Uncertainty Quantification in Deep Neural Network Models for the Determination of Gene Mutation Status in Gliomas

Dingkun Guo

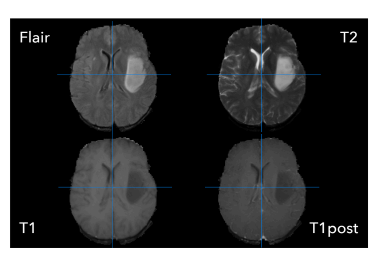
Instructor: Xun Huan

Department of Mechanical Enginnering

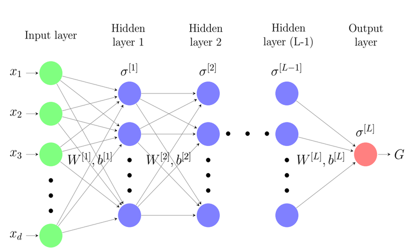
INTRODUCTION

Background

- Identifying **isocitrate dehydrogenase (IDH) mutation status** in brain tumors is critical for treatment planning; the presence of IDH mutation increases survival probability
- **Magnetic resonance imaging (MRI)** of the patient brain provides images useful for this assessment; four types of images, generated with different settings, are studied here
- **Deep neural network (DNN)**, a common *data-driven model*, is trained from existing patient images, and then used to make IDH status predictions for new images
- **Uncertainty quantification (UQ)** uses statistical methods to produce a measure of confidence on these predictions



Four Types of MRI images


An Example of Deep Neural Network


Maximize Fuel Efficient of Aircraft Engine

Model

- A residual convolutional neural network was developed [Chang et al. 2018] to predict IDH status from MRI scans
- **Input: MRI scans of the brain**
- **Output: prediction probability of positive IDH mutation**
- Current model reports single-value predictions, and does not offer prediction uncertainty induced by noisy and limited number of training data
- Over 22 million DNN weights were trained using only 496 data points: uncertainty quantification is crucial

Objective

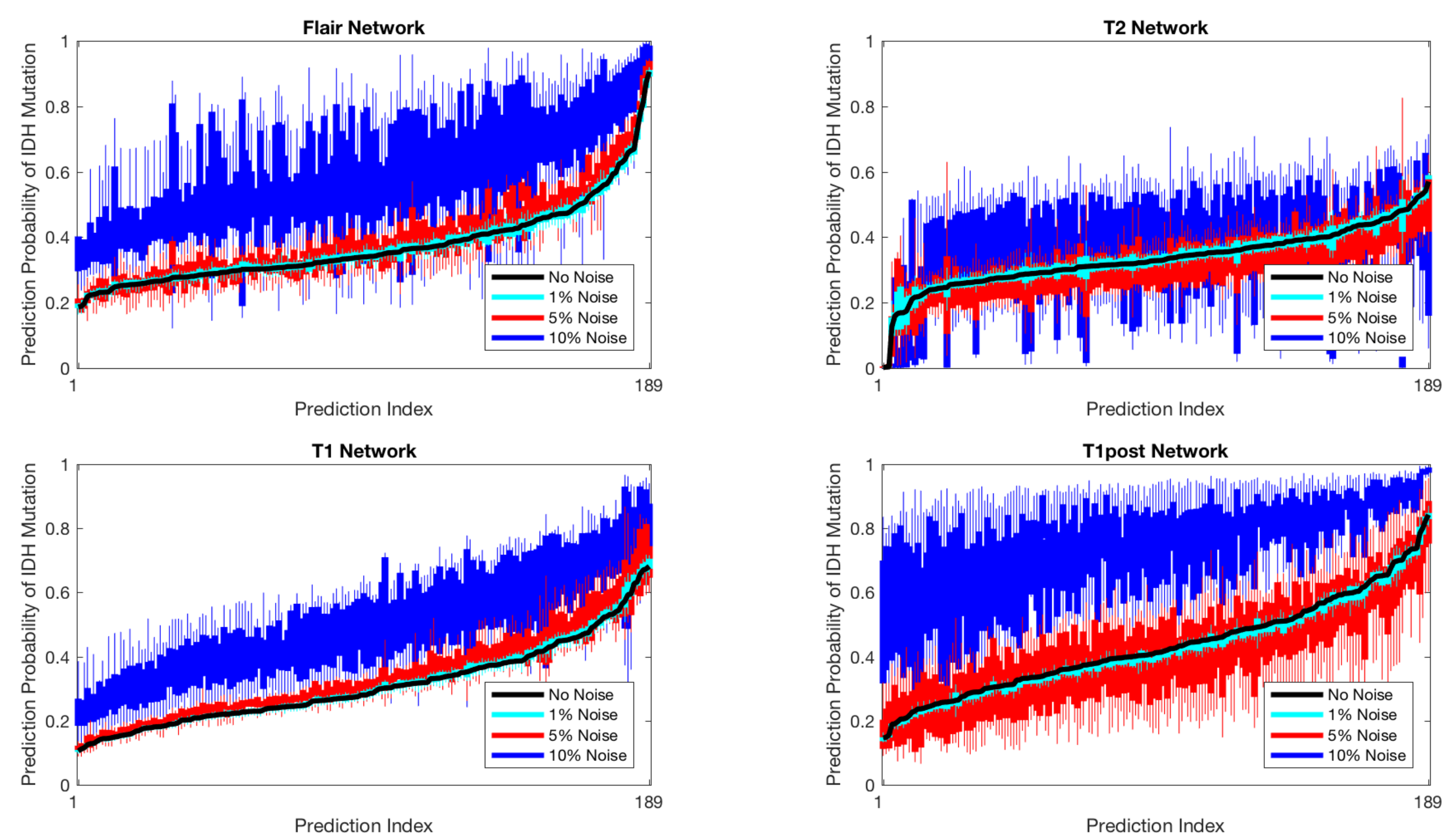
- To develop computational capability for quantifying uncertainty in DNN models systematically and rigorously
- To determine the most sensitive factors in a model
- To assess robustness of model predictions against noise

