Propagating Uncertainty Across Cascaded Medical Imaging Tasks For Improved Deep Learning Inference

Raghav Mehta [1*], Thomas Christinck [1*], Tanya Nair [1], Paul Lemaitre [1], Douglas L. Arnold [2,3], and Tal Arbel [1]

1 Centre for Intelligent Machines, McGill University, Montreal, Canada
2 Montreal Neurological Institute, McGill University, Montreal, Canada
3 NeuroRx Research, Montreal, Canada

UNSURE 2019: **Un**certainty for **S**afe **U**tilization of Machine Lea**r**ning in M**e**dical Imaging

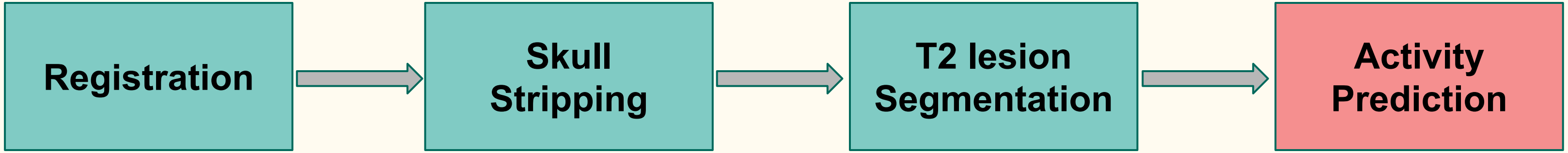
Introduction

- Medical Image analysis pipeline performs sequence of inference task before the task of interest

Introduction

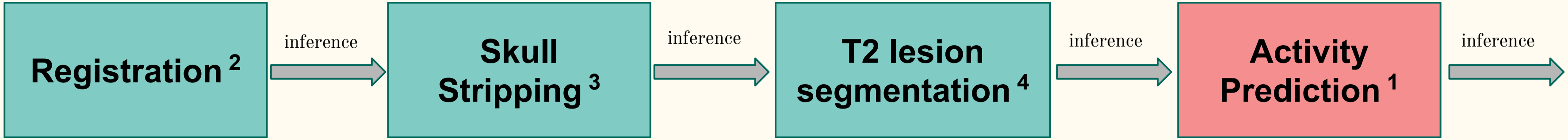
- Medical Image analysis pipeline performs sequence of inference tasks before the task of interest

- Ex. Multiple Sclerosis (MS) disease activity prediction [1]



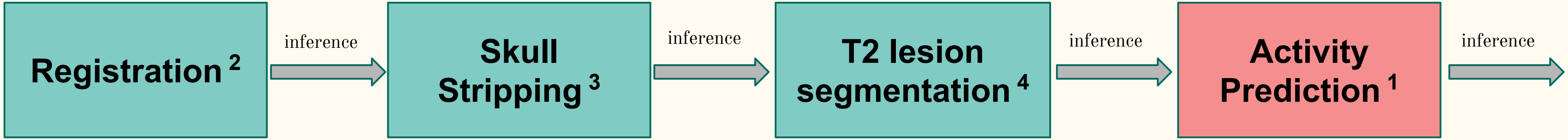
[1] Sepahvand et al. “CNN Prediction of Future Disease Activity for Multiple Sclerosis Patients from Baseline MRI and Lesion Labels.”, Brainlesion 2018

Introduction



[1] Sepahvand et al. "CNN Prediction of Future Disease Activity for Multiple Sclerosis Patients from Baseline MRI and Lesion Labels.", Brainlesion 2018.
[2] Fan et al., "Adversarial learning for mono-or multi-modal registration.", Medical image analysis 2019.
[3] Kleesiek et al. "Deep MRI brain extraction: a 3D convolutional neural network for skull stripping.", NeuroImage 2016.
[4] Nair et al., "Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.", Medical Image Analysis 2019.

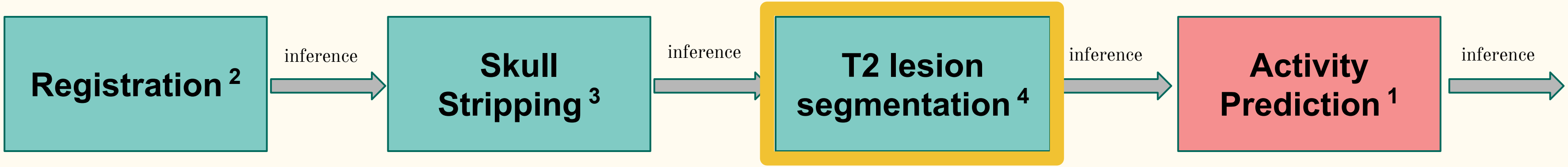
Introduction



- Deep Learning based solutions provide only deterministic output
- Errors can accumulate over sequence of tasks
- This can hinder downstream task

[1] Sepahvand et al. “CNN Prediction of Future Disease Activity for Multiple Sclerosis Patients from Baseline MRI and Lesion Labels.”, Brainlesion 2018.
[2] Fan et al., “Adversarial learning for mono-or multi-modal registration.”, Medical image analysis 2019.
[3] Kleesiek et al. “Deep MRI brain extraction: a 3D convolutional neural network for skull stripping.”, NeuroImage 2016.
[4] Nair et al., “Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.”, Medical Image Analysis 2019.

Introduction

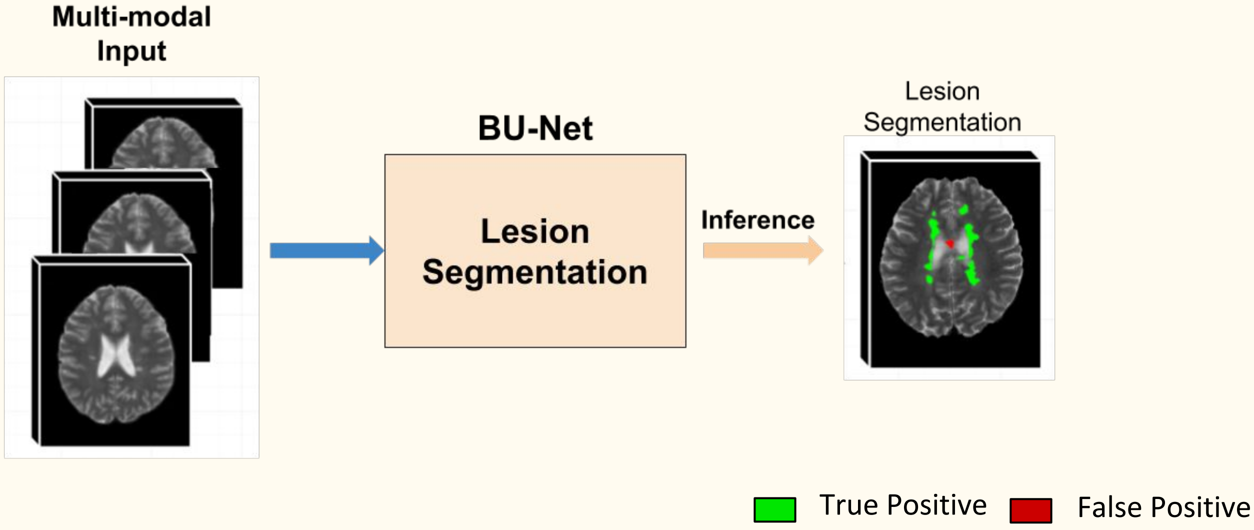


- Deep Learning based solutions provide only deterministic output
- Errors can accumulate over sequence of tasks
- This can hinder downstream task

[1] Sepahvand et al. “CNN Prediction of Future Disease Activity for Multiple Sclerosis Patients from Baseline MRI and Lesion Labels.”, Brainlesion 2018.
[2] Fan et al., “Adversarial learning for mono-or multi-modal registration.”, Medical image analysis 2019.
[3] Kleesiek et al. “Deep MRI brain extraction: a 3D convolutional neural network for skull stripping.”, NeuroImage 2016.
[4] Nair et al., “Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.”, Medical Image Analysis 2019.

Introduction

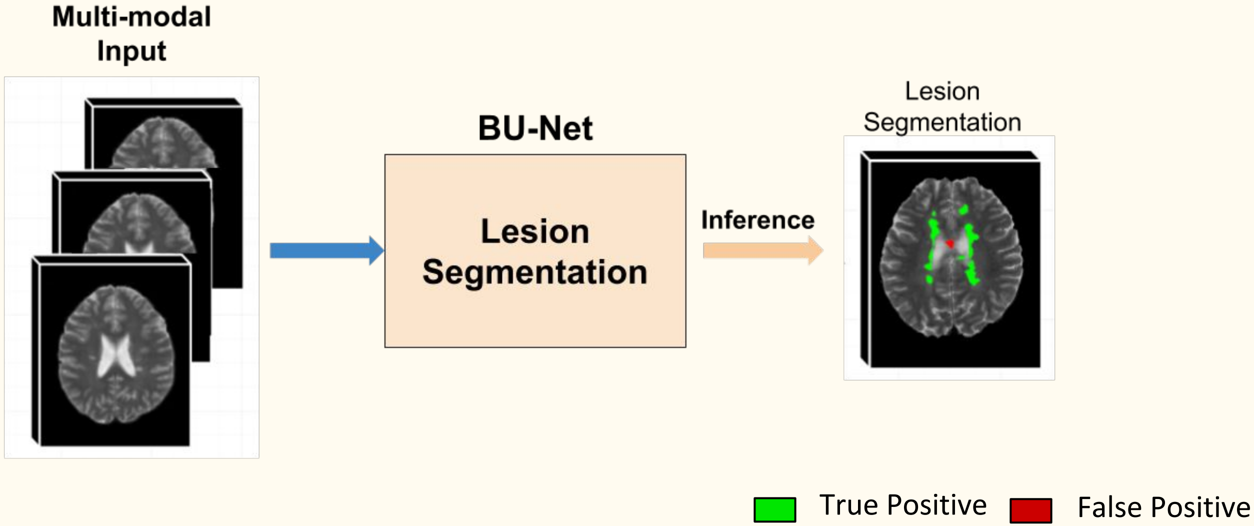
- MS T2 lesion segmentation using Bayesian U-Net (BU-Net) [4]



[4] Nair et al., "Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.", Medical Image Analysis 2019.

Introduction

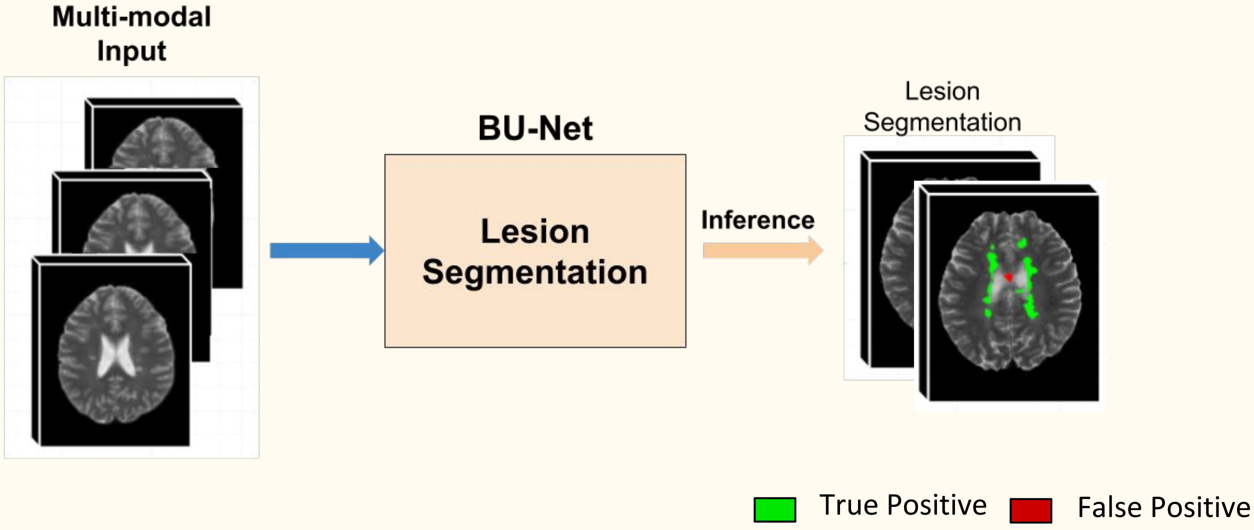
- MS T2 lesion segmentation using Bayesian U-Net (BU-Net) [4]
    - Uncertainty Estimation using Monte-Carlo Dropout (MC-Dropout) [5]



[4] Nair et al., "Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.", Medical Image Analysis 2019.
[5] Gal and Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning.", ICML 2016.

Introduction

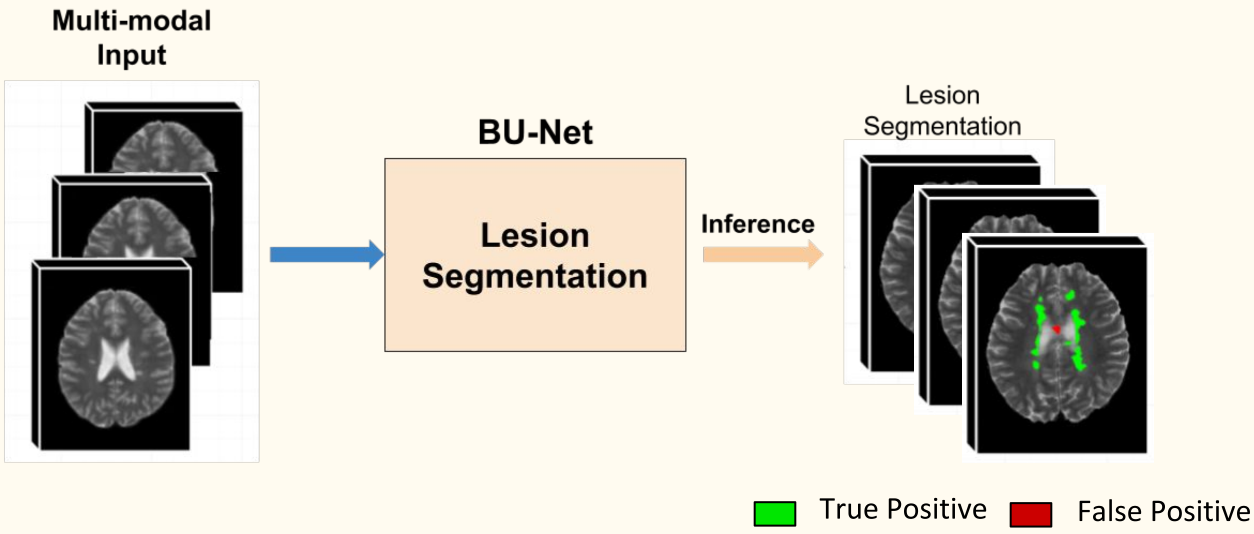
- MS T2 lesion segmentation using Bayesian U-Net (BU-Net) [4]
    - Uncertainty Estimation using Monte-Carlo Dropout (MC-Dropout) [5]



[4] Nair et al., "Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.", Medical Image Analysis 2019.
[5] Gal and Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning.", ICML 2016.

Introduction

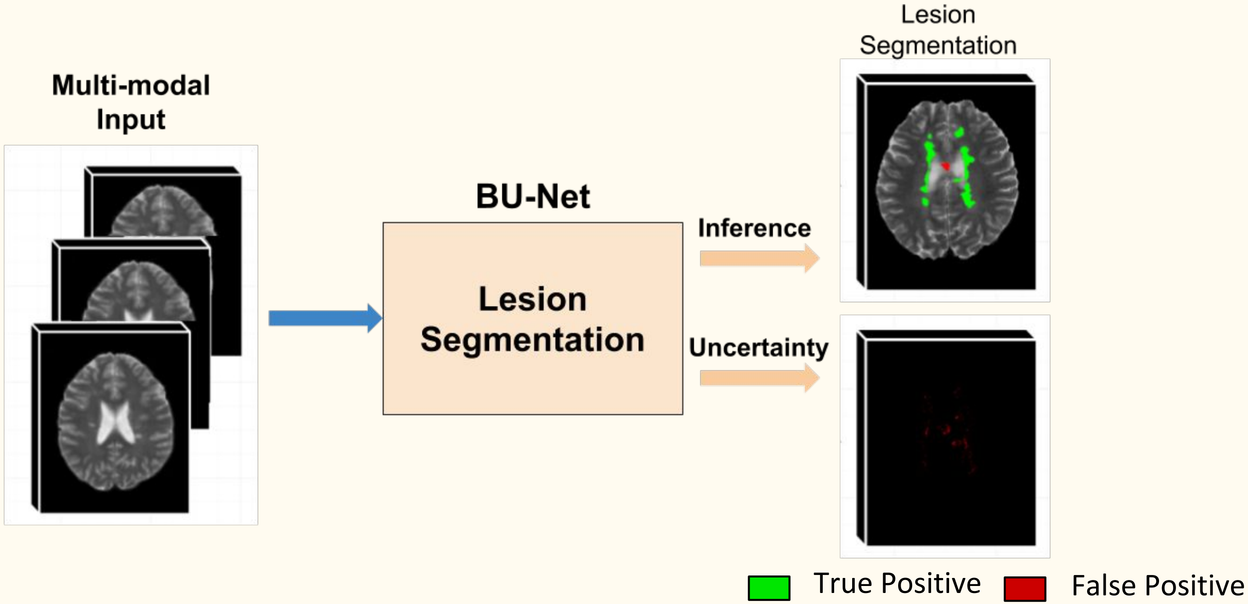
- MS T2 lesion segmentation using Bayesian U-Net (BU-Net) [4]
  - Uncertainty Estimation using Monte-Carlo Dropout (MC-Dropout) [5]



[4] Nair et al., "Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.", Medical Image Analysis 2019.
[5] Gal and Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning.", ICML 2016.

Introduction

- MS T2 lesion segmentation using Bayesian U-Net (BU-Net) [4]
  - Uncertainty Estimation using Monte-Carlo Dropout (MC-Dropout) [5]



[4] Nair et al., "Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.", Medical Image Analysis 2019.
[5] Gal and Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning.", ICML 2016.

Introduction

- Bayesian Deep Learning provides uncertainty estimation
  - Monte-Carlo (MC) Dropout [5]
  - Variational Dropout [6]
  - Probabilistic U-Net [7]
  - Deep Ensemble [8]
  - ...

[5] Gal and Ghahramani, “Dropout as a bayesian approximation: Representing model uncertainty in deep learning.”, ICML 2016.
[6] Kingma et al., “Variational dropout and the local reparameterization trick.”, NeurIPS 2015.
[7] Kohl et al., “A probabilistic u-net for segmentation of ambiguous images.”, NeurIPS 2018.
[8] Lakshminarayanan et al., “Simple and scalable predictive uncertainty estimation using deep ensembles.”, NeurIPS 2017.

Introduction

- Applied to different medical image analysis context contexts
  - MS T2 lesion segmentation and detection [4]
  - Lung cancer lesion segmentation [9]
  - Modality Synthesis [10]
  - dMRI Super-Resolution [11]
  - Brain structure segmentation [12]
  - MR registration [13]
  - Diabetic Ratinopathy Screening [14]
  - ...

[4] Nair et al., “Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.”, Medical Image Analysis 2019
[9] Hu et al., “Supervised uncertainty quantification for segmentation with multiple annotations.”, MICCAI 2019.
[10] Mehta et al., “RS-Net: Regression-Segmentation 3D CNN for Synthesis of Full Resolution Missing Brain MRI in the Presence of Tumours.”, SASHIMI 2018.
[11] Tanno et al., “Bayesian Image Quality Transfer with CNNs: Exploring Uncertainty in dMRI Super-Resolution.”, MICCAI 2017.
[12] Roy et al. “Bayesian QuickNAT: Model uncertainty in deep whole-brain segmentation for structure-wise quality control.”, NeuroImage 2019.
[13] Dalca et al., “Unsupervised Learning of Probabilistic Diffeomorphic Registration for Images and Surfaces.”, Medical Image Analysis 2019.
[14] Leibig et al. “Leveraging uncertainty information from deep neural networks for disease detection.”, Scientific reports 2017

Introduction

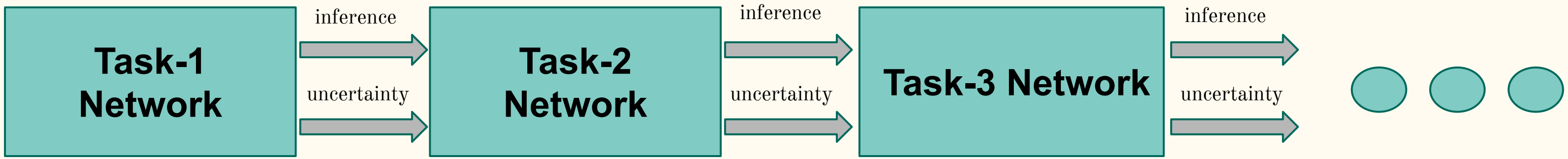
- Applied to different medical image analysis context contexts
  - MS T2 lesion segmentation and detection [4]
  - Lung cancer lesion segmentation [9]
  - Modality Synthesis [10]
  - dMRI Super-Resolution [11]
  - Brain structure segmentation [12]
  - MR registration [13]
  - Diabetic Ratinopathy Screening [14]
  - ...
- Papers report that
  - Areas where network is prone to error have higher uncertainty [10,11,13]
  - Improved performance when the network output is evaluated on its most certain predictions [4, 14]

[4] Nair et al., "Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.", Medical Image Analysis 2019
[9] Hu et al., "Supervised uncertainty quantification for segmentation with multiple annotations.", MICCAI 2019.
[10] Mehta et al., "RS-Net: Regression-Segmentation 3D CNN for Synthesis of Full Resolution Missing Brain MRI in the Presence of Tumours.", SASHIMI 2018.
[11] Tanno et al., "Bayesian Image Quality Transfer with CNNs: Exploring Uncertainty in dMRI Super-Resolution.", MICCAI 2017.
[12] Roy et al. "Bayesian QuickNAT: Model uncertainty in deep whole-brain segmentation for structure-wise quality control.", NeuroImage 2019.
[13] Dalca et al., "Unsupervised Learning of Probabilistic Diffeomorphic Registration for Images and Surfaces.", Medical Image Analysis 2019.
[14] Leibig et al. "Leveraging uncertainty information from deep neural networks for disease detection.", Scientific reports 2017

Can we use this uncertainty to improve downstream task?

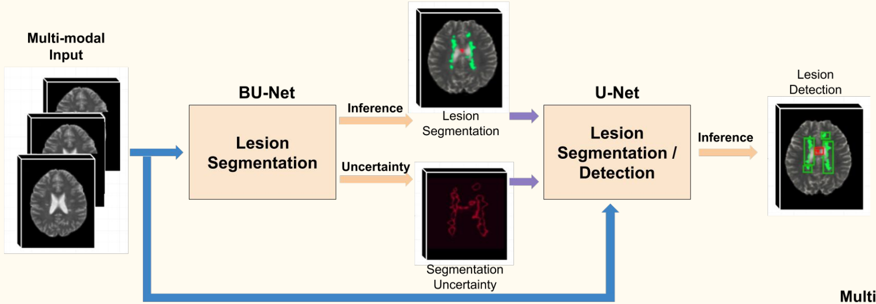
Proposed Framework

- Leveraging Uncertainty for improved inference in cascaded medical image analysis task



