NeurIPS 2021 Workshop on Deep Learning and Inverse Problems

Learning Lipschitz-Controlled Activation Functions in Neural Networks for Plug-and-Play Image Reconstruction Methods

Pakshal Bohra, Dimitris Perdios, Alexis Goujon, Sébastien Emery and Michael Unser

Biomedical Imaging Group, École polytechnique fédérale de Lausanne (EPFL)

Plug-and-Play (PnP) Image Reconstruction

- Recover image from measurements

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}$$

Measurements · Imaging system · Image · Noise

- Variational framework

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}} E(\mathbf{y}, \mathbf{H}\mathbf{x}) + \lambda R(\mathbf{x})$$

Data-consistency · Regularization

  - Can be solved using proximal algorithms such as forward-backward splitting (FBS)

- PnP: replace proximal operator with powerful denoiser D (e.g., state-of-the-art CNNs)

Challenge

- PnP-FBS converges to a fixed point if

$$\mathrm{D} = \beta\mathrm{CNN} + (1-\beta)\mathrm{Id}$$

1-Lipschitz · $\in (0,1)$

⚠ Constraining ReLU CNNs ➞ drop in performance

Contribution

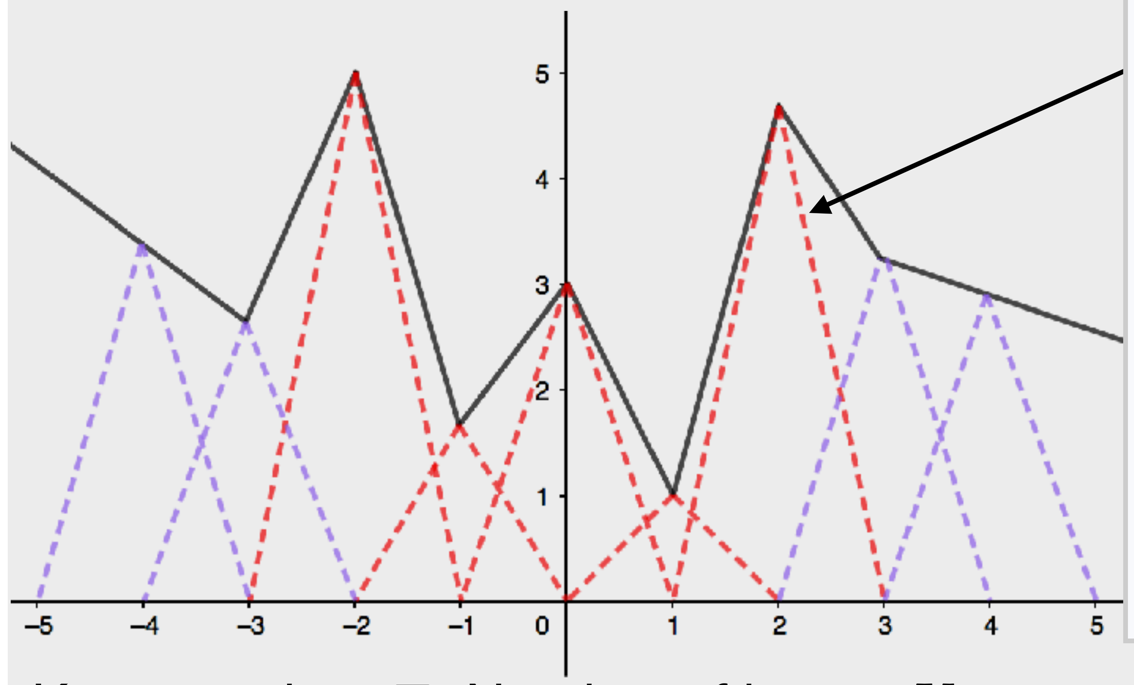
- 1-Lipschitz neural networks with learnable linear spline activation functions improve provably convergent PnP image reconstruction methods

1-Lipschitz Deep Spline CNNs

$$\mathrm{CNN}(\mathbf{x}) = \mathbf{C}_Q \circ \cdots \circ \boldsymbol{\sigma}_q \circ \mathbf{C}_q \circ \cdots \circ \boldsymbol{\sigma}_1 \circ \mathbf{C}_1(\mathbf{x})$$

Learnable linear spline nonlinearities (pointwise) [1]

Convolutional layer



Linear B-spline basis functions
- ✔ Compact support
- ✔ Efficient forward & backward pass
- ✔ Easy to compute Lipschitz constant (max. absolute derivative)

Knot spacing: T, Number of knots: K

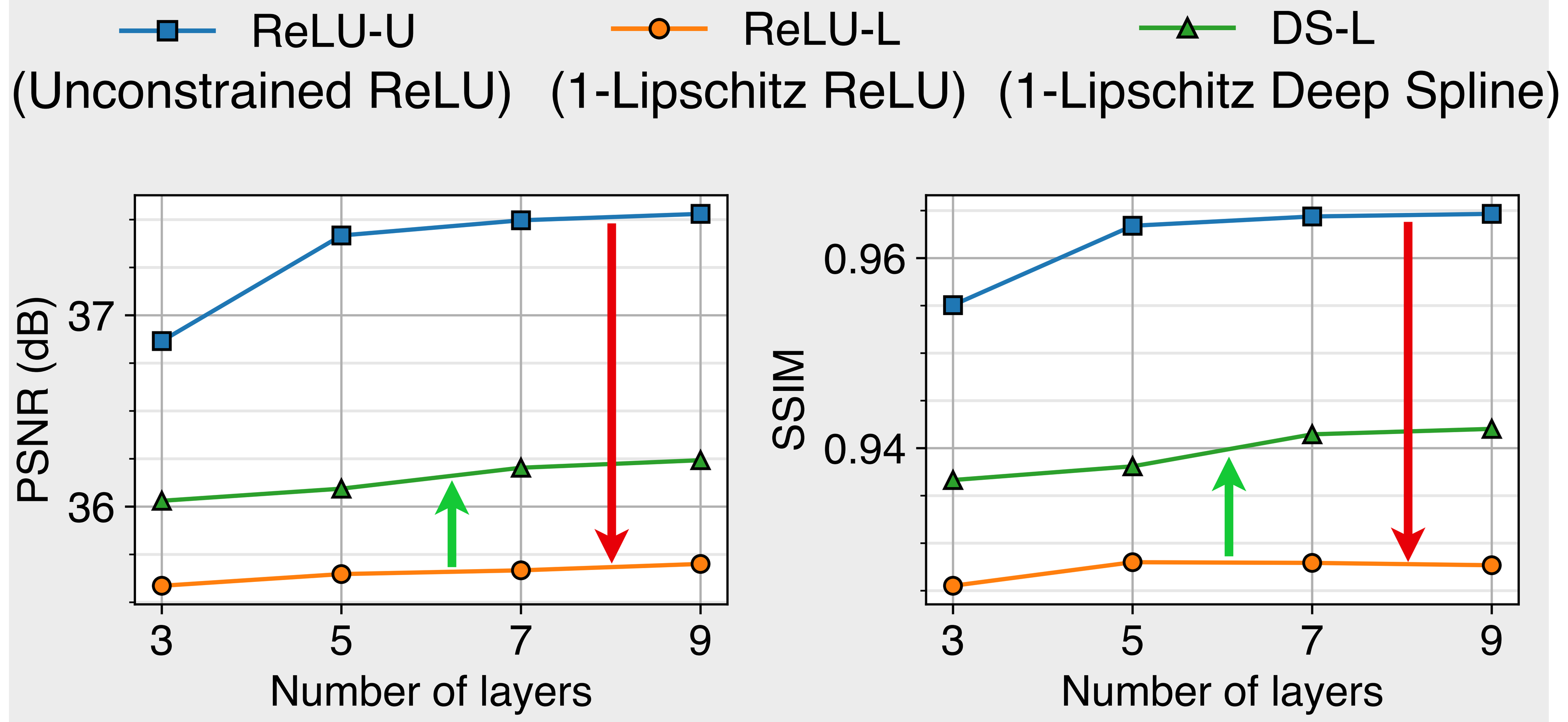
- Constrain the Lipschitz constant of each layer to be no greater than one
  - Convolutional layer: Spectral normalization [2]
  - Linear spline layer: Slope normalization

Experiments

- **Gaussian denoising**
  - ~240k 40x40 patches from BSD500 dataset
  - $\sigma = 5/255$ (standard dev. of Gaussian noise)
  - 3x3 convolution kernels, 32 channels
  - $T = 0.1, K = 51$
  - Number of layers = 3, 5, 7, 9
- **Compressed sensing MRI**
  - 256x256 ground-truth images
  - Subsampling ratio = 0.3
  - $\sigma_{\mathbf{n}} = 10/255$ (standard dev. of Gaussian noise)
  - Number of layers = 5

Results

- **Gaussian denoising**



↓ Drop in performance for constrained ReLU nets

↑ DS-L performs better than ReLU-L even with fewer parameters

- **Compressed sensing MRI**

Subsampling mask	Random		Radial		Cartesian	
Image type	Brain	Bust	Brain	Bust	Brain	Bust
Zero-filling	12.72	11.49	12.15	9.51	10.99	9.15
ReLU-L	20.25	17.05	19.02	16.22	13.70	12.85
DS-L	**20.78**	**17.76**	**19.51**	**17.00**	**14.23**	**13.53**

✔ DS-L systematically outperforms ReLU-L

References

[1] M. Unser. A representer theorem for deep neural networks. JMLR 2019.
[2] E. Ryu, J. Liu, S. Wang, X. Chen, Z. Wang, and W. Yin. Plug-and-play methods provably converge with properly trained denoisers. ICML 2019.