



IMPERIAL



Invertible Neural Networks and their Applications

Jun-Jie Huang and Pier Luigi Dragotti

Outline

1. Overview of Invertible Neural Networks

- Origin of INN and Normalizing flows
- INN for Inverse Problems

2. Wavelet-Inspired Invertible Neural Network

3. INN and diffusion models: INDigo

4. Other applications of INN

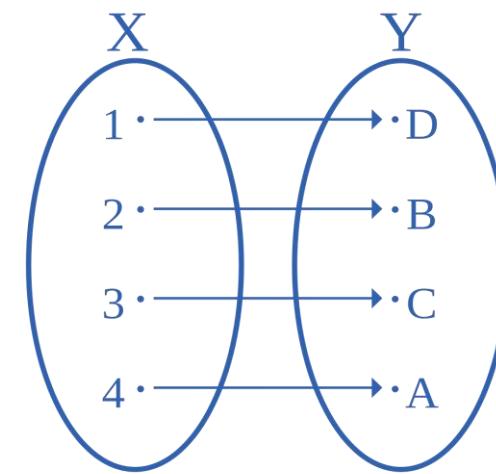
1. Overview of Invertible Neural Networks

Invertible Neural Networks (INNs) are bijective function approximators which have a forward mapping

$$F_{\theta}: \mathbb{R}^d \rightarrow \mathbb{R}^l$$
$$x \mapsto z$$

and inverse mapping

$$F_{\theta}^{-1}: \mathbb{R}^l \rightarrow \mathbb{R}^d$$
$$z \mapsto x$$



A bijective function (or
invertible function)

1. Overview of Invertible Neural Networks

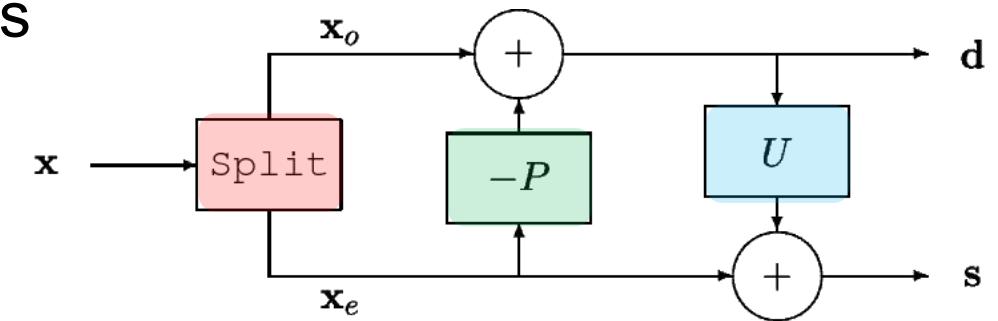
How to Achieve Invertibility?

Invertible via lifting scheme like architectures

- Signal splitting
- Alternative prediction and update

$$\text{Split} \rightarrow \begin{cases} d = x_o - P(x_e) \\ s = x_e + U(d) \end{cases}$$

Forward pass



$$\begin{cases} x_o = d + P(x_e) \\ x_e = s - U(d) \end{cases} \rightarrow \text{Merge}$$

Backward pass

Factoring wavelet transforms into lifting steps

I Daubechies, W Sweldens

Journal of Fourier analysis and applications 4 (3), 245-267

3900

1998

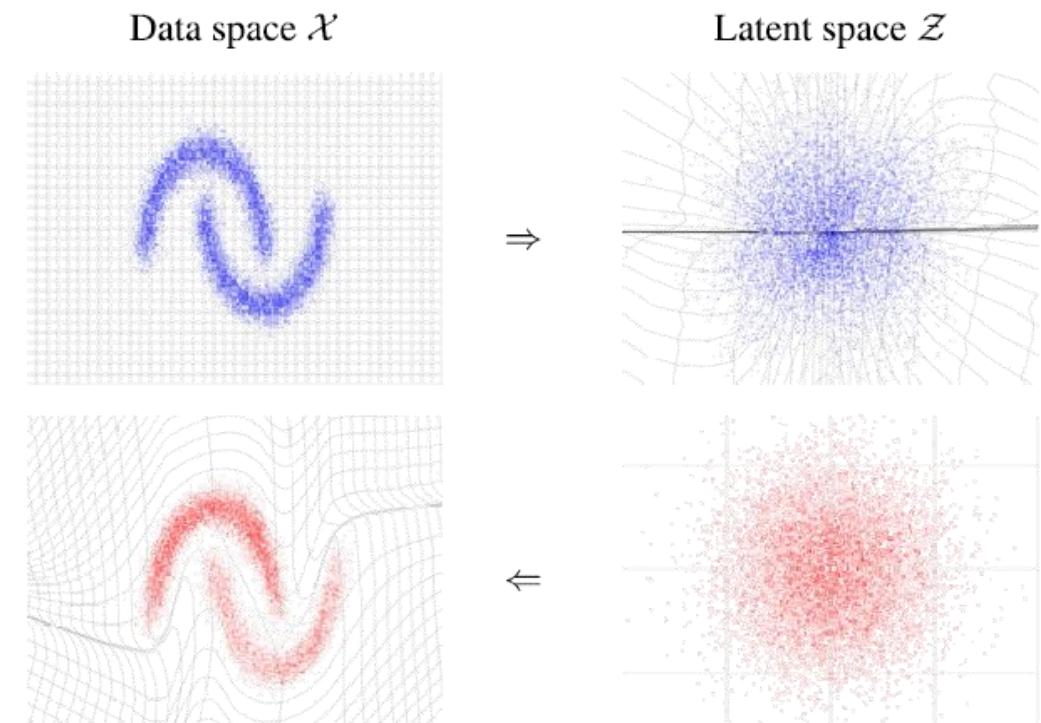
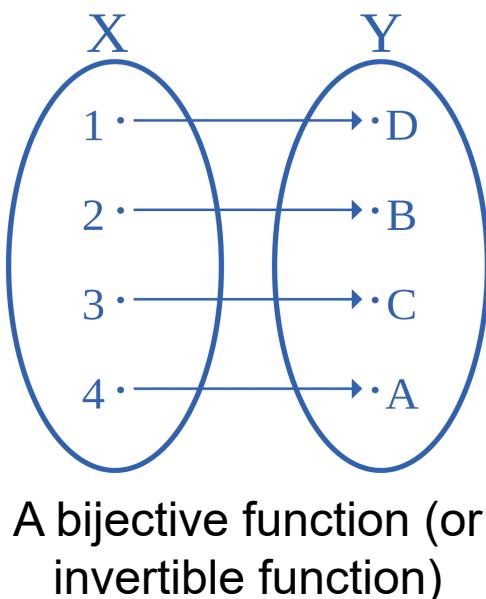
1. Overview of Invertible Neural Networks

Invertible Neural Networks are bijective function approximators
which have a forward mapping

$$F_{\theta}: \mathbb{R}^d \rightarrow \mathbb{R}^l$$
$$x \mapsto z$$

and inverse mapping

$$F_{\theta}^{-1}: \mathbb{R}^l \rightarrow \mathbb{R}^d$$
$$z \mapsto x$$



1. Overview of Invertible Neural Networks

Also known as **Normalizing Flow** for generative modeling

- Tractable Jacobian, allows explicit computation of posterior probability

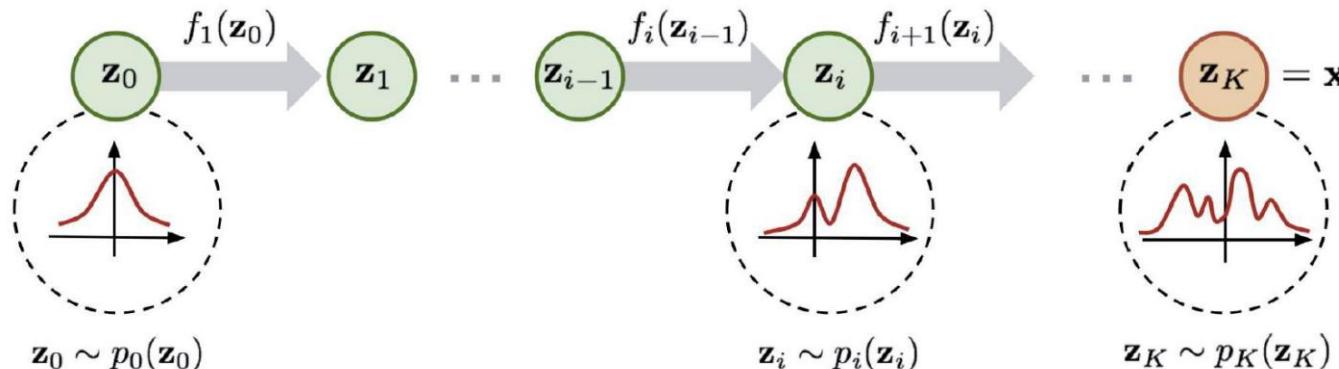


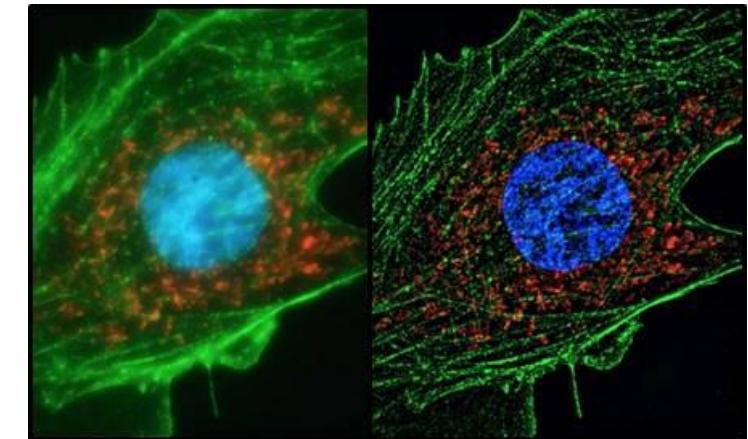
Figure 1: Synthetic celebrities sampled from our model; see Section 3 for architecture and method, and Section 5 for more results.

Kingma, Durk P., and Prafulla Dhariwal. "Glow: Generative flow with invertible 1x1 convolutions." in Proceedings of *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

1. Overview of Invertible Neural Networks

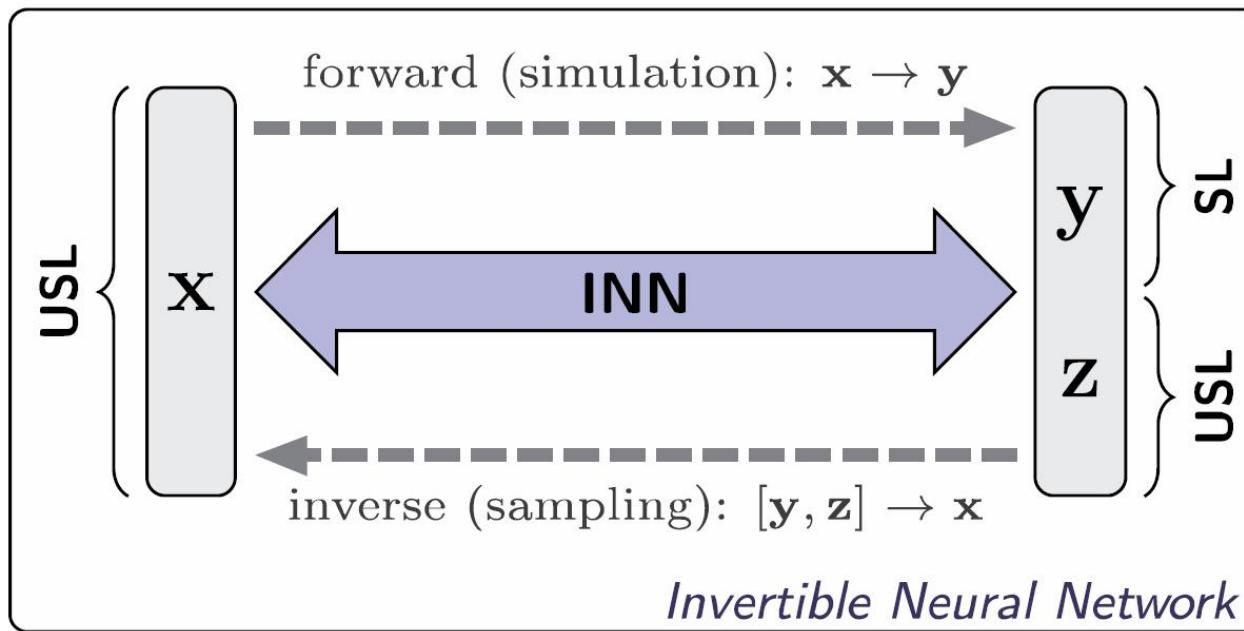
Inverse problems involve reconstructing unknown physical quantities from indirect measurements :

- denoising
- super-resolution
- deblurring
- inpainting
- ...