RESULT

Noise in Model

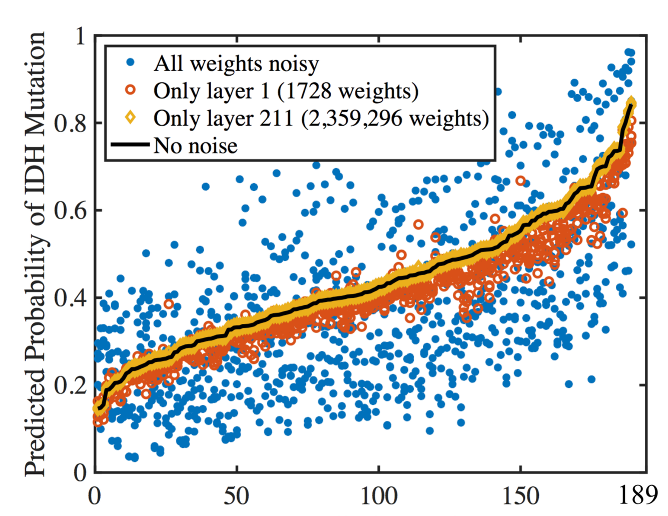
- **Adding Different Levels of Gaussian Noise to All Model Weights**

Prediction Probability of IDH Mutation with Different Noise in Model Weights (10 Runs)



- A 5% perturbation to the weights can alter predictions up to 20%
- T1post Network is most sensitive to the noise; larger uncertainty observed when prediction probability is small; T1 Network is less sensitive to large noise

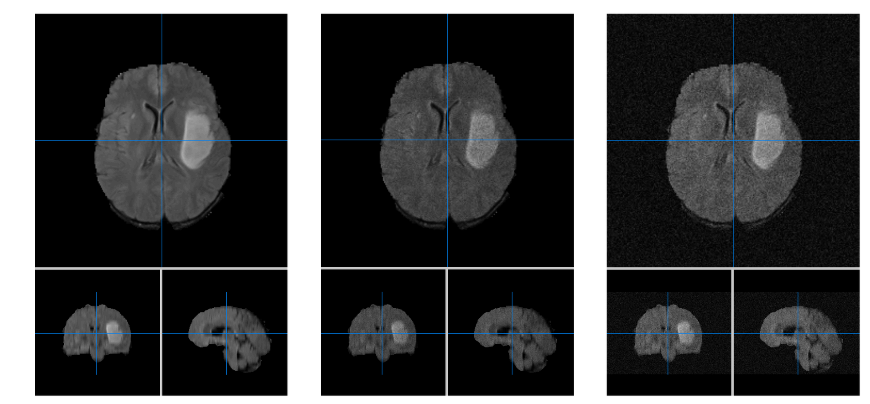
- **Adding Gaussian Noise to Weights in Different Layers**



- 5% Gaussian noise is added to different layers of T1post Network
- Although layer 211 has many more weights than layer 1, its noisy predictions have lower uncertainty
- Layer 1 contributes more to overall predictive uncertainty than layer 211
- This suggests some weights may be much more important than others

