

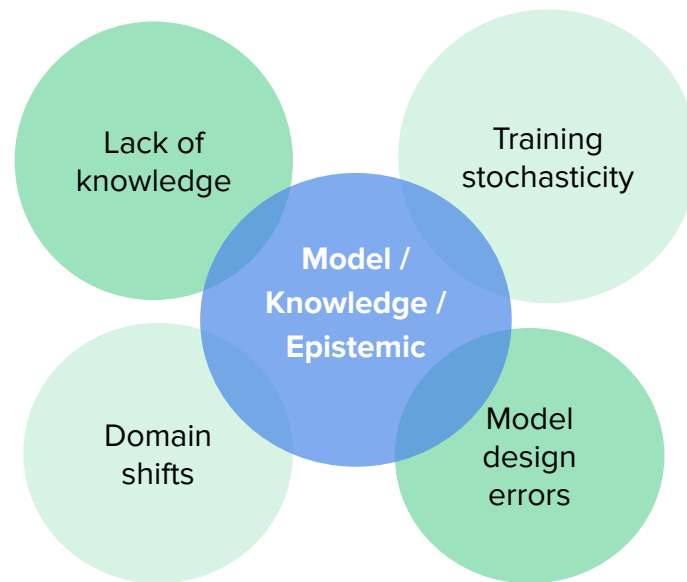
Uncertainty quantification in Medical image analysis

