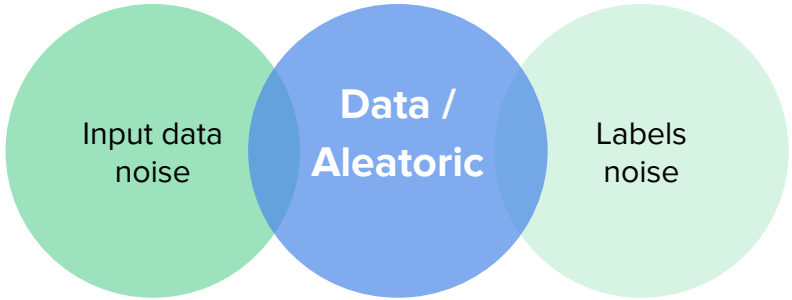
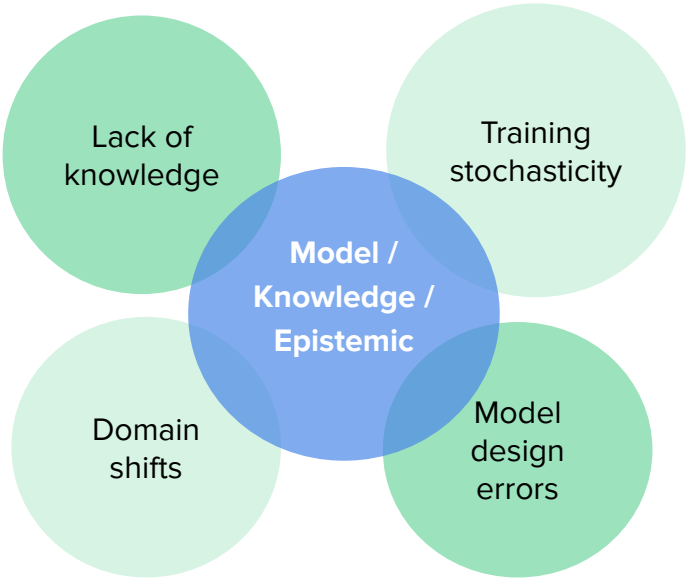
Uncertainty quantification in Medical image analysis

