

---

# Medically Informed Stable Diffusion (MISD)

— Laura Rojas, Nishanth Marer, —  
Hebron Taylor

---

# Introduction

---

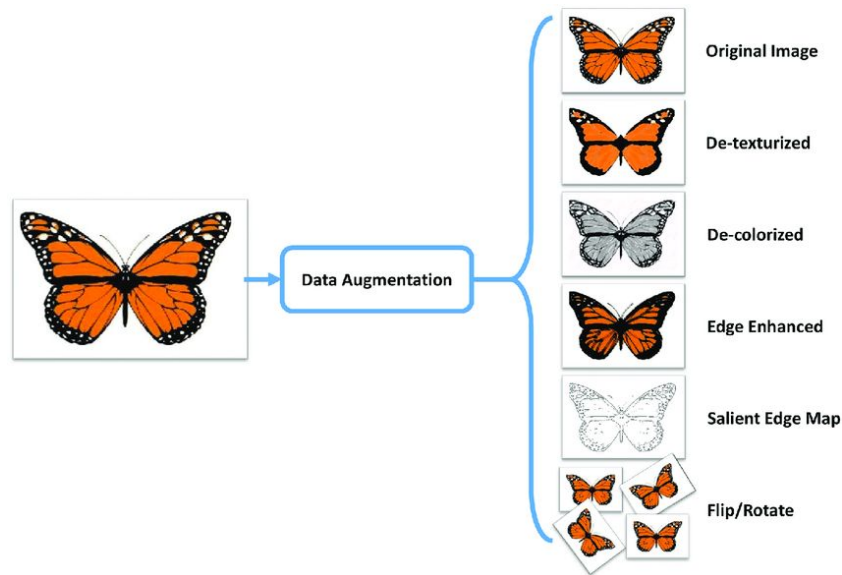
# Medical Interpretability

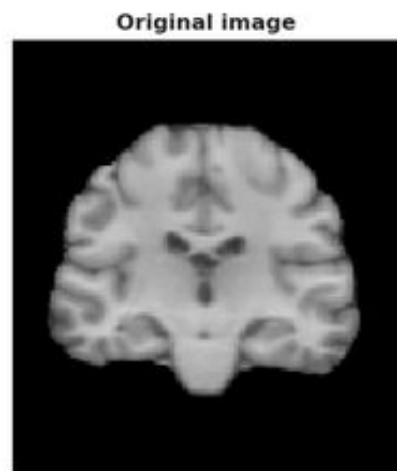
- Importance:
  - Clinical Relevance
  - Diagnostic Insights
  - Ethical Considerations
- Challenges:
  - Data Complexity
  - Biological Realism
  - Validation / Evaluation
  - Incorporating Domain Knowledge



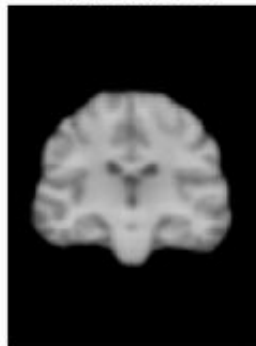
# Data Augmentation for Medical Imaging

- Data Transformation
  - Affine Transformation
  - Erasing Transformation
  - Elastic Transformation
  - Pixel-Level Transformation
- Generation of Artificial Data
  - Generative Adversarial Networks
  - Feature Mixing Methods
  - Model-based Modification
  - Reconstruction-based Methods

