Scale-Equivariant Imaging: Self-Supervised Learning for Image Super-Resolution and Deblurring

Jérémy Scanvic [1,2] Mike Davies [2] Patrice Abry [1] Julián Tachella [1]

[1]Laboratoire de Physique, ENS de Lyon,
[2]School of Engineering, University of Edinburgh

May 13, 2025
GdR IASIS, INRIA Paris

The problem of image deblurring

Sharp images

Unavailable

$$y = Ax + \varepsilon$$

???

Blurry images

Available

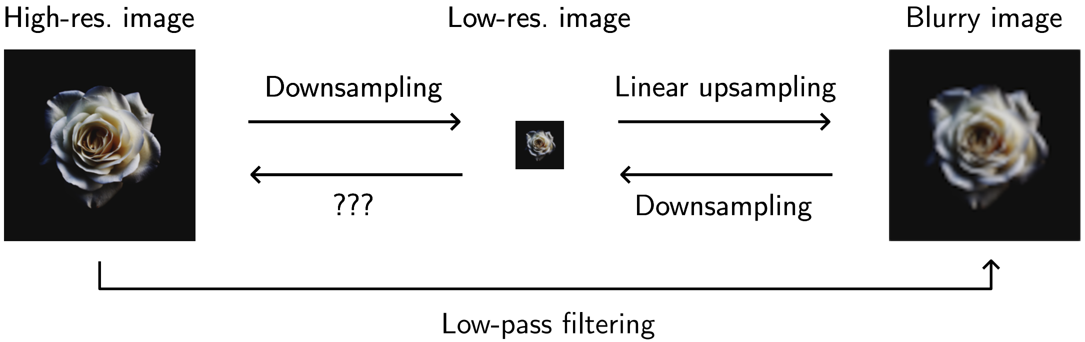
Blur filter

$$Ax = h * x \tag{1}$$

Additive Gaussian noise

$$\varepsilon \sim \mathcal{N}(0,\ \sigma^2 I) \tag{2}$$

Super-resolution as a deblurring problem



Ill-posedness of the problem

By the convolution theorem

$$\hat{y}(\xi) = \hat{h}(\xi)\hat{x}(\xi) + \hat{\varepsilon}(\xi) \tag{3}$$

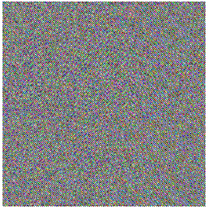
Inverse filtering

$$\frac{\hat{y}(\xi)}{\hat{h}(\xi)} = \hat{x}(\xi) + \frac{\hat{\varepsilon}(\xi)}{\hat{h}(\xi)} \tag{4}$$

Blurry image with noise

Inverse filter $\longrightarrow$

Extremely noisy reconstruction



Standard approaches

The model-based approach

$$\min_{x \in \mathbb{R}^n} \quad \frac{1}{2} \|Ax - y\|_2^2 + R(x) \tag{5}$$

- $+$ No ground truth image required
- $-$ Handcrafted for an image distribution

The supervised approach

$$\min_{\theta \in \mathbb{R}} \quad \mathbb{E}_{x,y} \|f_\theta(y) - x\|_2^2 \tag{6}$$

- $-$ Ground truth images required
- $+$ Data-driven instead of handcrafted

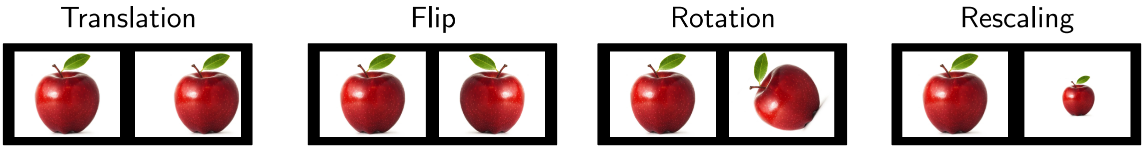
Invariance and image transformations

Insight. Natural image distributions are generally invariant to many transformations.

Definition. A set of images $\mathcal{X}$ is said to be **invariant** to a group of transformations if

$$x \in \mathcal{X} \implies T_g(x) \in \mathcal{X} \tag{7}$$

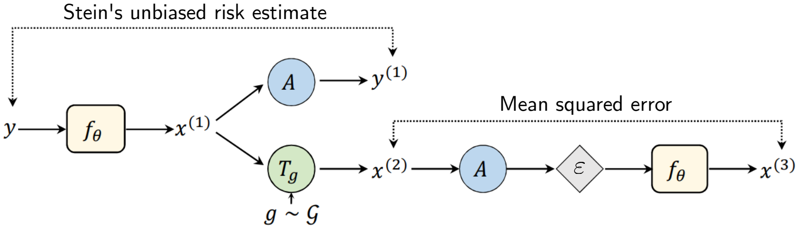
no matter the image x and the transformation T_g, e.g., a 90° rotation.