Nataliia Molchanova

Sources of uncertainty

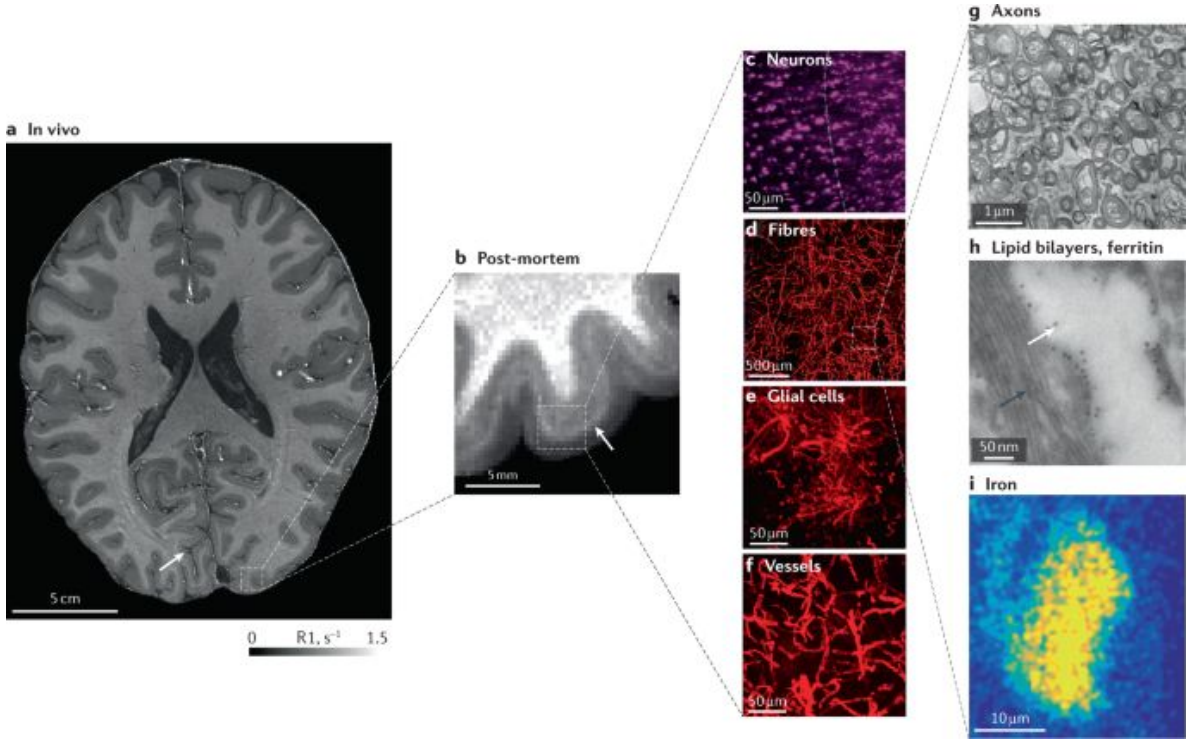




Gawlikowski et al., Artif. Intell. Rev., 2024

Noise in the input data

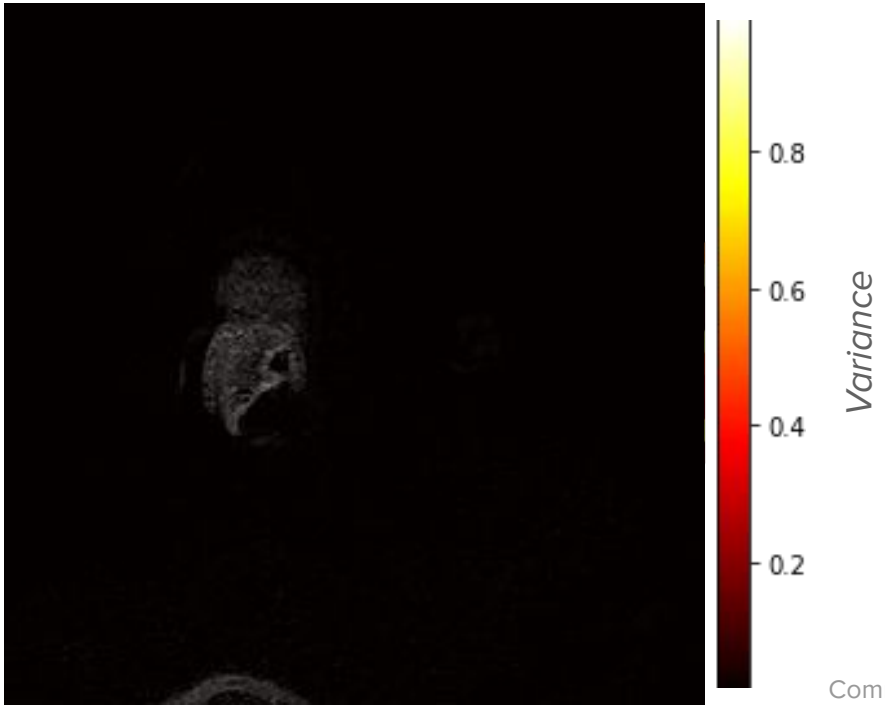
Limitations of the measurement systems (MRI, CT, PET, etc.)



Weiskopf et al., Nat. Rev. Phys., 2021

Inter-rater variability

Example of white matter lesions annotations from 7 raters



Commowick et al., Sci. Rep., 2018

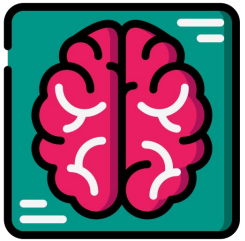
Noise in the labels

Expertise

Visual perception

Low quality

Artefacts

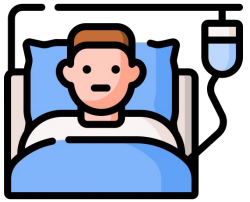


Guidelines ambiguities

Annotation tools

Rear pathologies

Demographic differences



Knowledge uncertainty in MIA

Low data regimes

Challenge Name	Task	Medical Imaging Modality	Number of Patients	Countries	Patients Worldwide
KiTS21 Challenge	Kidney Tumor Segmentation	CT	**300**	USA	**400k / Year**
MSSEG-1	Multiple Sclerosis Lesion Segmentation	MR	**53**	France	**2.9M**
LiTS	Liver Tumor Segmentation	CT	**130**	Germany	**800k / Year**
PROMISE12	Prostate MR Image Segmentation	MR	**50**	United States, Canada, Germany, France, UK	**1.4M / Year**

