Deep Generative Networks for Image Processing

Akshat Dave, Anil Kumar Vadathya, Kaushik Mitra

Introduction

- Data priors are essential for many signal processing problems
- Recently deep learning has led to amazing improvements in computer vision tasks via discriminative training

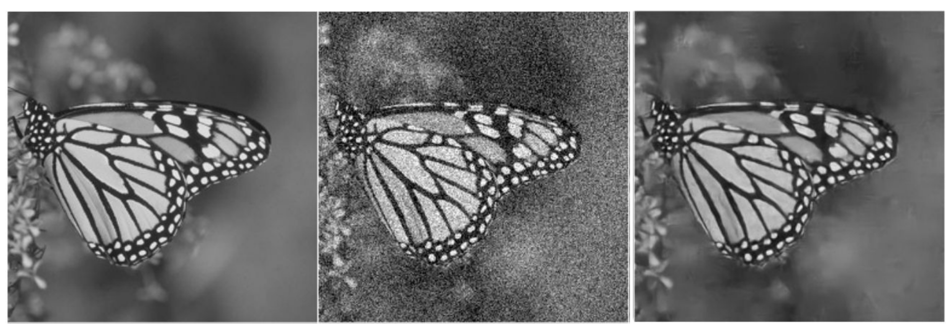
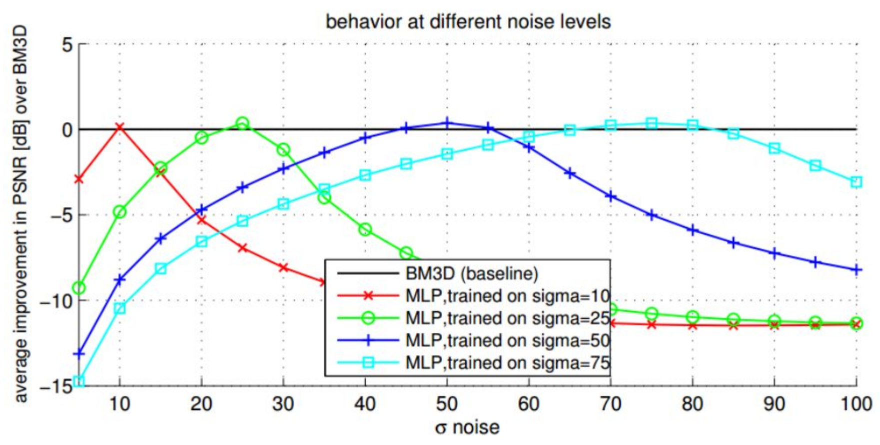


Image denoising by XJ Mao et al., 2016

Deconvolution using CNNs by Li Xu et al., 2014

- Discriminative vs Generative models
  - Discriminative training - Task specific networks
  - Generative model - same data prior for solving many related problems in image processing

MLP trained for denoising doesn't generalize for different noise levels.



Denoising using MLPs by HC Burger et al., 2012

- **Goal** - leverage deep architectures to model visual data priors facilitating inference in image processing problems

Prior Work: Deep generative models

- Traditional dictionary learning approach can't model long range dependencies in image data
- RIDE - image is modeled as causal 2d sequence with Spatial LSTMs
- Pixel RNN - extends the above approach with deep architectures

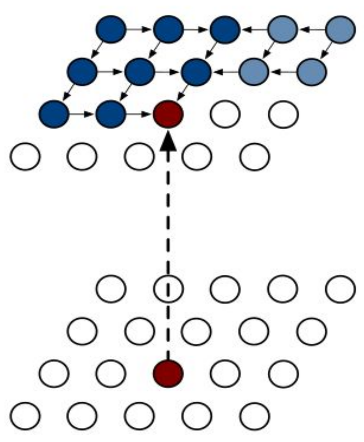
occluded completions original

In-painting using PixelRNN by Oord A.V.D. et al., 2016

Modeling spatial dependencies

$$p(\boldsymbol{x}) = p(x_1)p(x_2|x_1) \ldots p(x_i|\boldsymbol{x}_{<i})$$

- Each factor is now a sequence where current pixel distribution is modeled conditioned on past context represented by a Spatial LSTM



Spatial LSTMs by L Theis et al., 16

Modeling the prior:

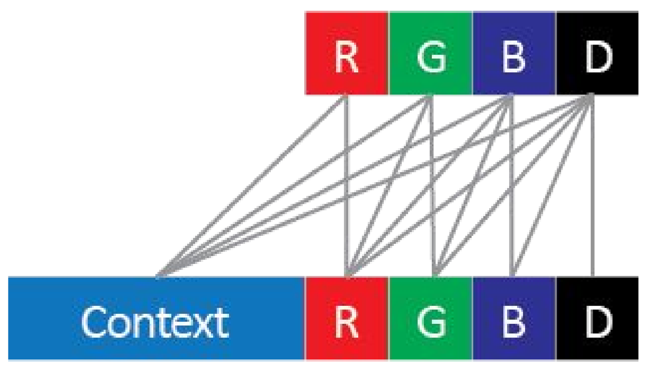
- Conditional GSM

$$p(\mathbf{y} \mid \mathbf{x}) = \sum_{cs} p(c, s \mid \mathbf{x}) p(\mathbf{y} \mid \mathbf{x}, c, s),$$

$$p(c, s \mid \mathbf{x}) \propto |\lambda_{cs}\mathbf{K}_c|^{\frac{1}{2}} \exp\left(-\frac{1}{2}\lambda_{cs}\mathbf{x}^\top\mathbf{K}_c\mathbf{x}\right)$$

$$p(\mathbf{y} \mid \mathbf{x}, c, s) = |\mathbf{M}_c|^{\frac{1}{2}} \exp\left(-\frac{1}{2}\lambda_{cs}(\mathbf{y} - \mathbf{A}_c\mathbf{x})^\top\mathbf{M}_c(\mathbf{y} - \mathbf{A}_c\mathbf{x})\right) / (2\pi)^{\frac{D}{2}},$$

- Use a deep recurrent network for representing context **x** while modeling pixel 'y'
  - Captures long range dependencies
- A simple extension to RGB and RGBD priors



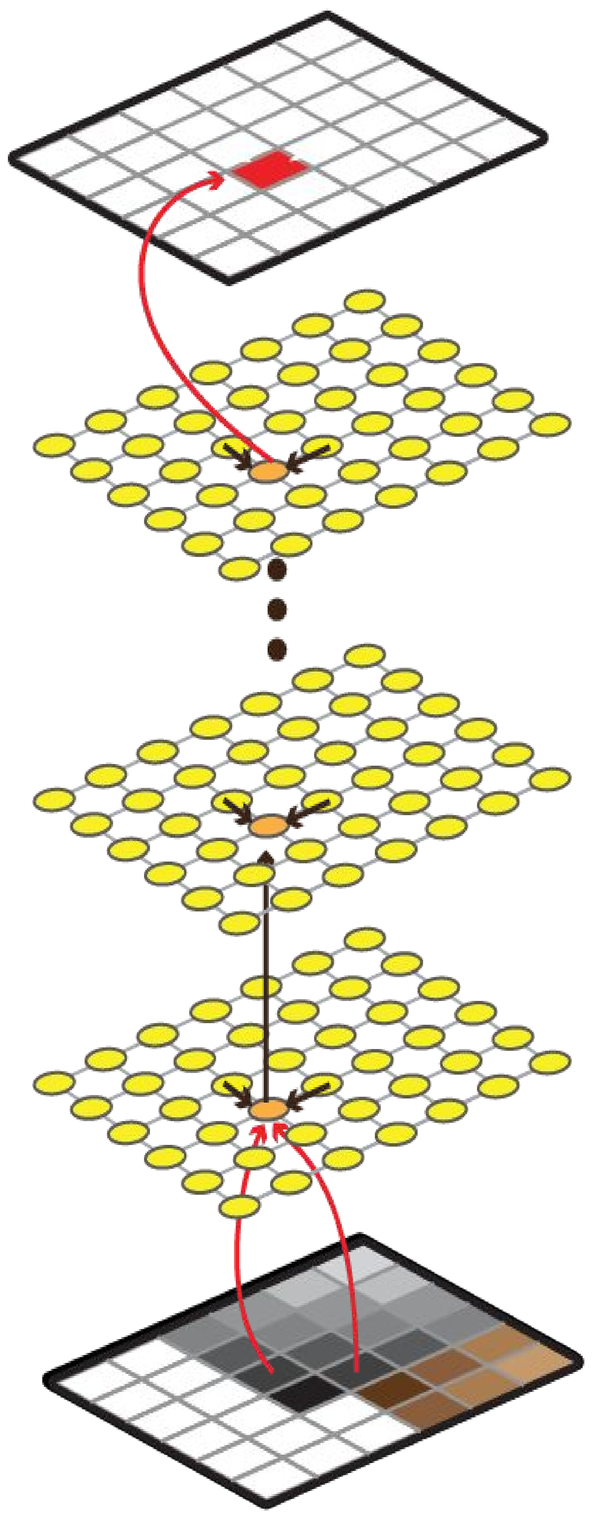


Image modeling using Stacked SLSTMs by L Theis et al., 2015

Inference

- Forward model of image degradation

$$Y = AX + n$$

- Inference can be done using Maximum-a-Posteriori (MAP) Principle

$$\arg\max_X \; p(Y|X)p(X)$$

Applications

- Image Processing
  - Denoising, Inpainting, Super-resolution
- Compressive Sensing Systems
- Computational Photography

