

Physics-Guided Conditional Flow Matching for Lensless Imaging

Qiwen Xiao Edison Liu

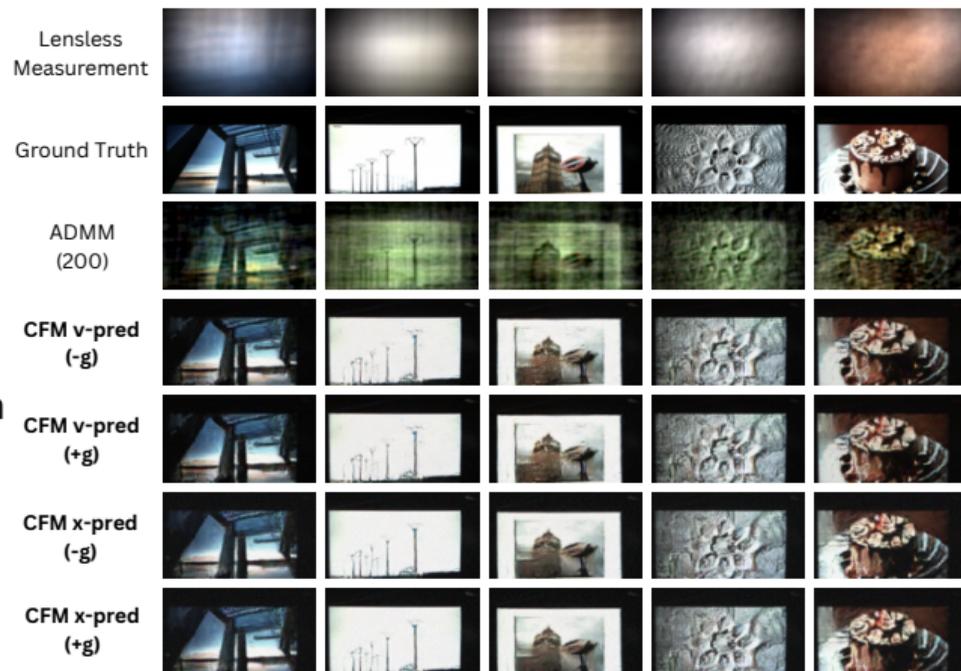
Department of Electrical Engineering and Computer Science (EECS)
University of California, Irvine

March 17, 2026

Code: <https://github.com/Charley-xiao/lensless-flow>

Problem: diffuser-based lensless reconstruction

- Measurement y is a caustic pattern; image x is latent.
- Calibrated forward model:
 $y \approx H_\phi(x)$.
- Ill-posed + mismatch \Rightarrow structured artifacts.
- Goal: reconstruct high-fidelity x from one y .



(Examples from our report)

- **ADMM:** interpretable; often artifact-prone.
- **U-Net:** fast supervised regression; can oversmooth textures.

Mindset shift

Instead of solving for one \hat{x} , we learn the conditional distribution $p_{\theta}(x | y)$ and sample reconstructions from it.

Approach: Conditional Flow Matching (CFM) as a continuous-time conditional generative model.

Flow matching

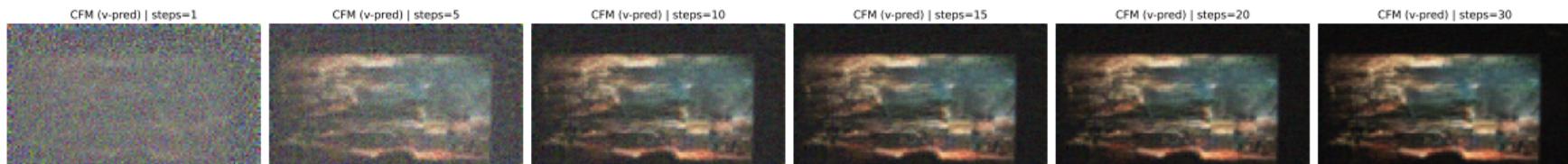
- Learn an ODE that transports **noise** \rightarrow **image** *conditioned on* y .
- Training is **supervised regression** on a chosen path (no backprop through ODE solvers).

Linear path and target velocity

Sample $x_1 \sim \mathcal{N}(0, \sigma_0^2 I)$ and $t \sim \mathcal{U}[t_{\min}, t_{\max}]$:

$$x_t = (1 - t)x_1 + tx, \quad v^* = \frac{d}{dt}x_t = x - x_1.$$

Train $v_\theta(x_t, y, t)$ by minimizing $\mathbb{E}\|v_\theta - v^*\|_2^2$. [TFM⁺23, LCBH⁺22]



Coarse structure appears early; refinements saturate with more steps.

Two model parameterizations

v-prediction

Predict velocity directly:

$$v_{\theta} = f_{\theta}(x_t, y, t).$$

- Simple and effective.

x-prediction (best)

Predict clean image then induce velocity:

$$\hat{x}_{\theta} = f_{\theta}(x_t, y, t), \quad v_{\theta} = \frac{\hat{x}_{\theta} - x_t}{\max(1 - t, \delta)}.$$

- Slightly cleaner details in practice.

- Sample $z_0 \sim \mathcal{N}(0, \sigma_0^2 I)$, integrate $\frac{dz}{dt} = v_\theta(z, y, t)$ from $t \in [0, 1]$.
- Solver: Heun (RK2) for a good accuracy/compute trade-off.

Optional data-consistency (DC) refinement

With $\ell_{\text{dc}}(z; y) = \|H_\phi(z) - y\|_2^2$, apply after each step:

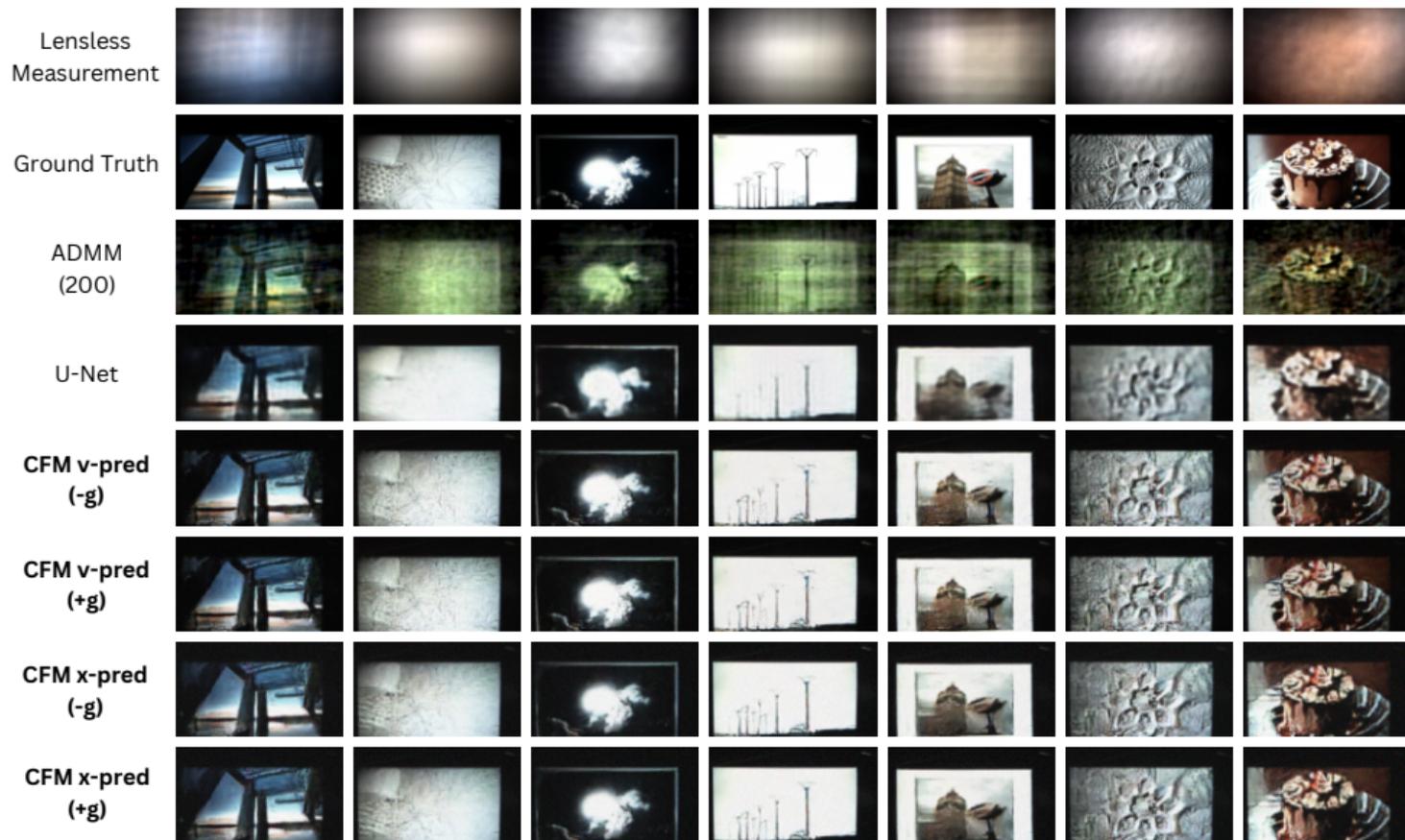
$$z \leftarrow z - 2\eta H_\phi^\top (H_\phi(z) - y).$$

- Dataset: DiffuserCam MirFlickr, downsample $\times 4$, normalized to $[0, 1]$.
- Baselines: ADMM, supervised U-Net.
- Our models: CFM v -pred / x -pred, each with and without DC guidance.
- Metrics: PSNR / SSIM / MSE on test set.

Method	SSIM \uparrow	PSNR \uparrow	MSE \downarrow
ADMM	0.15	7.26	0.1879
U-Net	0.76	20.67	0.0086
CFM v-pred (-g)	0.79	22.26	0.0062
CFM v-pred (+g)	0.80	22.27	0.0062
CFM x-pred (-g)	0.81	23.00	0.0053
CFM x-pred (+g)	0.81	23.01	0.0051

- CFM improves over U-Net and far outperforms ADMM.
- **x-pred** is consistently best.
- Guidance helps modestly (largest effect in MSE).

Qualitative comparison



- CFM provides a high-fidelity conditional generative approach for lensless reconstruction.
- x-prediction yields better fidelity than v-prediction in our study.
- Physics guidance is an *optional refinement* with modest gains under conservative settings.
- Next: robustness to PSF mismatch; higher resolution.

-  Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matt Levine, *Flow matching for generative modeling*, arXiv preprint arXiv:2210.02747 (2022).
-  Alexander Tong, Kilian Fatras, Nikolay Malkin, Guillaume Huguet, Yanlei Zhang, Jarrid Rector-Brooks, Guy Wolf, and Yoshua Bengio, *Improving and generalizing flow-based generative models with minibatch optimal transport*, arXiv preprint arXiv:2302.00482 (2023).