

Noise Frequencies Let Pixel-Space Diffusion Priors Detect Native-Resolution Images

Bumsoo Kim^{1†} · Hyun-Jun Jin^{2†} · Mi Young Lee¹ · Seungmin Rho¹ · Sanghyun Seo^{1★}

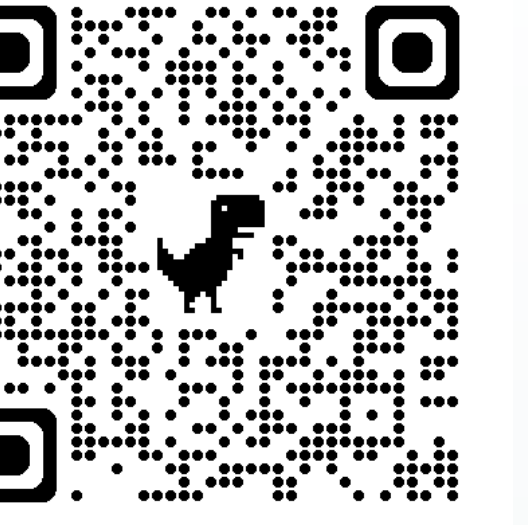
¹Chung-Ang University, South Korea. ²Korea University, South Korea.

† Now with Smilegate ★ Corresponding author (sanghyun@cau.ac.kr)



KOREA
UNIVERSITY

PixRes



Abstract

PixRes detects whether an image is native at its stated resolution — training-free, from the noise a pre-trained pixel-space diffusion prior predicts along PF-ODE inversion. 0.953 vs 0.918 AUC across eight datasets, with the largest margins where resampling is subtlest.

Motivation and Introduction

High pixel count $\neq$ high information content.

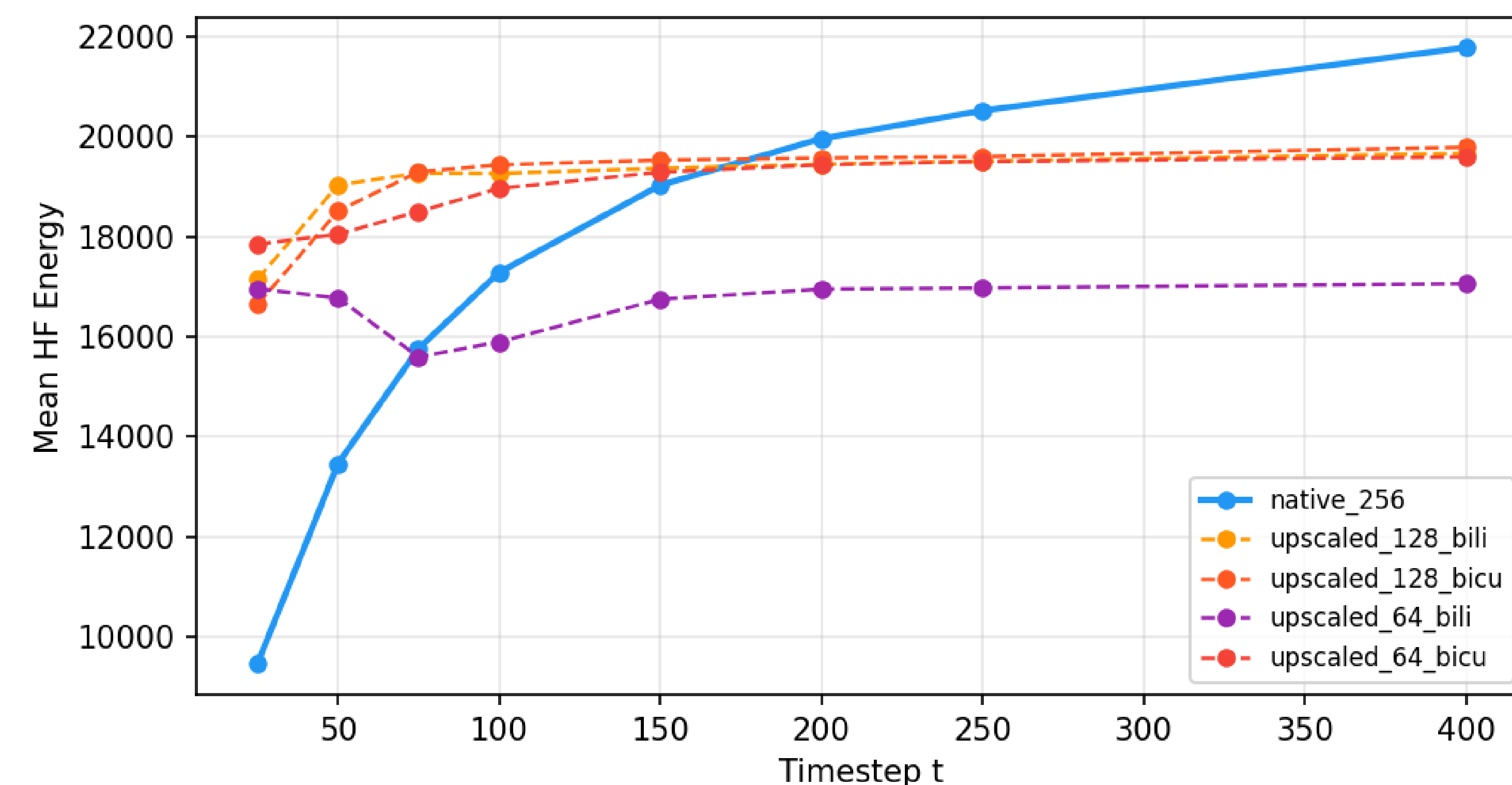
- IQA scores **perceptual degradation** — blur, noise, compression. Nateness is intrinsic, and IQA cannot reach it.
- Learned estimators (SS-ERE) need **domain-specific training** on a fixed degradation pipeline — and retraining for every new domain.
- We instead ask a **pre-trained generative prior** whether it agrees with the image. No training, no reconstruction.

Contributions:

- Generative priors work as **zero-shot low-level image assessors**.
- One frequency-aware scalar, read off **PF-ODE inversion in pixel space**.
- Extends from binary detection to **continuous effective-resolution estimation**.

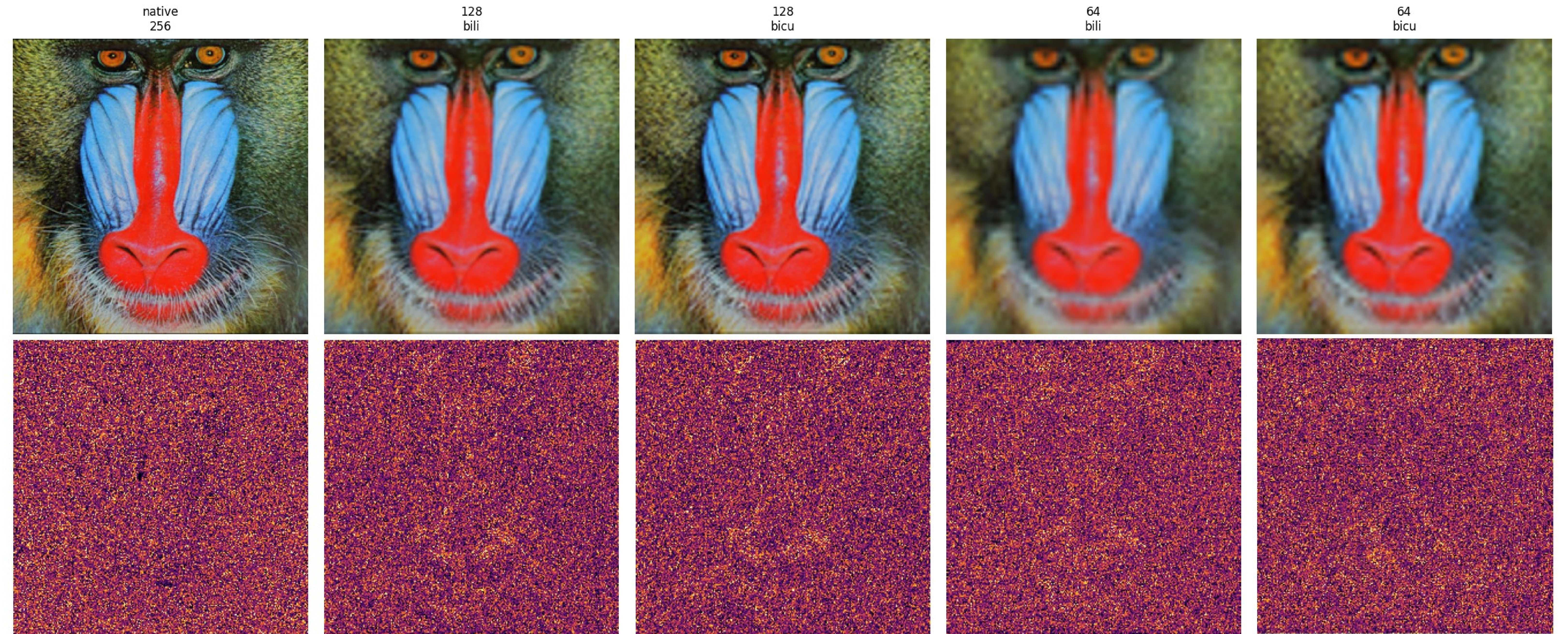
Pre-analysis

- Pixel-space diffusion keeps the high-frequency detail a latent model discards.
- Discriminability peaks on a mid-noise plateau — weak at low noise (content-dominated) and at high noise (collapses to white).



High resolution does not always imply high quality.

All five images below share the same width and height, yet the information they actually carry is not the same. But how do we detect it?

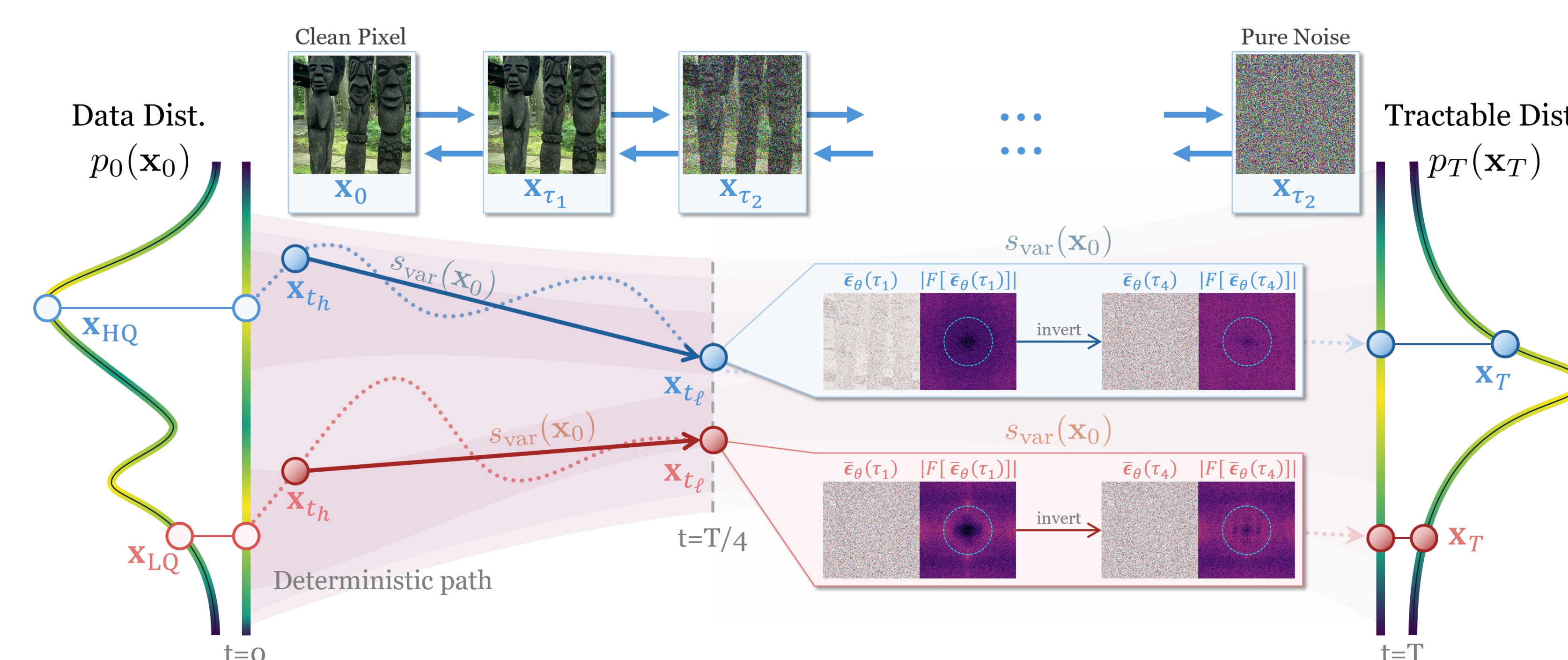


Top: inputs resampled from different resolutions.

Bottom: noise predicted along PF-ODE inversion.

Method

Native detection with no re-training



- **Invert** — deterministic DDIM in pixel space, $T/4$ steps.
- **Read** — radial power spectrum of noise, high-frequency band only.
- **Score** — dispersion of that spectrum across timesteps.
- **Decide** — native if score < threshold value (fixed on a small held-out set).

Results

Dataset	Method	vs 96	vs 128	vs 160	vs 192	vs 224
FFHQ	SS-ERE	1.000	0.999	0.980	0.920	0.814
	Ours	1.000	1.000	1.000	0.996	0.989
BSD100	SS-ERE	0.992	0.982	0.937	0.840	0.719
	Ours	0.998	0.990	0.996	0.952	0.807
Urban100	SS-ERE	0.996	0.978	0.899	0.753	0.637
	Ours	0.990	0.983	0.989	0.964	0.695
DIV2K	SS-ERE	0.992	0.985	0.953	0.870	0.745
	Ours	1.000	0.992	0.996	0.962	0.771
COCO	SS-ERE	0.983	0.978	0.945	0.857	0.677
	Ours	1.000	0.999	1.000	0.995	0.858
WikiArt	SS-ERE	0.989	0.975	0.939	0.860	0.750
	Ours	0.984	0.965	0.985	0.959	0.816

Detection AUC. Each column is a downscale anchor.

Conclusion

A generative prior already knows enough low-level statistics to judge nativeness — no labels, no degradation pipeline, no per-domain tuning. **PixRes** reads that evidence off a single deterministic inversion pass.