Uncertainty Quantification for Inverse Problems with Generative Priors under Distribution Shift

Namhoon Kim and Sara Fridovich-Keil

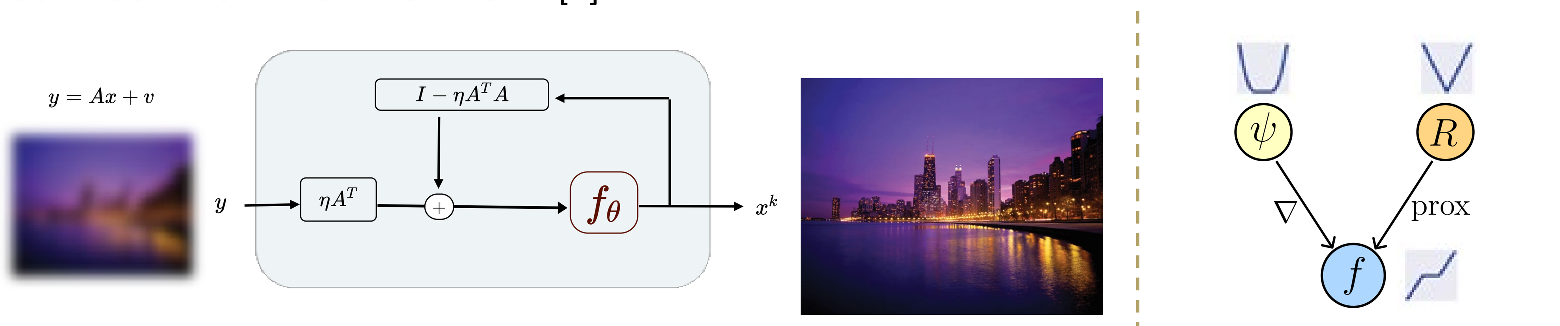
School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia

Abstract

- Generative models for inverse problems improve reconstruction quality and reduce required measurements but may hallucinate when the target image lies outside their training distribution.
- Existing uncertainty quantification methods often rely on unavailable in-distribution calibration data, use heuristic rather than statistical estimates, or only address uncertainties from model complexity or limited data—ignoring uncertainty from distribution shifts.
- We highlight the need for instance-level uncertainty quantification in the presence of distribution shift and propose a strategy to provide it.
- Our hypothesis is that, with different limited sets of random measurements, reconstruction will be more stable for in distribution targets. Therefore, we propose reconstruction variation across different random measurements as a way to quantify distribution shift uncertainty.

Background

Learned Proximal Network [1]



$f_\theta(x) \triangleq \arg\min_z \frac{1}{2}\|z - x\|^2 + R_\theta(z)$ via *Learned Proximal Networks*

$R_\theta(x) \to -\log p_x(x)$ via *Proximal Matching Training*

The LPN learns an input-convex function ψ whose gradient f is guaranteed to be the proximal operator for some regularizer R.

A learned proximal network (LPN) replaces the traditional proximal operator with a neural network trained to fit the complex data distribution. Similar to a diffusion model, this method imposes rich priors on the solution and enables higher-quality reconstructions in inverse problems.