Compressive Sensing Camera, RICE

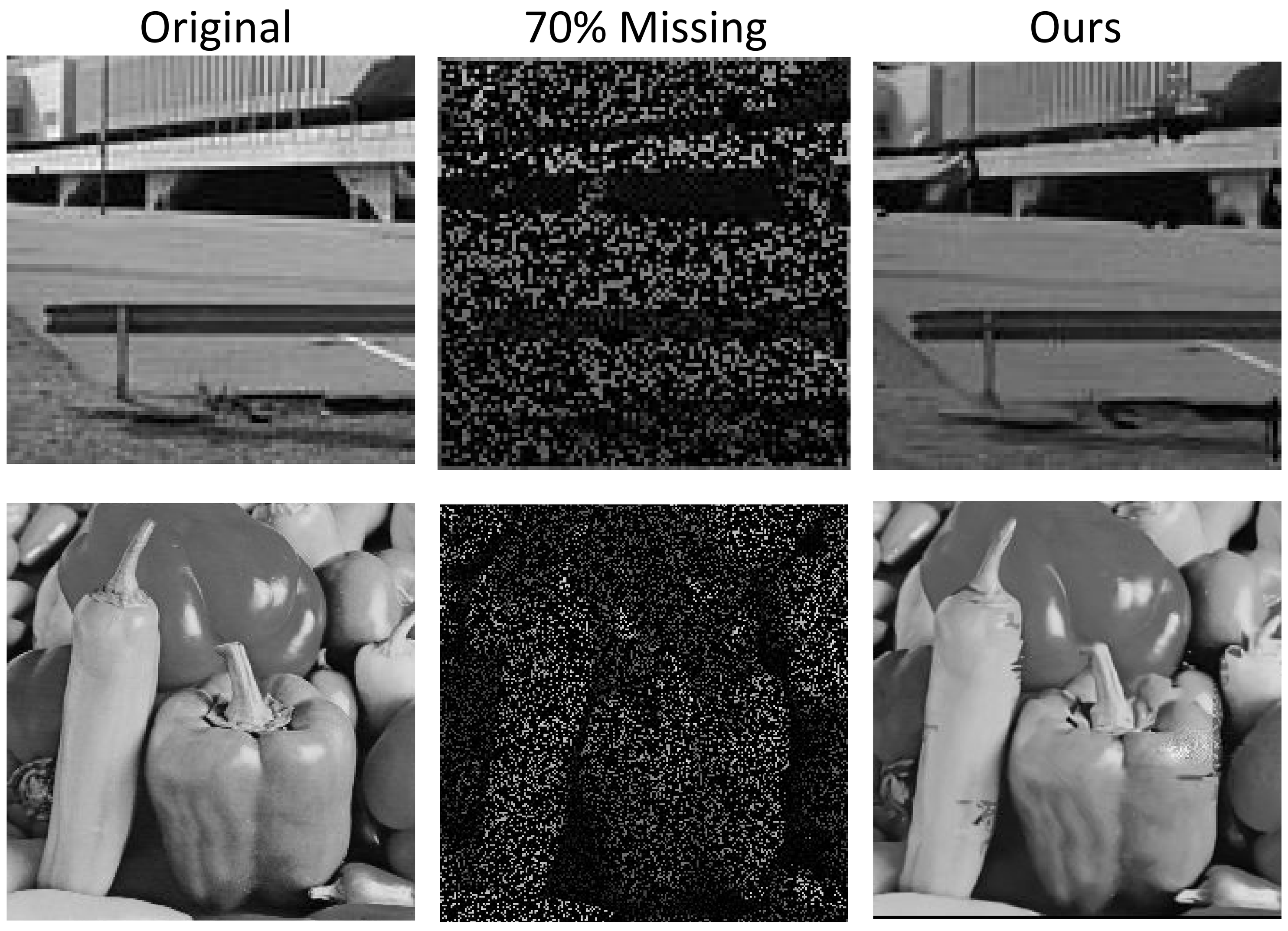
Conclusion & Future Work

- Signal priors are essential for solving ill posed problem
- We want to leverage the deep models for modeling such powerful priors
- Adapting the priors to the specific image contents
- Making priors amenable for efficient inference

Results

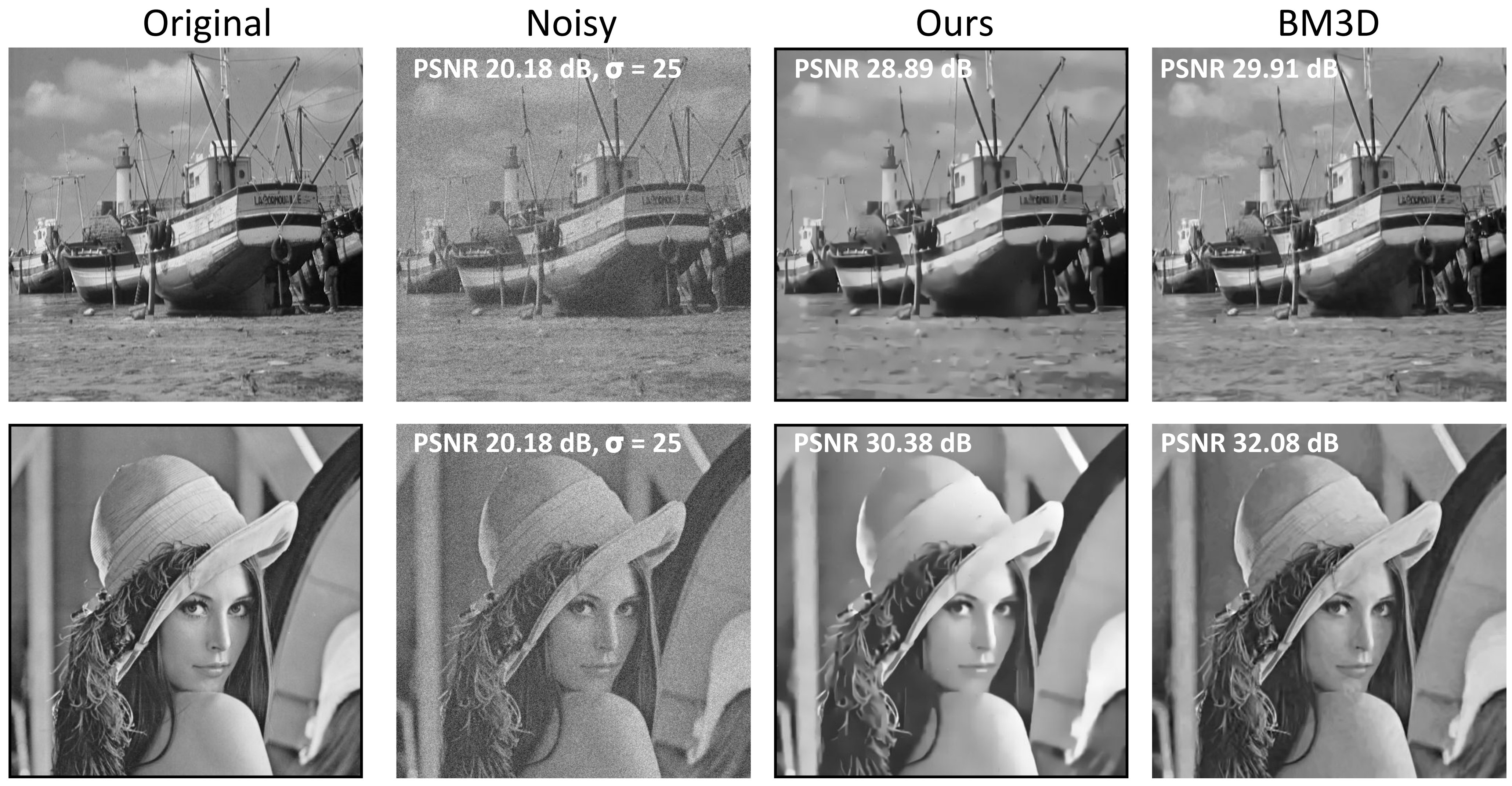
Inpainting

Original | 70% Missing | Ours



Denoising

Original | Noisy | Ours | BM3D



iitmadras

Computational Imaging Lab
Department of Electrical Engineering
IIT Madras

QUALCOMM