Histological Brain Imaging Super-Resolution with Frequency-guided Diffusion Models

Giovanni Casari, Federico Bolelli, Costantino Grana

University of Modena and Reggio Emilia, Italy
{name.surname}@unimore.it

UNIMORE UNIVERSITÀ DEGLI STUDI DI MODENA E REGGIO EMILIA

IEEE ISBI 2026
2026 IEEE International Symposium on Biomedical Imaging
April 8-11, 2026 | London, UK

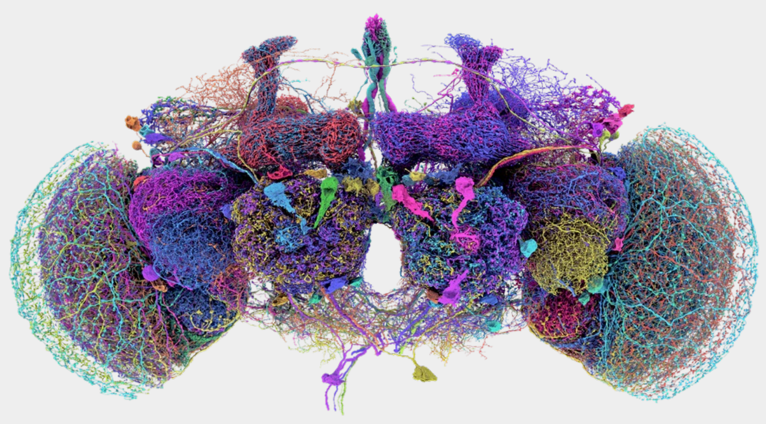
AImage Lab University of Modena and Reggio Emilia

Poster No. 112

Problem Statement

High-resolution histological brain imaging is crucial for modeling neuronal morphology and cytoarchitecture, enabling applications such as **large-scale scaffold models** of brain circuits.

