WINTER IN DATA SCIENCE
ANALYTICS CLUB IIT BOMBAY

Project Report

Student name: *Chaitanya Aggarwal - 210050038*

Course: *PyNET using Bokeh*
Due date: *January 2023*

Introduction

Bokeh effect is a very popular photography technique used to make the subject in the shot stand out sharply against a blurred background. DSLR cameras capture this affect naturally because they have wide aperture and shallow field of view. However, the smaller sensors on mobile camera are not capable to do so. We try to create a deep learning model that can be deployed on mobile architecture.

Approach

Bokeh effect simulation problem belongs to a group of tasks dealing with both global and local image processing. High-level image analysis is needed here to detect the areas on the photo where the bokeh effect should be applied, whereas low-level processing is used for rendering the actual shallow depth-of-field images and refining the results. Therefore, in this work we base our solution on the PyNET architecture designed specifically for this kind of tasks: it is processing the image at different scales and combining the learned global and local features together

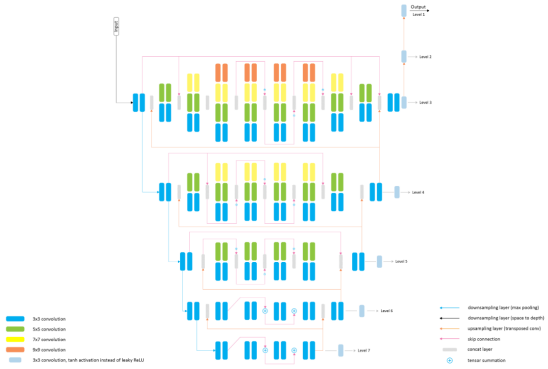
Implementation

While a large part of the implementation required for the project was part of the research paper itself, work needed to be done to train the model on our own machines. I used google colab for this purpose. To reduce the processing time to a practical amount, only 2000 images were randomly selected from the database of 4.5k+ images.

Learnings

Theoretical.

1. Understood the intuition behind inverted pyramid approach for designing the deep neural network. The lower layers are fed down sampled (low resolution) images from the higher layers. These layers are then trained on the data. The lower layers learn to map the most high level details of the picture.



2. The lower layers pass their output vectors onto the upper layers at different depths. The vector data from upper layer is concatenated/summed with the lower layer and passed onto the next CNN layer.

3. As we go higher and higher in the pyramid, the depth of each set of layer increases. The layers start capturing and learning the finer details of the image. This leads to a model that is highly accurate in identifying the subject from the background. When the model is trained on sufficient amount of data, it produces great results.

Practical / Implementation Based.

1. Got hands on experience on working on a large data-set of around 20Gb. Learned to resolve problems associated with such big sized data such as memory and GPU limitations. Resolved the issue by finally sampling from the data, instead of using the entire dataset for training.

2. Figured out ways of using Google Colan and Google Drive in a Local PC like environment, so that python scripts can be executed on Google's Cloud Hardware, while the user can still use local environment features (terminal like commands, straight from Jupyter Notebook in Google Colab)

3. Got better at writing and reading machine learning models using Keras and Tensorflow. Developed an understanding of how modular approach can be used to create the machine learning models for easier access in future, whether for retraining the model or using it on testset.