

Fast and Accurate Image Upscaling with Super-Resolution Forests

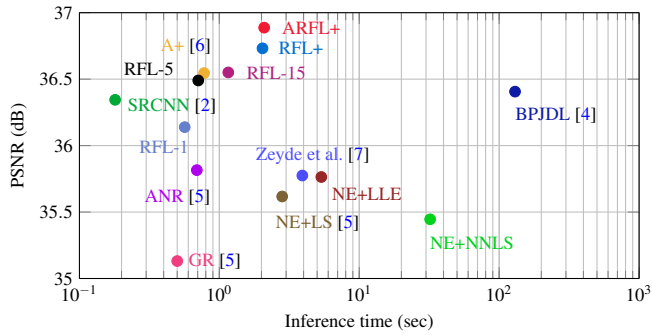
Samuel Schulter[†]Christian Leistner[‡]Horst Bischof[†][†]Institute for Computer Graphics and Vision, Graz University of Technology. [‡]Microsoft Photogrammetry, Austria.

Figure 1: Visualization of the trade-off between accuracy and inference time for different methods. RFL- $\{1, 5, 15\}$, RFL+, and ARFL+ are the proposed methods.

Single image super-resolution (SISR) [3] is a classical and important computer vision problem with many interesting applications, ranging from medical and astronomical imaging to law enforcement. The task in SISR is to generate a visually pleasing high-resolution output from a single low-resolution input image. Beside basic bicubic upsampling, more powerful methods exist that rely on statistical image priors [3] or use sophisticated machine learning techniques [2] to learn a mapping from low- to high-resolution patches. Among the best performing algorithms are (coupled) dictionary learning approaches [7] building on sparse-coding. Recently, Timofte *et al.* [5] highlighted the computational bottlenecks of these methods and proposed to replace the single dictionary with many smaller ones, thus avoiding the costly sparse-coding step during inference.

In this work, we observe that the efficient formulation from [5] can be naturally casted as a locally linear multivariate regression problem of the form

$$\hat{x}_H = W(x_L) \cdot x_L, \quad (1)$$

where \hat{x}_H and x_L are the high- and low-resolution data samples, respectively. Learning the data-dependent regression function $W(x_L)$ can be formulated as

$$\arg \min_{W(x_L)} \sum_{n=1}^N \|x_H^n - W(x_L^n) \cdot x_L^n\|_2^2, \quad (2)$$

which can be naturally solved with random forests [1]. Random forests establish the data dependence of the local regression functions by hierarchically splitting the input data in a tree structure. Each leaf node in all trees stores a linear regression function computed by solving a simple regularized least squares problem.

For learning the tree structure we introduce a novel objective function that not only operates on the output but also on the input data space. In general, the basic quality of a splitting function is computed as

$$Q(\sigma, \Theta, X_H, X_L) = \sum_{c \in \{Le, Ri\}} |x^c| \cdot E(x_H^c, x_L^c), \quad (3)$$

where $|\cdot|$ is the cardinality operator and $X_{\{H,L\}}^{\{Le,Ri\}}$ defines the low- and high-resolution data falling into left and right nodes, respectively. The function $E(X_H, X_L)$ aims at measuring the compactness or the purity of the data. The intuition is to have similar data samples falling into the same leaf nodes, thus, giving coherent predictions. The novel regularized compactness measure is defined over both the output and the input space as

$$E(X_H, X_L) = \frac{1}{|X|} \sum_{n=1}^{|X|} \left(\|x_H^n - m(x_L^n)\|_2^2 + \kappa \cdot \|x_L^n - \bar{x}_L\|_2^2 \right), \quad (4)$$



Figure 2: A qualitative result for an upscaling factor of 3. We compare the original image (top left) with 3 methods and report the PSNR and IFC scores: (top right) bicubic: 38.33, 4.61; (bottom left) BPJDL [4]: 40.79, 4.57; (bottom right) proposed random forest: 41.55, 5.77.

where $m(x_L^n)$ is the prediction for the sample x_L^n , \bar{x}_L is the mean over the samples x_L^n , and κ is a hyper-parameter steering the influence of the regularization.

The experimental evaluation reveals that our random forest approach achieves state-of-the-art results on several benchmarks, while being among the fastest methods during inference, see Figure 1. We also note that our super-resolution forests are relatively fast to train compared to sparse-coding or neural network based methods. In Figure 2 we present qualitative results for a single image. The paper (and also the poster) contains more detailed quantitative and qualitative evaluations.

- [1] Yali Amit and Donald Geman. Shape Quantization and Recognition with Randomized Trees. *NECO*, 9(7):1545–1588, 1997.
- [2] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Learning a deep convolutional network for image super-resolution. In *ECCV*, 2014.
- [3] William T. Freeman, Thouis R. Jones, and Egon C. Pasztor. Example-Based Super-Resolution. *CGA*, 22(2):56–65, 2002.
- [4] Li He, Hairong Qi, and Russell Zaretzki. Beta Process Joint Dictionary Learning for Coupled Feature Spaces with Application to Single Image Super-Resolution. In *CVPR*, 2013.
- [5] Radu Timofte, Vincent De Smet, and Luc Van Gool. Anchored Neighborhood Regression for Fast Example-Based Super-Resolution. In *ICCV*, 2013.
- [6] Radu Timofte, Vincent De Smet, and Luc Van Gool. A+: Adjusted Anchored Neighborhood Regression for Fast Super-Resolution. In *ACCV*, 2014.
- [7] Roman Zeyde, Michael Elad, and Matan Protter. On Single Image Scale-Up using Sparse-Representations. In *Curves and Surfaces*, 2010.