1. Overview of Invertible Neural Networks

Invertible Neural Networks are ideal architectures to address inverse problems



Ardizzone, Lynton, Jakob Kruse, Sebastian Wirkert, Daniel Rahner, Eric W. Pellegrini, Ralf S. Klessen, Lena Maier-Hein, Carsten Rother, and Ullrich Köthe. "Analyzing inverse problems with invertible neural networks." in Proceedings of International Conference on Learning Representations (ICLR), 2019.

Outline

1. Overview of Invertible Neural Networks

- Origin of INN and Normalizing flows
- INN for Inverse Problems

2. Wavelet-Inspired Invertible Neural Network

3. INN and diffusion models: INDigo

4. Other applications of INN

2. Wavelet-inspired Invertible Neural Network

Image Denoising

- Recover a clean image from noisy observations
- Raw image data is usually noisy

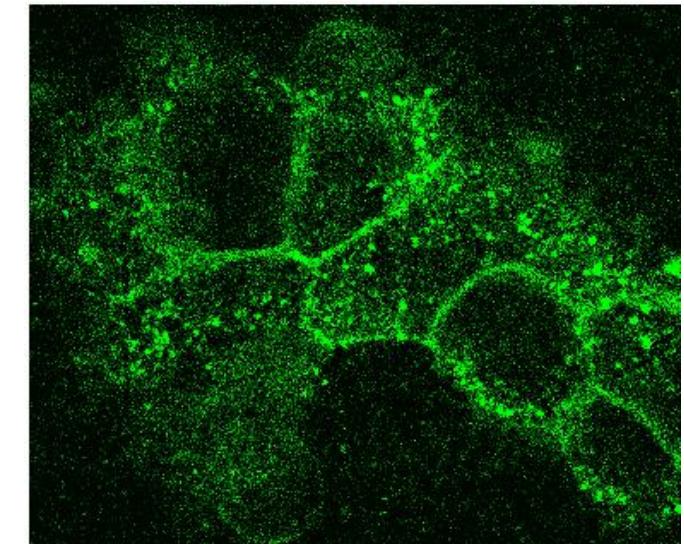


Denoising is the “simplest” inverse problem yet plays an important role in many applications

$$\mathbf{y} = \mathbf{x} + \mathbf{e}$$

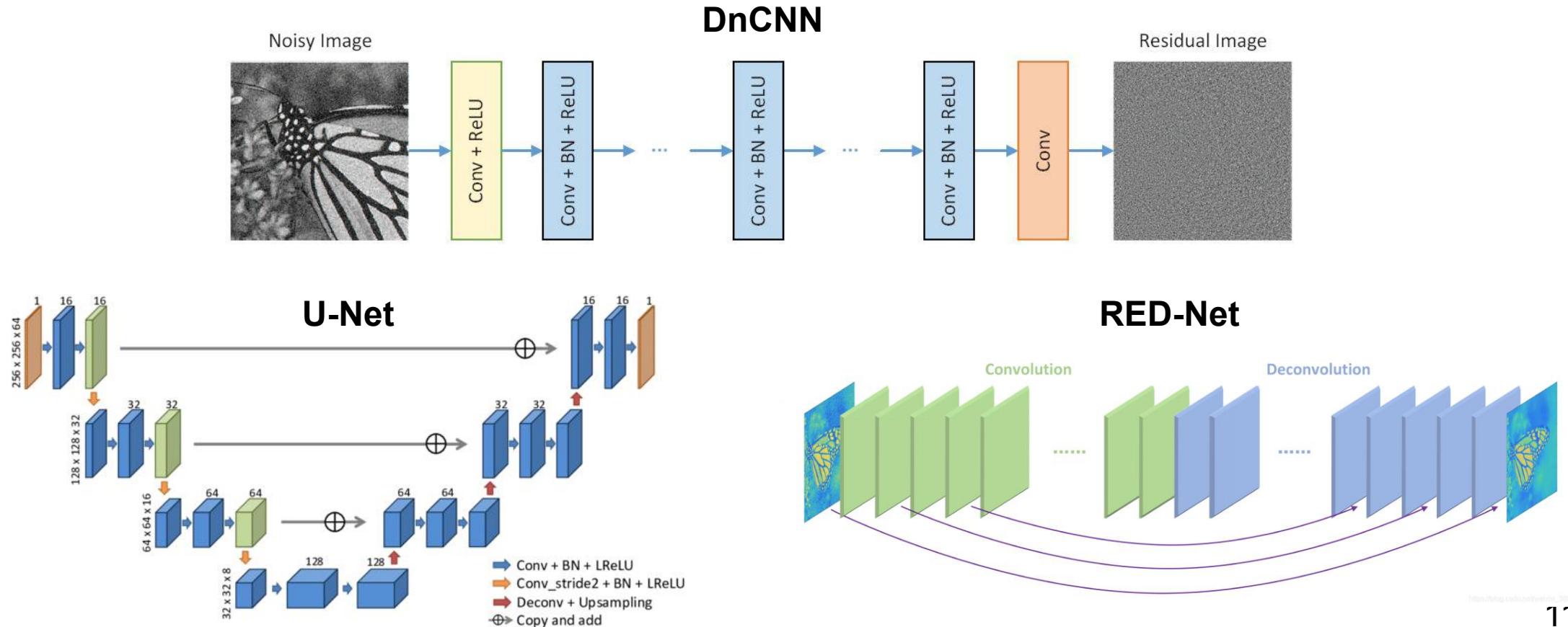
Measured Clean Noise

A diagram showing the equation $\mathbf{y} = \mathbf{x} + \mathbf{e}$. Three arrows point upwards from the words "Measured", "Clean", and "Noise" to the variables \mathbf{y} , \mathbf{x} , and \mathbf{e} respectively.



2. Wavelet-inspired Invertible Neural Network

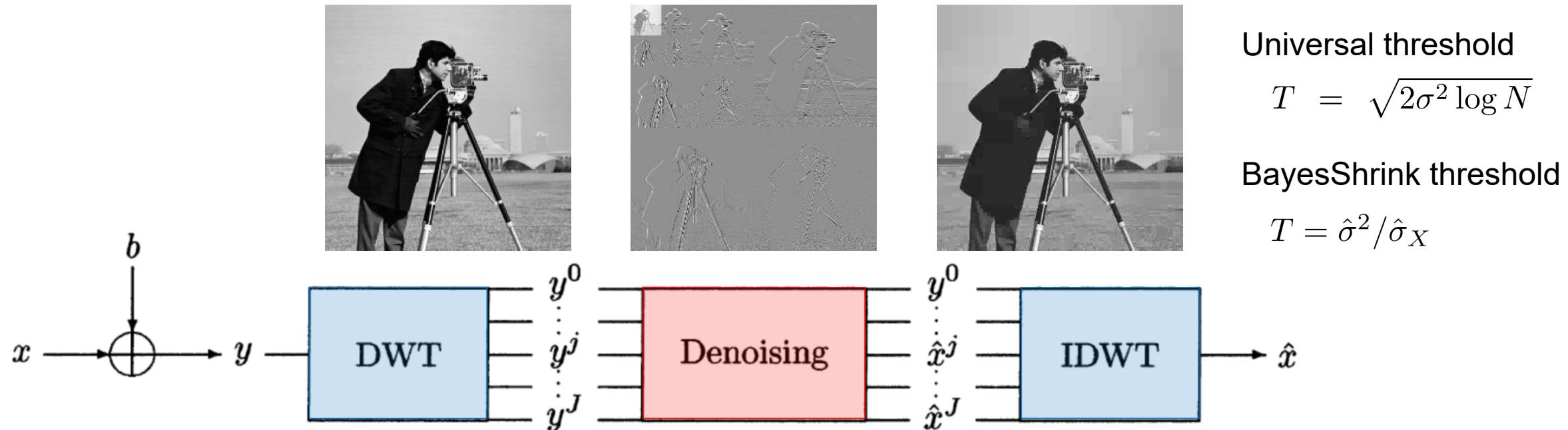
Deep Learning methods are effective while less interpretable and controllable



2. Wavelet-inspired Invertible Neural Network

Wavelet Thresholding is a widely used denoising approach

- **Wavelets** provide invertible sparse representations of piecewise smooth images



Orthonormal bases of compactly supported wavelets
I Daubechies
Communications on pure and applied mathematics 41 (7), 909-996

12504 1988

2. Wavelet-inspired Invertible Neural Network

Motivation:

- Whether it is possible to **combine the merits of Wavelet Thresholding and DNNs** for image denoising and other image restoration tasks?

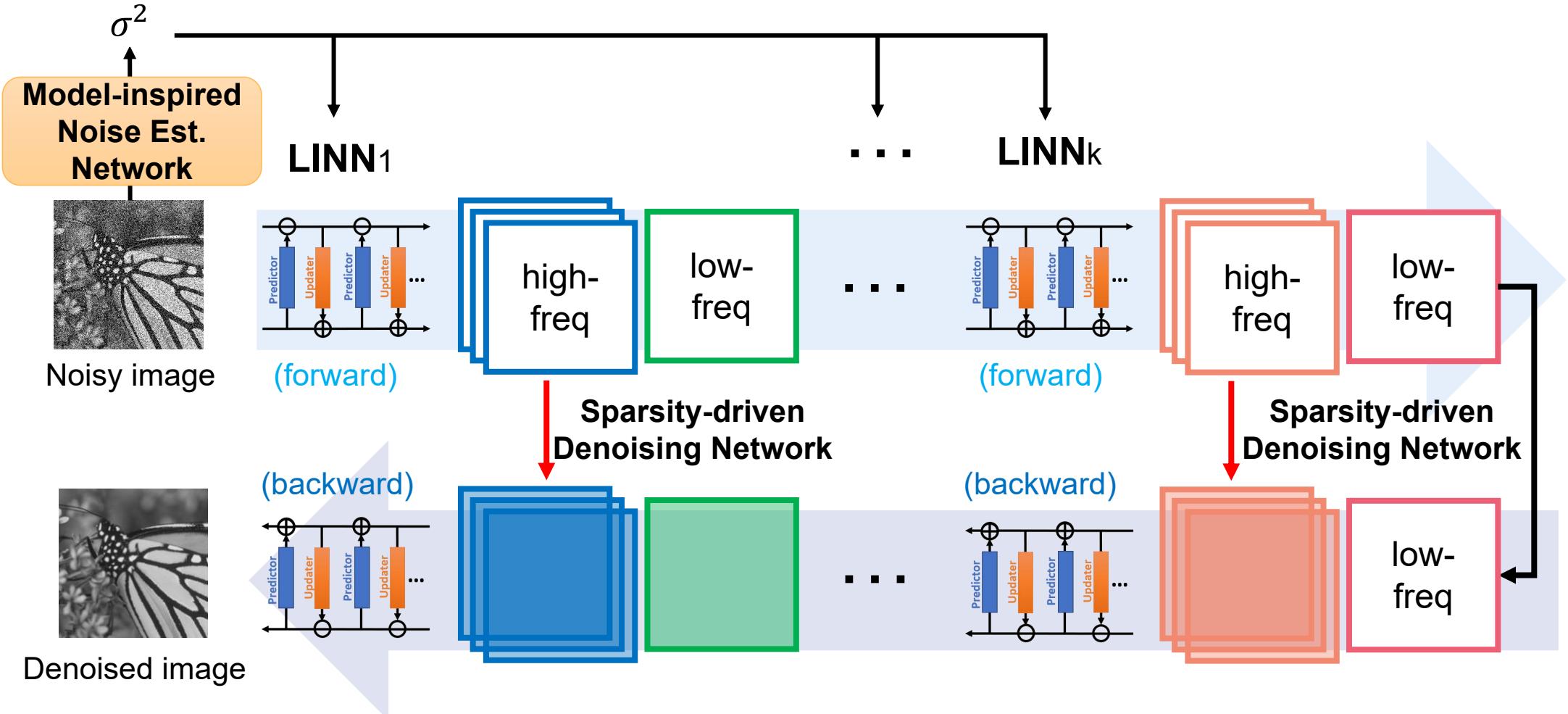
Idea:

- Learning a redundant transform with perfect reconstruction property using **a Wavelet-inspired INvertible Network (WINNet)**



2. Wavelet-inspired Invertible Neural Network

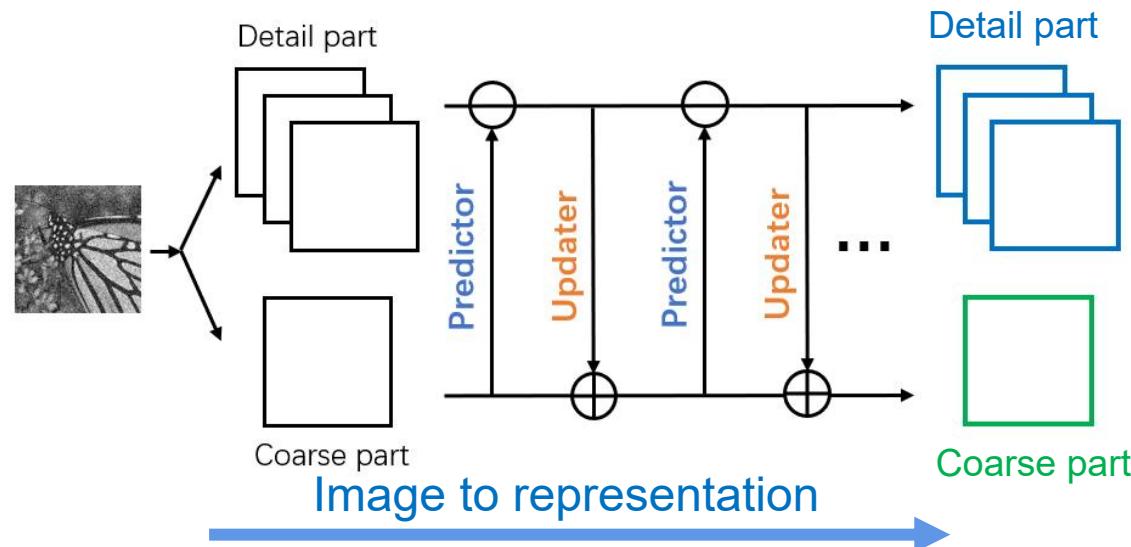
Overall architecture



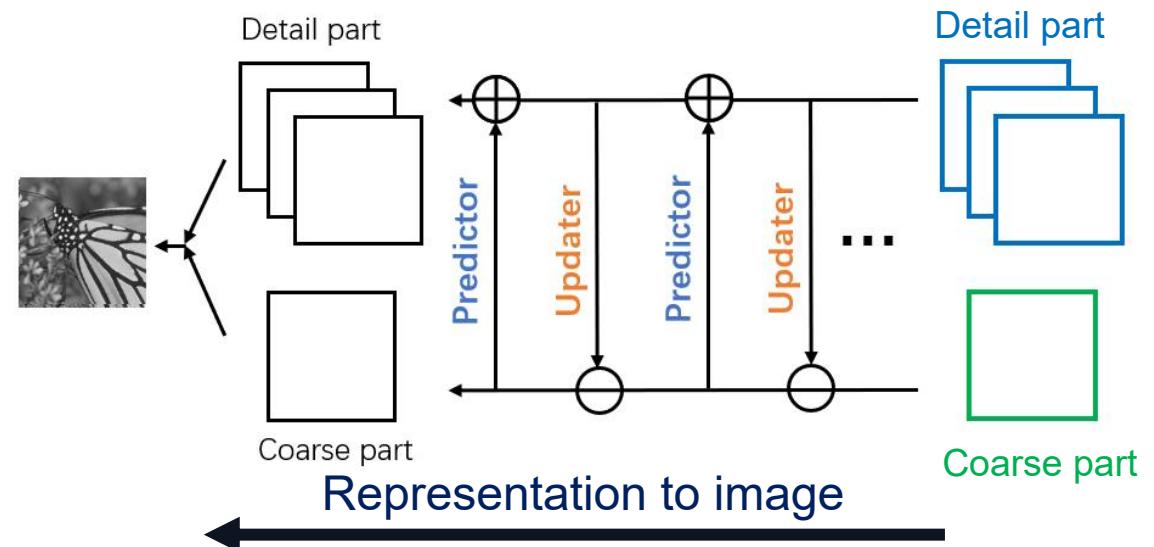
2. Wavelet-inspired Invertible Neural Network

Lifting inspired Invertible Neural Network (LINN)

- Forward pass



- Backward pass

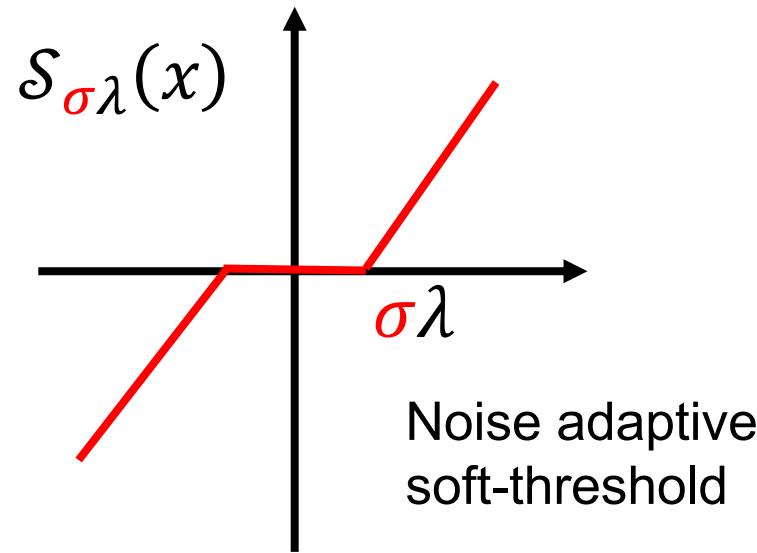
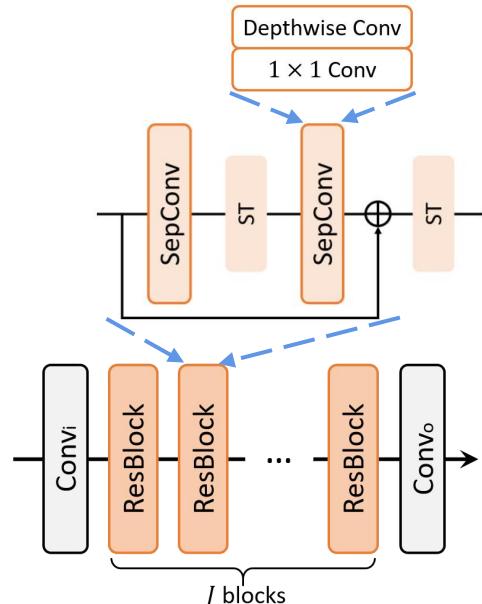


When no operation is applied on the representation, perfect reconstruction can be achieved using the backward pass.

2. Wavelet-inspired Invertible Neural Network

Lifting inspired Invertible Neural Network (LINN)

- Predictor/Updater networks
 - Separable convolutional networks with soft-thresholding non-linearity
 - Noise adaptive soft-threshold

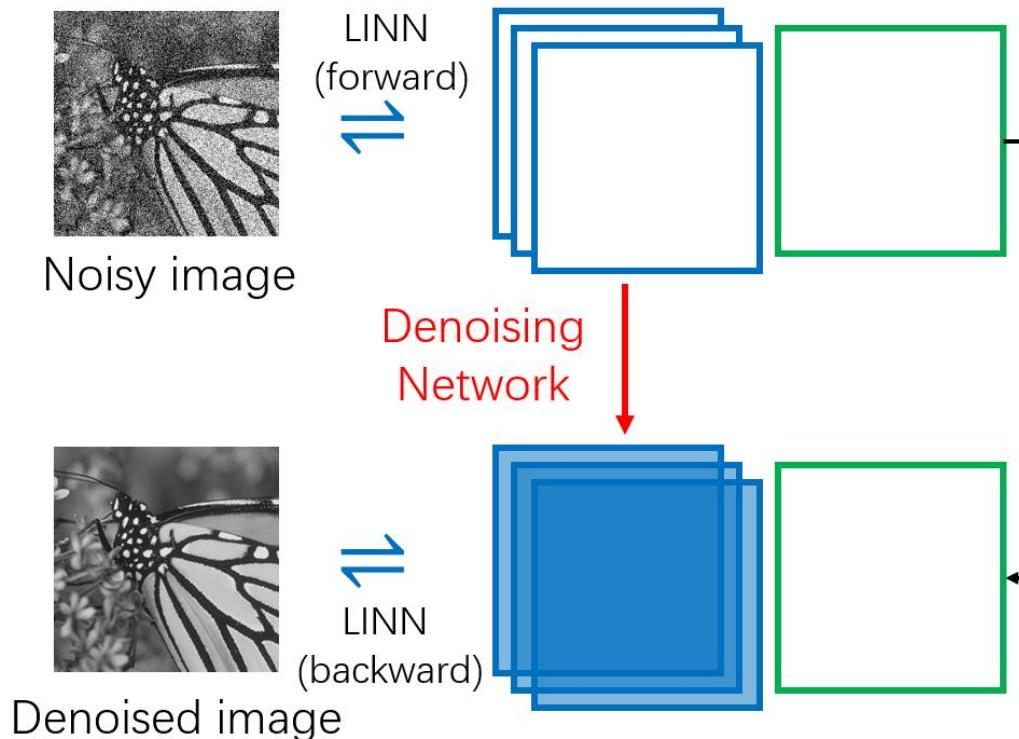


Noise adaptive
soft-threshold

2. Wavelet-inspired Invertible Neural Network

Sparsity-driven Denoising Network

- Non-invertible component
- A well-understood denoising network can lead to enhanced interpretability



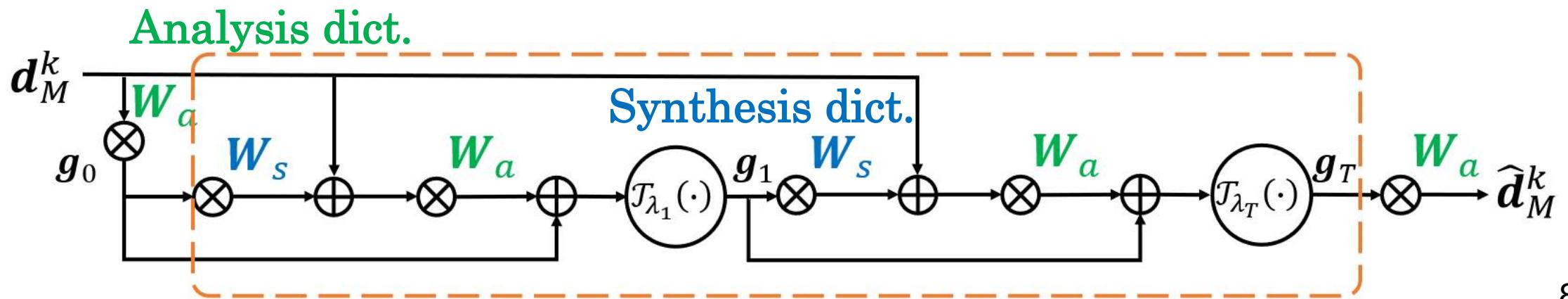
2. Wavelet-inspired Invertible Neural Network

