O A T M L

McGill
Centre for Intelligent Machines

Uncertainty Evaluation Metric for Brain Tumour Segmentation

Raghav Mehta [1], Angelos Filos [2], Yarin Gal [2], and Tal Arbel [1]

1. Probabilistic Vision Group & Medical Imaging Lab, Centre for Intelligent Machines, McGill University, Canada
2. Oxford Applied and Theoretical Machine Learning Group, University of Oxford, England

@QUBraTS (Quantification of Uncertainty in Brain Tumour Segmentation)

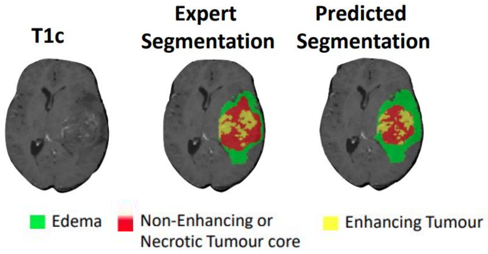
MIDL 2020

Brain Tumour Segmentation

- Automatic tumour segmentation is of clinical importance
  - Diagnose and staging
  - Outcome prediction
  - Surgical planning

Brain Tumour Segmentation

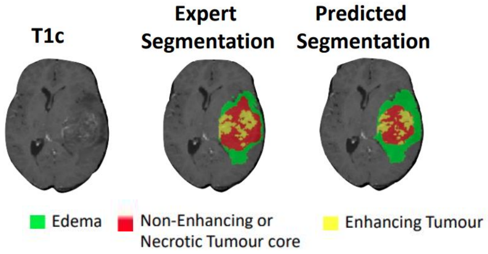
- Automatic tumour segmentation is of clinical importance
  - Diagnose and staging
  - Outcome prediction
  - Surgical planning
- Deep learning models outperform other methods on popular MICCAI BraTS (brain tumour segmentation) challenge



Ronneberger et. al., MICCAI 2015; Cicek et. al., MICCAI 2016; Kamnitsas et. al., MedIA 2016, BrainLes 2016; Isensee et. al., BrainLes 2018; Havaei et. al., MedIA 2017

Brain Tumour Segmentation

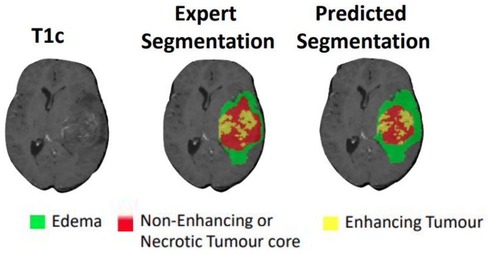
- Automatic tumour segmentation is of clinical importance
  - Diagnose and staging
  - Outcome prediction
  - Surgical planning
- Deep learning models outperform other methods on popular MICCAI BraTS (brain tumour segmentation) challenge
- Tumour Segmentation problem is hard:
  - Large variability in size, shape, position;
  - Subtle boundaries, tumours look like other structures;
  - Sub-tissues can be small (e.g. enhancements);



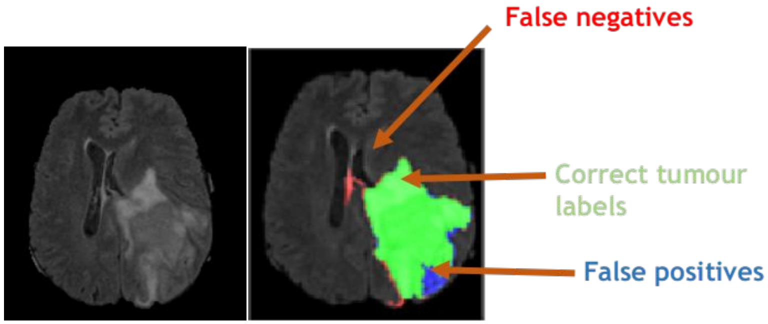
Ronneberger et. al., MICCAI 2015; Cicek et. al., MICCAI 2016; Kamnitsas et. al., MedIA 2016, BrainLes 2016; Isensee et. al., BrainLes 2018; Havaei et. al., MedIA 2017

