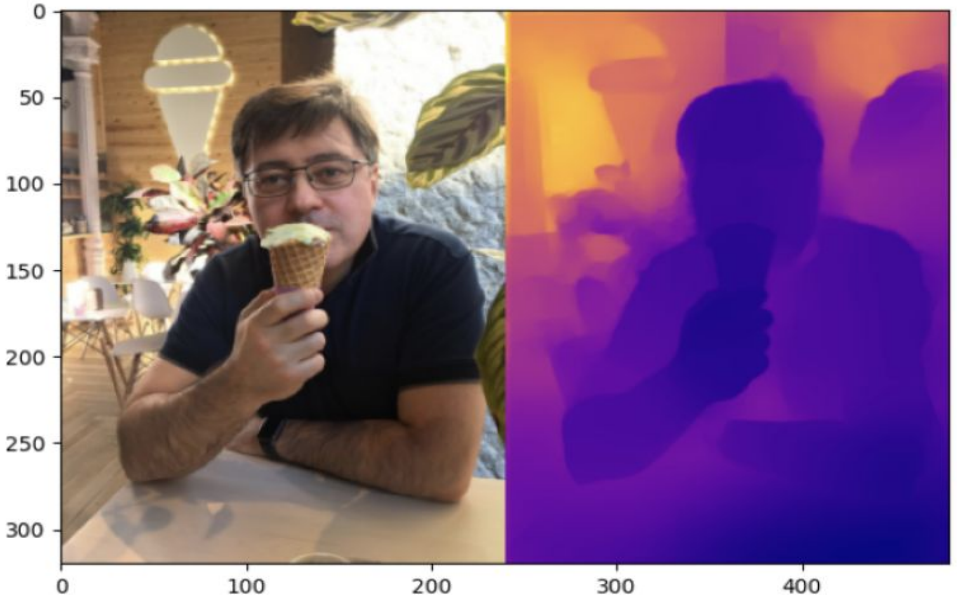
Cut small objects out for more uniform background

Auto Cropping

Gradual Blur



Improve Mask quality



Stuff I’m Working On

- Eye Level Tilt

Thanks