# MRI Super-Resolution Using Fine-Tuned Diffusion Models

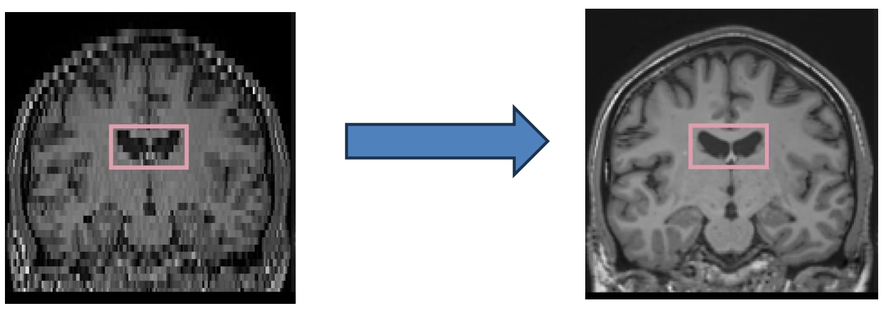
Jonas Vinson
University of Toronto

## Motivation

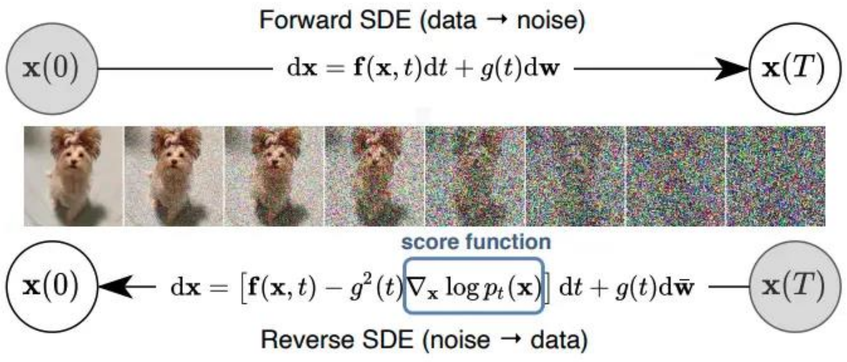
**Problem:** MRIs are the primary tool for the diagnosis and management of various brain disorders. Clinicians often acquire low-resolution scans due to the long acquisition times required for high-resolution imaging.

**Super-resolution (SR):** These are techniques aim to reconstruct high-quality images from low-resolution inputs.

**LR image** **HR image**



Recent advances in diffusion-based models have shown promising results for MRI super-resolution. These methods, however, are typically trained from scratch on 3D volumetric MRIs which is computationally expensive and difficult to scale.




## Related Work

**DisC-Diff:**
This method concatenates HR LR, and an auxiliary contrast image into a single latent representation, which is then processed by a diffusion model to produce a multi-contrast SR reconstruction [1]

**d3T*:**
Their model employs conditional latent diffusion. During the reverse diffusion process, the latent representation of the LR MRI serves as conditioning to guide the denoising trajectory, preserving anatomical structure.[2]

**TSR-LDM:**
This work combines a VQGAN-based encoder–decoder with a latent diffusion model trained on 2D cardiac MRI slices.[3]

While many studies adopt similar latent diffusion architectures for MRI SR, no prior work (from what I have found) has explored adapting a large, pretrained natural-image diffusion model (such as Flux) for medical image super-resolution.