Brain Tumour Segmentation

- Automatic tumour segmentation is of clinical importance
    - Diagnose and staging
    - Outcome prediction
    - Surgical planning



- Deep learning models outperform other methods on popular MICCAI BraTS (brain tumour segmentation) challenge
- Tumour Segmentation problem is hard:
    - Large variability in size, shape, position;
    - Subtle boundaries, tumours look like other structures;
    - Sub-tissues can be small (e.g. enhancements);
- **Deep learning models can make mistakes!**



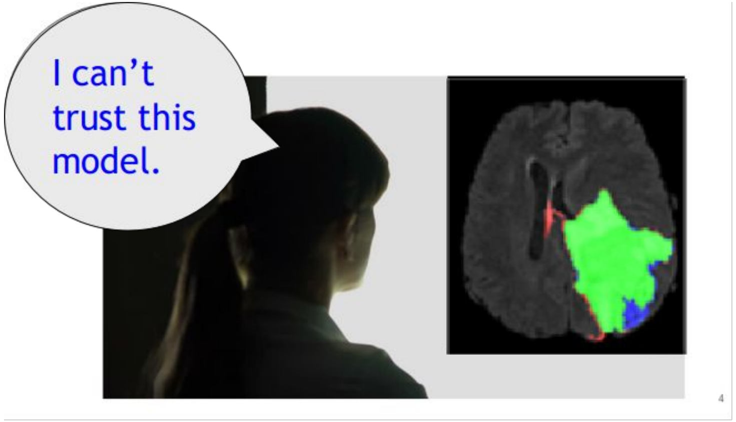
Whole Tumour Segmentation

Ronneberger et. al., MICCAI 2015; Cicek et. al., MICCAI 2016; Kamnitsas et. al., MedIA 2016, BrainLes 2016; Isensee et. al., BrainLes 2018; Havaei et. al., MedIA 2017

@QUBraTS

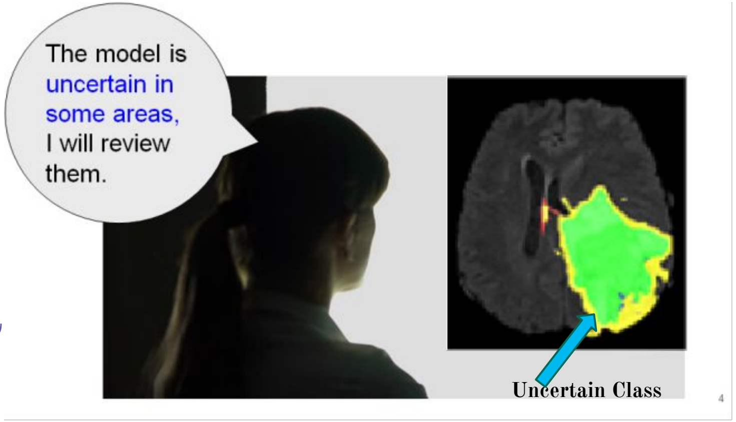
Segmentation of Brain Tumours - Uncertainty

- **Errors in** results of **machine learning** algorithms for segmentation of brain tumours can lead to
  - distrust by clinicians,
  - hesitation in inclusion of machine learning models into clinical workflow



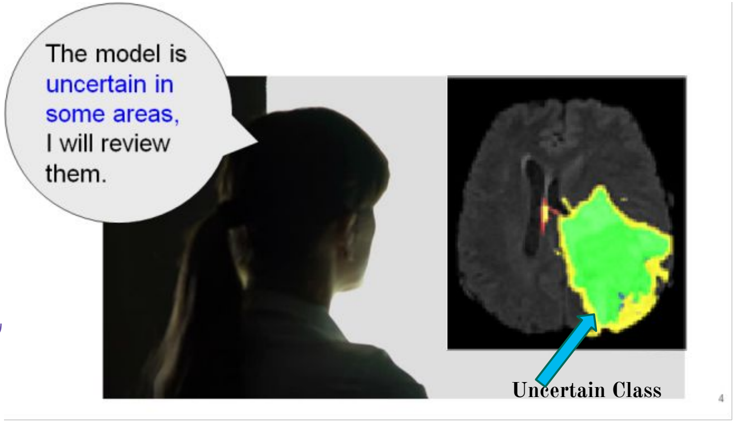
Segmentation of Brain Tumours - Uncertainty