Human Hippocampus Scaffold Model

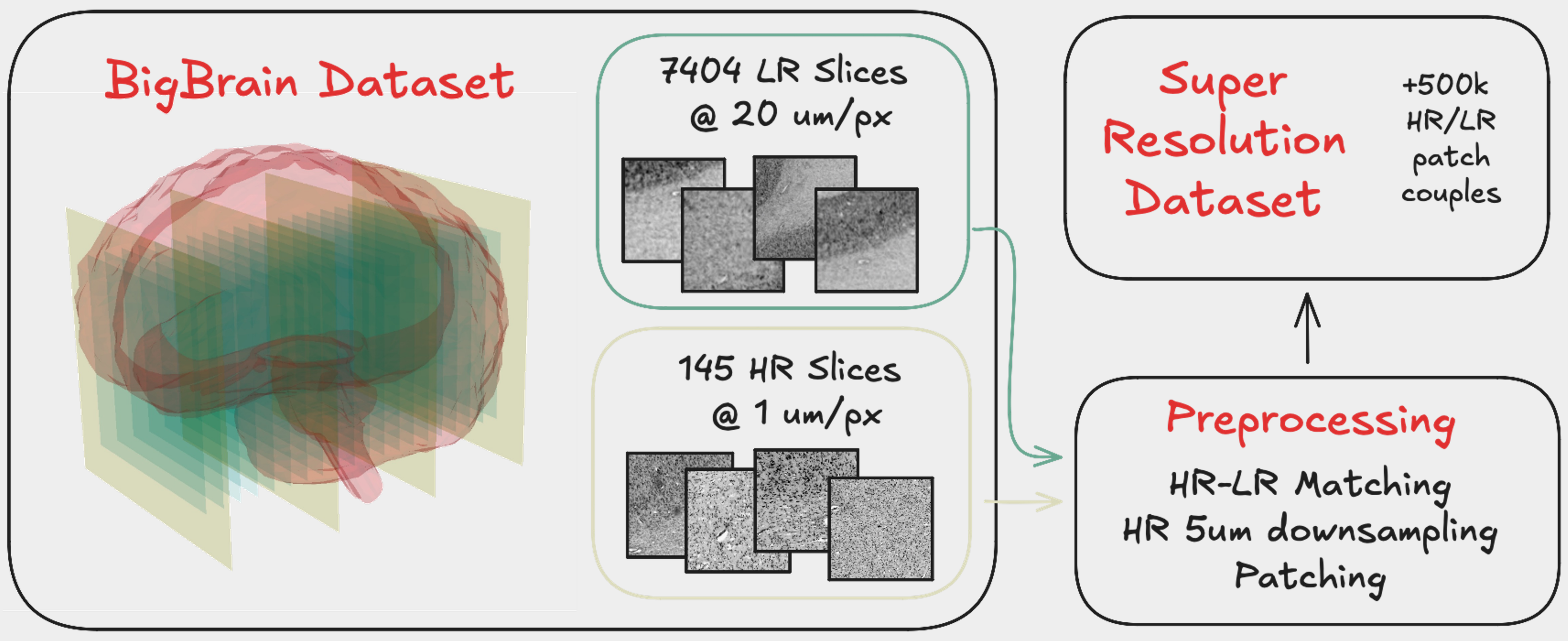


Drosophila Whole-Brain Model

However, acquiring **whole-brain data at micrometer resolution** is technically and economically infeasible, with most datasets available only at much lower resolution.

Super-resolution enables reconstruction of fine cellular detail from low-resolution data.

Dataset



We use the **BigBrain dataset** [1], leveraging 145 sections re-scanned at 1 μm as high-resolution (HR) references, paired with their original low-resolution (LR) sections.
To reduce the spectral gap, HR images are smoothed and downsampled to 5 μm, consistent with a 4x scale factor.

- LR: 20 μm sections → 128×128 patches
- HR: 1 μm (→ 5 μm) → 512×512 patches

Proposed Architecture

Brain-SR builds on the InvSR [2] paradigm, where a frozen pretrained **latent diffusion model** is steered by a lightweight **Noise Predictor network**.

Given a Low-Resolution patch:

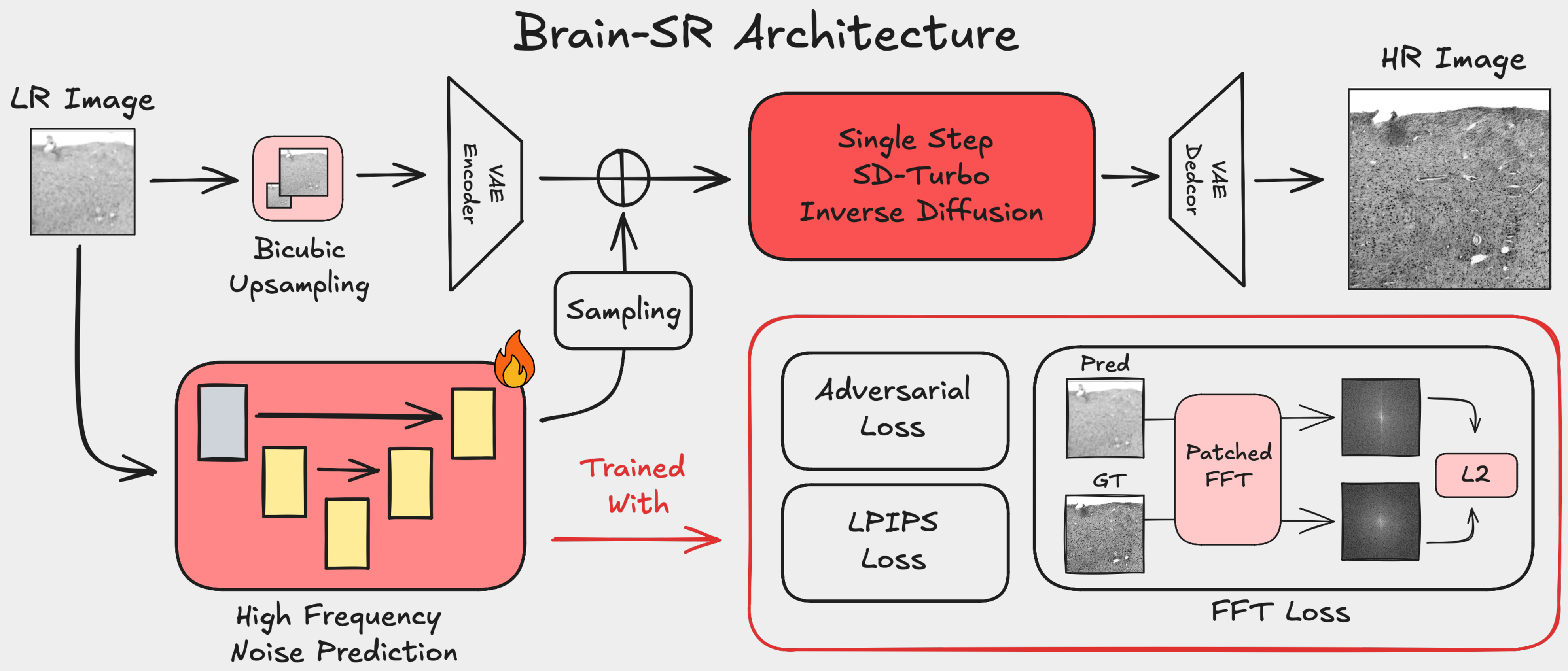
- LR (x_{LR}) is bicubically **upsampled** and encoded into the VAE latent space (z_{LR});
- The Noise Predictor (NP) estimates a latent residual encoding the **missing high-frequency content**;
- This residual is scaled and added to the LR latent to initialize the diffusion process:

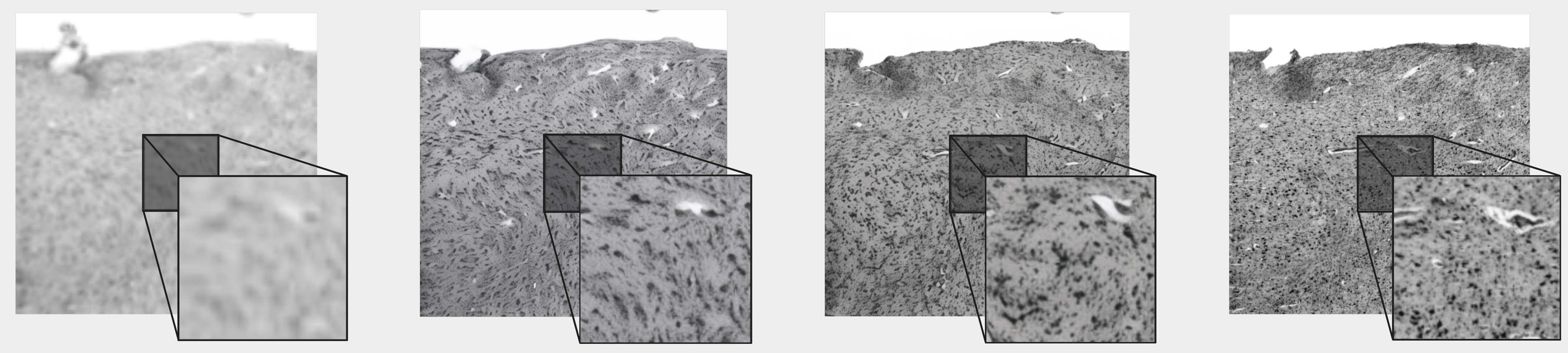
$$z_t = z_{LR} + \sigma_t \cdot S(NP(x_{LR}));$$

- A **single-step denoising** (SD-Turbo) refines the latent image into a clean latent representation;
- The latent is **decoded** by the VAE to obtain the final high-resolution image.

The training objective combines perceptual and adversarial terms with a new **patch-wise FFT loss**.

This matches local Fourier magnitude spectra, enforcing the distribution of correct frequency content while avoiding phase sensitivity.





Results

Model	λ_{L2}	λ_{ADV}	λ_{LPIPS}	λ_{FFT}	FFT_{conf}	L2 ↓	PSNR ↑	SSIM ↑	LPIPS ↓	FID ↓
Brain-SR-b1	1.0	0.05	1.5	–	–	**0.015**	**19.98**	0.303	0.550	139.4
Brain-SR-b2	1.0	0.10	2.0	–	–	0.017	19.40	**0.322**	0.412	98.7
Brain-SR-fft	0.3	0.08	0.7	0.2	{16,8}	0.019	18.90	0.319	0.401	83.2
Brain-SR-p16	–	0.08	0.7	0.4	{16,8}	0.024	17.86	0.280	0.412	63.7
Brain-SR-p8	–	0.08	0.7	0.4	{8,4}	0.023	18.06	0.282	0.412	62.0
Brain-SR	–	0.08	0.7	0.4	{4,2}	0.021	18.32	0.296	**0.398**	**57.5**
*Brain-SR**	–	*0.08*	*0.7*	*0.4*	*{4,2}*	*0.020*	*18.67*	*0.317*	*0.386*	*33.4*

References

[1] K. Amunts et al., "BigBrain: An Ultrahigh-Resolution 3D Human Brain Model," Science, vol. 340, no. 6139, pp. 1472 – 1475, 2013.
[2] Zongsheng Yue et al., "Arbitrary-steps Image Super-resolution via Diffusion Inversion," in IEEE/CVF CVPR, 2025.