Knowledge uncertainty in MIA

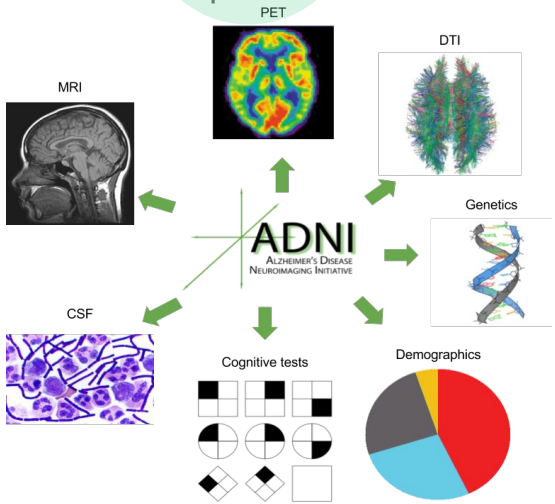
Worldwide prevalence

60M patients

TADPOLE dataset

≃ 1.5k patients

Low data regimes

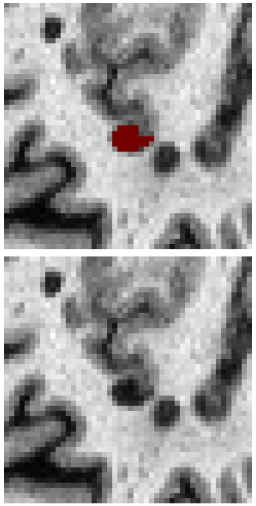


Knowledge uncertainty in MIA

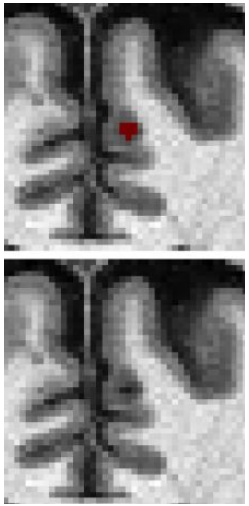
Subtypes of cortical brain lesions

Hidden data biases

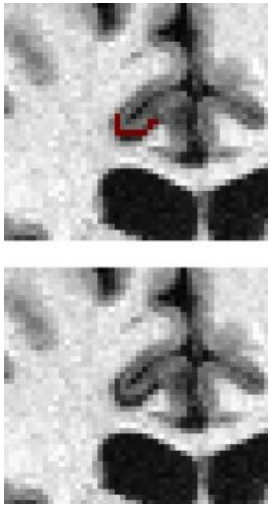
Leukocortical 60%



Intracortical 30%



Subpial <5%