Computed Tomography

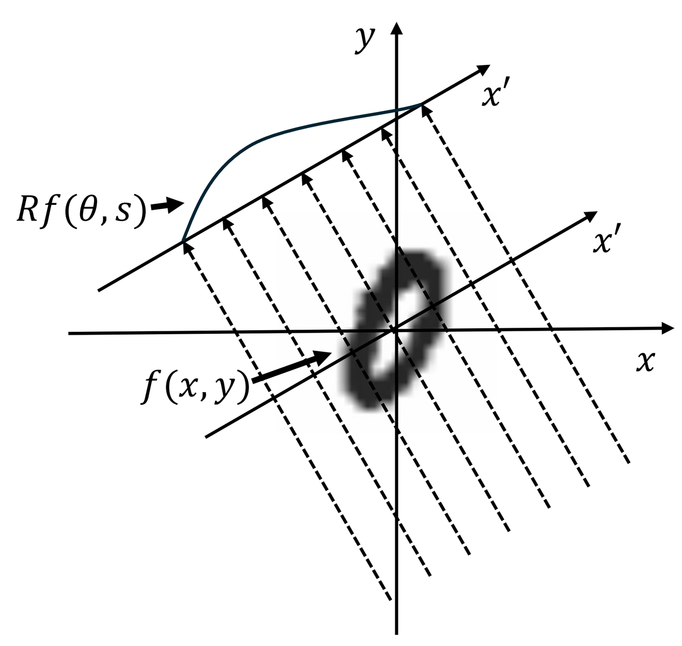
Radon Transform

$$Rf(\theta, s) = \int_{-\infty}^{\infty} f(x', y')dl$$

where, $x'cos\theta + y'sin\theta = s$

CT Forward Model

$$p(\theta, s) = Rf(\theta, s)$$



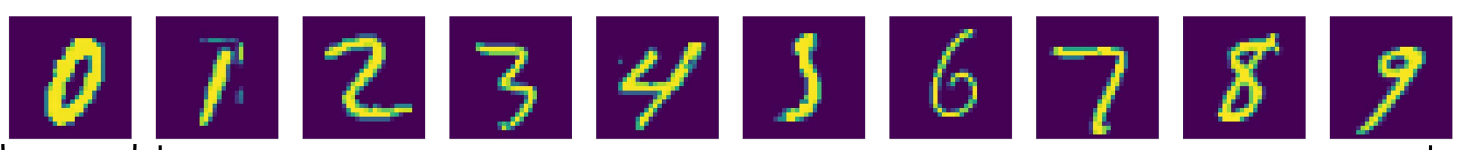
CT uses a rotating, narrow X-ray beam and computer processing to create detailed cross-sectional "slice" images of the body, offering more information than standard X-rays. The amount of information in a CT scan depends on the detector resolution and the number of angles used. Higher resolution captures finer spatial details in each projection, and more angles provide a wider variety of views of the object. When the total number of measurements (detector resolution × number of angles) is less than the number of unknowns (image height × width), the reconstruction problem is underdetermined, and we rely on information encoded in the prior.

Randomness

Existing methods for quantifying reconstruction uncertainty without calibration data often leverage randomness over models, by training multiple copies of a learned prior with different random seeds and then treating these as an ensemble. However, this process (1) is computationally expensive due to retraining, and (2) accounts for uncertainty due to limited training data and overparameterization but doesn't directly quantify distribution shift uncertainty. Instead, we propose to quantify instance-level reconstruction uncertainty by measuring sensitivity to the random measurements used at inference time, to detect distribution shift relative to a learned prior, without retraining.

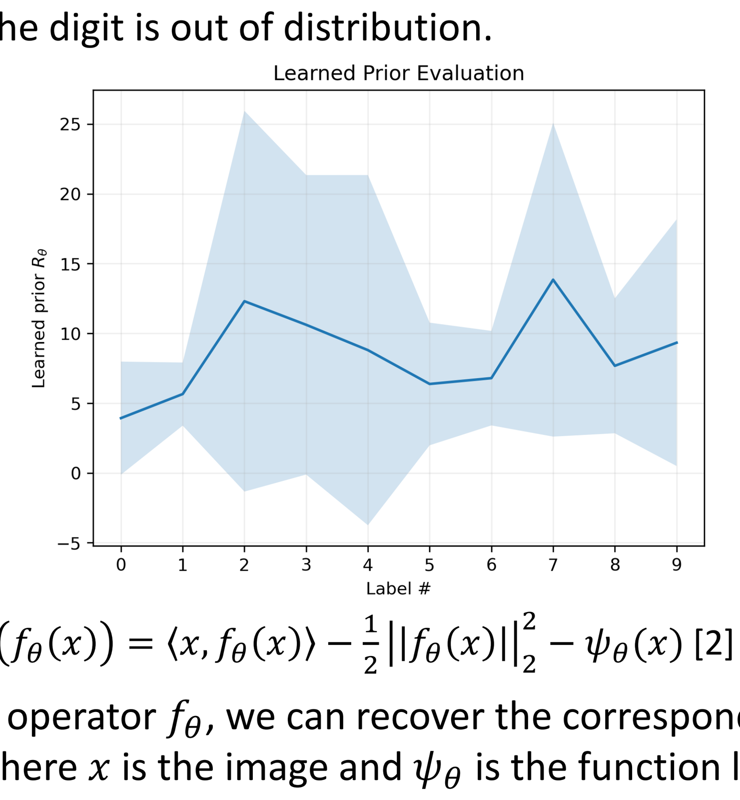
Experimental Setup

Experiment Setting



In Distribution — Out of Distribution

For the MNIST Dataset, we use label 0 as our in-distribution data and labels 1-9 as our out-of-distribution data. We trained the LPN with only 0 and evaluated with all labels to see if the LPN distinguishes if the digit is out of distribution.