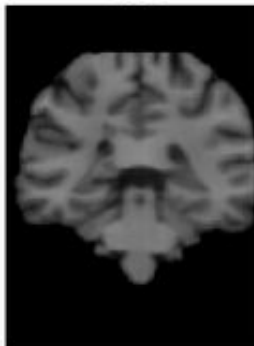




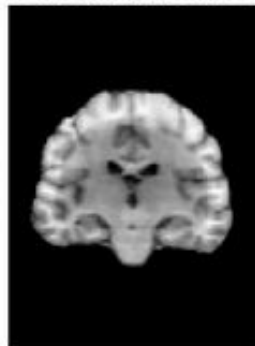
Gaussian Blur



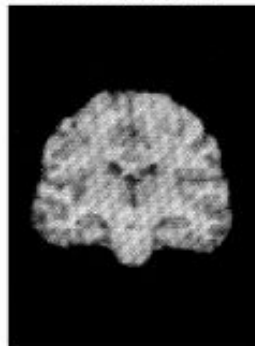
Crop



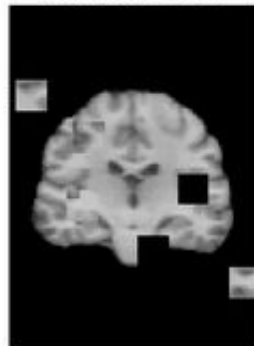
Ghosting Artefact



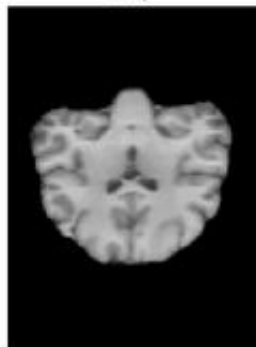
Spike Artefact



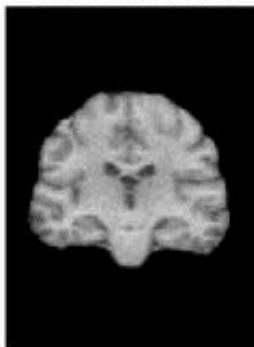
Random Swap



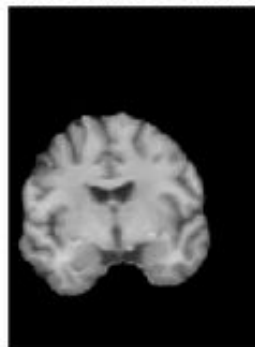
Flip



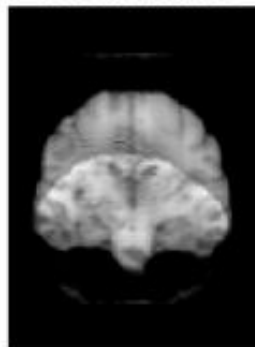
Gaussian Noise



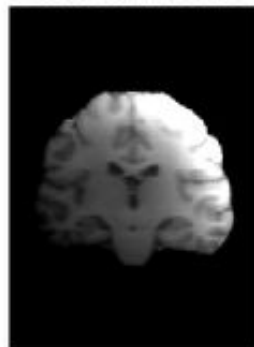
Affine transformation



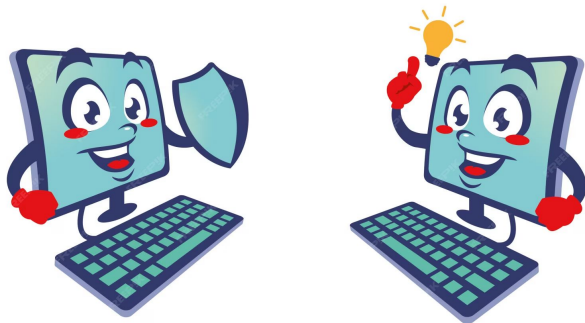
Motion Artefact



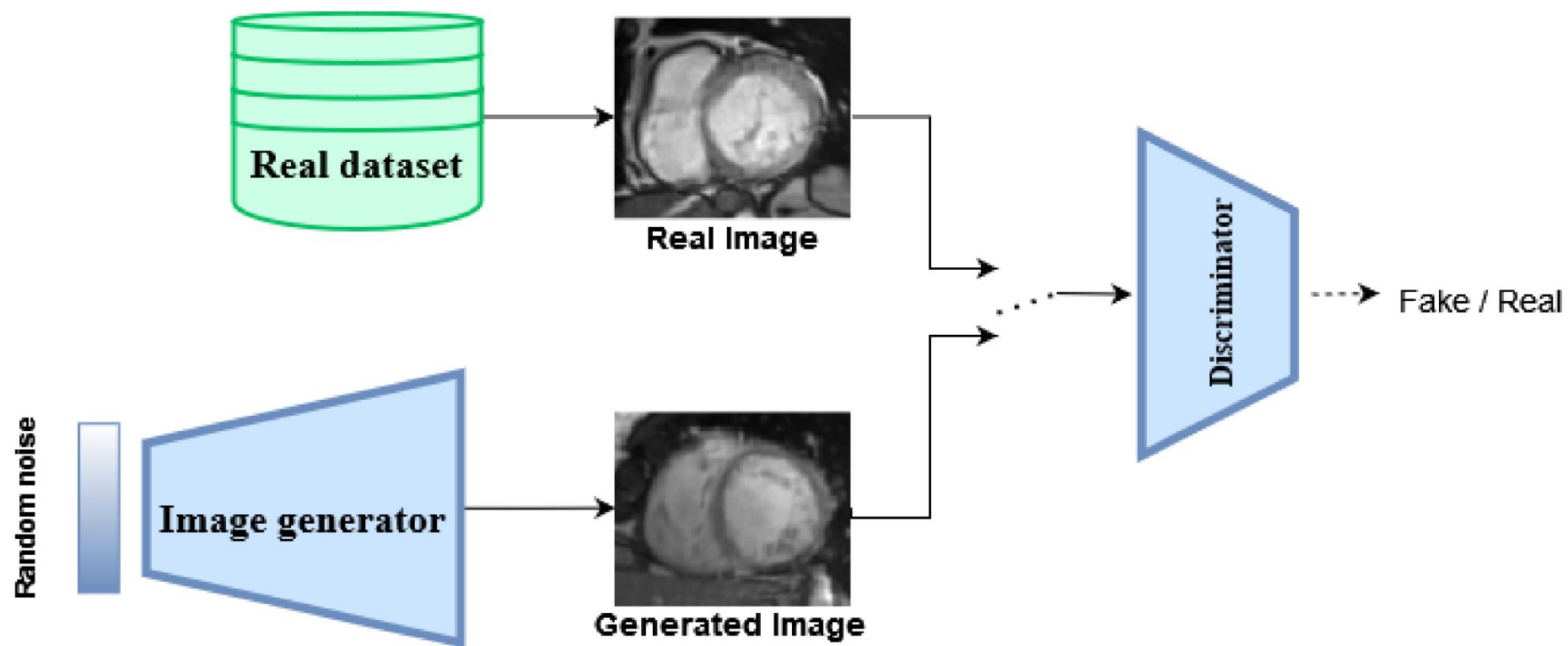