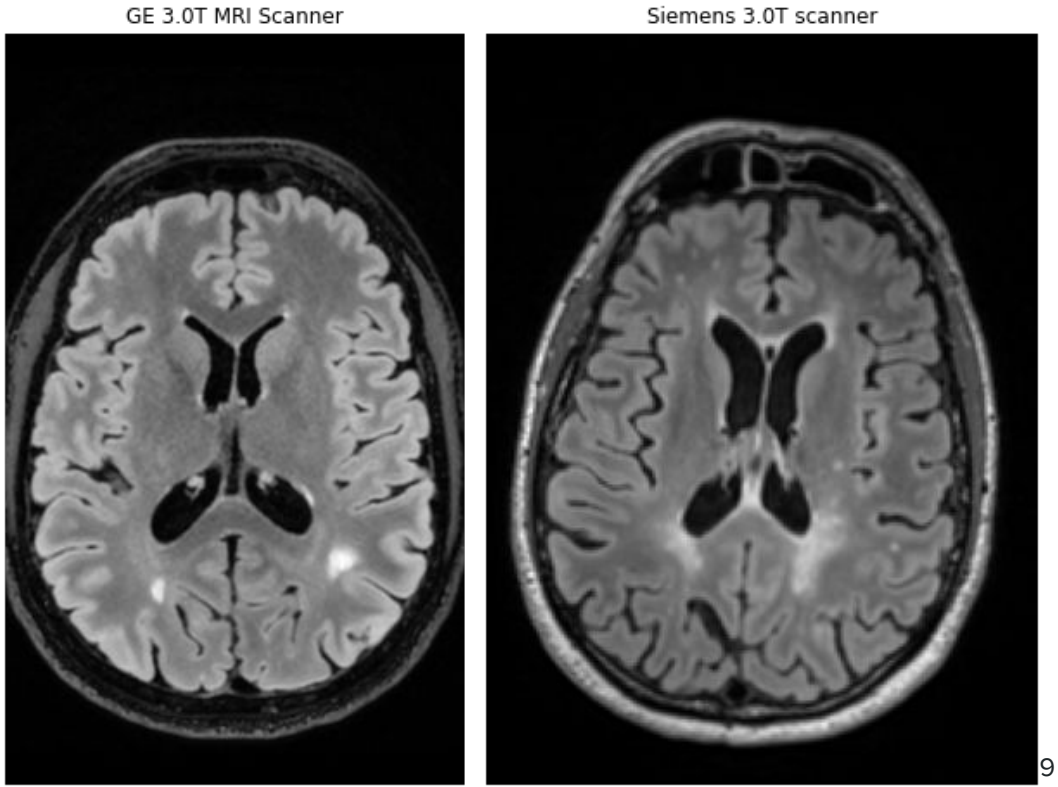


Molchanova et al., MICCAI iMIMIC, 2024

Knowledge uncertainty in MIA

Domain shifts

- Covariate (input data)
- Label
- Concept drift (reality)

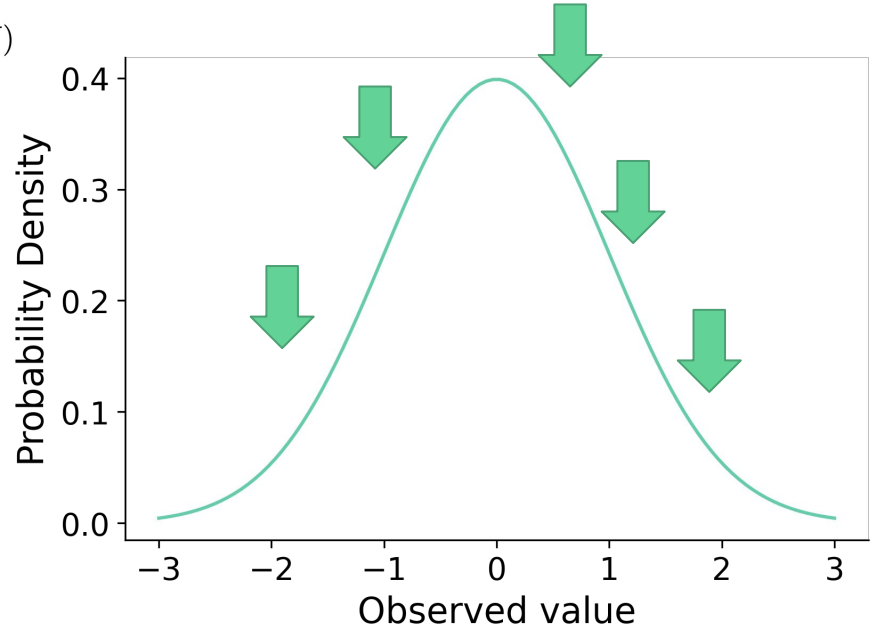


UQ methods in MIA

Bayesian framework

$$p(\mathbf{y}^*|\mathbf{x}^*, \mathbf{X}, \mathbf{Y}) = \int p(\mathbf{y}^*|\mathbf{x}^*, \boldsymbol{\omega}) p(\boldsymbol{\omega}|\mathbf{X}, \mathbf{Y}) \mathrm{d}\boldsymbol{\omega}$$