$$R_\theta(f_\theta(x)) = \langle x, f_\theta(x)\rangle - \frac{1}{2}\|f_\theta(x)\|_2^2 - \psi_\theta(x) \text{ [2]}$$

For the learned proximal operator f_θ, we can recover the corresponding learned prior using this equation from [2], where x is the image and ψ_θ is the function learned by the prox network, with the prox operator f_θ as its gradient. The plot shows mean and +/- standard deviation of the resulting priors over 100 images per label. We see that the trained prior does recognize the digit 0 as in distribution by giving it the lowest regularization score on average, while the more visually different digits usually receive higher regularization scores under the learned prior.

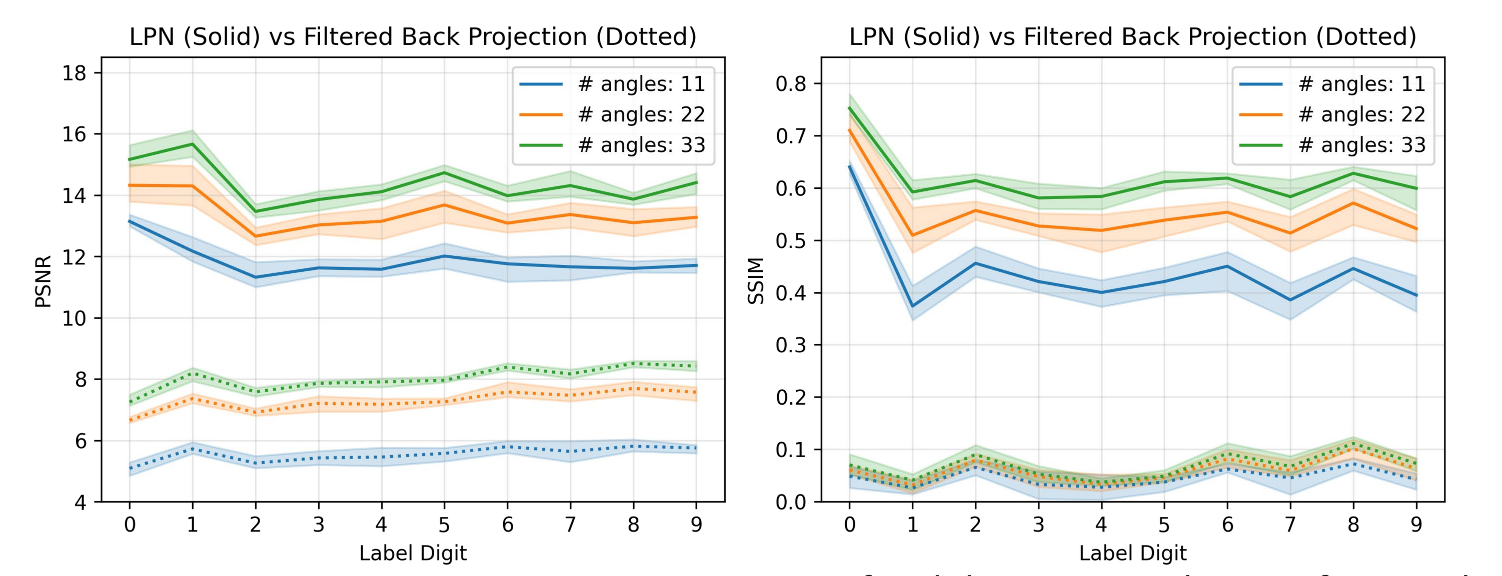
References

[1] Fang, Zhenghan, Sam Buchanan, and Jeremias Sulam. "What's in a Prior? Learned Proximal Networks for Inverse Problems." *Proceedings of the Twelfth International Conference on Learning Representations (ICLR)*, 2024

[2] Gribonval, R., Nikolova, M. A "Characterization of Proximity Operators." *J Math Imaging Vis* 62, 773–789 (2020).

