BIGPrior: Towards Decoupling Learned Prior Hallucination and Data Fidelity in Image Restoration

Majed El Helou Sabine Süsstrunk

School of Computer and Communication Sciences, EPFL, Switzerland

Abstract

Classical restoration algorithms use a variety of priors, implicitly or explicitly. Their priors are hand-designed and their weights are heuristically assigned. Thus, deep learning methods often produce superior restoration quality. Deep networks are, however, capable of strong and hardly-predictable hallucinations. Networks jointly and implicitly learn to be faithful to the observed data while learning an image prior, and the separation of original and hallucinated data is then not possible. This limits their wide-spread adoption in restoration applications. Furthermore, it is often the hallucinated part that is victim to degradation-model overfitting. We present an approach with decoupled network-prior hallucination and data fidelity. We refer to our framework as the Bayesian Integration of a Generative Prior (BIGPrior). It is rooted in a Bayesian restoration framework, and tightly connected to classical restoration methods. In fact, our approach can be viewed as a generalization of a large family of classical restoration algorithms.

Key Take-aways

Our paper presents a learning-based restoration framework that forms a generalization of various families of classical methods. It is both tightly connected with Bayesian estimation upon which it builds, and also to classical dictionary methods. Our BIGPrior makes the explicit integration of learned-network priors possible, notably a generative-network prior. Its biggest advantage is that, by decoupling data fidelity and prior hallucination, it structurally provides a per pixel fusion metric that determines the contribution of each. This can be important both for end users and for various downstream applications. We show results on **image colorization, blind image denoising, and image inpainting**. We hope this work will foster future learning methods with clearly decoupled network hallucinations, for the sake of interpretability, reliability, and to safeguard against the hazards of black-box restoration.

MAP and Classical Restoration

$$\arg\min_x -\log(P_{Y|X}(y|x)P_X(x))$$

$$\arg\min_x \underbrace{\psi_d(f'(x),y)}_{\text{data fidelity}} + \beta \cdot \underbrace{\psi_p(f''(x))}_{\text{prior}}$$

The priors are *hand-designed heuristics*; we want to exploit **learned priors**.

The weight is *only adaptive to the observed signal's* quality, not adaptive to the confidence in the prior's fitness; we want it to be **doubly-adaptive**.

The data fidelity and prior hallucination cannot be *explicitly distinguished*; we want to **decouple them**.

Our Proposed BIGPrior

$$\hat{x} = \underbrace{(1-\phi(y;\theta_1)) \odot g^{-1}(y)}_{\text{data fidelity}} + \phi(y;\theta_1) \odot \underbrace{G(z^*;\theta_2)}_{\text{prior}}$$

Generative network space projection:

$$z^* = \arg\min_z \mathcal{L}_G(f(G(z)),y)$$

Guide-free learning with a fidelity bias term:

$$\mathcal{L}(x,y;\theta_1,\theta_2) = ||\,(1-\phi(y,\theta_1)) \odot g^{-1}(y) + \phi(y,\theta_1) \odot G(z^*;\alpha^*,\theta_2) - x||_2^2 + \boxed{\rho \cdot ||\phi(y,\theta_1)||_1}$$

Relation to MAP, Classical Methods

$$\hat{x}_i = \arg\max_{x_i} P_{X_i|Y_i}(x_i|y_i) = \frac{y_i}{1+1/S_i} + \frac{\bar{x}_i}{1+S_i} \qquad \phi(y_i) = \frac{1}{1+S_i}$$

M. El Helou and S. Süsstrunk, "Blind Universal Bayesian Image Denoising with Gaussian Noise Level Learning," in IEEE TIP, 29:4885-4897, 2020.

$$\hat{x} = \arg\min_x \underbrace{||x-y||_2^2}_{} + \beta \cdot \underbrace{\sum_{j=1}^{J} |x \otimes f_j|^\gamma}_{}$$

D. Krishnan and R. Fergus, "Fast image deconvolution using hyper-Laplacian priors", in NeurIPS, pages 1033-1041, 2009.

$$\hat{x} = \arg\min_{x,d(x,Dv)<\epsilon} \psi_d(x,y) + \beta \cdot \psi_p(v) \qquad x \in \mathrm{span}(D)$$