Nataliia Molchanova

Sources of uncertainty

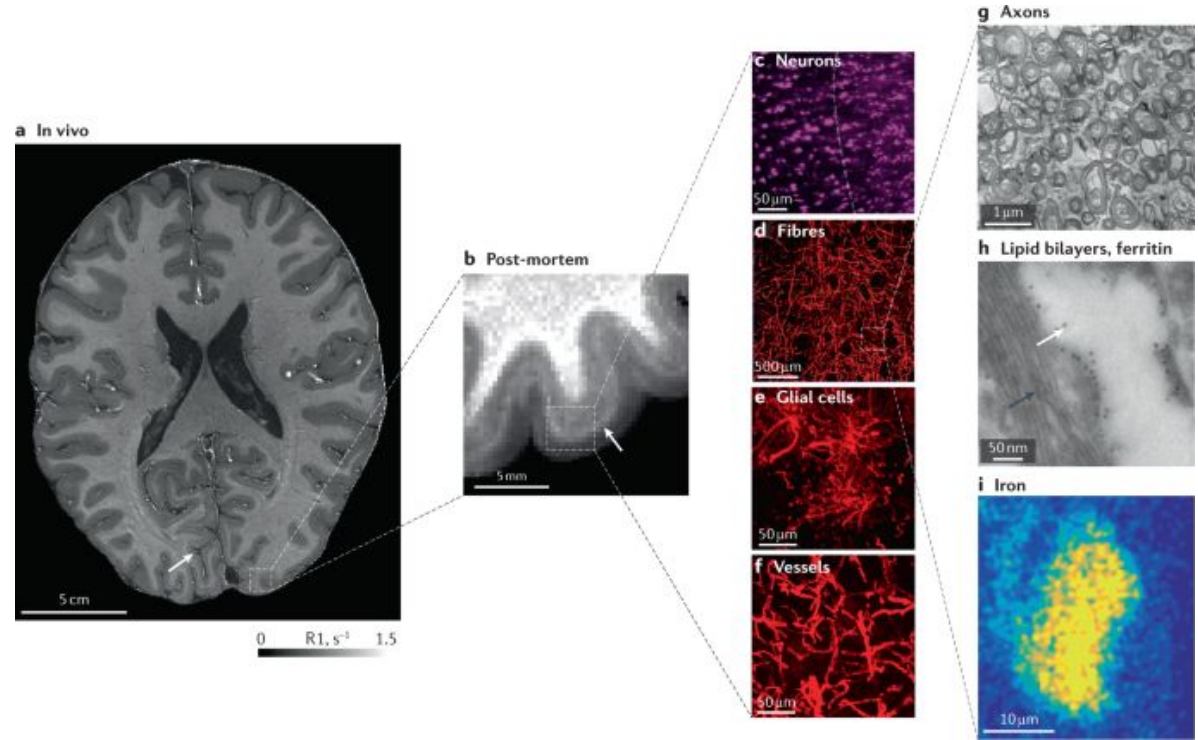


Knowledge / Model / Epistemic



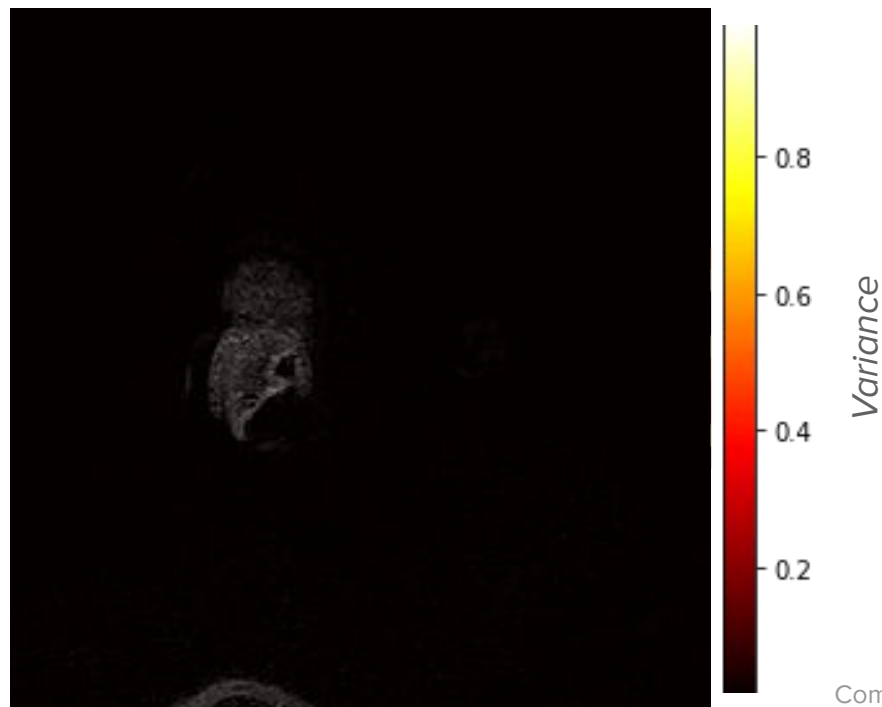
Noise in the input data

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



Inter-rater variability

Example of white matter lesions annotations from 7 raters



Noise in the labels

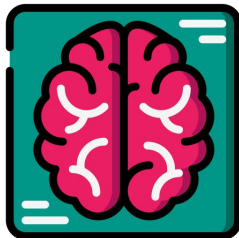


Expertise

Visual
perception

Low quality

Artefacts

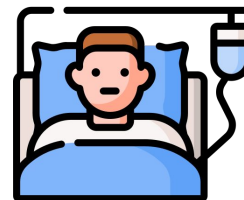


Guidelines
ambiguities

Annotation
tools

Rare
pathologies

Demographic
differences



Knowledge uncertainty in MIA

Low data regimes

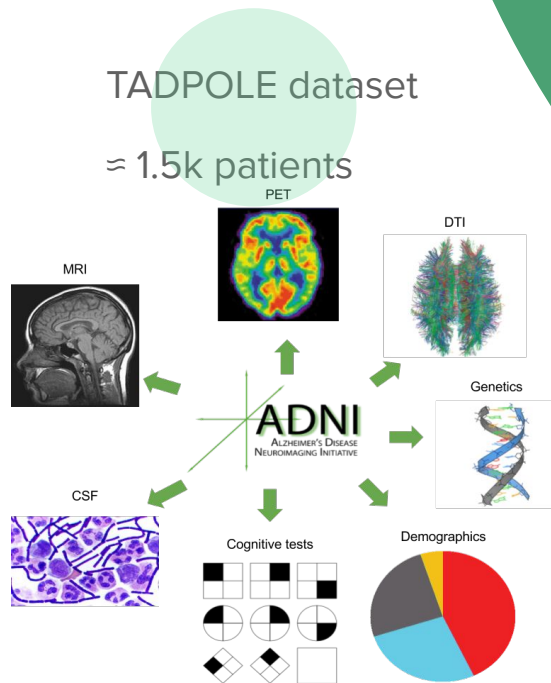
Challenge Name	Task	Medical Imaging Modality	Number of Patients	Countries	Patients Worldwide
KiTS21 Challenge	Kidney Tumor Segmentation	CT	300	USA	400k / Year
MSSEG-1	Multiple Sclerosis Lesion Segmentation	MR	53	France	2.9M
LITS	Liver Tumor Segmentation	CT	130	Germany	800k / Year
PROMISE12	Prostate MR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