Noise in Data

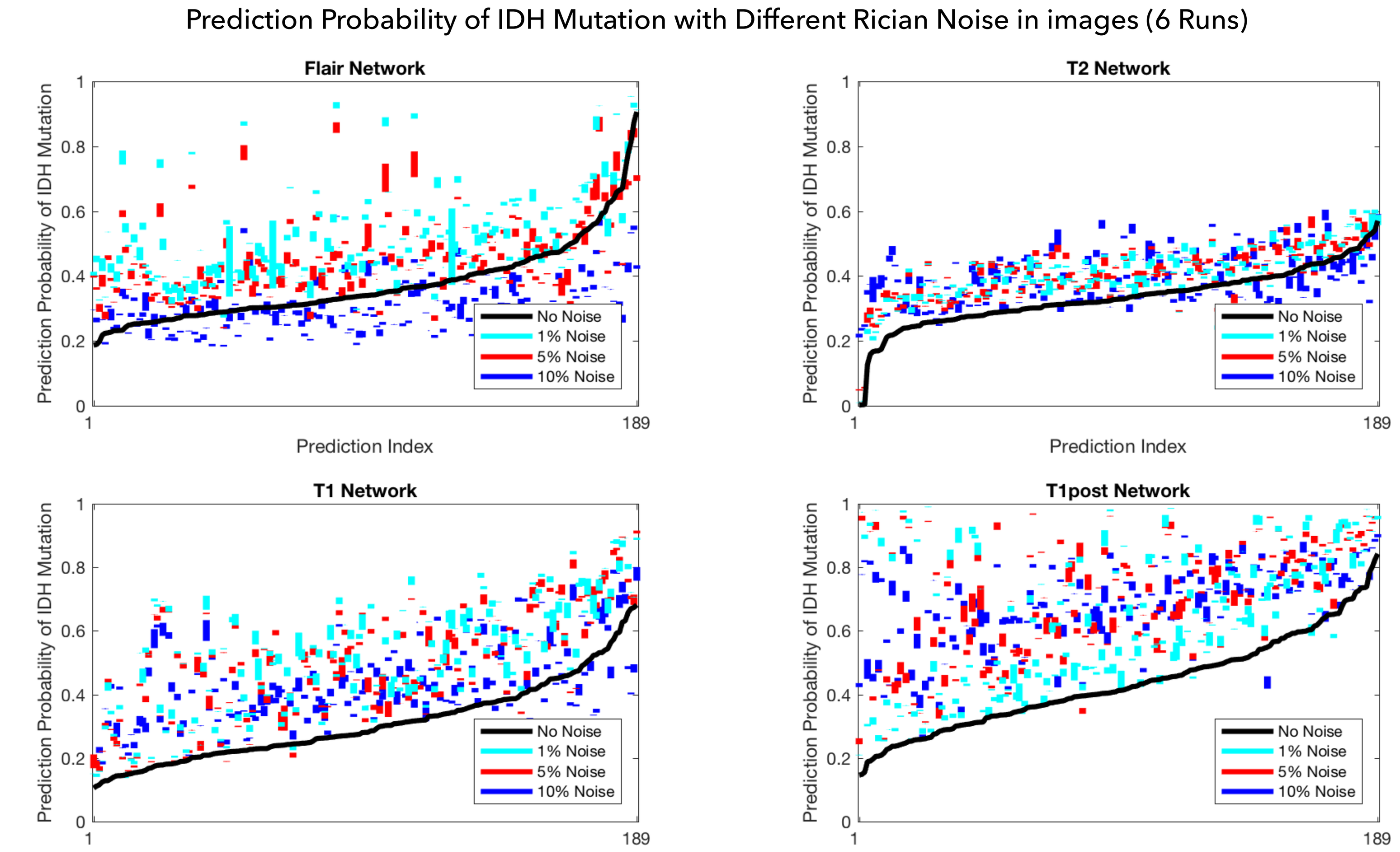
- **Comparison of Different Types of Noise in MRI images**


No Noise Image — 10% Gaussian Noise — 5% Rician Noise

- Different noise distributions in MRI images is hard to detect by human eyes, but can cause large difference in DNN predictions
- Rician noise is more commonly used in MRI applications

- **Adding Different Levels of Rician Noise to images**

Prediction Probability of IDH Mutation with Different Rician Noise in images (6 Runs)



- Rician noise in images shifts predictions but produces less uncertainty than noise in weights; small noise can have larger impact on predictions
- The model appears more robust against image noise, but becomes less accurate
- Noise in black areas of images may have large influence on predictions

METHODS

Model Evaluation

- Use the model to make predictions for 63 patients from The Cancer Genome Atlas (TCGA) database

Noise in Model

- Analyze the sensitivity of model predictions with respect to the trained model weights
- Use a Monte Carlo sampling approach, where random noise (1%, 5%, 10%) is added to the trained DNN model weights

Noise in Data

- Add Rician Noise to each pixel of MRI images
- $R = \sqrt{X^2 + Y^2}$ where $X \sim N(v, \sigma^2)$ and $Y \sim N(0, \sigma^2)$

DISCUSSION

Predictions from deep neural network models studied here can be quite sensitive to noise in model weights, which can be affected by the quality of training data and structure of the deep neural network. Uncertainty should be quantified and subsequently reduced from different aspects of models and data, to enable high-confidence predictions imperative for decision-making for patient treatments.

WHAT'S NEXT

- Analyze relationship between uncertainty and other features of tumor such as size/volume
- Conduct sensitivity analysis layer by layer to identify uncertainty contributions
- Provide a list of criteria for assessing model robustness and generalizability
- Repeat model evaluations to produce additional data for statistical analysis

REFERENCE

- Chang et al. (2018) Residual Convolutional Neural Network for the Determination of IDH Status in Low- and High- Grade Gliomas from MR Imaging, Clin Cancer Res, 24(5).
- Pedano et al. (2016) Radiology Data from The Cancer Genome Atlas Low Grade Glioma [TCGA-LGG] collection. The Cancer Imaging Archive.

ACKNOWLEDGMENT

- Srikanth Kuthuru, Nicholas Wang, and Prof. Arvind Rao (University of Michigan Department of Computational Medicine & Bioinformatics)