Biasfield Artefact



# Image to Image with GAN

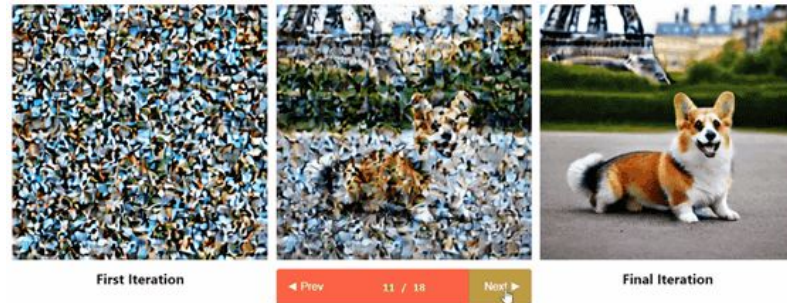


- Generative Adversarial Network
  - Generator vs Discriminator
- Source-Target Mapping
  - Capture characteristics of image collection
  - Translate characteristics into other images
- Synthetic Data
  - Cheaper
  - Cannot be re-identified



# Image to Image with Stable Diffusion

- Diffusion Model
  - Generates realistic, high-resolution images
  - Class-conditional image synthesis
- Trainable with limited data
  - Requires 200 images to fine-tune
- Synthetic Data
  - Cheaper
  - Cannot be re-identified