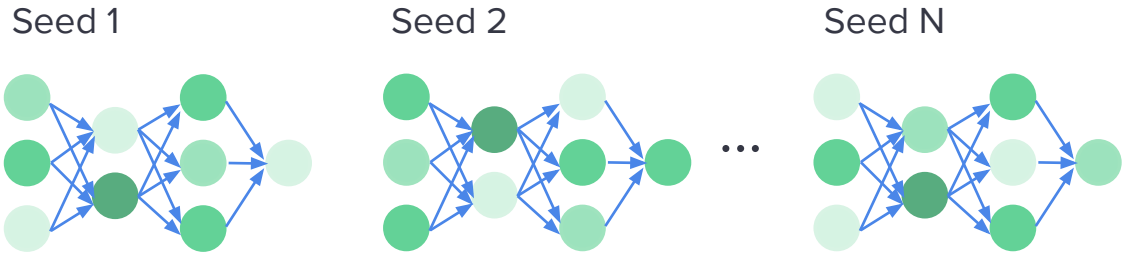
$p(\boldsymbol{\omega}|\mathbf{X}, \mathbf{Y})$



Gal, PhD Thesis, 2019

Deep ensemble

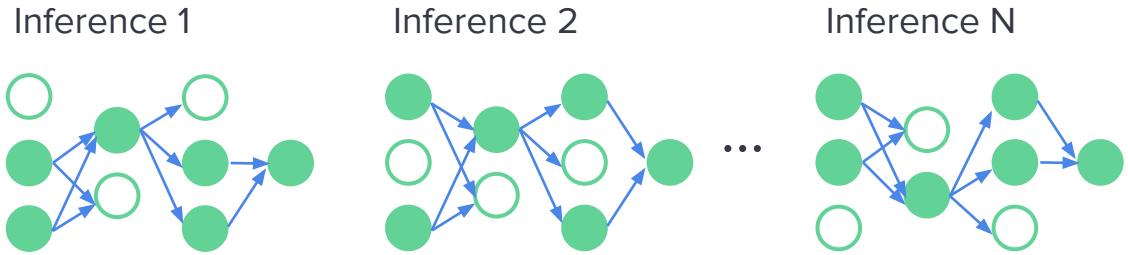
Train N identical neural networks with different random seeds



Lakshminarayanan et al, Neurips, 2017

Monte Carlo dropout

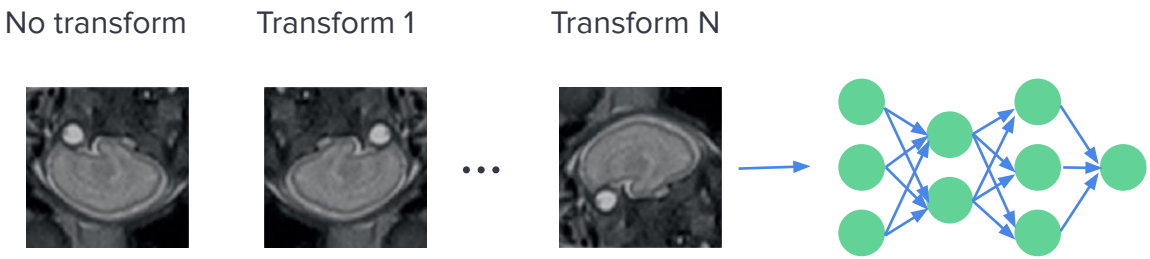
Dropout during the inference time induces a distribution over the weights and biases of the network