Results

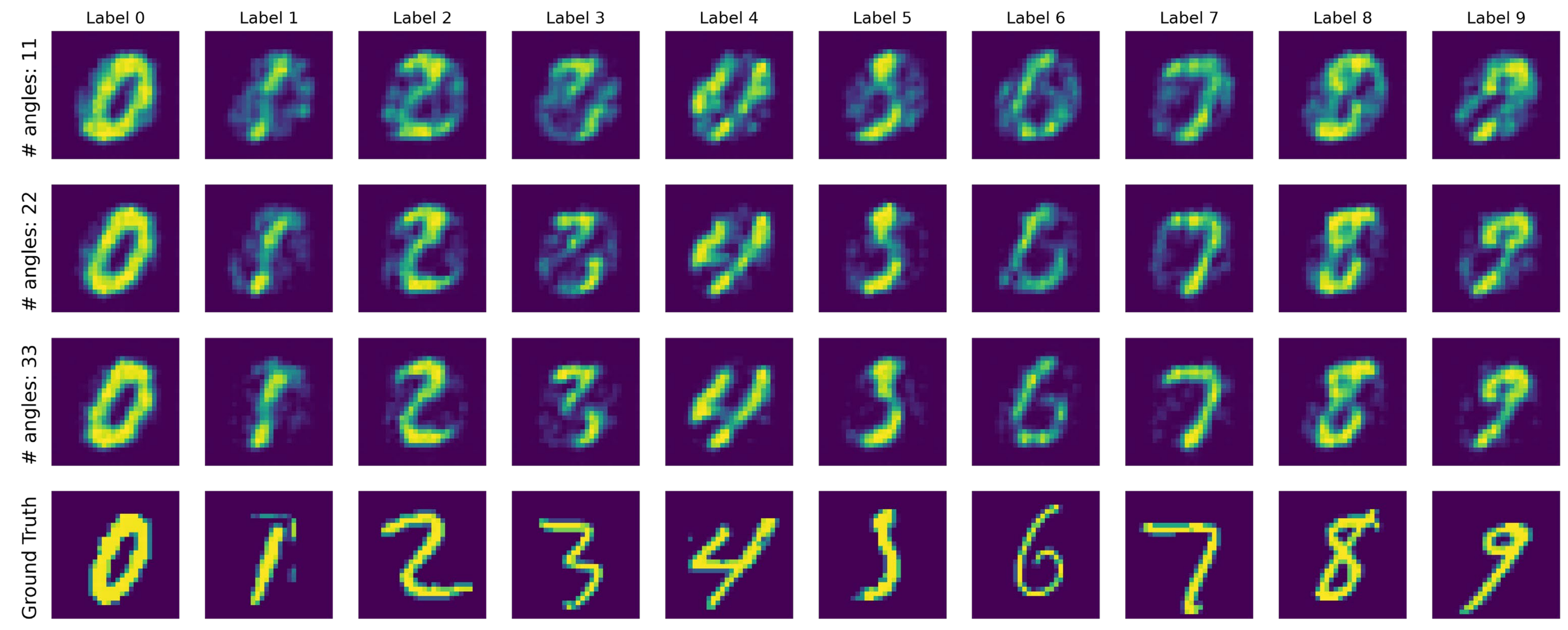
CT Reconstruction Results



In our CT reconstruction experiments, we maintain a fixed detector resolution of 22 pixels and vary the number of measurement angles. We repeat the procedure across 10 samples of each digit, and 10 random seeds to choose the random measurement angles for each CT scan. We average the reconstruction PSNRs across the 10 sample images of each digit, and plot the mean and range (min-max) of the resulting average PSNRs across the 10 random seeds for each digit. The results indicate that the learned prior is beneficial even for out of distribution digits, but much more effective for in-distribution digits, and especially so when the number of measurement angles is reduced.