Sparsity-driven Denoising Network

- We model the denoising process as Convolutional Sparse Coding

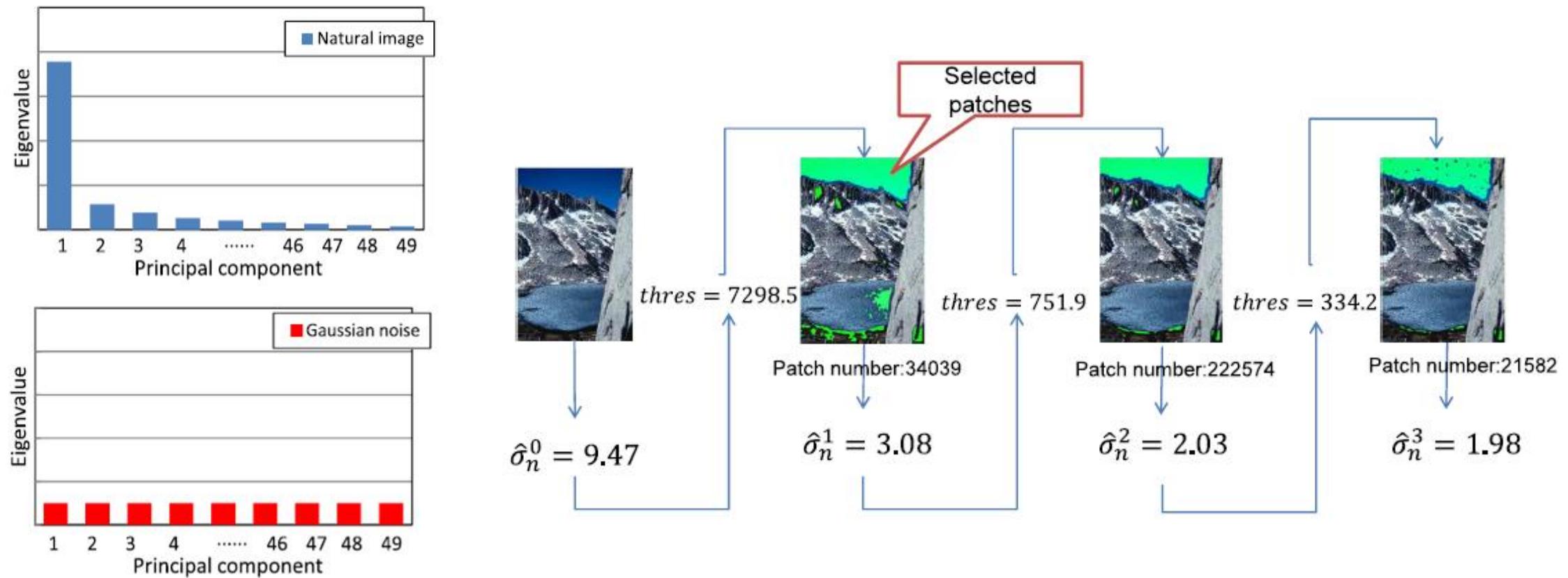
$$\mathbf{g} = \underset{\mathbf{g}}{\operatorname{argmin}} \frac{1}{2} \left\| \mathbf{z}_d^{(I)} - \sum_{i=1}^M \mathbf{D}_i \otimes \mathbf{g}_i \right\|_2^2 + \sum_{i=1}^M \lambda_i \|\mathbf{g}_i\|_1$$

- Unfold it in to CLISTA Network $\mathbf{G}_t = \mathcal{T}_{\lambda_t} (\mathbf{G}_{t-1} + \mathbf{W}_a \otimes (\mathbf{D}_M^k - \mathbf{W}_s \otimes \mathbf{G}_{t-1}))$



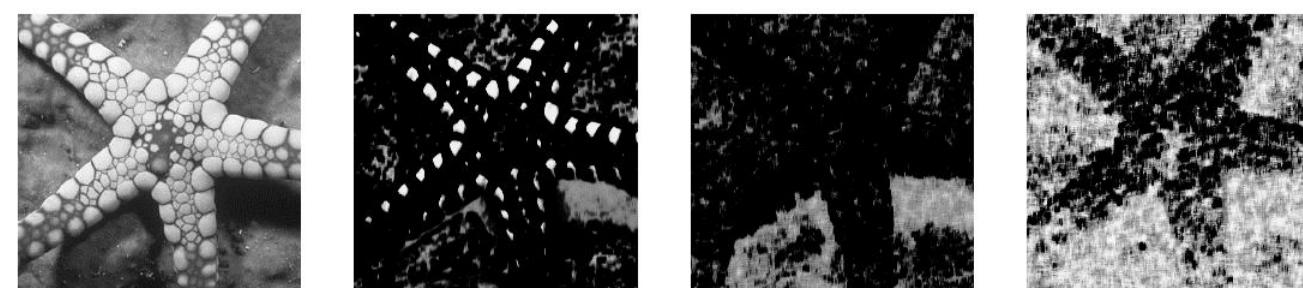
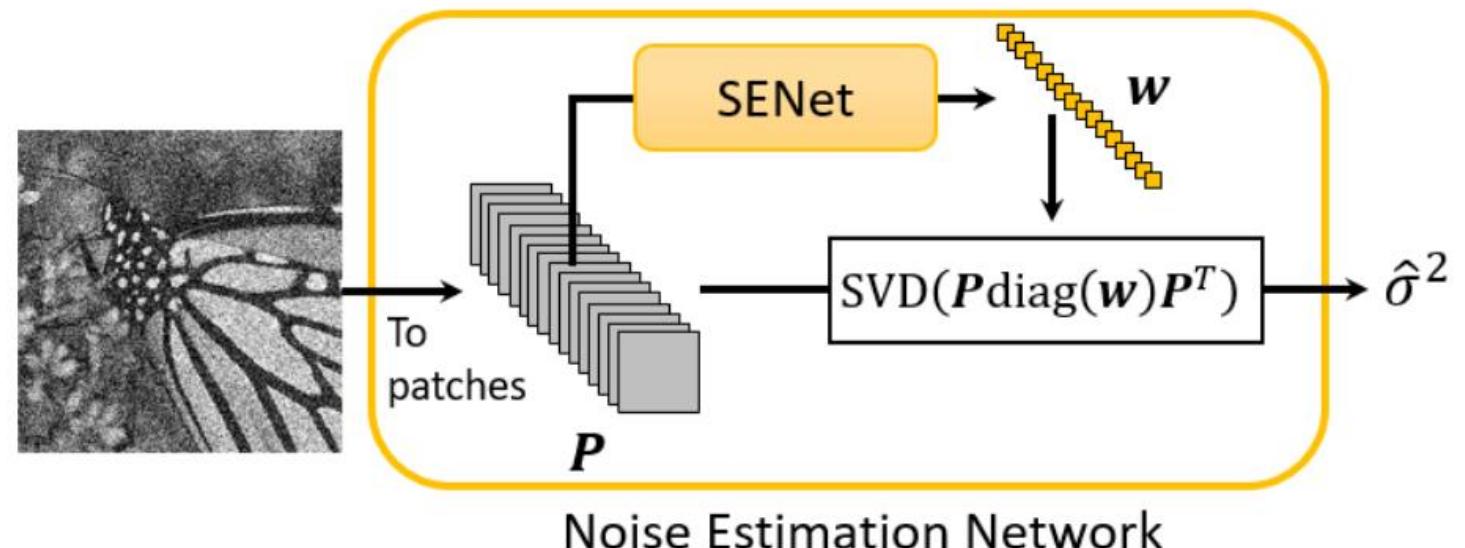
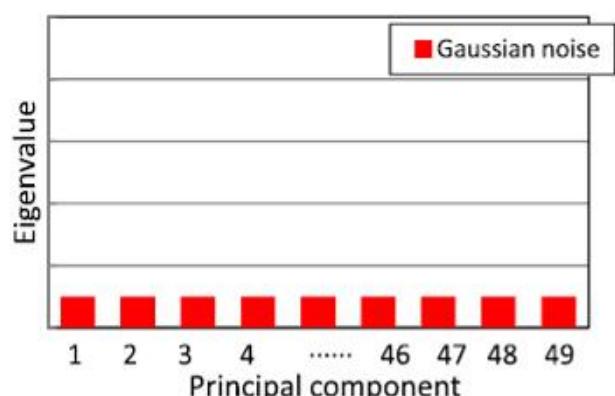
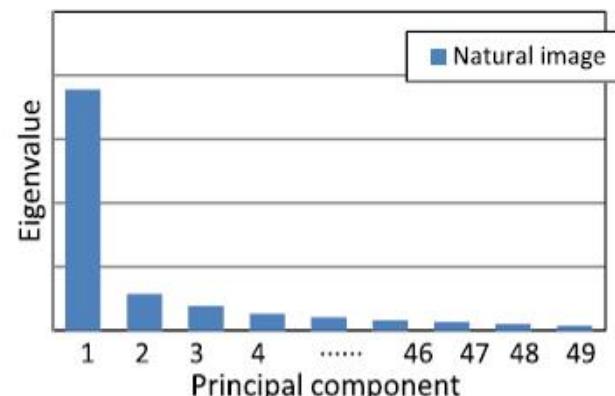
2. Wavelet-inspired Invertible Neural Network

Model-inspired Noise Estimation Network



2. Wavelet-inspired Invertible Neural Network

Model-inspired Noise Estimation Network



(a) Clean image. (b) $\sigma = 1$. (c) $\sigma = 20$. (d) $\sigma = 60$.

2. Wavelet-inspired Invertible Neural Network

Experimental Settings:

- Training loss:

- Mean squared error between restored image and clean image $\mathcal{L}_r = \frac{1}{2N} \sum_{i=1}^N \|X_i - \hat{X}_i\|_2^2$
- Spectral norm loss for LINN $\mathcal{L}_s = \frac{1}{K \cdot M \cdot J} \sum_{k=1}^K \sum_{m=1}^M \sum_{j=1}^J \|P_{m,j}^k\|_S + \|U_{m,j}^k\|_S$
- Orthogonal loss for CLISTA Network $\mathcal{L}_o = \|\mathbf{W}_s \otimes \mathbf{W}_a - \delta\|_F^2$

- Optimizer:

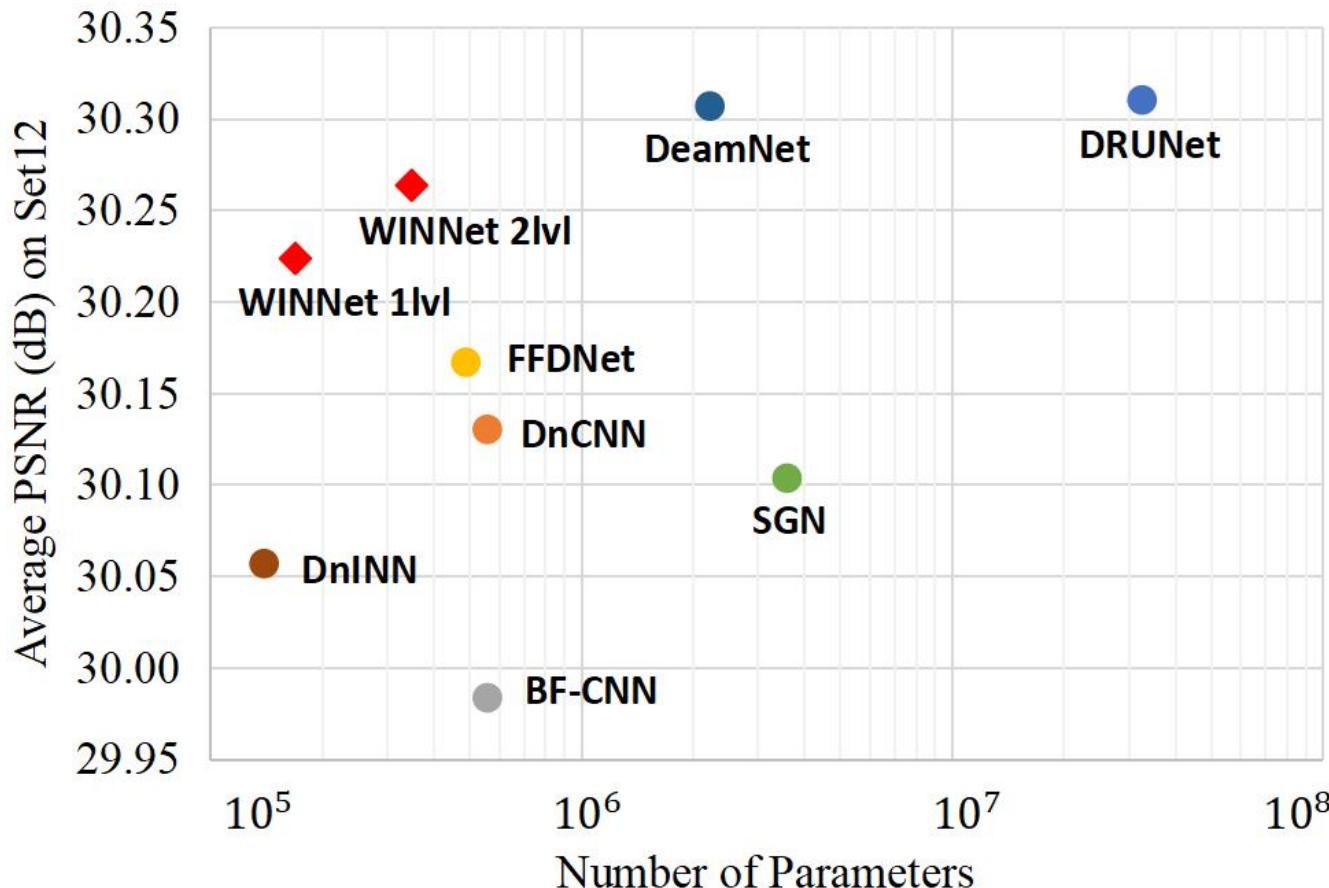
- Adam with learning rate 1×10^{-3} which decays to 1×10^{-4} at the 30-th epoch