Gal & Ghahramani, ICML, 2016

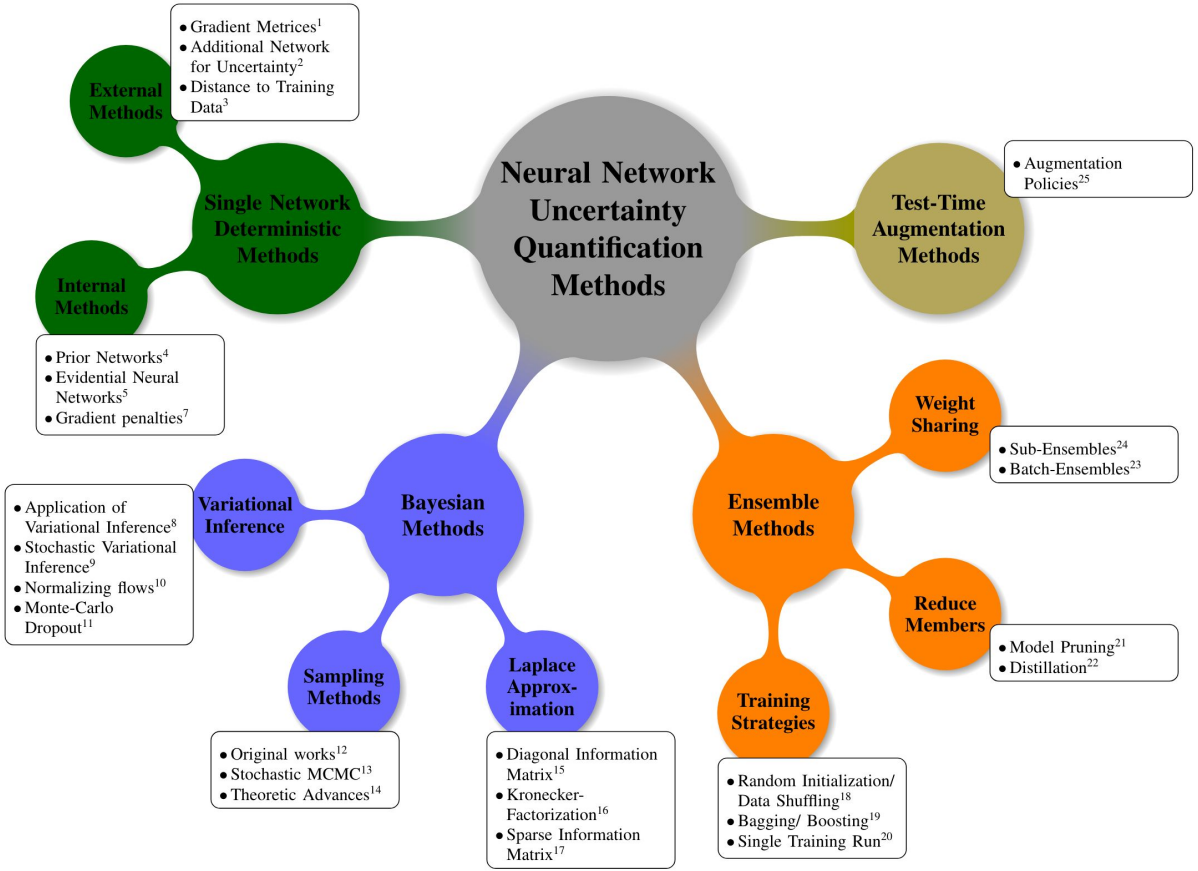
Test time augmentation

Perform several inferences with the same input, but transformed using an invertible transformation



Wang et al, Neurocomputing, 2019

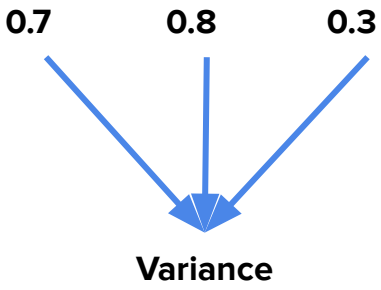
… and many more



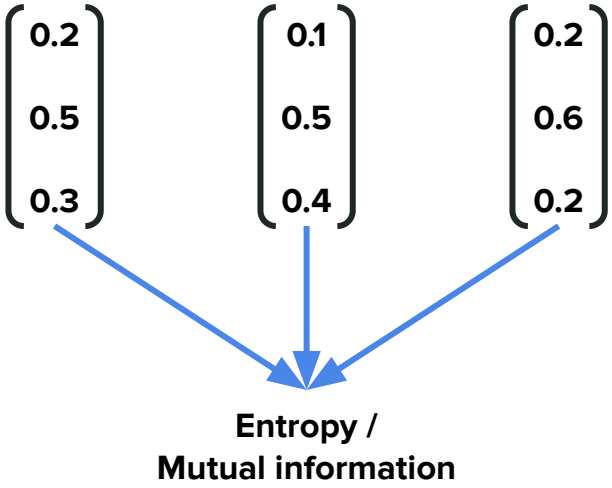
Gawlikowski et al., Artif. Intell. Rev., 2024