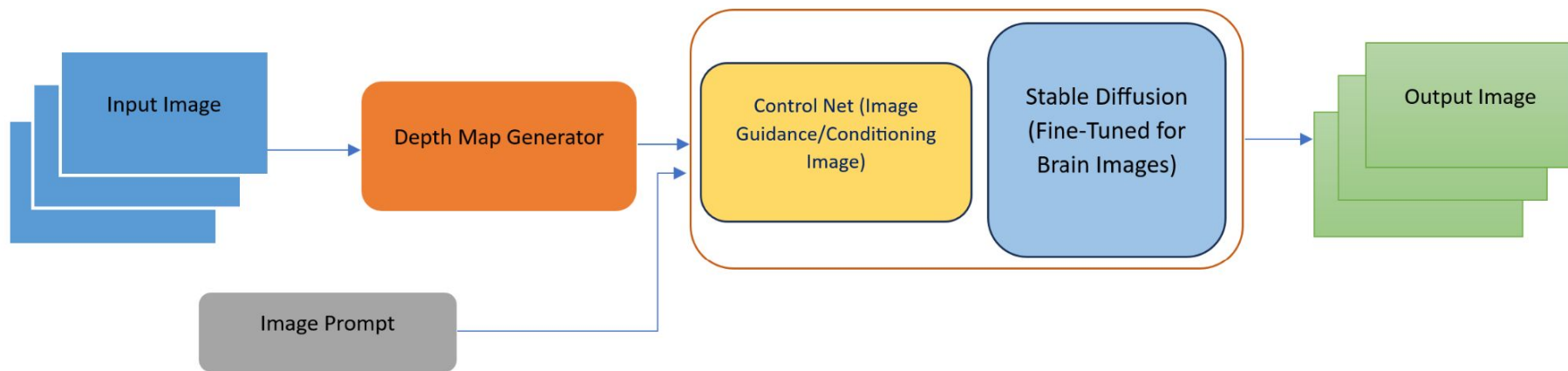
# Proposal

---

# Proof of Concept - Model Structure

1. ControlNet
  - “Conditioning Image” guides image generation by adding constraints
2. StableDiffusion
  - Uses prompt to generate image
3. UNet
  - Image to image training to refine generated image
4. Textual Inversion
  - Trains text embeddings

# Diagram

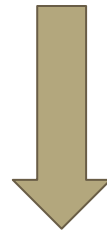
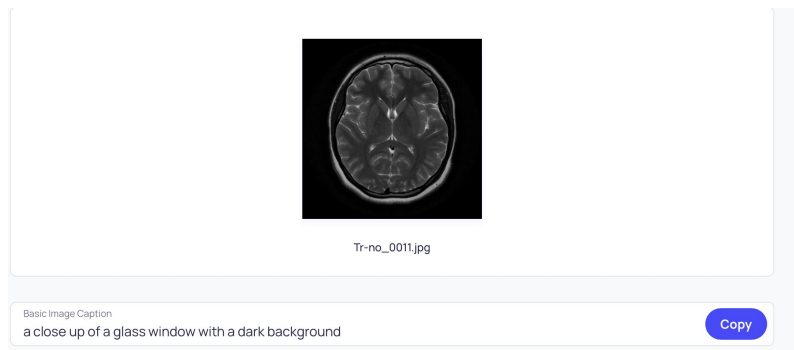


# Proof of Concept

---

# Prompt Engineering

- Training images require custom captions
- image-to-caption generators are not so good
- Must engineer effective & unique captions



- 1 Intricate brain MRI reveals neural networks in stunning detail
- 2 CT scan showcases brain structures with high-resolution clarity
- 3 Detailed MRI maps brain activity patterns for neurological insights
- 4 CT imaging unveils brain anatomy in precise cross-sectional views
- 5 Brain MRI offers unparalleled insights into neural connectivity
- 6 CT scan captures intricate brain structures in detailed cross-sections

# Stable Diffusion: prompt only

*"brain CT scan"*

- Abstract concept of 'brain'
- Cinematographic
- Not medically accurate
- Undesired details



