Intrinsic LoRA: A Generalist Approach for Discovering Knowledge in Generative Models
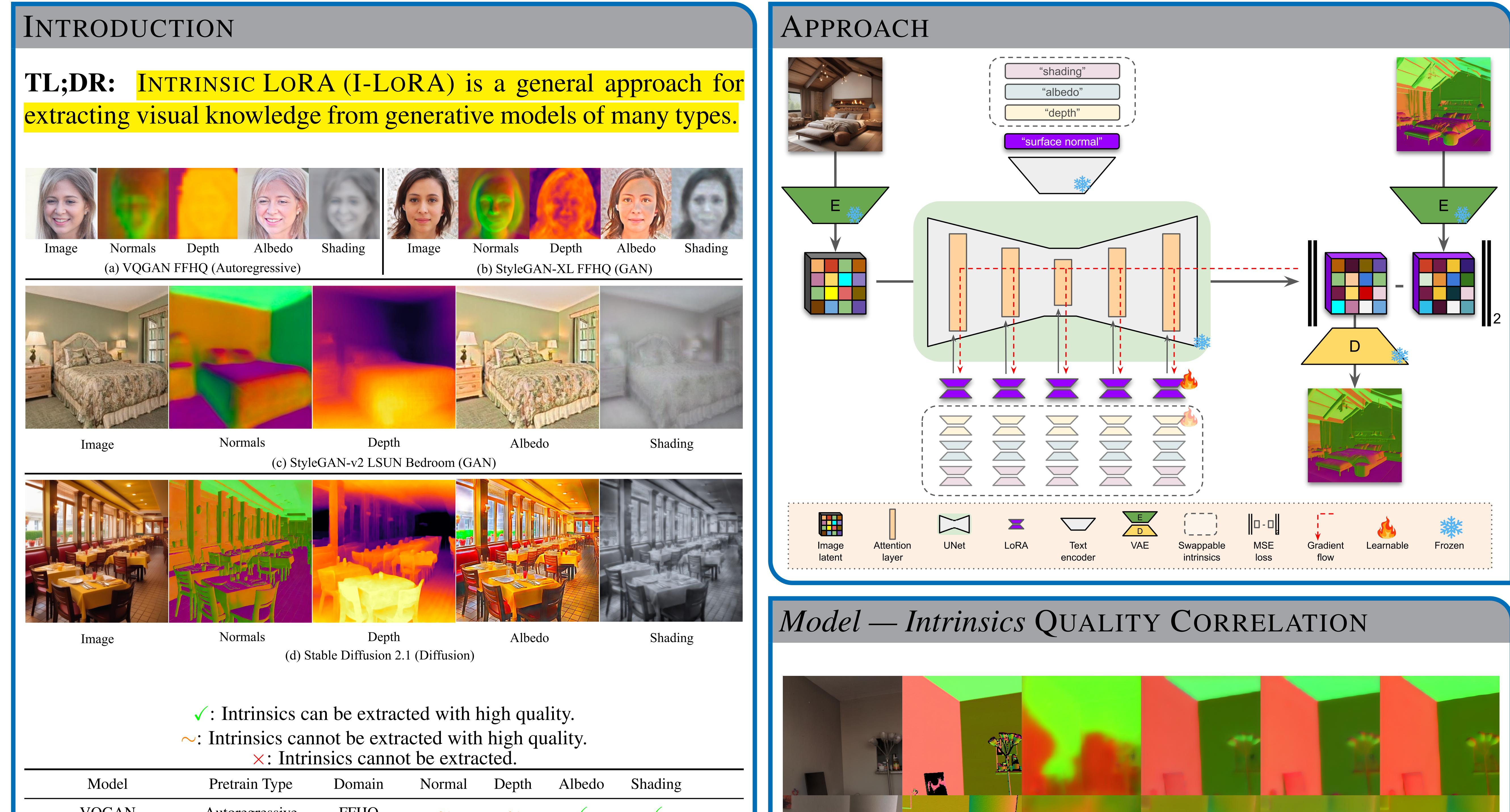