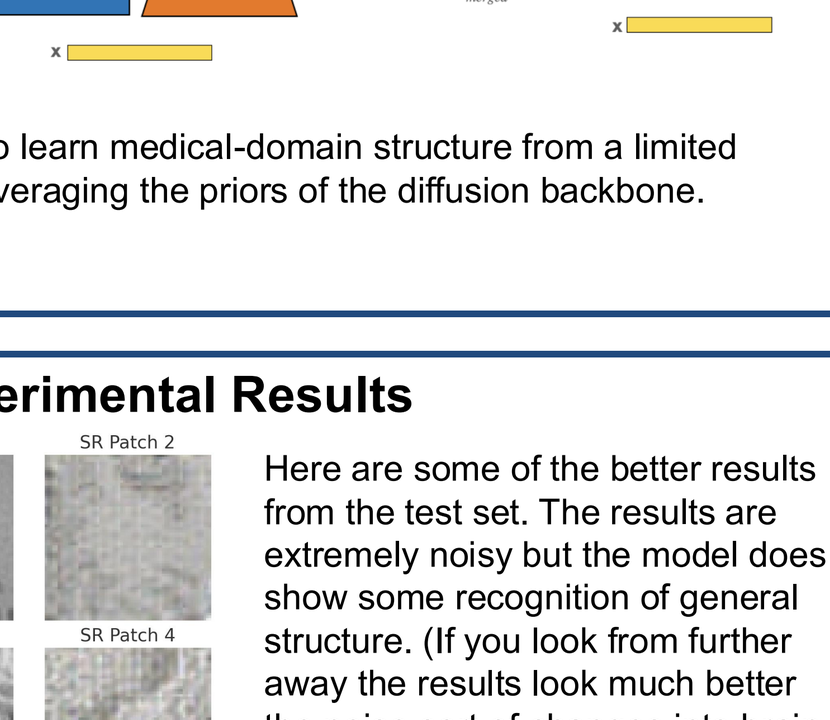
## References

[1] Y. Mao, L. Jiang, X. Chen, and C. Li, “Disc-diff: Disentangled conditional diffusion
model for multi-contrast mri super-resolution,” 2023.
[2] D. Yoon, Y. Myong, Y. G. Kim, Y. Sim, M. Cho, B.-M. Oh, and S. Kim, “Latent diffusion model-based mri superresolution enhances mild cognitive impairment prog-nostication and alzheimer’s disease classification,” NeuroImage, vol. 296, p. 120663,2024.
[3] V. Dubey, “Temporal and spatial super resolution with latent diffusion model in medical mri images,” 2024.
[4] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-resolution image synthesis with latent diffusion models,” 2022.

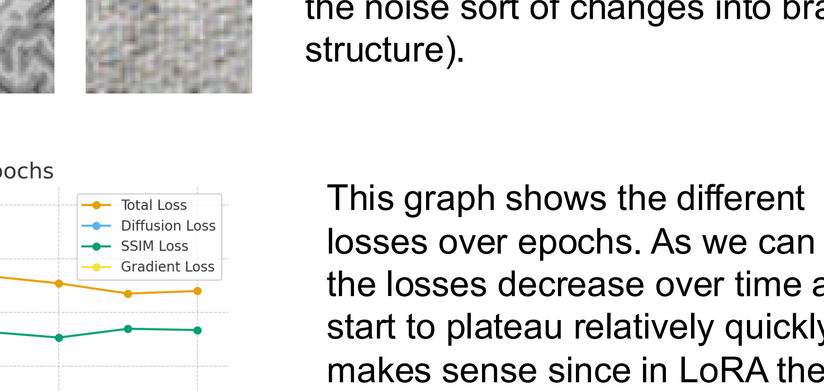
## New Technique

In my work I propose fine tunning a state-of-the-art large diffusion model Flux, for brain MRI super resolution. Such Models are trained on billions of parameters and massive natural image datasets giving them strong generative priors. My goal is to adapt these capabilities to the MRI domain.

Flux operates a latent space allowing for efficient high-res generation. By attaching a ControlNet I guide the diffusion process with structural information, ensuring anatomically correct output

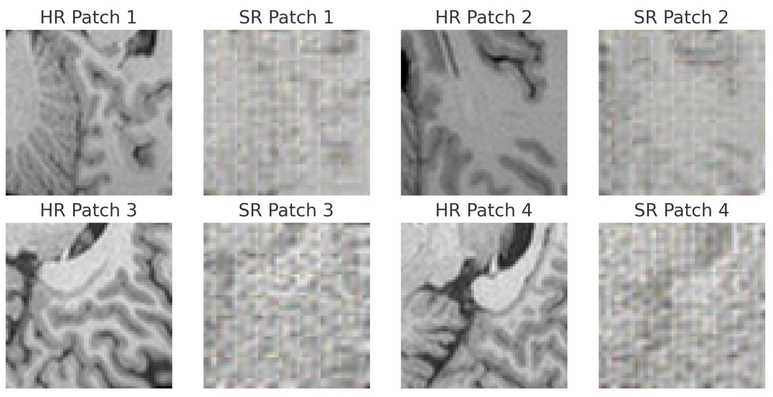



LoRA (Low-Rank Adaption) works by freezing the original model weights and inserting two small trainable matrices into certain layers. Updating these low-rank adapters reduces GPU memory while keeping the strong priors of the the pre-trained model.

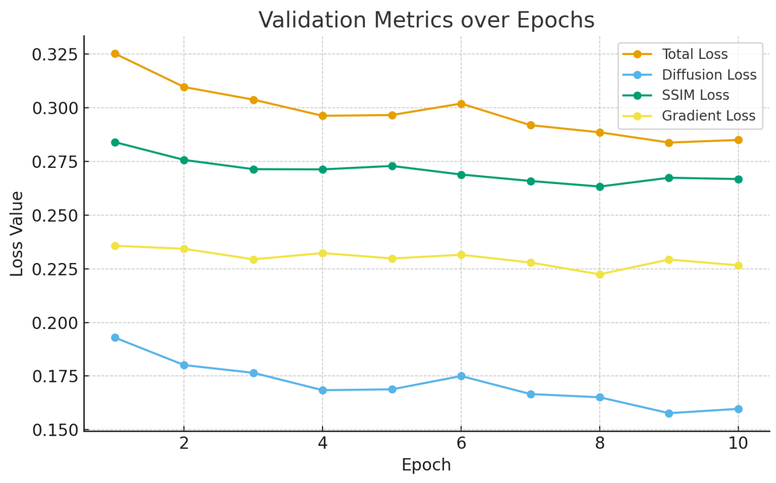



Using both techniques allows Flux to learn medical-domain structure from a limited dataset and training regime while leveraging the priors of the diffusion backbone.

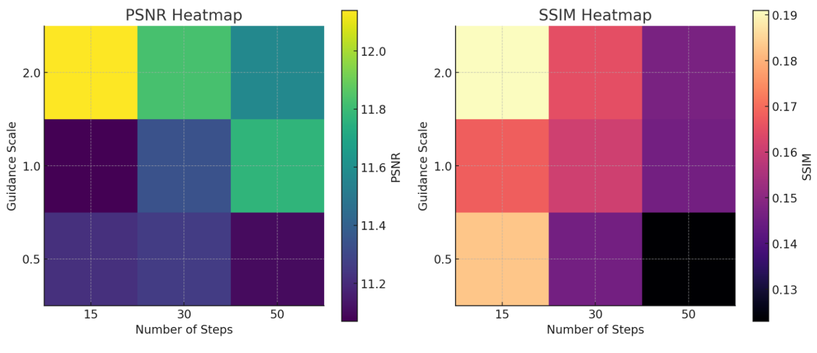
## Experimental Results




Here are some of the better results from the test set. The results are extremely noisy but the model does show some recognition of general structure. (If you look from further away the results look much better the noise sort of changes into brain structure).




This graph shows the different losses over epochs. As we can see the losses decrease over time and start to plateau relatively quickly this makes sense since in LoRA there are not many parameters to learn and they converge quite quickly




This is a heatmap of the SSIM and PSNR for varying diffusion steps and level of guidance from the control net. Higher guidance values appear to slightly improve the PSNR and SSIM, showing the controlNet provided some structural conditioning

The model demonstrates limited structural coherence but fails to recover fine anatomical details, and both the PSNR and SSIM and low. Further experiments could be done to train the controlNet individually to extract more meaningful features. Or maybe focus on a deeper LoRA adaption. However, it seems like generally the model have difficulty with this domain change.