Knowledge uncertainty in MIA

Worldwide prevalence

60M patients

Low data regimes



Knowledge uncertainty in MIA

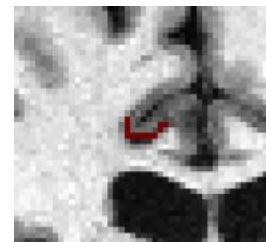
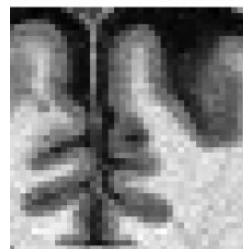
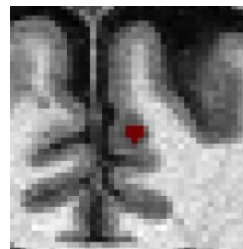
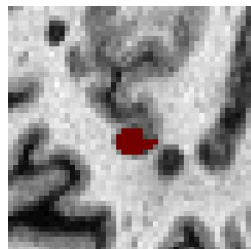
Subtypes of cortical brain lesions

Leukocortical 60%

Intracortical 30%

Subpial <5%

Hidden data biases

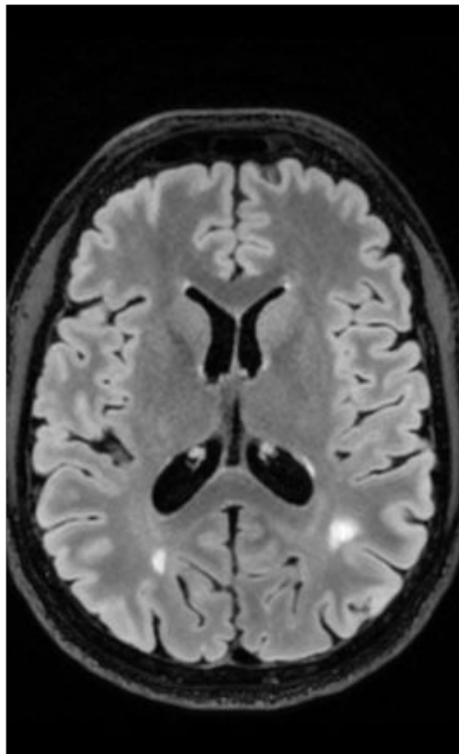


Knowledge uncertainty in MIA

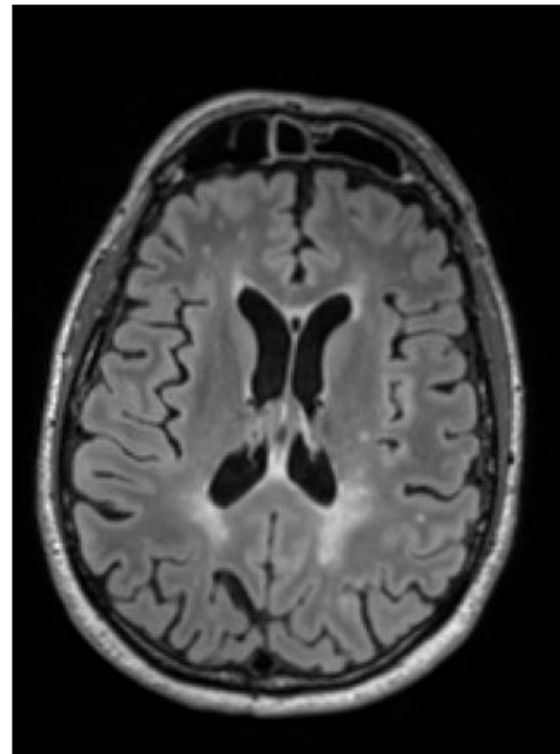
Domain shifts

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

GE 3.0T MRI Scanner



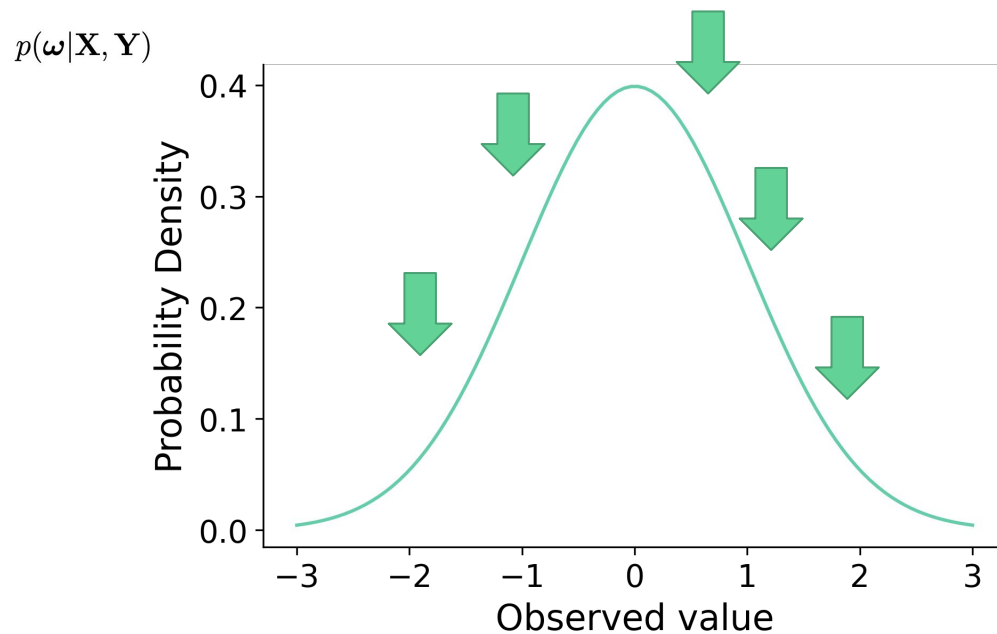
Siemens 3.0T scanner



