# Explainable Bilevel Optimization: an Application to the Helsinki Deblur Challenge

Silvia Bonettini, **Giorgia Franchini**, **Danilo Pezzi**, Marco Prato
Dipartimento di Scienze Fisiche, Informatiche e Matematiche, UNIMORE

## Introduction

The Helsinki Deblur Challenge 2021 (HDC) [1] had at its core a blind deconvolution problem. Given he observed data

$$f = u^{\text{true}} * h + \eta \in \mathbb{R}^n, \tag{1}$$

the goal is to recover a suitable approximation of $u^{\text{true}}$. Here $*$ denotes the convolution operator with the kernel $h$, which is also completely unknown, and $\eta$ is the noise, which is assumed to be additive white gaussian noise. The final quality of the image was evaluated through a Optical Character Recognition software (OCR).

## Method

We propose an approach within the algorithm unfolding framework [2, 3]. The fundamental idea can be summed up with the following problem (more details can be found in [4]):

$$\begin{cases} \underset{\theta\in\Theta}{\operatorname{argmin}} \ \mathcal{L}(u^*(\theta)) \\ \text{s. to } \ u^*(\theta) = \mathcal{A}_E^{(K)}(f;\theta) \end{cases} \tag{2}$$

In essence, we learn the parameters $\theta$ of the reconstruction algorithm $\mathcal{A}$ through the merit function $\mathcal{L}$ and a dataset $\mathcal{D} = \{(f_s; g_s) : s = 1, \dots, S\}$, where $f_s$ is a corrupted sample and $g_s$ is the corresponding ground truth.

**Lower Level Algorithm:**
We train the unrolling of a FISTA-like algorithm $\mathcal{A}_E^{(K)}(f;\theta)$, which starts from the data $f$ and runs for $K$ iterations, applied to the energy functional:

$$E_f(u;\theta) = \frac{1}{2}\|u * h(r) - f\|^2 + \gamma TV_\delta(u) + \lambda B(u) + \iota_{[0,1]^n}(u), \tag{3}$$

where
- The PSF is modelled as a disc convolution kernel of radius $r$, which is unknown.
- $B(u) = \sum_{p=1}^n u_p(1-u_p)$ is a concave bimodal function.

Overall, the parameters to learn are $\theta = (r, \gamma, \delta, \lambda, \alpha)$, which include the steplengths $\alpha = (\alpha_1, \dots, \alpha_K)$ of the inner steps of $\mathcal{A}$.

**Upper Level Penalty Function:**
We propose two different loss functions to train the algorithm $\mathcal{A}$.

**1) A supervised SSIM-based loss function:** it measures the similarity of patches between the reconstruction $u^*$ and the ground truth $g$:

$$\mathcal{L}(u^*(\theta)) = 1 - SSIM(u^*(\theta), g_s) \tag{4}$$

*Pros:*
- Performs well even with images that are $1/8$-th of their original size (the images provided by the HDC had $1460 \times 2360$ pixels!)
- Can handle images of different dimensions simultaneously.
- The loss itself (not $\mathcal{A}$ though) works with natural images as well.

*Cons:*
- Requires pre-processing of the "ground truths" provided.
- Is supervised and the ground truth may not be completely unavailable.

**2) An unsupervised OCR score predictor:**

$$\mathcal{L}(u^*(\theta,\alpha)) = \exp\left(-\frac{SVR(u^*(\theta,\alpha))}{100}\right), \tag{5}$$

where $SVR(u)$ is a performance predictor that is tasked with guessing the OCR score of $u$. The training of the SVR is *not* unsupervised.

*Pros:*
- Is unsupervised.
- Circumvents the issues with the ground truths provided.

*Cons:*
- The dimensions of the images needs to be fixed.
- Requires more images as the SVR needs be trained.
- Compared to the SSIM-based loss, it works best only with images $1/4$-th of their original size.

## Numerical Experience

**Experiment settings**

- Both penalty functions operated with only $4$ samples
- The extra dataset to train the SVR is built with three different kinds of images: 400 blurred samples provided directly from the HDC, 1230 synthetic images that emulates those of the HDC and 429 reconstructions obtained with random parameters configurations.
- The upper level optimization is solved with the Scaled Gradient Projection method.

| | $K=60$ | | $K=70$ | | $K=80$ | |
|---|---|---|---|---|---|---|
| | SSIM | SVR | SSIM | SVR | SSIM | SVR |
| Step 6 | 85.20 | 85.60 | 85.60 | 82.45 | 85.08 | 83.28 |
| Step 8 | 83.88 | 82.63 | 84.15 | 81.80 | 82.45 | 80.13 |
| Step 10 | 70.88 | 73.90 | 71.35 | 76.30 | 72.72 | 73.23 |
| Rec. Time | 0.37s | 1.15s | 0.44s | 1.29s | 0.50s | 1.57s |

Tab. 1: Average OCR scores obtained on the official 40 test image.


Fig. 1: from top to bottom: ground truth; corrupted sample; reconstruction (SVR merit function).

| | Ground truth | | Noisy sample | | Reconstruction | |
|---|---|---|---|---|---|---|
| | OCR | SVR | OCR | SVR | OCR | SVR |
| Times | 100 | 81.92 | 0 | 31.22 | 90 | 66.32 |
| Verdana | 100 | 81.96 | 0 | 30.67 | 100 | 68.44 |

Tab. 2: Comparison between the true OCR score and the prediction of the SVR for the above images.

## Future Work

- Application of the methodologes in different contexts.
- Extension of the unsupervised merit function to natural images: use of no-reference natural scene statistics quality measures
- Data augmentation: working with patches instead of full images.

## References

[1] M. Juvonen, S. Siltanen, and F. Silva de Moura. *Helsinki Deblur Challenge 2021 test dataset.* Zenodo, Nov. 2021. DOI: 10.5281/zenodo.5713637. URL: https://doi.org/10.5281/zenodo.5713637.

[2] S. Bonettini, G. Franchini, D. Pezzi, and M. Prato. “Learning the Image Prior by Unrolling an Optimization Method”. In: *2022 30th European Signal Processing Conference (EUSIPCO).* 2022, pp. 952–956. DOI: 10.23919/EUSIPCO55093.2022.9909852.

[3] J. Domke. “Generic Methods for Optimization-Based Modeling”. In: *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics.* Ed. by N. D. Lawrence and M. Girolami. Vol. 22. Proceedings of Machine Learning Research. PMLR, 2012, pp. 318–326.

[4] S. Bonettini, G. Franchini, D. Pezzi, and M. Prato. “Explainable Bilevel Optimization: An Application to the Helsinki Deblur Challenge”. In: *Inverse Problems and Imaging* 17.5 (2023), pp. 925–950. DOI: 10.3934/ipi.2022055.