Project Practice for Deep Learning Project1: Evaluating Image Super-Resolution in the Context of Downstream Tasks

Prof. Yujiao Shi

October 8th 2024

Instructions:

1. Deadline: **2024-10-18 23:59:59**
2. You need to write LaTeX. Word is ok, but PDF! Give your solution in English.
3. Your report should be submitted in PDF format and packed with your code, and the naming format of the file is studentID-name-project1.zip
4. Please submit your report through email to *yangjw2023@shanghaitech.edu.cn* with the subject line "studentID-name-project1"

1 Introduction

This project extends the discussion on image super-resolution (SR) in the context of downstream tasks. In the following, we will explore and evaluate a classical SR work[1].

In modern digital camera pipelines, the transformation from a Bayer pattern to an RGB image involves intersting operations, as shown in Figure 1. Initially, light is captured as a Bayer matrix, a semi-colored image requiring interpolation, known as demosaicing, to reconstruct full color at each pixel. This process can be seen as applying convolution kernels to estimate missing colors. Subsequently, the image is adjusted for exposure through a scalar multiplication (gain adjustment) and corrected for accurate color representation using a Color Transformation Matrix (CTM), a 3x3 matrix altering color values linearly. The final step, gamma correction, applies a non-linear transformation ($I_{out} = I_{in}^{\gamma}$) to adjust luminance, producing the standard RGB output. This streamlined description highlights the mathematical foundation underpinning the conversion from raw sensor data to a display-ready image.

For our exploration today, the work we're looking into, Deep Burst SR, takes a Bayer image as input and accomplishes the super-resolution tasks mentioned in class. This work has made its code [2] available and provides a rich dataset for training and

[1]Deep Burst Super-Resolution [Bhat et al., 2021]
[2]https://github.com/goutamgmb/deep-burst-sr

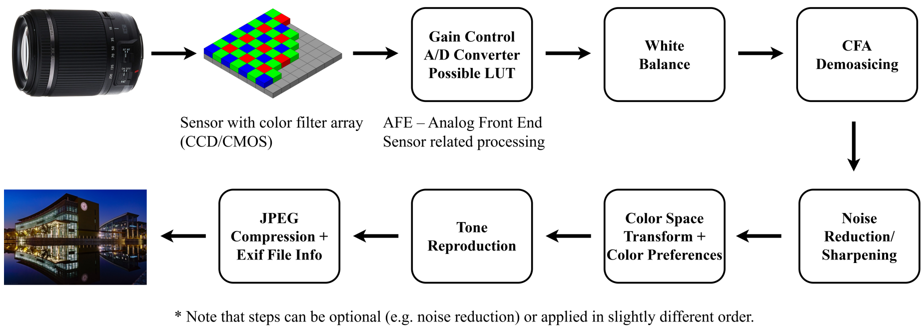

Figure 1: Modern digital camera pipelines

testing. As the first project of this course, we do not wish to expend a significant amount of computational power on training. Instead, we aim to enhance our ability to quickly become proficient with an open-source deep learning framework. You are required to run the open-source code from this work, record your own dataset, perform super-resolution tasks, and then carry out downstream tasks using the original and super-resolved images for evaluation and comparison.

2 Project Requirements

1 **Testing on Real Data(Basics):**

- **Take an HR image by your own phone:**
  - create LR burst images similar to the synthetic data shown in the paper;
  - super-resolve the burst images;
  - change the Bayer pattern of LR images and super-resolve them again;
  - compare the similarity between the predicted HR and GT HR images;

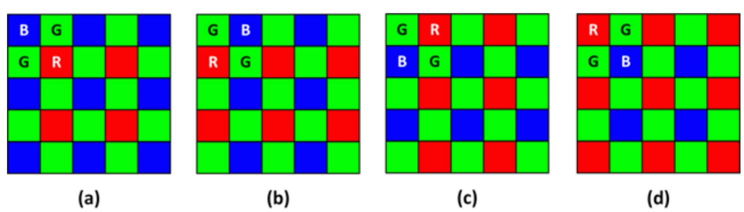

Figure 2: Different bayer patterns.

Capture and evaluate 20 images (repeat the above steps), including both indoor and outdoor scenes; compare and analyze the performance both quantitatively and qualitatively.

2 **Application to Downstream Tasks(Basics):**

* **Choose a downstream task (semantic segmentation or object detection):**
    - apply your LR & HR images to this downstream task;
    - Record and compare the performance difference.

3 **Collect training data by your own phones/cameras(Advanced):**

- Fine-tune the trained model;
- Record and compare the performance difference qualitatively and quantitatively.

4 **Others(Advanced):**

3 Submission Files

Please include your report (studentID-name-project1.pdf) and the codes folder in the package (studentID-name-project1.zip) attached to your submission email. PDF limited to 20M, PDF + zip limited to 50M.

In the codes folder, you also need to include a file named README.txt/md to describe the function of each file. Ensure that your code is runnable and consistent with your project.

The captured dataset should just be represented in the REPORT, it doesn't need to be submitted in the zip.

References

[Bhat et al., 2021] Bhat, G., Danelljan, M., Van Gool, L., and Timofte, R. (2021). Deep burst super-resolution. In *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 1