Uncertainty measures

Regression

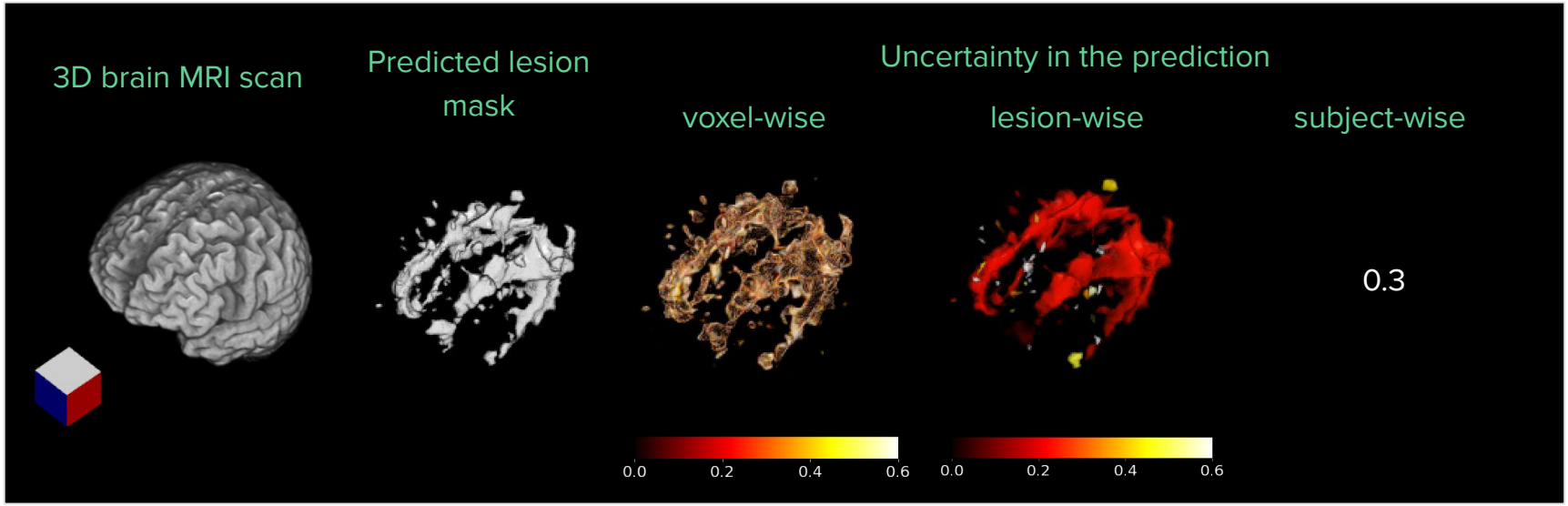


Classification



Uncertainty measures

Reconstruction and segmentation have structured outputs



Molchanova et al., ArXiv, 2023

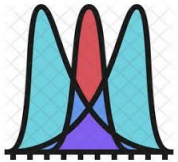
Applications and UQ evaluation

Quality assessment [1]

Trustworthiness [2]

Boosting performance [3]

Domain adaptation [4]



OOD detection [5,6]

Active learning [7,8]

[1] Roy et al., NeuroImage, 2019
[2] Evans et al., FGCS, 2022
[3] Nair et al., Med. Image Anal., 2020
[4] Xia et al., Med. Image Anal., 2020
[5] Linmans et al., Med. Image Anal., 2023
[6] Hong et al., Arxiv, 2024
[7] Budd et al., Med. Image Anal., 2021
[8] Wang et al., Med. Image Anal., 2023

Q&A

UQ in MIA cookbook:

Presentation and hands-on: