Visual Analytics for Super Resolution

Giovanni Ficarra

1 Introduction

During my PhD, I'm dealing with image super resolution. In this field, the most used metrics to evaluate models' performances are SSIM (Structural SIMilarity) and PSNR (Peak Signal to Noise Ratio), but there isn't a common agreement about their reliability. Thus, it is hard to establish which are the best models.

In the supplementary material [2] of [3] by Galasso et al. (2013), the authors rely on **scatter plots** to detect strange patterns in the results of neural networks for video segmentation: in each plot, a point is an image described by a pair of metrics (e.g. boundary precision-recall VS volume precision-recall). In this way, all the pairs of metrics are analyzed to discover when they agree and when they disagree, to find out if they are considering different but useful points of view or just making some errors. In particular, if a set of images are aligned along one direction, but distant along the other, we can inspect those image to analyze the different behavior of the two metrics, and find a reason for their inconsistency.

2 Dataset and Libraries

The 960 original images used for this projects were downloaded from Pexels[1], Nautilus Live[2] and [1][3]. Super-resoluted versions of such images were obtained with 19 experiments based on BSRGAN[4] [4] and evaluated with SSIM, PSNR_Y, PSNR_RGB and LPIPS [5]. More experiments and metrics may be added later, so the index for this project is $AS \geq 960 \times 19 \times 4 = 72'960$.

The preprocessing was performed using Python libraries such as PyTorch, Pandas and Numpy, while the visualizations will be realized using Plotly[5] and its Dash[6].

3 Visualizations and Analytics

Since our interest is in evaluating both metrics *and* SR models, other visualizations with respect to the scatter plots proposed in [3] can be useful too:

- **Parallel coordinates** with the average scores of each models, to show the performances of the networks according to the various metrics and compare them;
- A **box plot** with the average scores obtained by each image, to analyze differences among the images which are consistent among the models;
- Another **scatter plot** with two coordinates computed via **PCA**, to explore the possibility of finding new representative metrics, from the combination of the old ones;
- Pairs of original and super-resoluted **images**, shown when a point on a scatter plot is selected, to link the behavior of a model on that image with the score it obtained.

Brushing in the parallel coordinates allows to restrict the number of considered models, while in the scatter plots or the box plot it filters the set of images.

References

[1] Diva J. Amon, Brian R. C. Kennedy, Kasey Cantwell, Kelley Suhre, Deborah Glickson, Timothy M. Shank, and Randi D. Rotjan. Deep-sea debris in the central and western pacific ocean. *Frontiers in Marine Science*, 7, 2020.

[2] Fabio Galasso, Naveen Shankar Nagaraja, Tatiana Jiménez Cárdenas, Thomas Brox, and Bernt Schiele. A unified video segmentation benchmark: Annotation, metrics and analysis supplementary material.

[3] Fabio Galasso, Naveen Shankar Nagaraja, Tatiana Jiménez Cárdenas, Thomas Brox, and Bernt Schiele. A unified video segmentation benchmark: Annotation, metrics and analysis. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 3527–3534, 2013.

[4] Kai Zhang, Jingyun Liang, Luc Van Gool, and Radu Timofte. Designing a practical degradation model for deep blind image super-resolution, 2021.

[5] Richard Zhang, Phillip Isola, Alexei A. Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric, 2018.

[1] `https://www.pexels.com/it-it/`
[2] `https://nautiluslive.org/`
[3] `https://www.ncei.noaa.gov/access/ocean-exploration/video/`
[4] `https://github.com/cszn/BSRGAN`
[5] `https://plotly.com`, `https://plotly.com/python/`
[6] `https://plotly.com/dash/`, `https://dash.plotly.com/`