\+ constraints

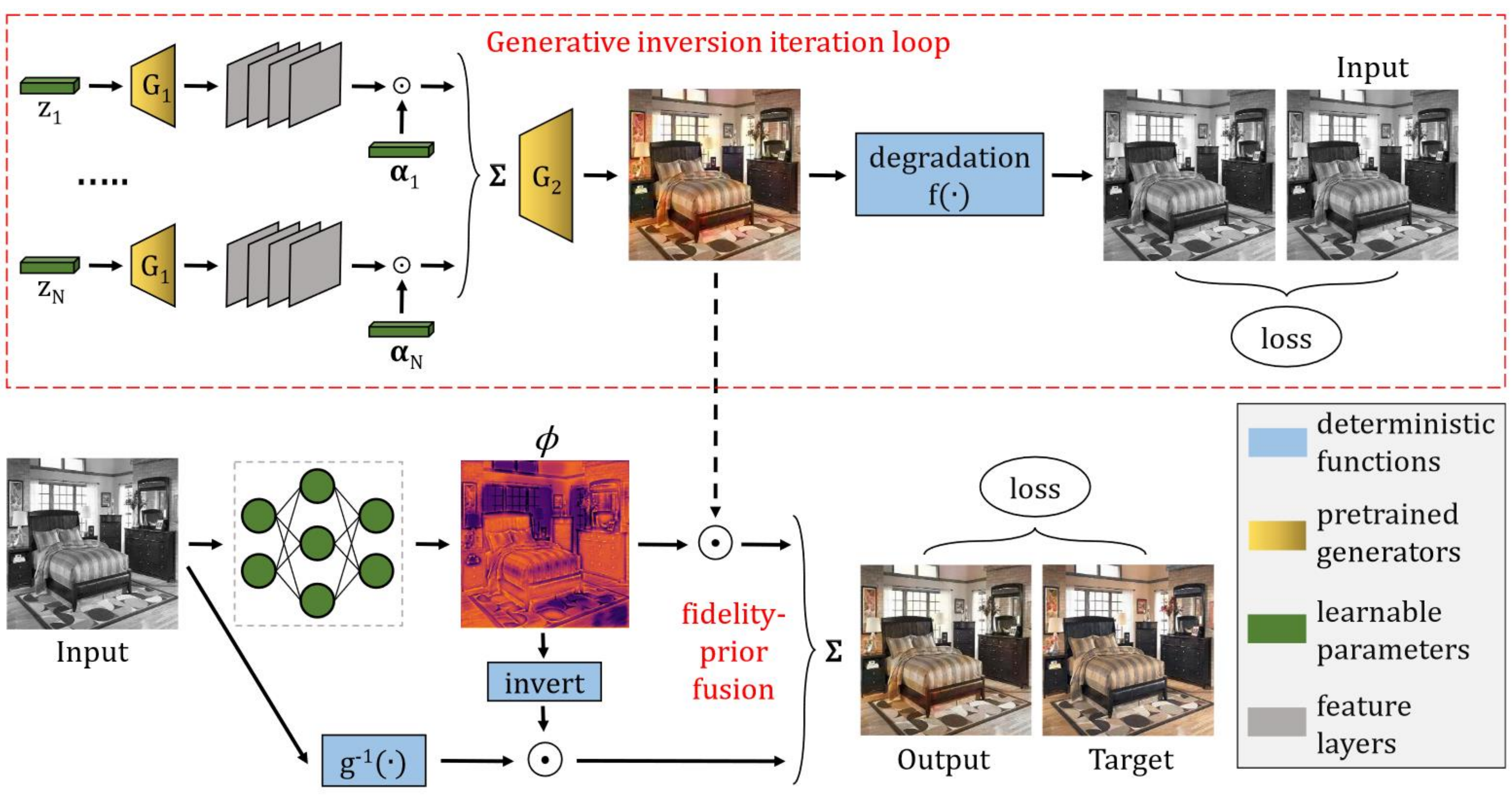
R. Giryes and M. Elad, "Sparsity-based Poisson denoising with dictionary learning", in IEEE TIP, 23(12):5057-5069, 2014.