UQ methods in MIA

Bayesian framework

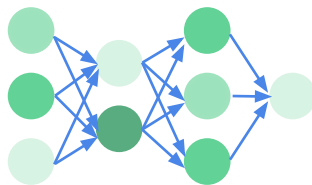
$$p(\mathbf{y}^*|\mathbf{x}^*, \mathbf{X}, \mathbf{Y}) = \int p(\mathbf{y}^*|\mathbf{x}^*, \omega)p(\omega|\mathbf{X}, \mathbf{Y})d\omega$$



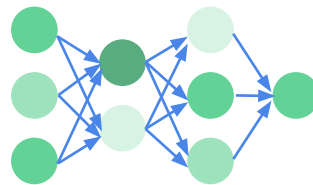
Deep ensemble

Train N identical neural networks with different random seeds

Seed 1

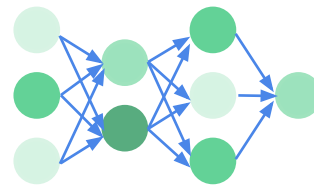


Seed 2



...

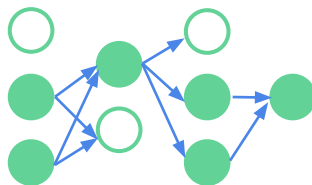
Seed N



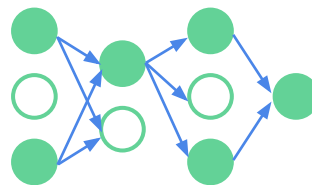
Monte Carlo dropout

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

Inference 1

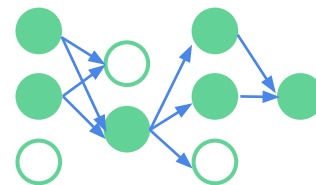


Inference 2



...