- Training data:

- 400 images of size 180×180

2. Wavelet-inspired Invertible Neural Network

Experimental Results — Non-blind denoising



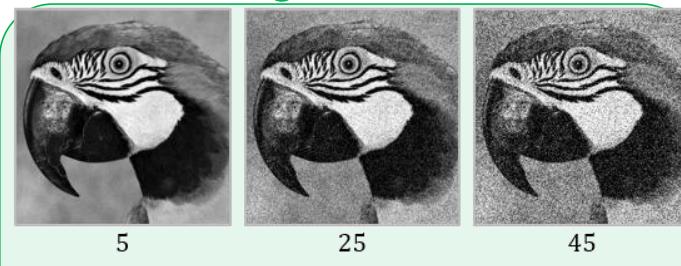
Comparison of average PSNR (dB) and number of parameters of different methods.
The testing dataset is *Set12* with noise level $\sigma = [15, 25, 50]$.

2. Wavelet-inspired Invertible Neural Network

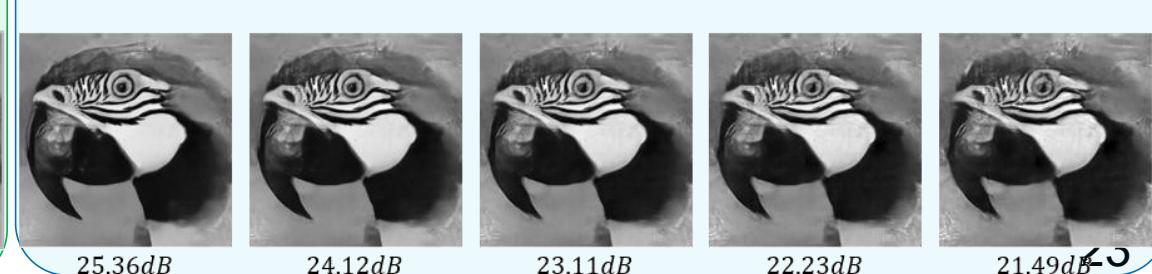
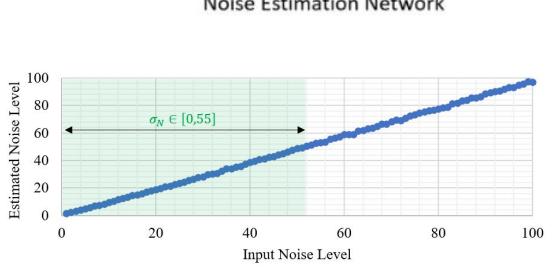
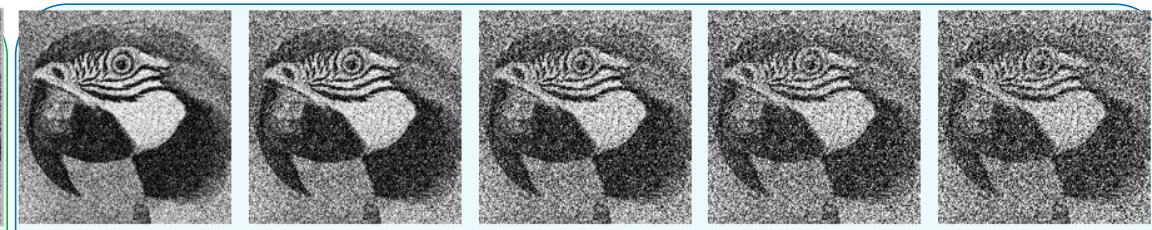
Experimental Results — Blind denoising

Dataset	Methods	$\sigma = 5$	$\sigma = 25$	$\sigma = 45$	$\sigma = 65$	$\sigma = 85$	$\sigma = 105$	$\sigma = 125$	$\sigma = 145$
BSD68	DnCNN-B [21]	37.75	29.15	26.62	23.00	16.07	13.19	11.68	10.79
	BUIFD [29]	37.41	28.76	25.61	23.07	18.81	15.98	14.45	13.52
	BF-CNN [28]	37.73	29.11	26.58	<u>25.12</u>	<u>24.10</u>	<u>23.33</u>	<u>22.70</u>	<u>22.18</u>
	WINNet (1-scale)	37.82	<u>29.13</u>	26.66	25.23	24.23	23.46	22.81	22.23
Set12	DnCNN-B [21]	37.88	30.38	27.68	23.52	15.95	13.18	11.78	10.92
	BUIFD [29]	37.34	30.18	27.01	24.27	19.41	16.28	14.66	13.73
	BF-CNN [28]	37.81	<u>30.33</u>	27.58	<u>25.83</u>	<u>24.54</u>	<u>23.55</u>	<u>22.74</u>	<u>22.07</u>
	WINNet (1-scale)	38.22	30.33	27.72	26.03	24.77	23.76	22.94	22.24

Training noise levels



Unseen noise levels



2. Wavelet-inspired Invertible Neural Network

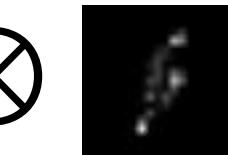
Application on Image Deblurring



Blur Image



Sharp Image



Noise

$$\mathbf{x} = \arg \min_{\mathbf{x}} \frac{1}{2\sigma^2} \|\mathbf{y} - \mathbf{k} \otimes \mathbf{x}\|_2^2 + \lambda \Phi(\mathbf{x})$$



$$\begin{cases} \mathbf{x}_k = \arg \min_{\mathbf{x}} \|\mathbf{y} - \mathbf{k} \otimes \mathbf{x}\|_2^2 + \frac{\lambda\sigma^2}{\beta^2} \|\mathbf{x} - \mathbf{z}_{k-1}\|_2^2 \\ \mathbf{z}_k = \arg \min_{\mathbf{z}} \frac{1}{2\beta^2} \|\mathbf{z} - \mathbf{x}_k\|_2^2 + \Phi(\mathbf{z}) \end{cases}$$

2. Wavelet-inspired Invertible Neural Network

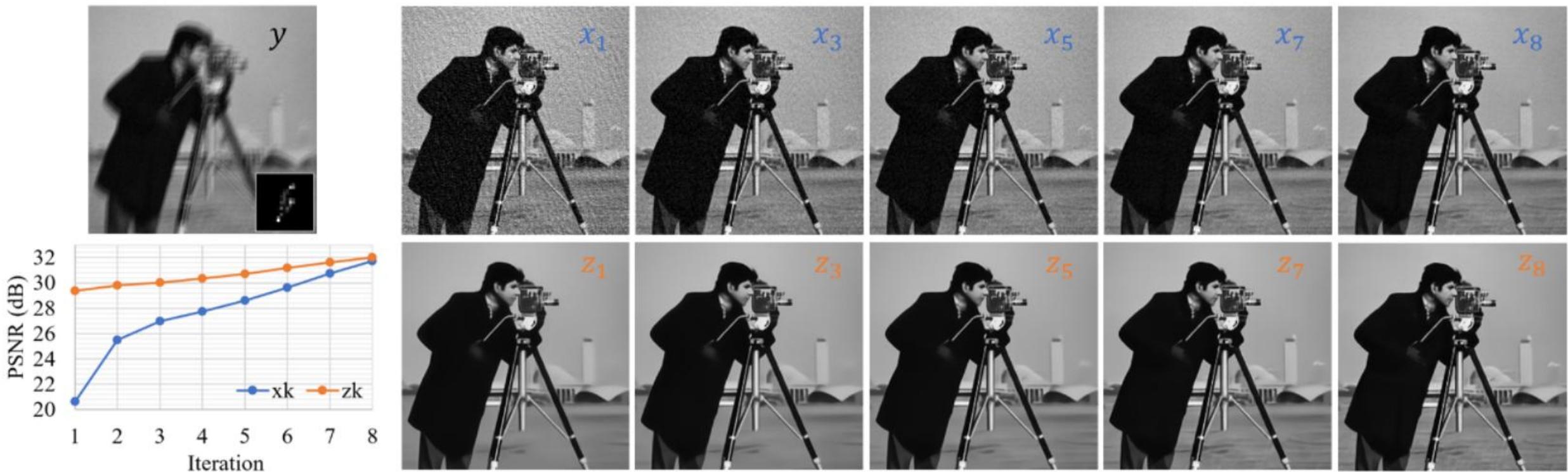
Image Deblurring with WINNet

Algorithm 1: Plug-and-Play image deblurring with blind WINNet.

- 1 **Input:** Input image y , kernel k , parameter λ ;
- 2 **Initialize:** $z_0 = y$, $\beta_0 = \text{NENet}(z_0)$, $\beta_1 = 10 \times \beta_0$,
 $k = 1$;
- 3 **while** $\beta_k > \beta_0$ **do**
- 4 $x_k = \arg \min_x \|y - k \otimes x\|_2^2 + \frac{\lambda \beta_0^2}{\beta_k^2} \|x - z_{k-1}\|_2^2$; %Auxiliary Update
- 5 $\beta_{k+1} = \text{NENet}(x_k)$;
- 6 $z_k = \text{WINNet}(x_k, 2\beta_{k+1})$;
- 7 $k = k + 1$;
- 8 **end**
- 9 **Output:** Deblurred image $x = z_{k-1}$.

2. Wavelet-inspired Invertible Neural Network

Experimental Results on Image Deblurring



2. Wavelet-inspired Invertible Neural Network

Take home message:

- With proper nonlinear over-parameterization, Wavelet-inspired network architecture can achieve **good performance, strong controllability, generalization ability and high interpretability**

J.-J. Huang, and P.L. Dragotti, "**WINNet: Wavelet-inspired Invertible Network for Image Denoising**," in *IEEE Transactions on Image Processing (TIP)*, 2022.

J.-J. Huang, and P.L. Dragotti, "**LINN: Lifting Inspired Invertible Neural Network for Image Denoising**," in Proceedings of *29th European Signal Processing Conference (EUSIPCO)*, Ireland, 2021.