Pipeline Overview

- The **prior** is extracted by projecting on the space of a pretrained generative network.
- The **data fidelity** is enforced by restricting to a bijective manipulation.
- The **fusion** between the decoupled terms is learned by a network.

The pixel-wise fusion map is readily available for downstream users or applications.

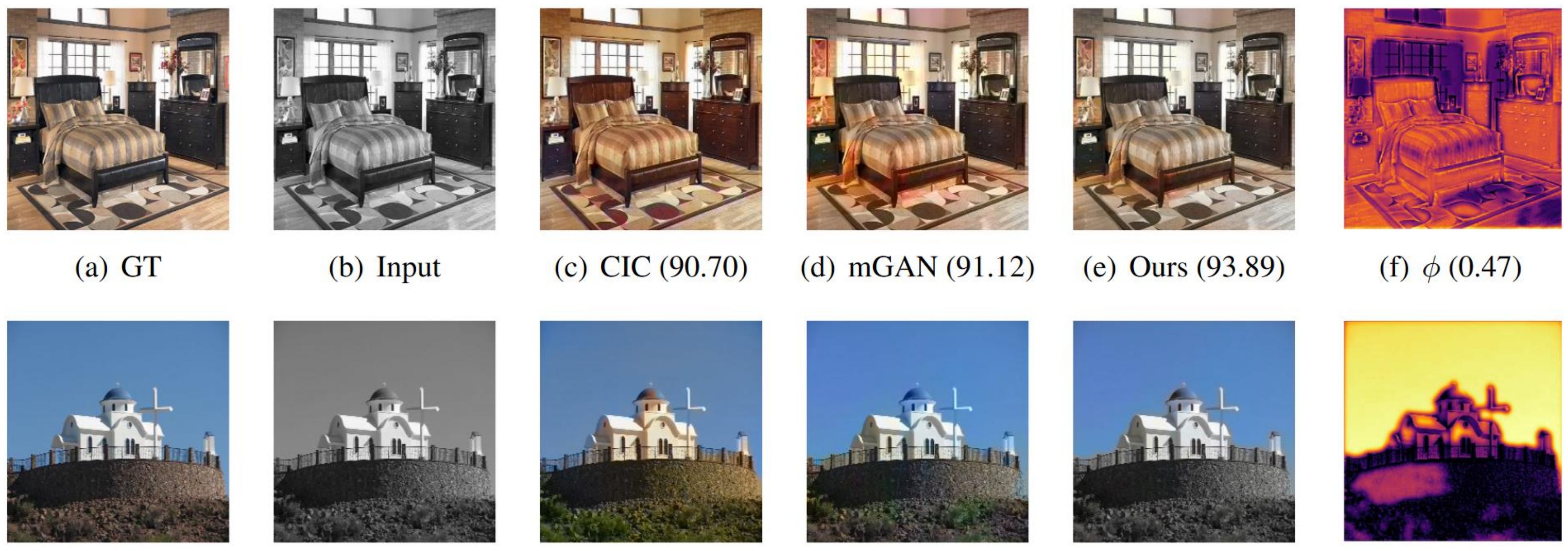
Future research can exploit this map to improve the robustness of downstream vision tasks, and for reliability and interpretability.



Sample Experimental Results

Image Colorization

Method	Bedroom set AuC [155] ↑	Church set AuC [155] ↑
Colorful colorization [155]	88.55	89.13
Deep image prior [131]	84.33	83.31
Feature map opt. [10]	85.41	86.10
mGAN prior [55]	88.52	89.69
Ours	**89.27**	**90.64**



(a) GT (b) Input (c) CIC (90.70) (d) mGAN (91.12) (e) Ours (93.89) (f) ϕ (0.47)

(g) GT (h) Input (i) CIC (94.17) (j) mGAN (93.87) (k) Ours (94.75) (l) ϕ (0.52)

Paper: https://arxiv.org/abs/2011.01406 **Code:** https://github.com/majedelhelou/BIGPrior