- **Errors in** results of **machine learning** algorithms for segmentation of brain tumours can lead to
    - distrust by clinicians,
    - hesitation in inclusion of machine learning models into clinical workflow
- Uncertainty defining **confidence in results permit clinical review** - bring clinician into the workflow



Segmentation of Brain Tumours - Uncertainty

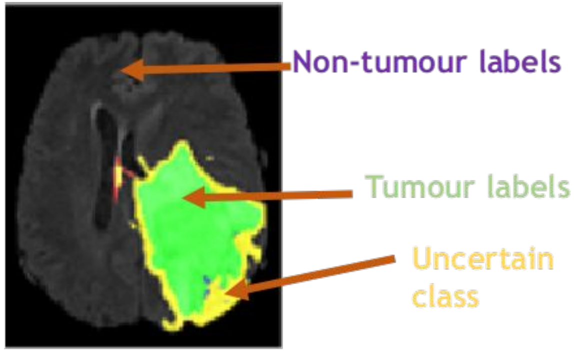
- **Errors in** results of **machine learning** algorithms for segmentation of brain tumours can lead to
  - distrust by clinicians,
  - hesitation in inclusion of machine learning models into clinical workflow
- Uncertainty defining **confidence in results permit clinical review** - bring clinician into the workflow
- **Bayesian Deep Learning** is useful for getting uncertainty [1,2,3]



[1] Gal and Ghahramani, "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning.", ICML 2016.
[2] Kohl et al., "A probabilistic u-net for segmentation of ambiguous images.", NeurIPS 2018.
[3] Lakshminarayanan et al., "Simple and scalable predictive uncertainty estimation using deep ensembles.", NeurIPS 2017.

Uncertainty Analysis: Clinical Adoption

Goal: Uncertainty to enable clinicians, radiologists, surgeons to focus on **reviewing the most uncertain predictions** and **trusting the most confident predictions**



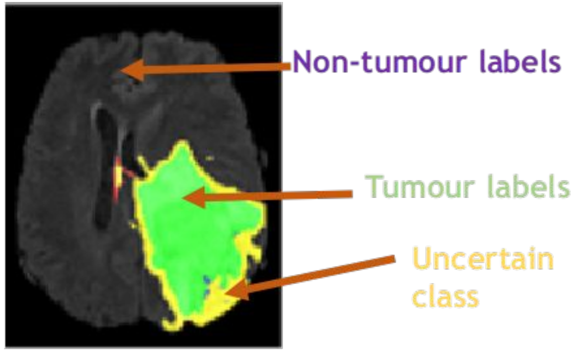
Uncertainty Analysis: Clinical Adoption

Goal: Uncertainty to enable clinicians, radiologists, surgeons to focus on **reviewing the most uncertain predictions** and **trusting the most confident predictions**

- Uncertainty metric **must** have the following properties:

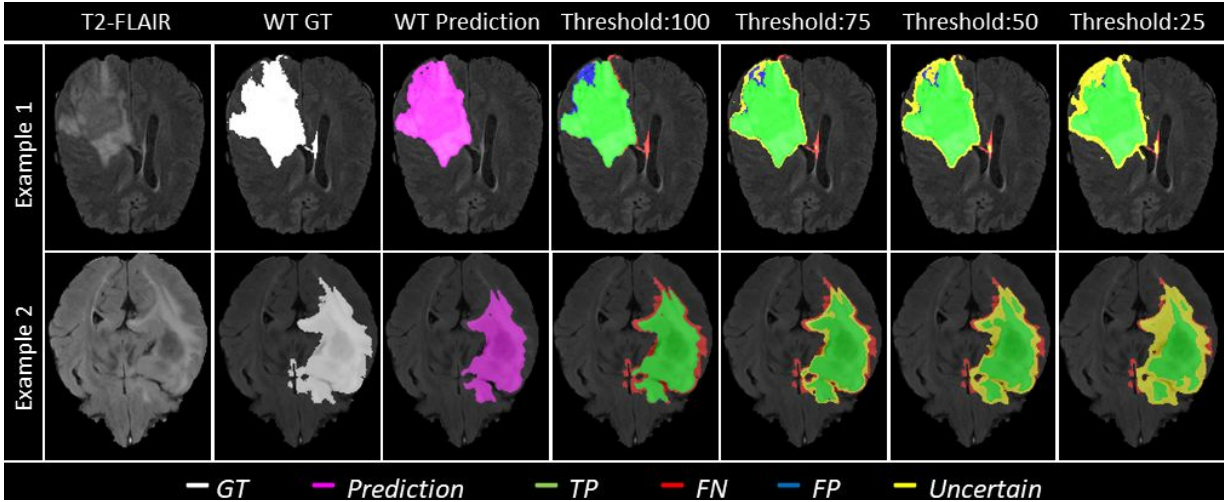
Confident predictions → Correct predictions

Incorrect predictions → Higher Uncertainties



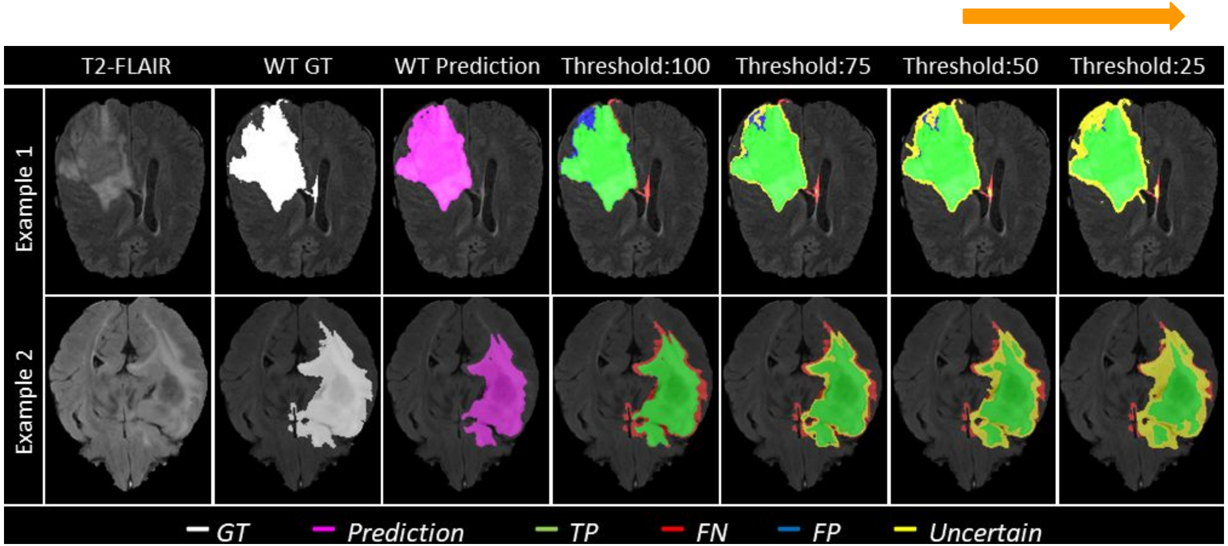
Quantification of Uncertainty for BraTS

- Compute the **uncertainty** of a model **at each voxel**



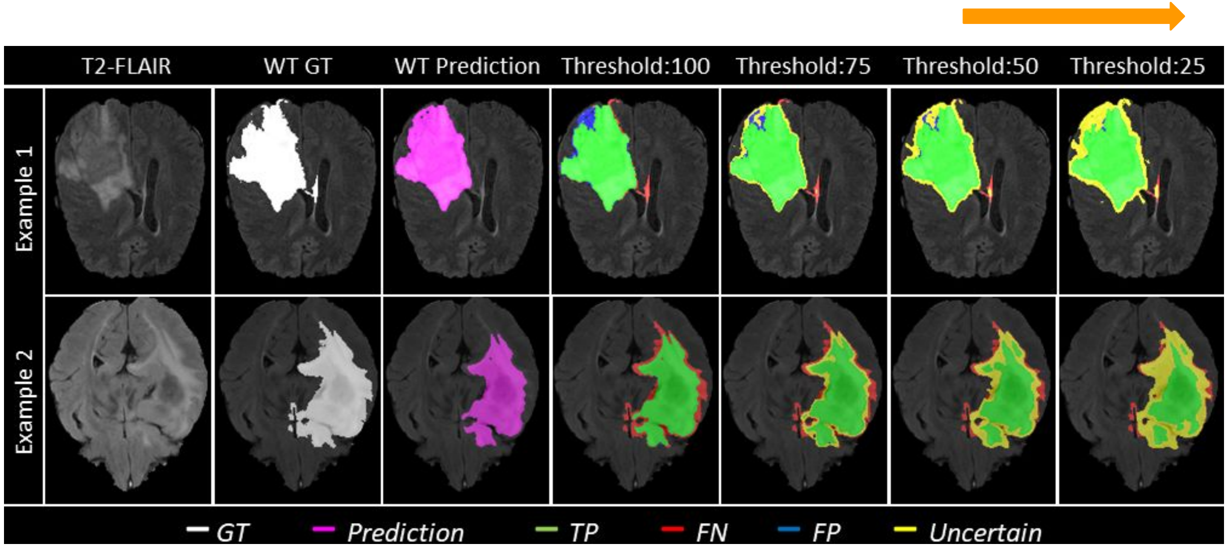
Quantification of Uncertainty for BraTS

- Compute the **uncertainty** of a model **at each voxel**
- Filter most uncertain voxels, calculate the metric of interest (e.g. **Dice**) on the remaining one. **Should Improve!**



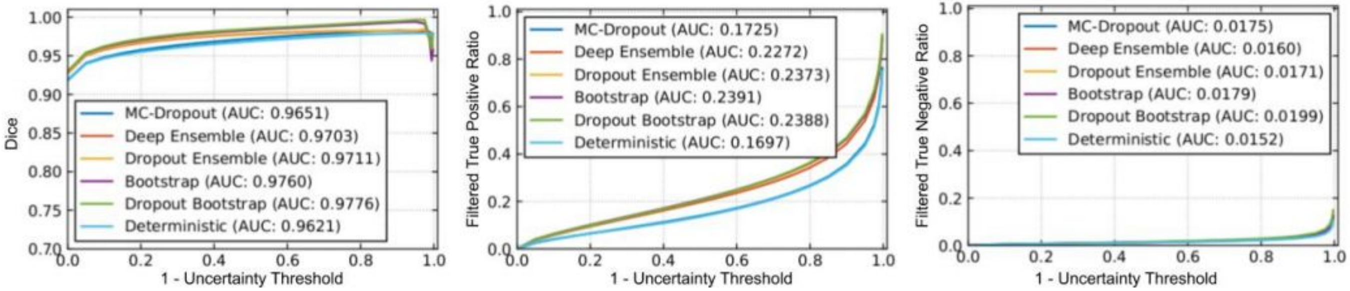
Quantification of Uncertainty for BraTS

- Compute the **uncertainty** of a model **at each voxel**
- Filter most uncertain voxels, calculate the metric of interest (e.g. **Dice**) on the remaining one. **Should Improve!**
- Not at the **expense of filtering out correct predictions!**
  - Penalize methods for higher filtering of correct predictions.



Benchmark Results (Entropy - whole tumour)

- 3D U-Net architecture [1]
- Brain Tumour Segmentation (BraTS) 2019 [2] Training set (335):
- Performances of whole tumour segmentation with the Entropy uncertainty measure [3]
- Comparison of various uncertainty generation methods:
  - MC-Dropout [4]
  - Deep Ensemble [5]
  - Dropout Ensemble [6]
  - Bootstrap
  - Bootstrap Ensemble



1. Cicek et al., MICCAI 2016
2. Bakas et al., arXiv:1811.02629, 2018
3. Gal et al., ICML 2017
4. Gal and Ghahramani, ICML 2016
5. Lakshminarayanan et al., NeurIPS 2017
6. Smith and Gal, arXiv:1803.08533

CIFAR | ICRA
CANADIAN INSTITUTE FOR ADVANCED RESEARCH | INSTITUT CANADIEN DE RECHERCHES AVANCÉES

NSERC
CRSNG

SYNAPTIVE MEDICAL

Thank You

@QUBraTS