Inference N



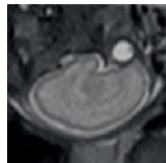
Test time augmentation

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

No transform

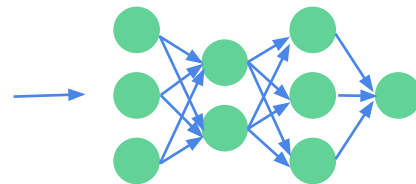
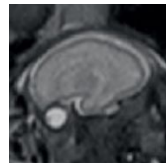


Transform 1

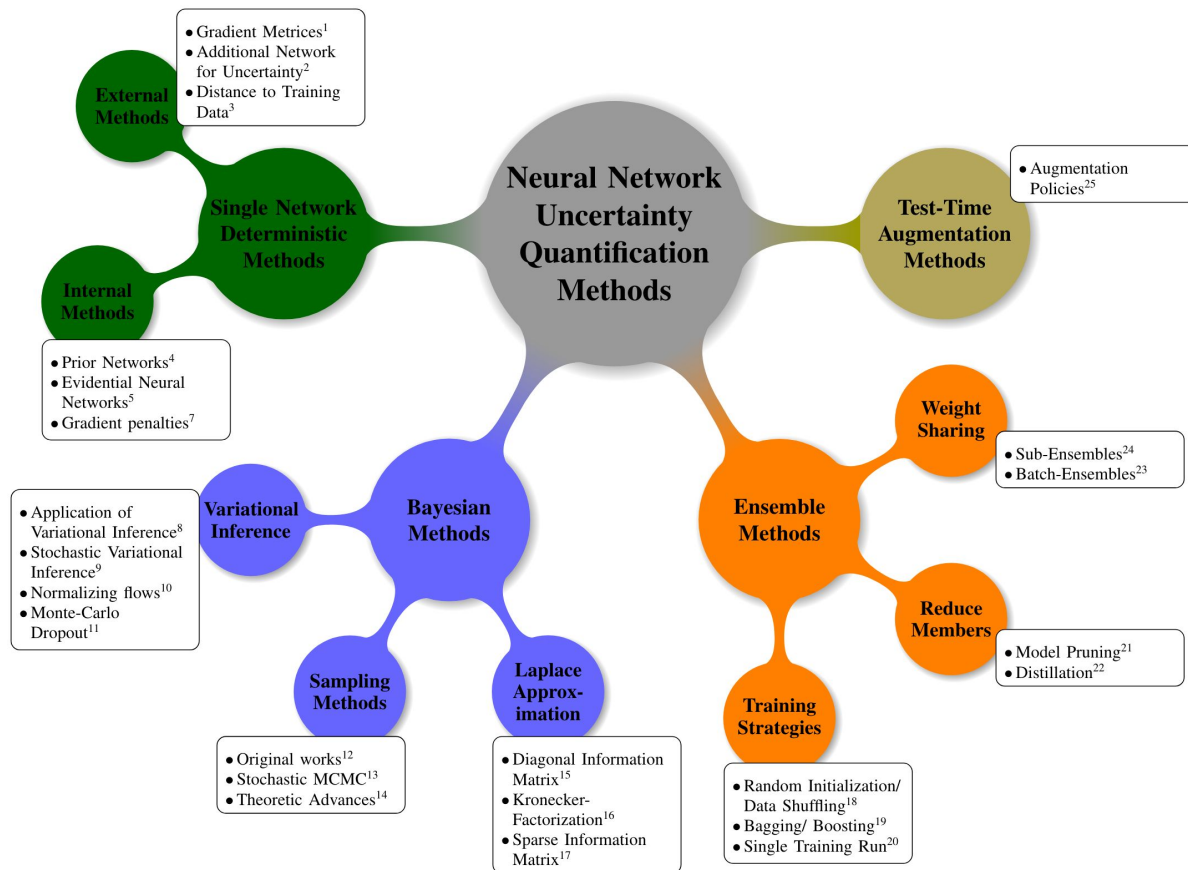


...