Equivariant imaging [Chen et al., 2021, 2022]

Insight. Image distribution invariances help regularize ill-posed inverse problems.

$$\min_{\theta} \quad \overbrace{\mathop{\mathbb{E}}_{x,y} \|Af_\theta(y) - Ax\|_2^2}^{\text{Measurement consistency}} + \overbrace{\mathop{\mathbb{E}}_{x,\varepsilon,g} \|f_\theta(AT_g x + \varepsilon) - T_g f_\theta(Ax + \varepsilon)\|_2^2}^{\text{Equivariance loss}} \tag{8}$$



Why scaling transforms for deblurring?

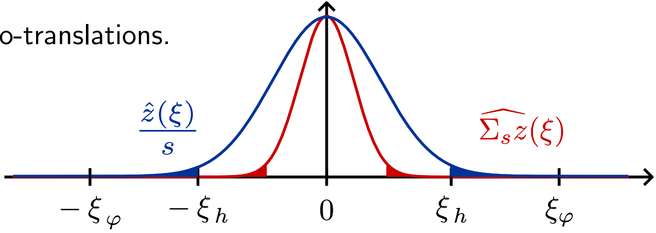
Theorem (Informal). [Scanvic et al., 2025]
A set of (sharp) images $\mathcal{Z} \subseteq \mathcal{S}$ is uniquely determined[1] by the blurry images

$$\mathcal{Y} = \{ \ h * z, \ z \in \mathcal{Z} \ \}, \tag{9}$$

obtained through a bandlimiting filter $h \in \mathcal{S}$,

- as long as $\mathcal{Z}$ is scale-invariant;
- but **not** if it is invariant to roto-translations.

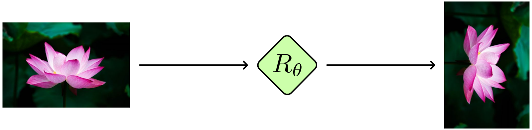
Transformations	PSNR
Scaling transformations	26.4
Translations & rotations	24.7



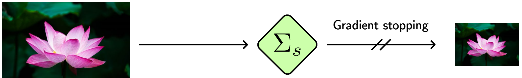
[1] up to an arbitrarily high frequency

Gradient stopping

Equivariant Imaging



Scale-Equivariant Imaging (Ours)

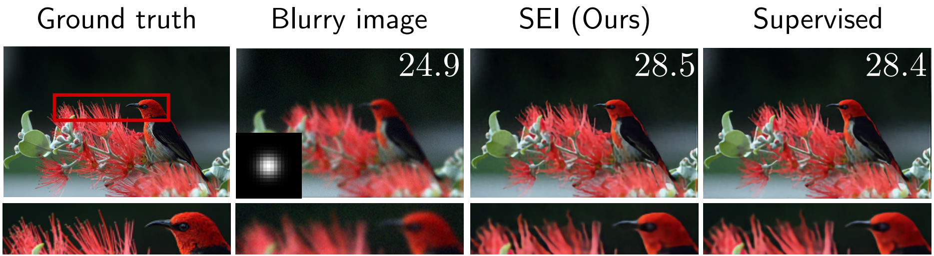


	Gaussian filter (medium)			Box filter (medium)			Downsampling ($\times 2$)		
GS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
✓	**25.9**	**0.795**	**0.254**	**27.0**	**0.818**	**0.178**	**28.1**	**0.850**	**0.197**
×	22.9	0.632	0.506	22.6	0.601	0.551	26.8	0.810	0.305

Impact of gradient stopping (GS) on test metrics

Gap with supervised training

Method	Gaussian filter (small)			Gaussian filter (medium)			Gaussian filter (large)			Box filter (small)			Box filter (medium)			Box filter (large)		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
Supervised	**30.7**	**0.923**	0.064	**25.9**	**0.803**	**0.249**	**23.7**	**0.694**	**0.361**	**29.1**	**0.879**	**0.107**	**27.3**	**0.825**	**0.171**	**25.4**	**0.755**	**0.255**
SEI (Ours)	30.3	0.917	**0.062**	25.9	0.795	0.254	23.7	0.687	0.379	28.8	0.874	0.112	27.0	0.818	0.178	25.0	0.739	0.309
Blurry images	26.4	0.794	0.269	22.8	0.597	0.461	21.2	0.494	0.604	23.4	0.629	0.359	21.9	0.528	0.464	21.0	0.465	0.552



Closing remarks

Preprint available on

arXiv

Code available on

GitHub

Powered by

Deep Inverse

References

Dongdong Chen, Julián Tachella, and Mike E. Davies. Equivariant Imaging: Learning Beyond the Range Space, August 2021.

Dongdong Chen, Julian Tachella, and Mike E. Davies. Robust Equivariant Imaging: A fully unsupervised framework for learning to image from noisy and partial measurements. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5637–5646, New Orleans, LA, USA, June 2022. IEEE. ISBN 978-1-6654-6946-3. doi: 10.1109/CVPR52688.2022.00556.

Jérémy Scanvic, Mike Davies, Patrice Abry, and Julián Tachella. Scale-Equivariant Imaging: Self-Supervised Learning for Image Super-Resolution and Deblurring, March 2025.