FMitF: Track I: Generative Neural Network Verification in Medical Imaging Analysis

NSF Award # 2220401

Taylor T. Johnson (PI), Ipek Oguz (Co-PI), Meiyi Ma (Co-PI), Daniel Moyer (Collaborator)

Vanderbilt University

https://github.com/verivital/nnv/, http://www.taylortjohnson.com/, https://meiyima.github.io/

Overview: This FMitF project develops robustness specification and verification methods for generative computer vision tasks using machine learning (ML), to enable trustworthy analysis of medical images.

Motivation

1. **Ensuring Robustness in Healthcare AI:** Deep neural networks (DNNs) are widely used in medical imaging, but are not robust and subject to adversarial perturbations, which can cause significant errors in model predictions.
2. **Formal Guarantees in Safety-Critical Domains:** While formal verification methods have been applied to other critical industries (e.g., aviation, autonomous vehicles), there is a gap in healthcare regarding **formal verification of AI models in Healthcare.**
3. **Pioneering Formal Verification in Medical Imaging:** The project seeks to establish a **foundation for certifying deep learning models in medical imaging**, particularly for complex tasks like 3D semantic segmentation.

Major Objectives

Objective A: Develop formal specification framework to describe robustness of computer vision tasks beyond classification tasks, such as in segmentation and image synthesis, as well as automatically infer (mine) these specifications.

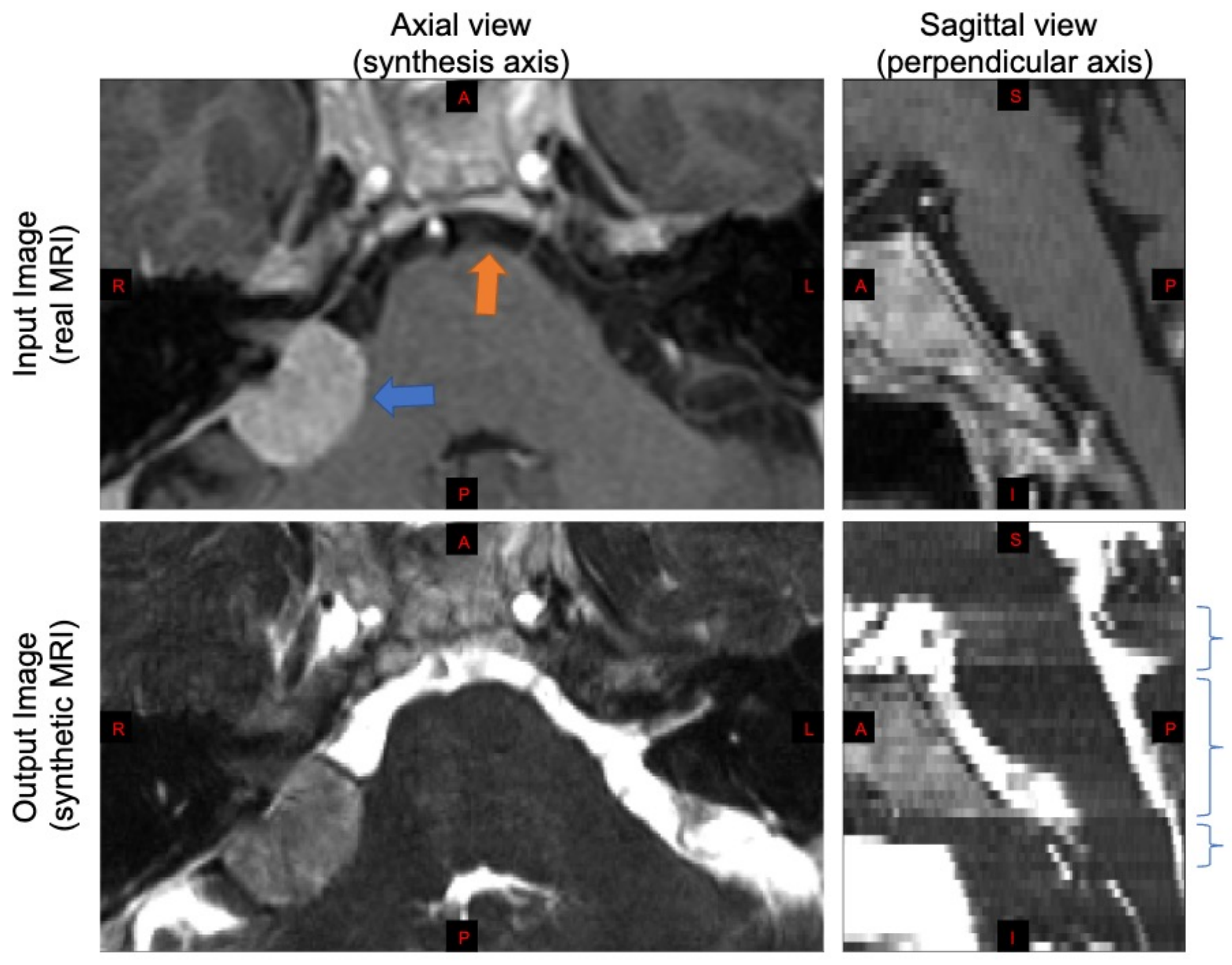
Objective B: Develop robustness verification & falsification methods in medical imaging analysis computer vision tasks