Transform N

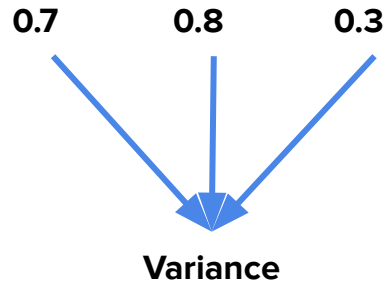


... and many more

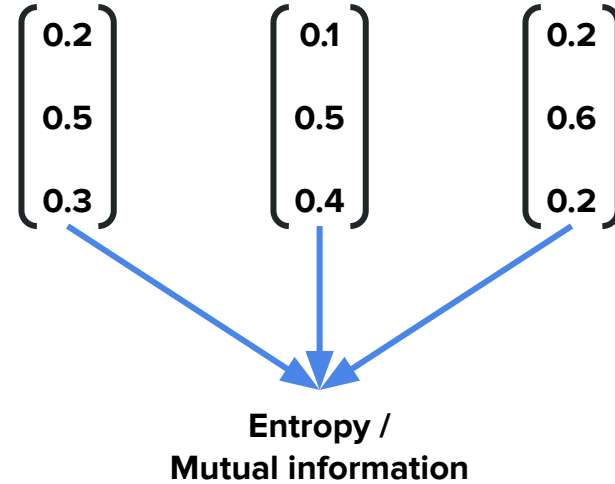


Uncertainty measures

Regression



Classification

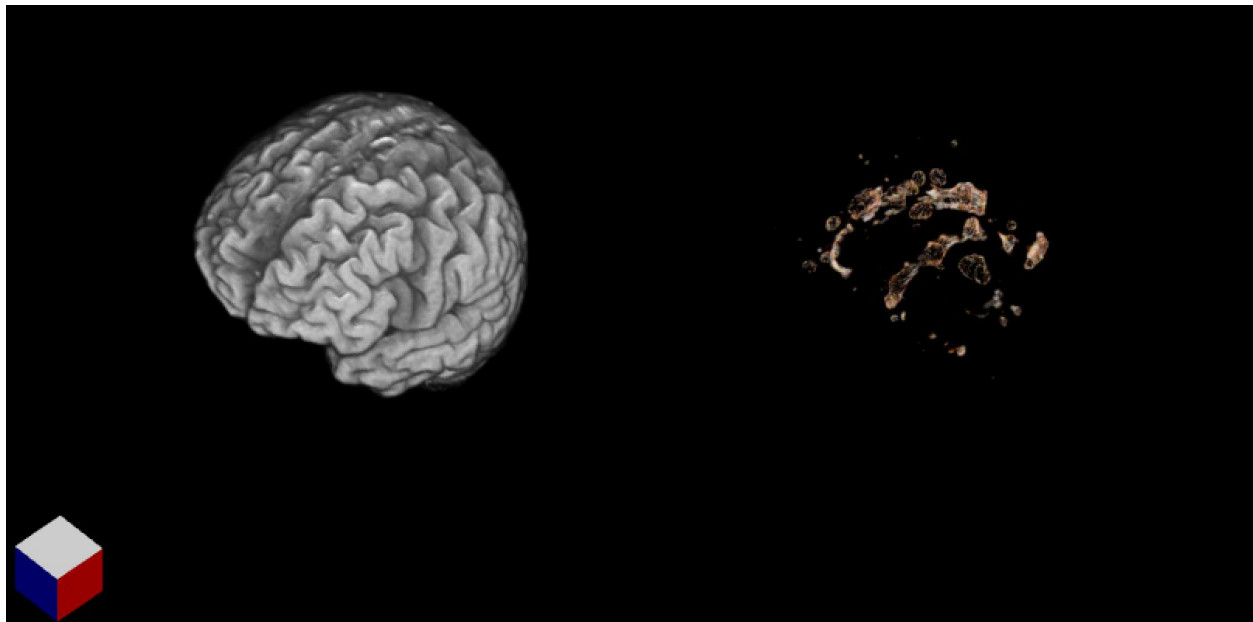


Uncertainty measures

Reconstruction and segmentation have structured outputs

3D brain MRI scan

Voxel-wise uncertainty of a segmentation output



Applications and UQ evaluation



Quality assessment ¹



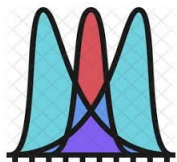
Trustworthiness ²



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:

