Frequency Regularized Deep Convolutional Dictionary Learning

Nikola Janjušević[†], Amir Khalilian[†], Yao Wang[†]

[†]New York University, Tandon School of Engineering

I. Introduction

Convolutional Dictionary Learning:

- Find optimal and data-driven basis for sparse representation
- Convolutional structured dictionary

$$\min_{\{d^j\}, Z_i} \sum_{i=1}^{N} \frac{1}{2} \|y_i - \sum_{j=1}^{M} d^j * z_i^j\|_2^2 + \lambda \sum_{j=1}^{M} \|z_i^j\|_1$$

- **Solution via iterating between:**
  - Sparse Coding
  - Dictionary Update
- **Slow!**
  - Sparse Code every signal in dataset
  - Dictionary changes every step

- $d^j \to$ dictionary filter
- $z_i^j \to$ sparse code
- $y_i \to$ noisy image

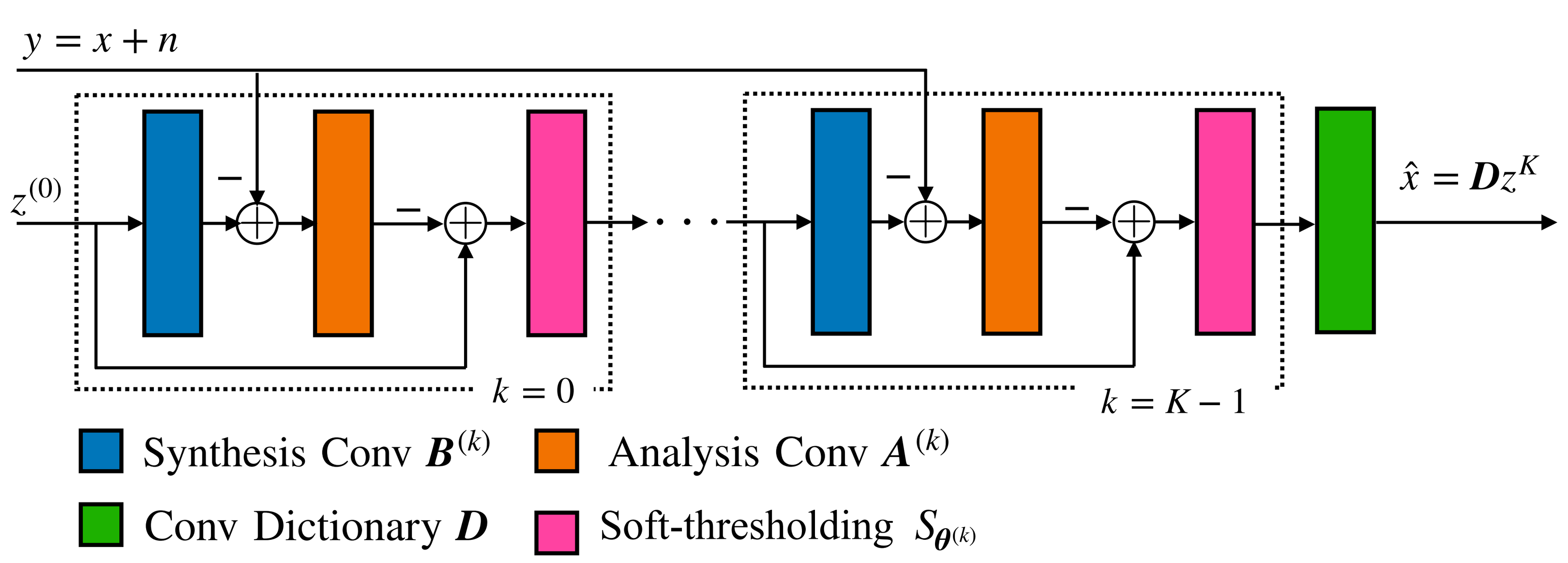
Use learnt convolutional sparse coding + dictionary learning to speed-up!

II. Convolutional Dictionary Learning Network (CDLNet)

- Approximate sparse coding: unrolled ISTA for K steps.

$$z^{(0)} = 0, z^{(k+1)} = S_{\theta^{(k)}} \left(z^{(k)} - A^{(k)} (B^{(k)} z^{(k)} - y) \right)$$

- Linear convolutional dictionary (decoder):
  - Use small-strided convolution.



- Training loss function:

$$\min_{\Theta} \|x - \hat{x}(y; \Theta)\|_2^2 \quad \text{s.t.:} \quad \theta^{(k)} \geq 0, \quad \|d^j\|_2^2 \leq 1.$$

- Network parameters: $\Theta = \{\{A^{(k)}, B^{(k)}, \theta^{(k)}\}_{k=0}^{K-1}, \{d^j\}_{j=1}^{M}\}$

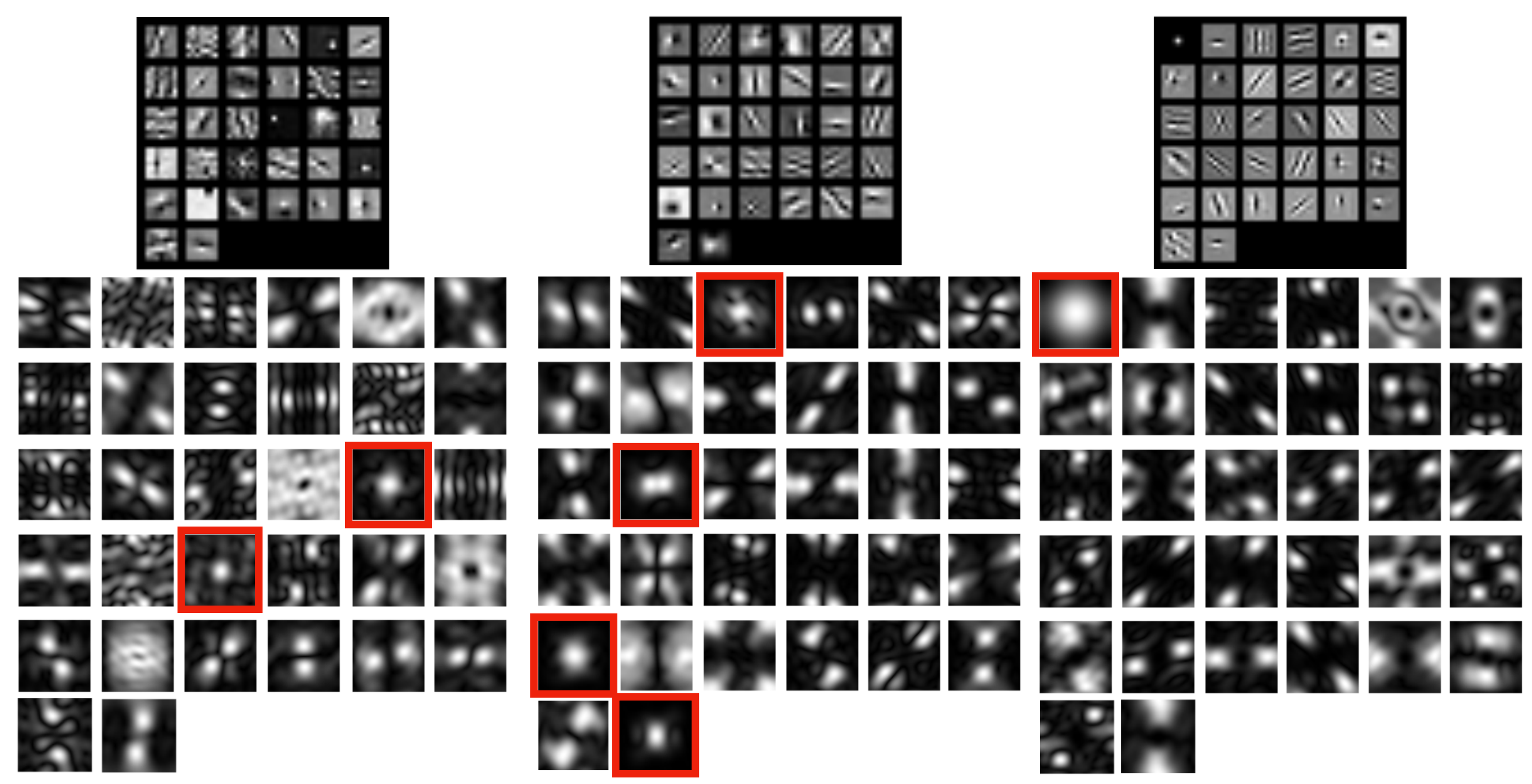
III. Frequency Regularization of CDLNet

- Mutual coherence of the dictionary:
  - Decreased mutual coherence → better representation.
  - Low-pass atoms increase the mutual coherence.
  - Natural images: piecewise smooth → require smooth filters.

- FCDLNet
  - Fix the low-pass channel.
  - Small stride: decrease mutual coherence while no need for shift processing.

$$d^1 = h, \quad d^j = g * \tilde{d}^j, \quad \|\tilde{d}^j\|_2 \leq 1, \ \forall j \in \{2, \cdots, M\}$$

- $h \to$ lowpass filter
- $g = \delta - h$



(a) Stride=1, No Freq. reg. $\mu(D) = 0.87$ (b) Stride=2, No Freq. reg. $\mu(D) = 0.77$ (c) Stride=2, with Freq. reg. $\mu(D) = 0.59$

Table 1: Denoising performance (PSNR) on BSD68 testset ($\sigma = \sigma_n^{\text{train}} = \sigma_n^{\text{test}}$).

σ	BM3D	FFDNet	DnCNN	CSCNet	FCDLNet	Big FCDLNet
15	31.07	31.63	**31.72**	31.40	31.45	31.66
25	28.57	29.19	**29.22**	28.93	28.99	**29.22**
50	25.62	26.29	26.23	26.04	26.11	**26.30**
Params	-	486k	556k	64k	66k	510k
CPU time (sec)	17.06	-	-	14.76	0.76	9.93
GPU time (sec)	-	-	-	0.34	0.03	0.14

- FCDLNet: $M = 64, K = 10$
- Big FCDLNet: $M = 169, K = 30$

IV. Blind Denoising: Noise-adaptive Thresholds

ISTA

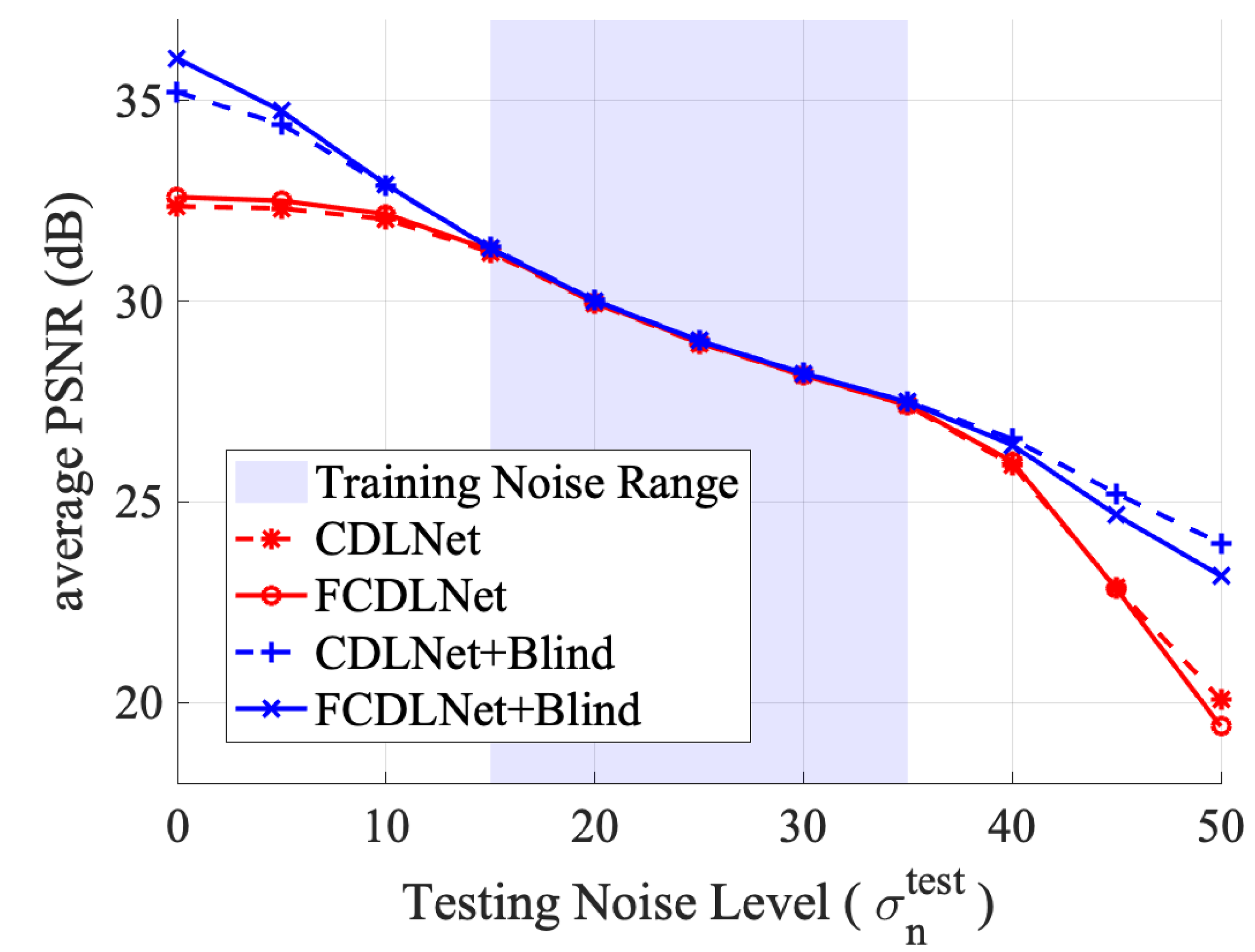
- Soft-thresholding:
  - $\theta \propto$ Expected Sparsity
  - $\theta \propto$ Noise Level

LISTA

- To get generalization across a noise range:

$$\theta^{(K)} = \nu^{(K)} \hat{\sigma}_n^2 \qquad \hat{\sigma}_n \approx \frac{\text{Median}(|c|)}{0.6745}$$

- $c \to$ the diagonal-detail Wavelet subband of an input image



(b) $\sigma_n^{\text{train}} \in [15, 35]$, $\sigma_n^{\text{test}} \in [0, 50]$

Table 2: Generalization of the network

	$\sigma_n^{\text{train}} = 20$			$\sigma_n^{\text{train}} = [18, 22]$	
σ_n^{test}	**ACSC**	**ConFirmNet**	**FCDLNet**	**FCDLNet**	**FCDLNet + Blind**
5	32.02 (0.02)	32.23 (0.01)	32.17 (0.01)	32.81 (0.01)	**34.25** (0.02)
15	31.88 (0.03)	32.04 (0.03)	32.06 (0.03)	32.30 (0.04)	**32.45** (0.03)
30	22.89 (0.03)	23.13 (0.04)	23.70 (0.05)	24.51 (0.05)	**25.31** (0.06)

VI. Conclusion

- Investigated unrolled convolutional sparse coding and dictionary learning frameworks.
- Proposed small-strided convolutional dictionary with a fixed low-pass channel, and a set of learned frequency regularized filters.
- Showed improved Denoising performance compared to other CDL methods.
- Showed better generalization using the parameterization of thresholds.