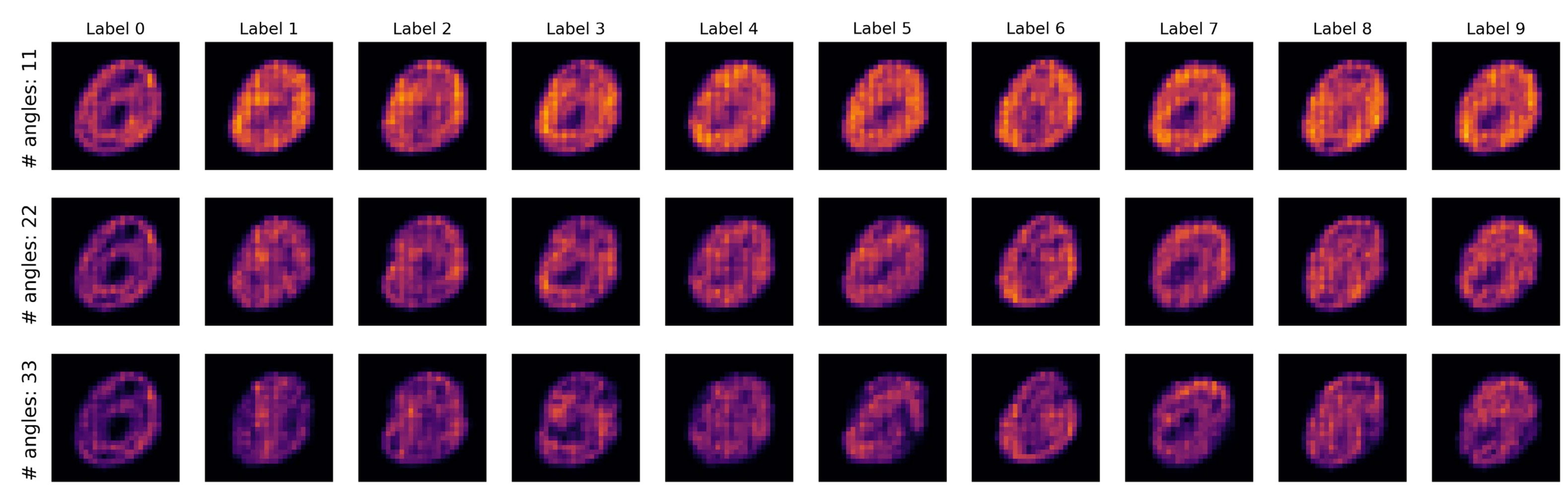
Mean & Standard Deviation Comparison

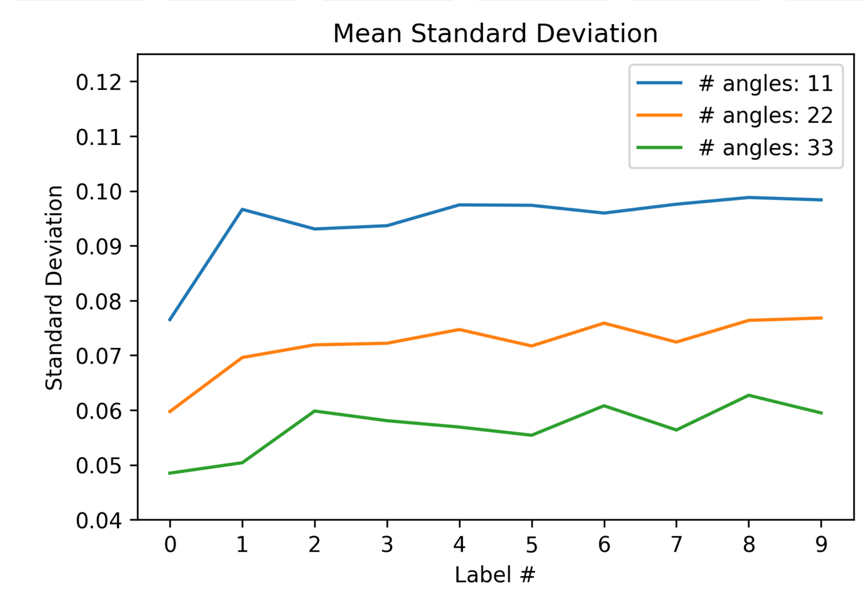
- **Mean**



Mean reconstructions over 10 different seeds are shown for one sample of each digit. As the number of angles increases, the quality of the reconstructed images also improves. However, the reconstruction quality of the in-distribution data is comparably better than the out-of-distribution data, especially when the number of measurement angles is smallest and thus the learned prior plays a larger role.

- **Standard Deviation**





If the target image is in the distribution learned by the prior network, we expect it to produce consistent predictions even as the set of random measurement angles changes. In contrast, if the target image is out of distribution for the prior, we expect higher variance of the reconstructions when the random measurement angles change. This is exactly what we find: **reconstruction variance over random measurements detects distribution shift**. The effect is most prominent when the number of measurement angles is small, which aligns with the setting when the learned prior has the most influence on the reconstruction and thus distribution shift poses the greatest risk.

Conclusion

Generative models have shown great promise as data-driven priors in solving inverse problems like CT reconstruction, enhancing image quality and reducing measurements. However, data-driven priors pose risks of hallucination under distribution shift, when the target image differs from the distribution used to train the prior. Here we validate the simple hypothesis that this distribution shift fragility can be detected without extensive computational or data-collection burden, by evaluating how consistent the reconstruction is across random subsets of the available measurements. Though preliminary, our work suggests a simple strategy to detect and mitigate distribution shift by collecting additional measurements until reconstruction stability crosses a desired threshold.