Visual Analytics for Underwater Super Resolution

Giovanni Ficarra

1   Introduction

During my PhD, I am dealing with underwater image super resolution. In this field, the most used metrics to evaluate the performances of the proposed neural networks are SSIM (Structural SIMilarity) and PSNR (Peak Signal to Noise Ratio), but there is not a common agreement about their reliability. In fact, often a visual inspection shows that images with higher scores do not appear better than images with lower scores. Thus, it is hard to establish which are the best models.

It is fundamental to find out which metrics are the most meaningful to proceed in the research of better SR models. To do this, it may be interesting to plot *all* the results of our experiments and visually analyze the differences and similarities among them with respect to the scores they obtains according to the considered metrics. Moreover, it may be useful to check what images with analogous results across different experiments have and do not have in common, for example to build a more significant benchmark dataset.

2   Related Works

In the supplementary material [1] of [2] 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 subset of images is aligned along one direction, but distant along the other, we can inspect those images to analyze the different behaviors of the two metrics and find a reason for their inconsistency, or discard one of the metrics.

Our intent is to extend this approach with an interactive application which allows to select images from the scatter plots to easily compare them and exploit also other visualizations, such as a parallel coordinates plot and a box plot, to combine different insights.

3   Dataset and Libraries

For this project, I used a dataset composed of frames extracted from videos downloaded from Pexels[1], Nautilus Live[2], and the website of the Ocean Exploration Video Portal of the National Centers for Environmental Information of the USA [3][3].

The resulting dataset was splitted in three parts, for training, validation and testing. Here we will use the test set, composed of 960 images, together with their super-resolved counterparts, obtained through about fifty experiments, where we tried different compression methods and training datasets, and many model parameters.

---

[1] https://www.pexels.com/it-it/
[2] https://nautiluslive.org/
[3] https://www.ncei.noaa.gov/access/ocean-exploration/video/

The dataset used for this project also includes, for each super-resolved image, its evaluation with SSIM, PSNR_Y, PSNR_RGB and LPIPS [5], already computed and stored in CSV files. Thus, the total amount of data is

$$960 \text{ images} \times 51 \text{ experiments} \times 4 \text{ metrics} = 195'840 \text{ values}.$$

To better visualize such amount of data, our application allows to see only information relative to a subset of the dataset, based on which compression method and training dataset were used.

This preprocessing was performed using Python libraries such as PyTorch, Pandas and Numpy, while the visualizations will be realized using Plotly[4] and its Dash[5].

4   Visualizations and Analytics

Since our interest is in evaluating both metrics *and* SR models, other visualizations with respect to the scatter plots proposed in [2] can be useful too:

- **Parallel coordinates** plot with the average scores of each model, to show the performances of the networks according to the various metrics and compare them;

- 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.

- A **box plot** with statistics of a selected subset of images, to analyze differences among the images across all the models.

4.1   The Parallel Coordinates Plot

.

Brushing in the parallel coordinates plot allows to restrict the number of considered models.

4.2   The Scatter Plots

.

4.3   The Box Plot

To enforce the idea that with the box plot we want to analyze the behavior of some images across all models, this diagram is kept separated from the others, so that filtering via buttons or brushing does not affect it.

---

[4]https://plotly.com, https://plotly.com/python/

[5]https://plotly.com/dash/, https://dash.plotly.com/

References

[1] Fabio Galasso et al. "A Unified Video Segmentation Benchmark: Annotation, Metrics and Analysis Supplementary Material". In: ().

[2] Fabio Galasso et al. "A unified video segmentation benchmark: Annotation, metrics and analysis". In: *Proceedings of the IEEE International Conference on Computer Vision*. 2013, pp. 3527–3534.

[3] Diva J. Amon et al. "Deep-Sea Debris in the Central and Western Pacific Ocean". In: *Frontiers in Marine Science* 7 (2020). ISSN: 2296-7745. DOI: 10.3389/fmars.2020.00369. URL: https://www.frontiersin.org/article/10.3389/fmars.2020.00369.

[4] Kai Zhang et al. *Designing a Practical Degradation Model for Deep Blind Image Super-Resolution*. 2021. DOI: 10.48550/ARXIV.2103.14006. URL: https://arxiv.org/abs/2103.14006.

[5] Richard Zhang et al. *The Unreasonable Effectiveness of Deep Features as a Perceptual Metric*. 2018. DOI: 10.48550/ARXIV.1801.03924. URL: https://arxiv.org/abs/1801.03924.