Outline

1. Overview of Invertible Neural Networks

- Origin of INN and Normalizing flows
- INN for Inverse Problems

2. Wavelet-Inspired Invertible Neural Network

3. INN and diffusion models: INDigo

4. Other applications of INN

3. INN and Diffusion Models

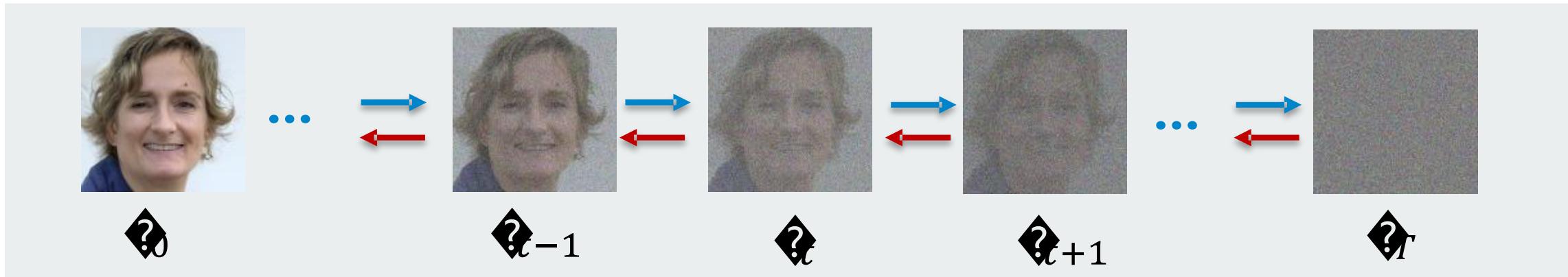
$$\hat{x} = \min_x \|H(x) - y\|^2 + \lambda \rho(x)$$

consistency term prior

- Impose consistency using the forward part of the INN
- Impose the prior using diffusion models
- Iterate

3. INN and Diffusion Models

Diffusion Models are good for “unconditional” generation of new samples (e.g., Denoising Probabilistic Diffusion Models)



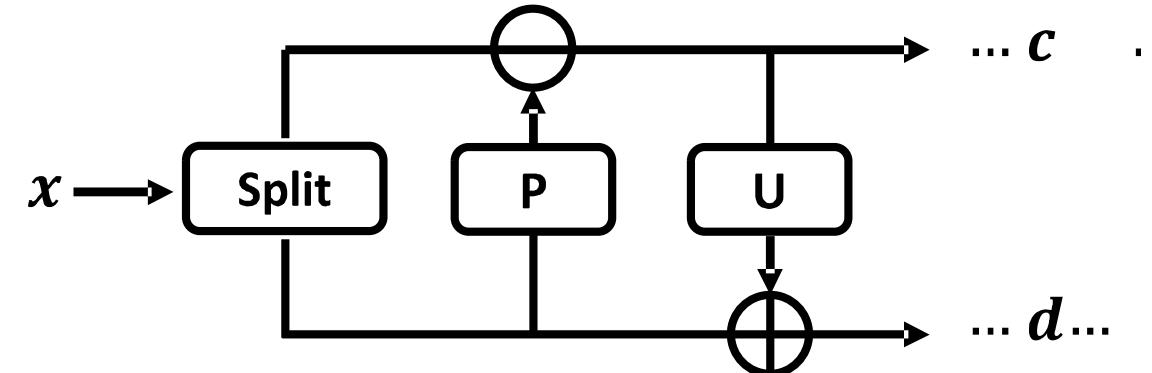
Motivation: Can we use a pretrained “unconditional” diffusion model for inverse problems?

J. Ho, J. Ajay and P. Abbeel. "Denoising diffusion probabilistic models." in Proceedings of *Advances in Neural Information Processing Systems (NeurIPS) 2020*.

3. INN and Diffusion Models

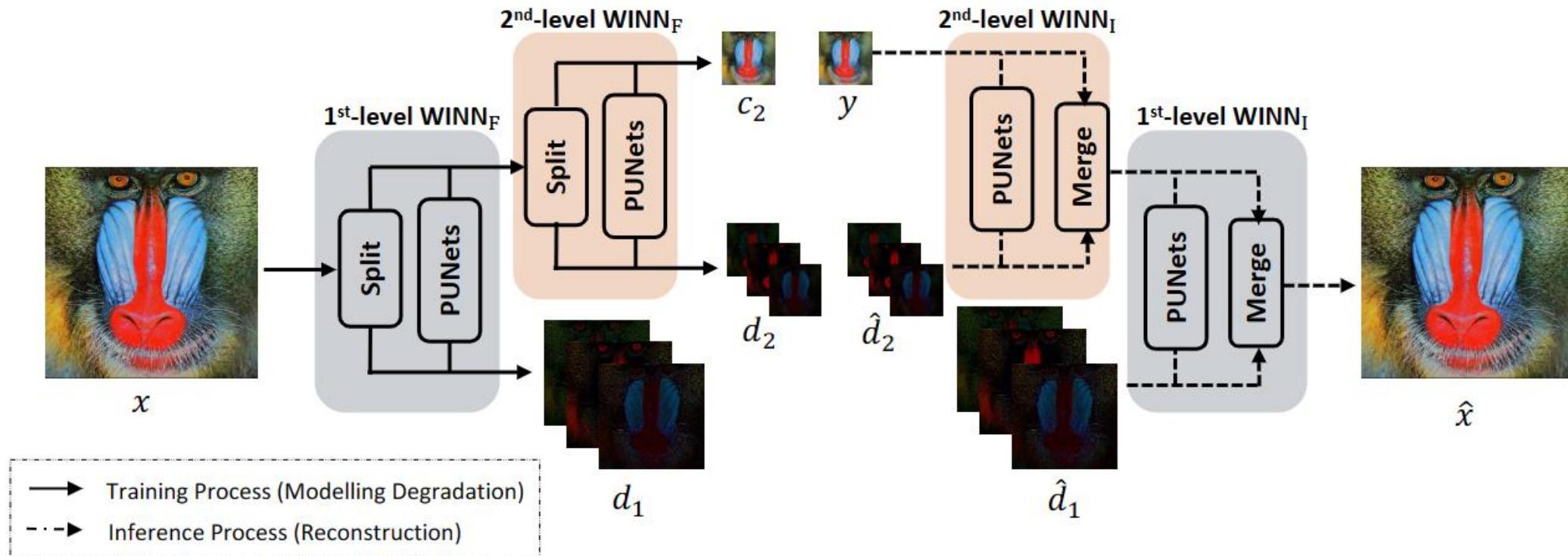
- Given a training set $\{x_i, y_i\}$ which contains N high-quality images and their low-quality counterparts, we learn the forward part of the INN using the following loss:

$$L(\Theta) = \frac{1}{N} \sum_{i=1}^N \|\mathbf{c}^i - \mathbf{y}^i\|_2^2,$$

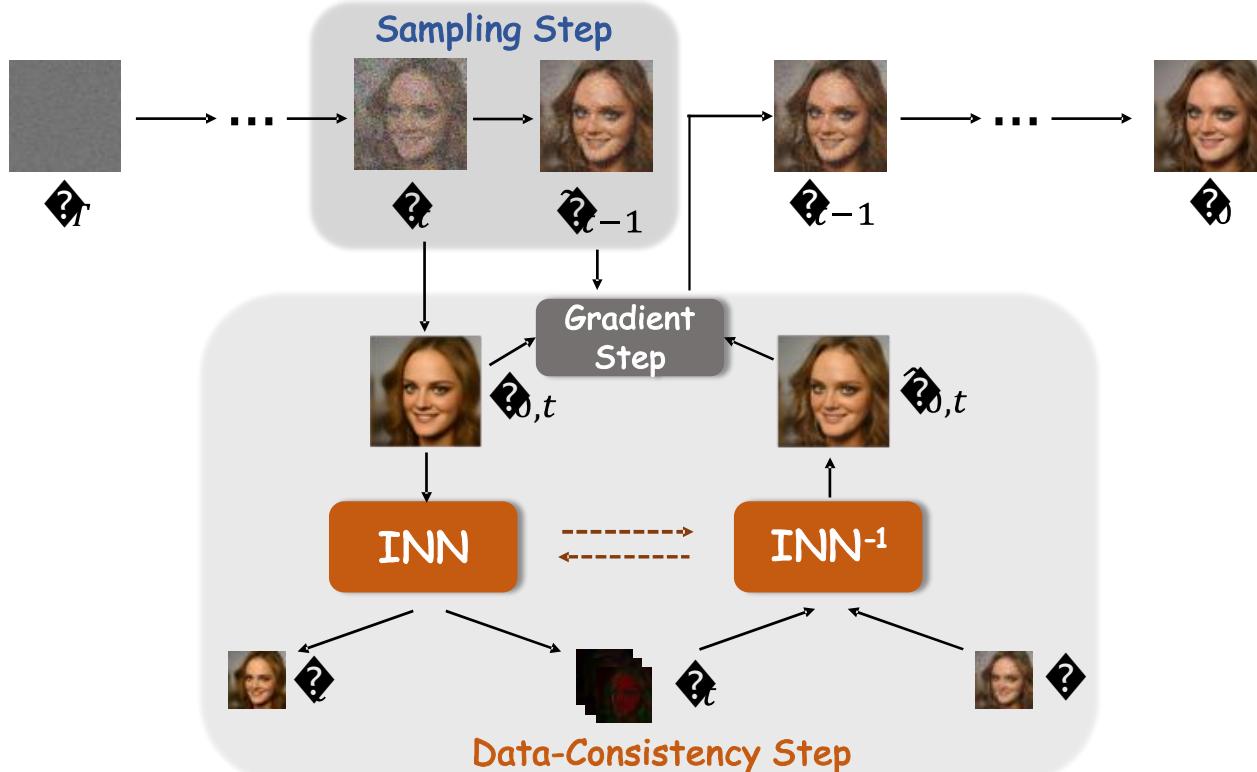


- Consequently, d models the lost details that need to be recovered with the diffusion model

3. INN and Diffusion Models



3. INN and Diffusion Models



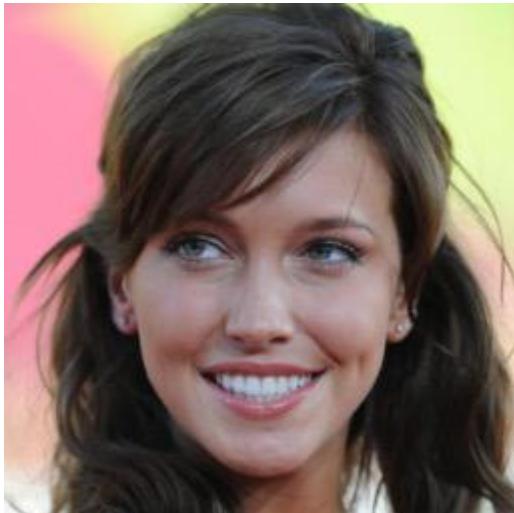
Algorithm 1 INDigo

```

1:  $\mathbf{x}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ 
2: for  $t = T, \dots, 1$  do
3:    $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$  if  $t > 1$ , else  $\mathbf{z} = \mathbf{0}$ 
4:    $\mathbf{x}_{0,t} = \frac{1}{\sqrt{\bar{\alpha}_t}} (\mathbf{x}_t - \sqrt{1 - \bar{\alpha}_t} \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t))$ 
5:    $\tilde{\mathbf{x}}_{t-1} = \frac{\sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} \mathbf{x}_t + \frac{\sqrt{\bar{\alpha}_{t-1}} \beta_t}{1 - \bar{\alpha}_t} \mathbf{x}_{0,t} + \sigma_t \mathbf{z}$ 
6:    $\mathbf{c}_t, \mathbf{d}_t = f_\phi(\mathbf{x}_{0,t})$ 
7:    $\hat{\mathbf{x}}_{0,t} = f_\phi^{-1}(\mathbf{y}, \mathbf{d}_t)$ 
8:    $\mathbf{x}_{t-1} = \tilde{\mathbf{x}}_{t-1} - \zeta \nabla_{\mathbf{x}_t} \|\hat{\mathbf{x}}_{0,t} - \mathbf{x}_{0,t}\|_2^2$ 
9: end for
10: return  $\mathbf{x}_0$ 