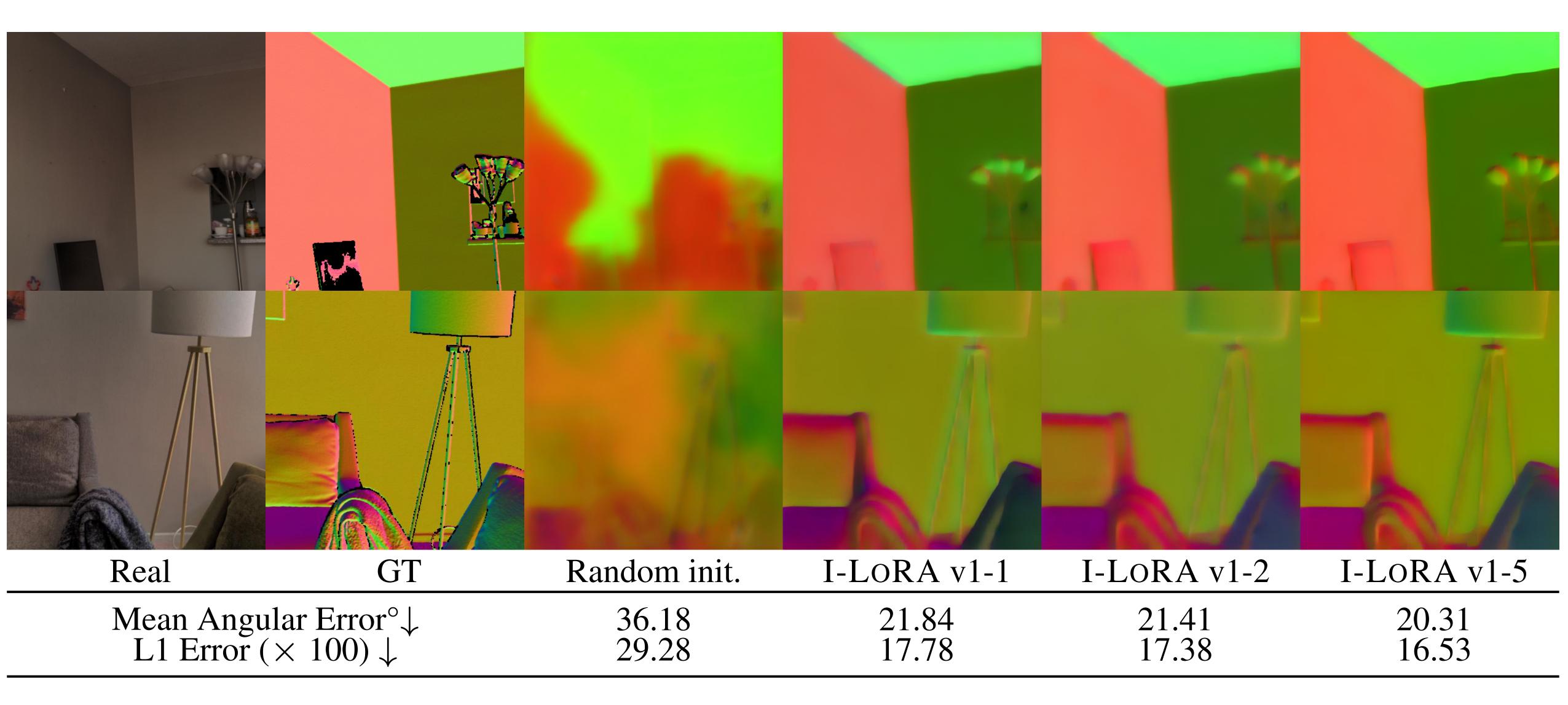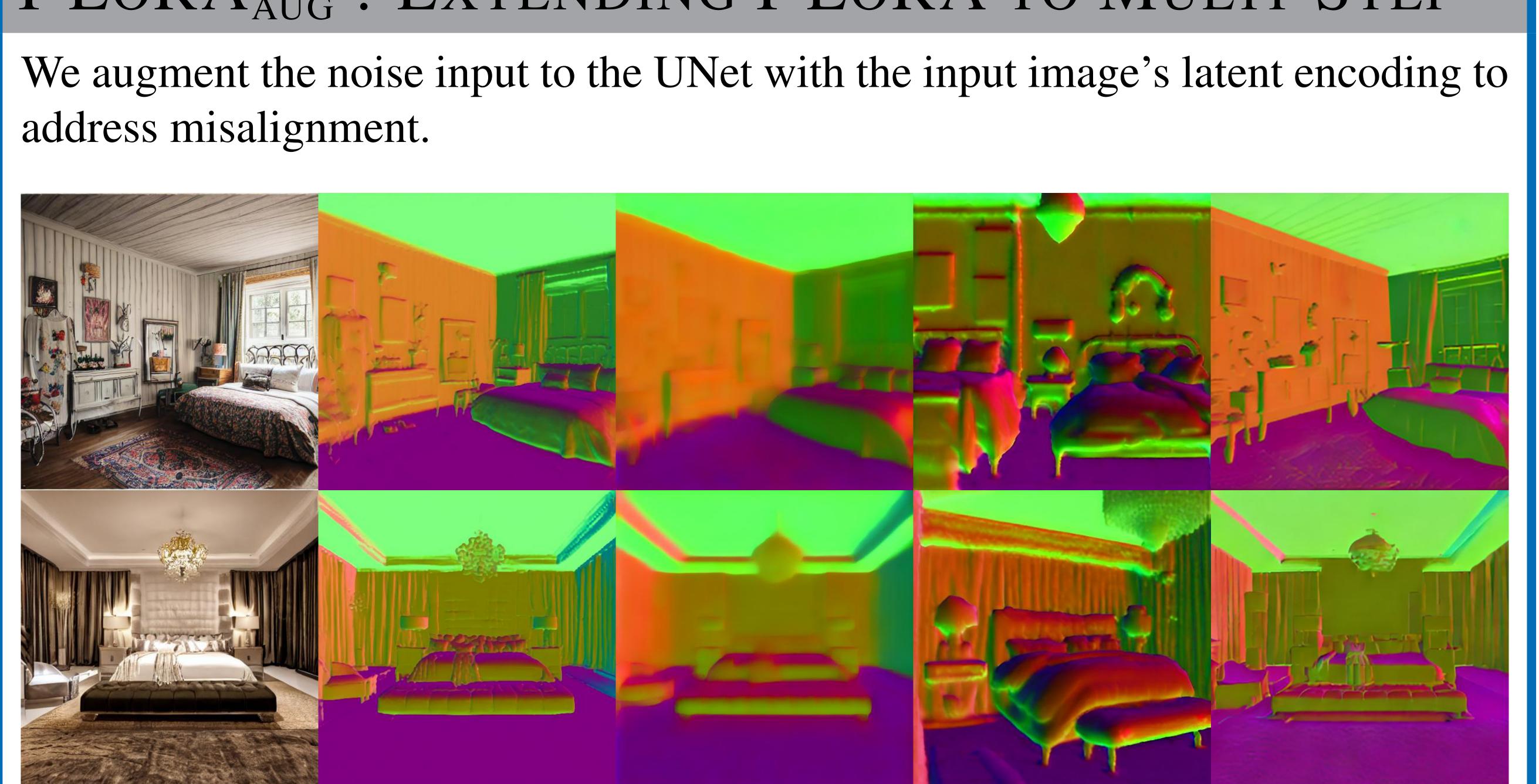
Xiaodan Du[1]   Nicholas Kolkin[2]   Greg Shakhnarovich[1]   Anand Bhattad[1]

[1]Toyota Technological Institute at Chicago   [2]Adobe Research

Introduction

TL;DR: Intrinsic LoRA (I-LoRA) is a general approach for extracting visual knowledge from generative models of many types.

(a) VQGAN FFHQ (Autoregressive)   (b) StyleGAN-XL FFHQ (GAN)

Image   Normals   Depth   Albedo   Shading

(c) StyleGAN-v2 LSUN Bedroom (GAN)

Image   Normals   Depth   Albedo   Shading

(d) Stable Diffusion 2.1 (Diffusion)

Image   Normals   Depth   Albedo   Shading

✓: Intrinsics can be extracted with high quality.
~: Intrinsics cannot be extracted with high quality.
×: Intrinsics cannot be extracted.

Model	Pretrain Type	Domain	Normal	Depth	Albedo	Shading
VQGAN	Autoregressive	FFHQ	~	~	✓	✓
SG-v2	GAN	FFHQ	✓	~	✓	✓
SG-v2	GAN	LSUN Bed	✓	✓	✓	✓
SG-XL	GAN	FFHQ	✓	~	✓	✓
SG-XL	GAN	ImageNet	×	×	×	×
SD-UNet (single-step)	Diffusion	Open	✓	✓	✓	✓
SD (multi-step)	Diffusion	Open	✓	✓	✓	✓

Summary of scene intrinsics found across different generative models.

Approach

"shading"   "albedo"   "depth"   "surface normal"

| Image latent | Attention layer | UNet | LoRA | Text encoder | VAE | Swappable intrinsics | MSE loss | Gradient flow | Learnable | Frozen |

Model — Intrinsics Quality Correlation

	Real	GT	Random init.	I-LoRA v1-1	I-LoRA v1-2	I-LoRA v1-5
Mean Angular Error° ↓			36.18	21.84	21.41	20.31
L1 Error (× 100) ↓			29.28	17.78	17.38	16.53

I-LoRA_AUG: Extending I-LoRA to Multi-Step

We augment the noise input to the UNet with the input image's latent encoding to address misalignment.

Image   Pseudo GT   I-LoRA   SDv1-5 (multi)   I-LoRA_AUG

We apply Zero SNR strategy on SDv2.1 to address color shift.

Image   GT   I-LoRA   I-LoRA1-5_AUG   I-LoRA_AUG

Limitation and Future Work

- Room for further reduction of training requirements.
- Despite I-LoRA_AUG's sharper results, its accuracy still lags behind single-step.
- Incorporating the extracted scene intrinsics into the learning process.
- Developing an evaluation of generative models based on physical properties.