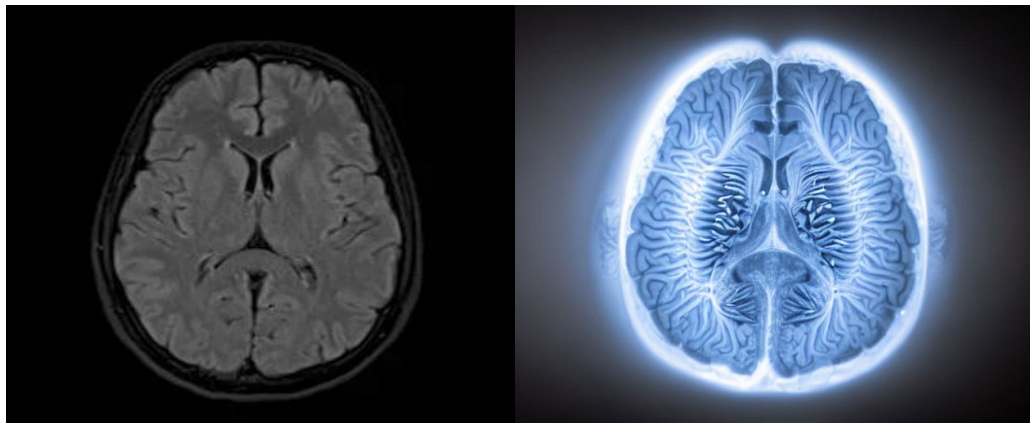
No Style

Default - Cinematographic

# ControlNet: image-guided with prompt

*"brain CT scan"*

- No concept of 'brain'
- Follows depth-map
- Not medically accurate
- Undesired details



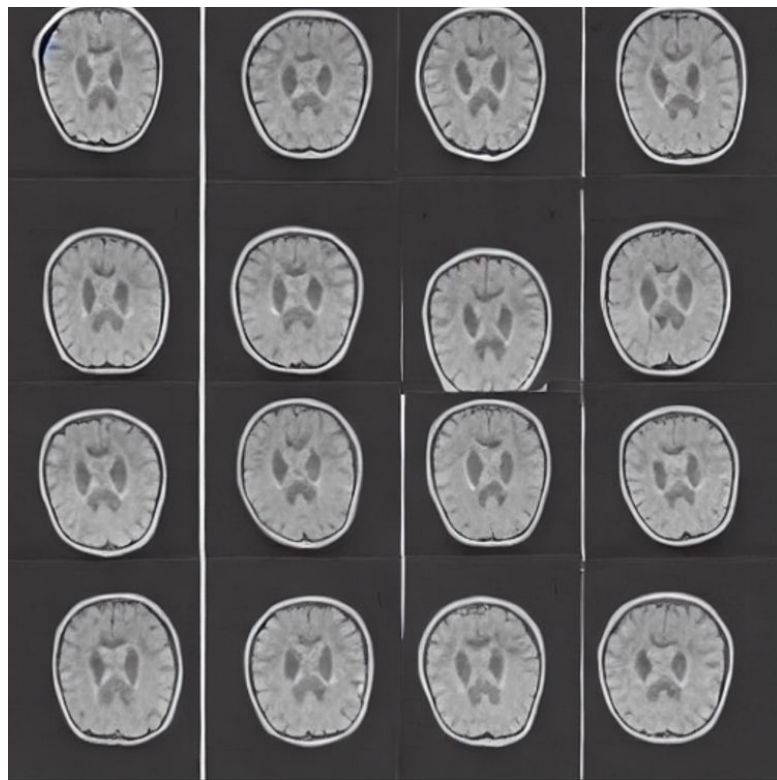
Original

ControlNet

# Fine Tuned: pretrained model with prompt only

*"brain CT scan"*

- Fine-tuned concept of "brain"
- Ability to create multiple images with slight variations in the internal structure while maintaining overall integrity
- Needs more guidance in terms of more finer details for better image generation

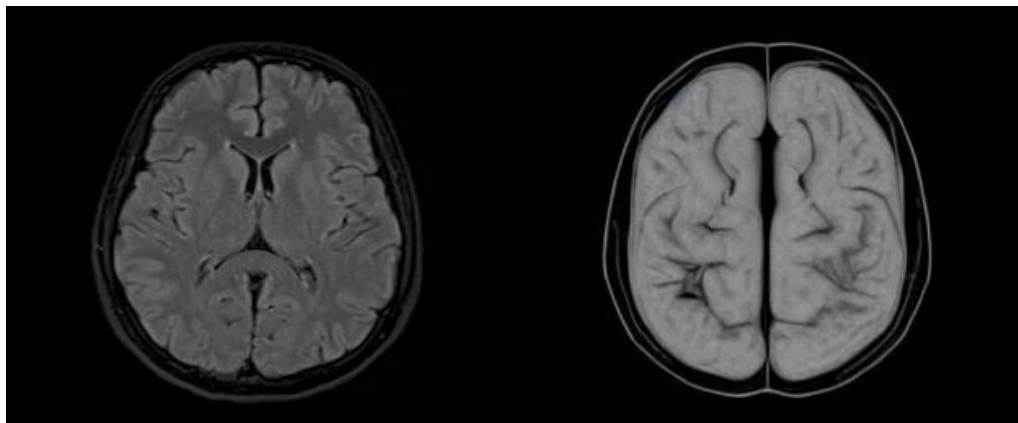




# Fine Tuned + ControlNet: image and prompt

*"brain CT"*

- Fine-tuned concept of "brain"
- Guided image generation
- Enforces constraints



Original

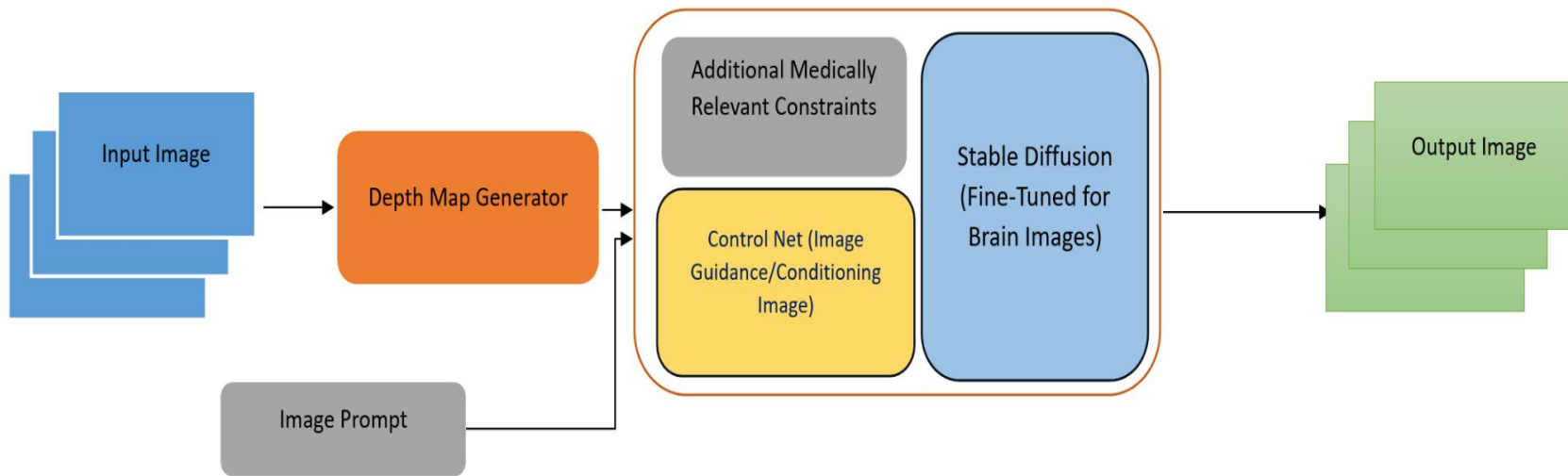
Fine Tuned + ControlNet

# Next Steps

---

# Biology-Informed Neural Network

- Add constraints to optimization algorithm
- Constraints inform model how to properly create images



# Conclusion

---

**Thank You**

# References

- <https://hai.stanford.edu/news/could-stable-diffusion-solve-gap-medical-imaging-data>
- <https://ieeexplore.ieee.org/document/10049010>
- <https://arxiv.org/abs/2303.13430>
- <https://www.sciencedirect.com/science/article/pii/S001048252201099X#b34>
- [https://openaccess.thecvf.com/content\\_iccv\\_2017/html/Zhu\\_Unpaired\\_Image-To-Image\\_Translation\\_ICCV\\_2017\\_paper.html](https://openaccess.thecvf.com/content_iccv_2017/html/Zhu_Unpaired_Image-To-Image_Translation_ICCV_2017_paper.html)
- <https://www.kaggle.com/datasets/masoudnickparvar/brain-tumor-mri-dataset>