```

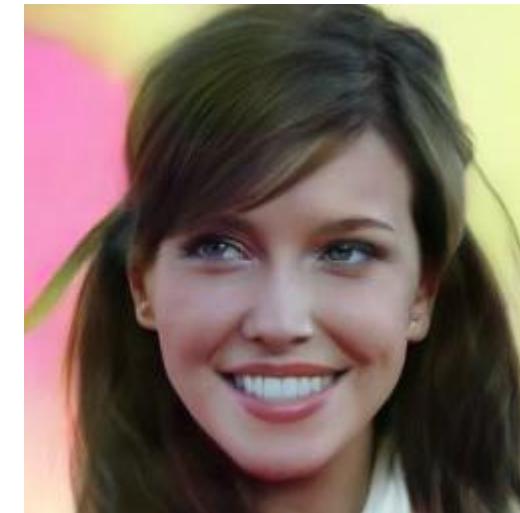
3. INN and Diffusion Models



Ground Truth



Degraded

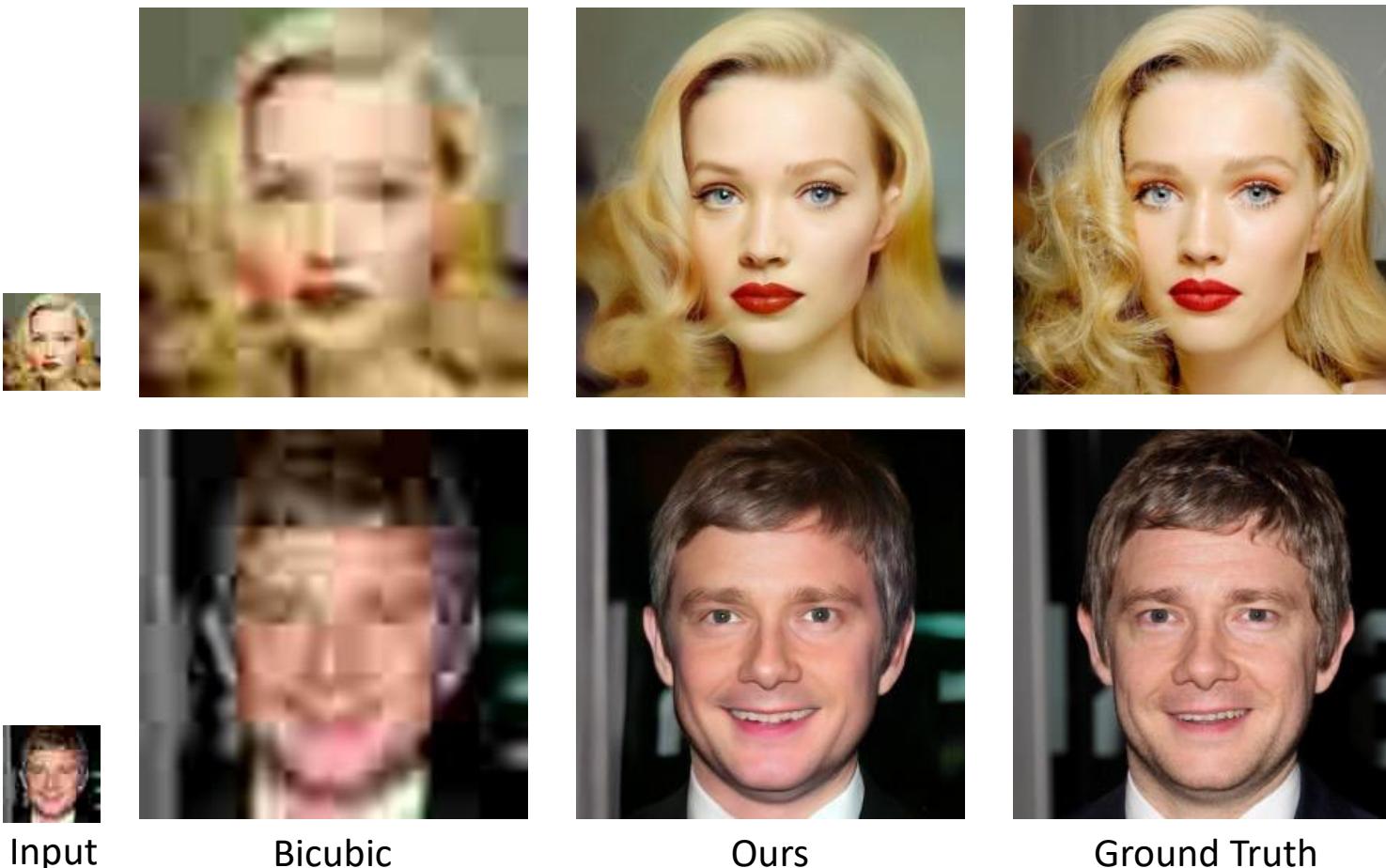


Reconstructed

- This approach is simple, flexible and effective
 - No-need to know the degradation process
 - The degradation process can be highly non-linear
 - No need to retrain the diffusion model for every new degradation (just need to train the INN)

3. INN and Diffusion Models

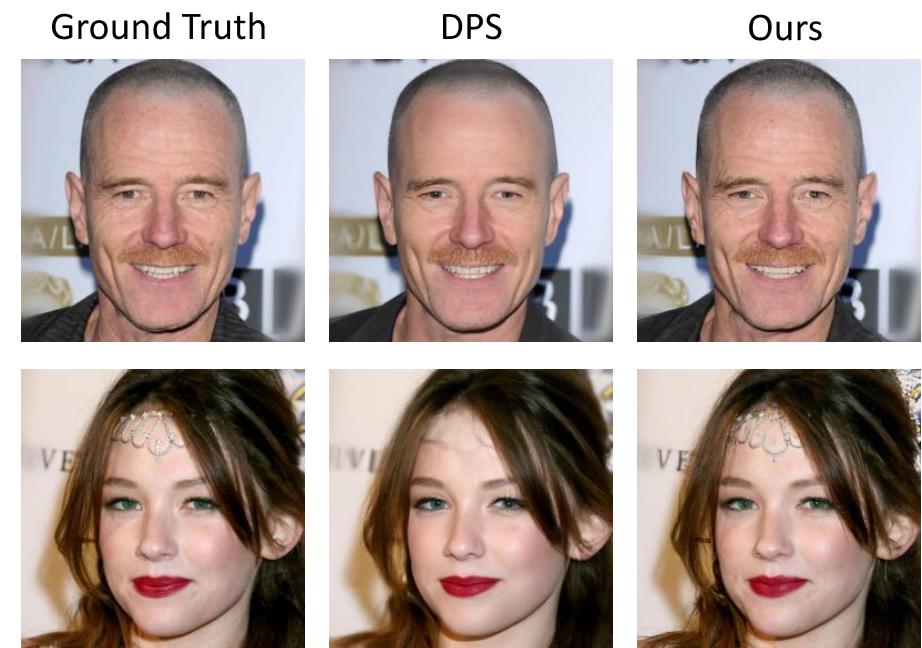
Results for non-linear degradation models



3. INN and Diffusion Models

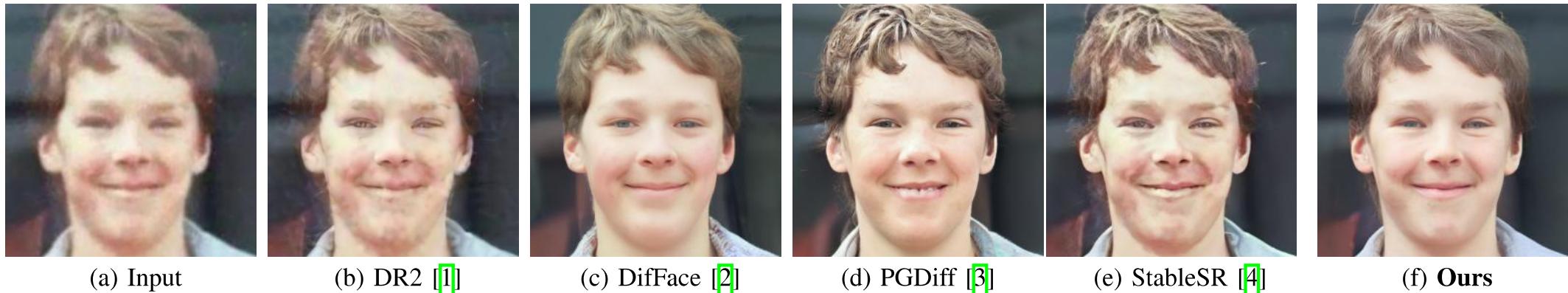
Results on 4x super-resolution

Method	Noise σ	PSNR ↑	FID ↓	LPIPS ↓	NQE ↓
LVR	0	27.43	44.04	0.2123	5.4689
DDRM	0	28.08	65.80	0.1722	4.4694
DPS	0	26.67	32.44	0.1370	4.4890
Ours	0	28.15	22.33	0.0889	4.1564
LVR	0.05	26.42	60.27	0.3045	4.6527
DDRM	0.05	27.06	45.90	0.2028	4.8238
DPS	0.05	25.92	31.71	0.1475	4.3743
Ours	0.05	27.16	26.64	0.1215	4.1004
LVR	0.10	24.60	88.88	0.4833	4.4888
DDRM	0.10	26.16	45.49	0.2273	4.9644
DPS	0.10	24.73	31.66	0.1698	4.2388
Ours	0.10	26.25	28.89	0.1399	3.9659



3. INN and Diffusion Models

Results on blind unsupervised deconvolution



D. You, F. Andreas, and P.L. Dragotti. "INDigo: An INN-Guided Probabilistic Diffusion Algorithm for Inverse Problems." in Proceedings of *IEEE 25th International Workshop on Multimedia Signal Processing (MMSP)*, 2023.
(Best Paper Award)

Outline

1. Overview of Invertible Neural Networks

- Origin of INN and Normalizing flows
- INN for Inverse Problems

2. Wavelet-Inspired Invertible Neural Network

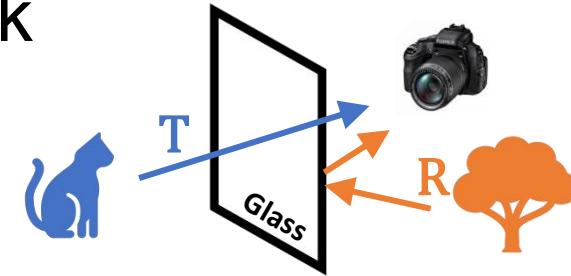
3. INN and diffusion models: INDigo

4. Other applications of INN

4. Other Applications of INN: Blind Source Separation

DURRNet: Deep Unfolded Reflection Removal Network

- Overparameterize the **wavelet transform** as a learnable **INN**

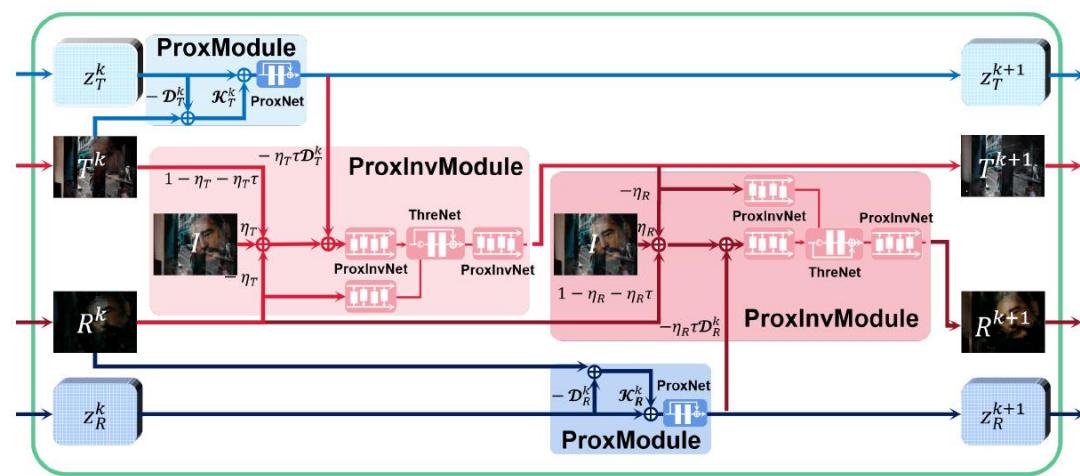


$$\min_{\mathbf{z}_T, \mathbf{z}_R} \frac{1}{2} \left\| \mathbf{I} - \sum_{i=1}^N \mathbf{D}_T^i \otimes \mathbf{z}_T^i - \sum_{i=1}^N \mathbf{D}_R^i \otimes \mathbf{z}_R^i \right\|_F^2 + \lambda_T p_T(\mathbf{z}_T)$$
$$+ \lambda_R p_R(\mathbf{z}_R) + \kappa \mathcal{E} \left(\sum_{i=1}^N \mathbf{D}_T^i \otimes \mathbf{z}_T^i, \sum_{i=1}^N \mathbf{D}_R^i \otimes \mathbf{z}_R^i \right)$$

Exclusion Prior:

$$\mathcal{E}(\mathbf{T}, \mathbf{R}) = \sum_{m=1}^M \|(\mathbf{W}_m \otimes \mathbf{T}) \odot (\mathbf{W}_m \otimes \mathbf{R})\|_1$$

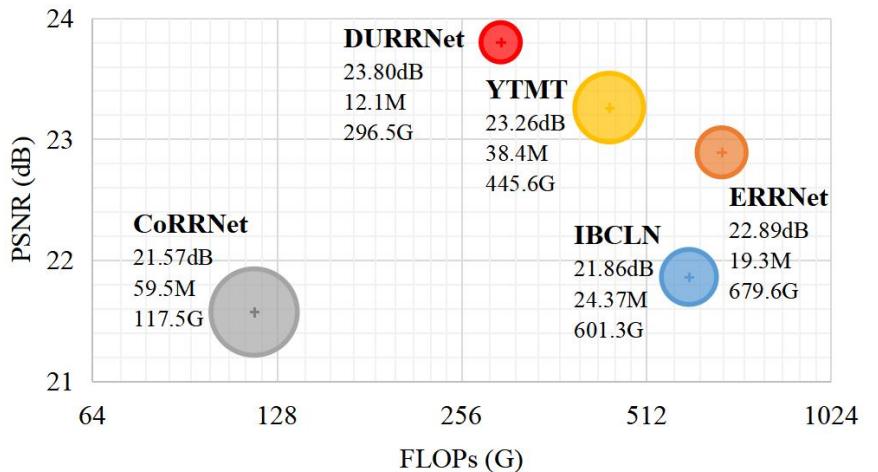
where \mathbf{W} denotes **wavelet transform**.



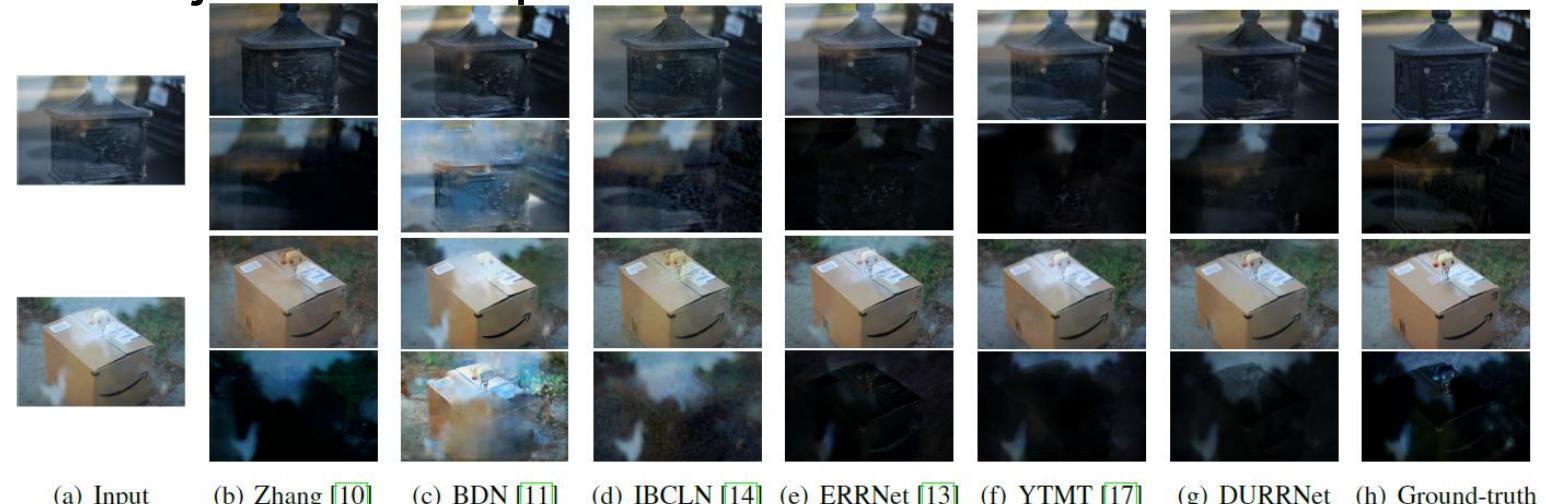
4. Other Applications of INN: Blind Source Separation

DURRNet: Deep Unfolded Reflection Removal Network

- PSNR v.s. FLOPS and #Params



- Subjective comparisons

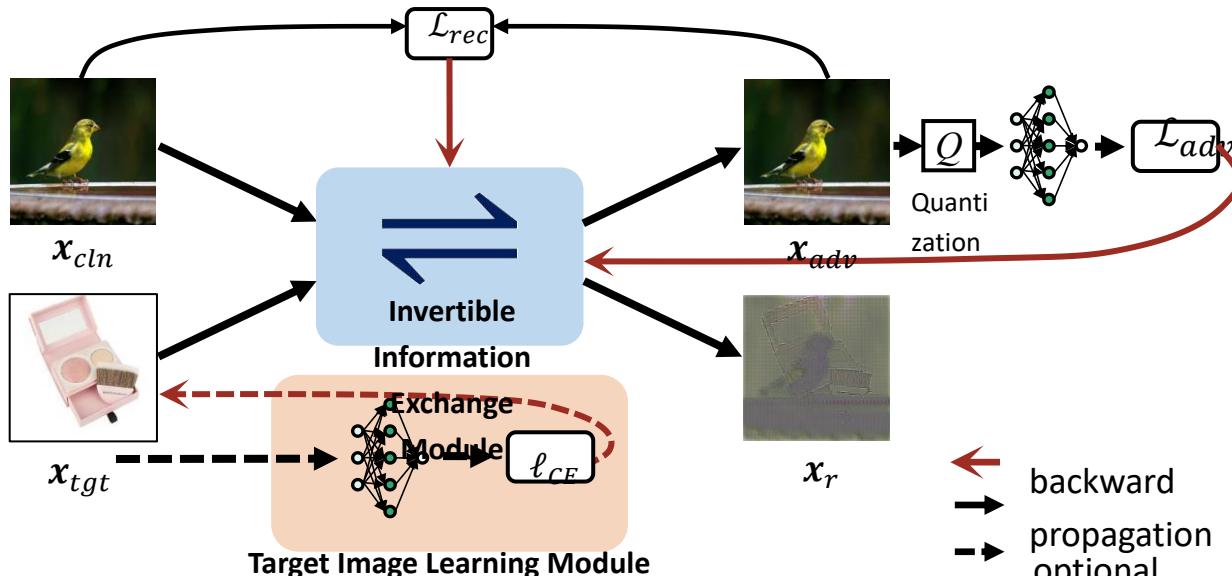


- Objective comparisons:

Dataset	Metrics	CEILNet [8]	Zhang <i>et al.</i> [10]	BDN [11]	IBCLN [14]	CoRRN [22]	ERRNet [13]	YTMT [17]	DURRNet (proposed)
Real20 (20)	PSNR (\uparrow)	18.45	22.55	18.41	21.86	21.57	22.89	23.26	23.80
	SSIM (\uparrow)	0.690	0.788	0.726	0.762	0.807	0.803	0.806	0.814
Nature (20)	PSNR (\uparrow)	19.33	19.56	18.92	23.57	21.84	20.60	23.85	24.24
	SSIM (\uparrow)	0.745	0.736	0.737	0.783	0.805	0.755	0.810	0.812

4. Other Applications of INN: Adversarial Attacks

AdvINN: Adversarial Attack via Invertible Neural Networks



Visualization & Interpretation

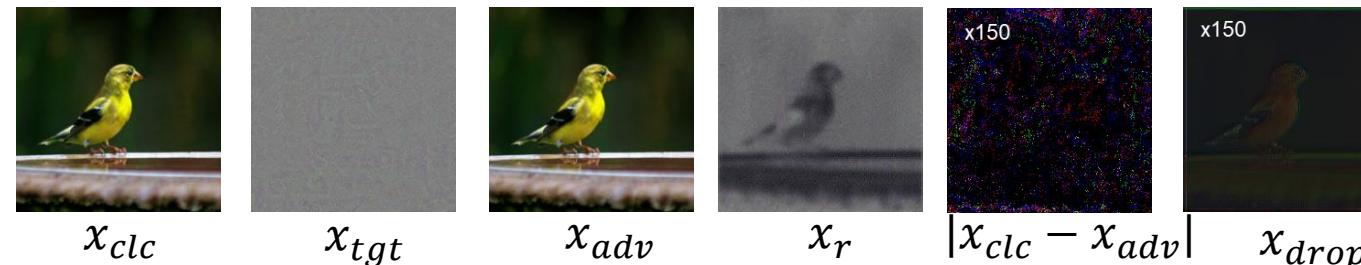


Table 1: Accuracy and evaluation metrics on different methods. All methods use $\epsilon = 8/255$ as the adversarial budget. ASR donates the accuracy of adversarial attacks. ↑ means the value is higher the better, and vice versa. (The best and the second best result in each column is in bold and underline.)

Dataset	Methods	$l_2 \downarrow$	$l_\infty \downarrow$	SSIM↑	LPIPS↓	FID↓	ASR(%)↑
ImageNet-1K	StepLL	26.90	0.04	0.948	0.1443	25.176	98.5
	C&W	10.33	0.07	0.977	0.0617	11.515	91.7
	PGD	64.42	0.04	0.881	0.2155	35.012	90.2
	PerC-AL	1.93	0.10	0.995	0.0339	5.118	100.0
	AdvDrop	18.47	0.07	0.977	0.0639	9.687	100.0
	SSAH	6.97	0.03	0.991	0.0352	5.221	99.8
	AdvINN-HCT	5.73	0.03	0.991	0.0206	3.661	100.0
	AdvINN-UAP	5.84	0.03	0.990	0.0212	2.900	100.0
	AdvINN-CGT	<u>2.66</u>	0.03	0.996	0.0118	1.594	100.0

Less perceptible adversarial examples with 100% attacking success rate!

Conclusions

- The perfect reconstruction property of the **Invertible Neural Networks** is intriguing
- Designing INN using wavelets/lifting leads to more **interpretable and simpler architectures**
- Good **generalization ability**
- Invertible neural networks have the **potential** for many image/signal processing applications

Q&A

Thanks for listening!

Related Publications

- J.-J. Huang and P.L. Dragotti, "WINNet: Wavelet-inspired Invertible Network for Image Denoising", in *IEEE Transactions on Image Processing (TIP)*, 2022.
- J.-J. Huang and P.L. Dragotti, "LINN: Lifting Inspired Invertible Neural Network for Image Denoising", in Proceedings of *29th European Signal Processing Conference (EUSIPCO)*, 2021.
- Y. Di, A. Floros, and P.L. Dragotti. "INDigo: An INN-Guided Probabilistic Diffusion Algorithm for Inverse Problems", in Proceedings of *IEEE 25th International Workshop on Multimedia Signal Processing (MMSP)*, 2023.
- Z. Chen et al. "Invertible Mosaic Image Hiding Network for Very Large Capacity Image Steganography", in Proceedings of *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024.
- J.-J. Huang et al. "DURRNNet: Deep Unfolded Single Image Reflection Removal Network with Joint Prior", in Proceedings of *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024.
- Z. Chen et al. "Imperceptible Adversarial Attack Via Invertible Neural Networks", in Proceedings of *AAAI Conference on Artificial Intelligence*, 2023.