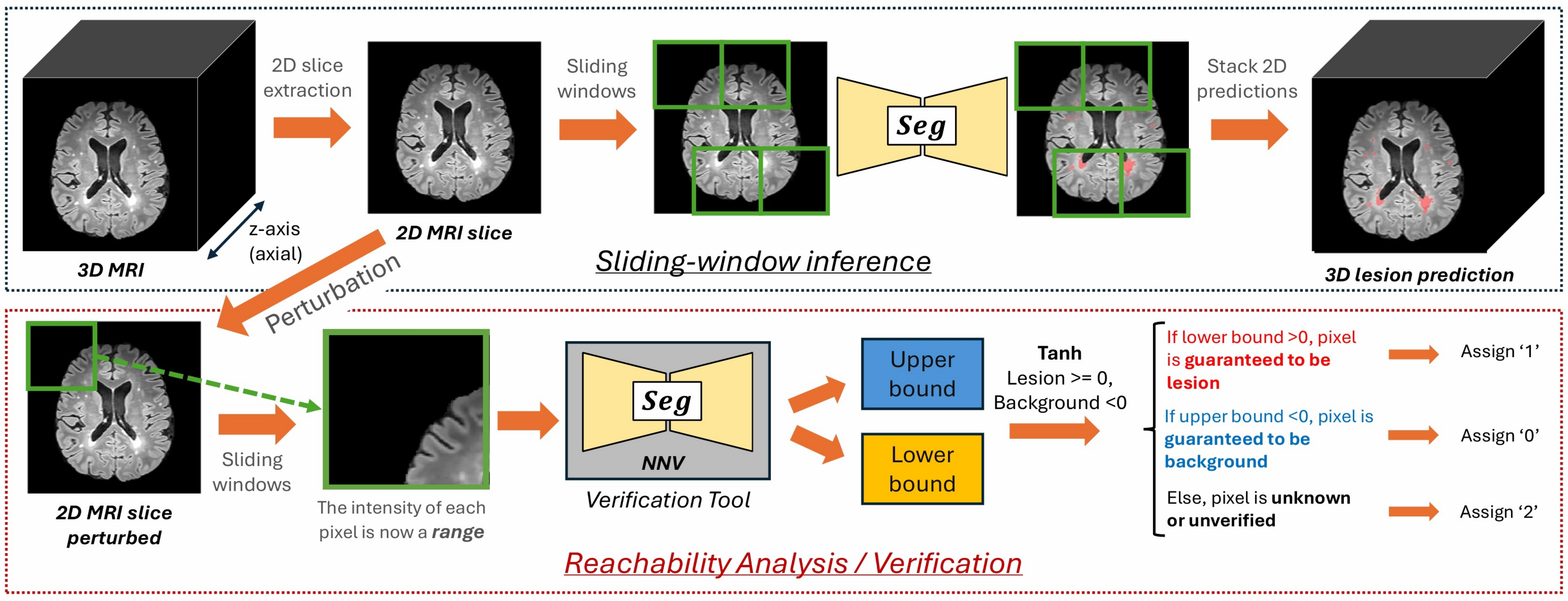
Objective C: Explore robustness in generative tasks and models (e.g., GANs) for tasks like image synthesis

Objective D: Evaluate verification methods on medical scans of different types (e.g., MRIs, CTs, etc.) from different sites (e.g., labs or hospitals)

Scientific Impact (Key results)

1. **First Formal Verification for Medical Image Segmentation:** This work is the **first to apply formal verification to medical image segmentation models.**
2. **Verification of 3D MRI Volumetric Data:** The study extends formal verification techniques to **high-dimensional 3D MRI data**, tackling a complex task like MS lesion segmentation.
3. **Demonstrating Worst-Case Guarantees:** The project provides formal guarantees on **worst-case performances** for segmentation models under different adversarial transformations


Axial view (synthesis axis)
Sagittal view (perpendicular axis)
Input Image (real MRI)
Output Image (synthetic MRI)
A
P
R
L
S
I



3D MRI

2D slice extraction

z-axis (axial)

2D MRI slice

Sliding windows

Seg

Stack 2D predictions

3D lesion prediction

Sliding-window inference

Perturbation

2D MRI slice perturbed

Sliding windows

The intensity of each pixel is now a *range*

Seg

NNV

Verification Tool

Upper bound

Lower bound

Tanh
Lesion >= 0,
Background <0

If lower bound >0, pixel is **guaranteed to be lesion** → Assign '1'

If upper bound <0, pixel is **guaranteed to be background** → Assign '0'

Else, pixel is **unknown or unverified** → Assign '2'

Reachability Analysis / Verification

Broader Impact on Society

1. As AI is increasingly used in medicine, but AI problems are well-known, all involved may not trust results and safety may be impacted.
2. Characterization of medical imaging analysis models with formal methods may help improve such trust as a verification & validation methodology.

Educational Broader Impact

1. Undergraduate research internships and Immersive Projects at Vanderbilt in summers 2023 and 2024 (Gloria Zhang, Seojin Lee, Lana Cartailler).
2. Research integrated into undergraduate and graduate courses (AI, Projects in AI, Automated Verification).

Community Engagement

1. Co-organization of International Verification of Neural Networks Competition (VNN-COMP), held with CAV, https://sites.google.com/view/vnn2024
2. Tutorials on NNV at EMSOFT'23, IAVVC'23, and DSN'24, and Upcoming tutorial with medical imaging community (SPIE'25)

The NSF Formal Methods in the Field PI Meeting (2024 FMitF PI Meeting)
November 12-13, 2024 | The University of Iowa | Iowa City, Iowa