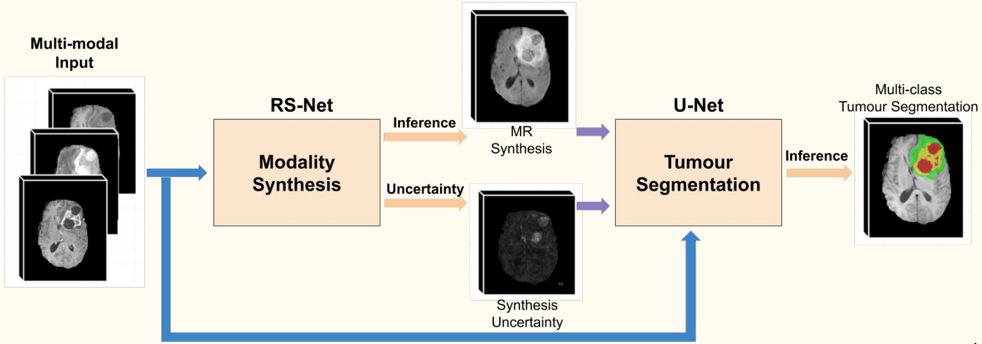
Proposed Framework

- Leveraging Uncertainty for improved inference in cascaded medical image analysis task



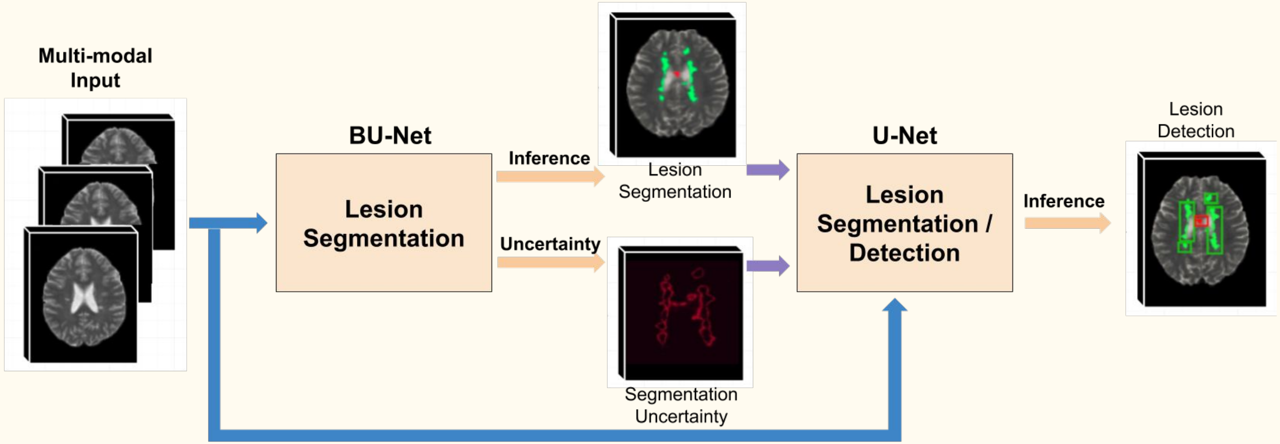
MS T2 lesion segmentation detection

Brain Tumour Segmentation



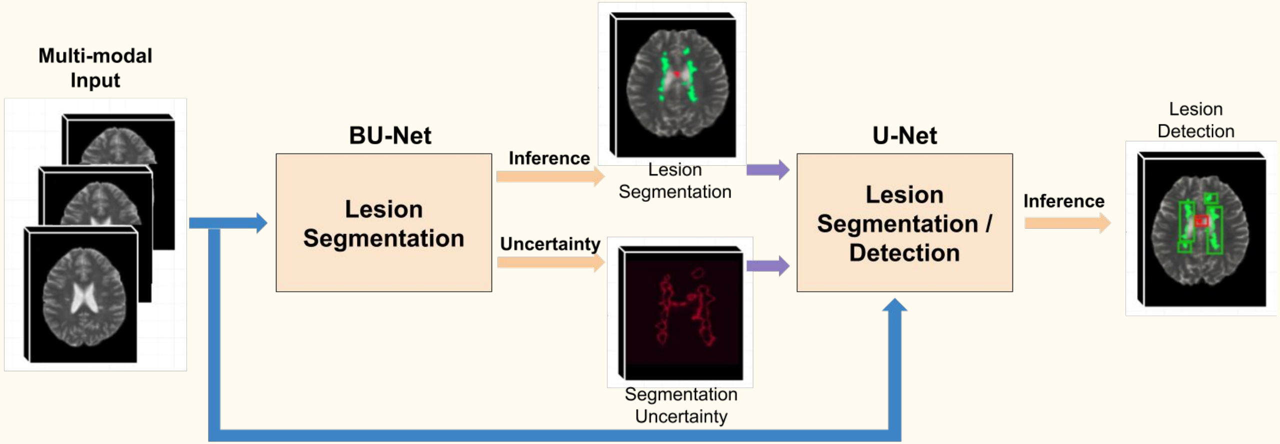
MS T2 Lesion Segmentation/Detection Pipeline

- **Task:** Accurate MS T2 lesion segmentation/detection



MS T2 Lesion Segmentation/Detection Pipeline

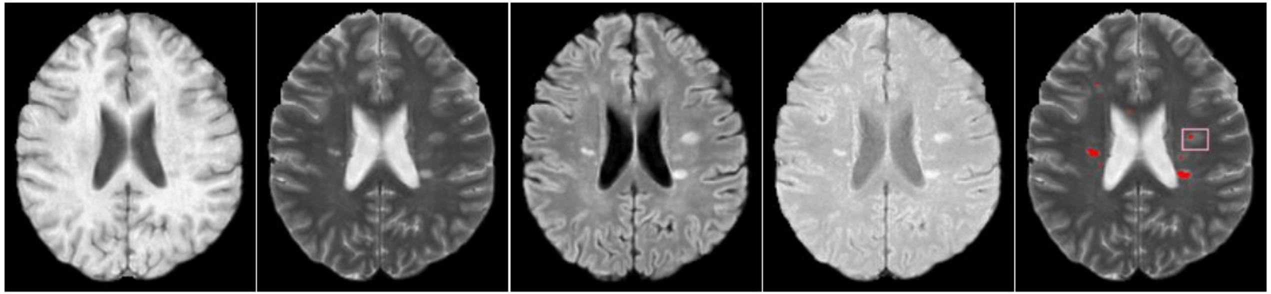
- **Task:** Accurate MS T2 lesion segmentation/detection
- **Task-1 Network:** Bayesian U-Net (BU-Net) [4] for lesion segmentation
- **Task-2 Network:** 3D U-Net [15] for segmentation
- MC-Dropout [5] to estimate uncertainty in BU-Net



[4] Nair et al., "Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.", Medical Image Analysis 2019.
[5] Gal and Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning.", ICML 2016.
[15] Cicek et al., "3D U-Net: learning dense volumetric segmentation from sparse annotation.", MICCAI 2016.

MS T2 Lesion Segmentation/Detection Pipeline

- **Dataset**
  - Proprietary multi-site, multi-scanner patient MRI from 2 clinical trials of patients with relapsing-remitting MS (RRMS)
  - 5800 multi-modal MRI (T1,T2, FLAIR, PD)
  - Expert T2 lesion labels
    - 40% of the available data used to train/validate BU-Net
    - 50% of the remaining to train 3D U-Net
    - 10% to test 3D U-Net



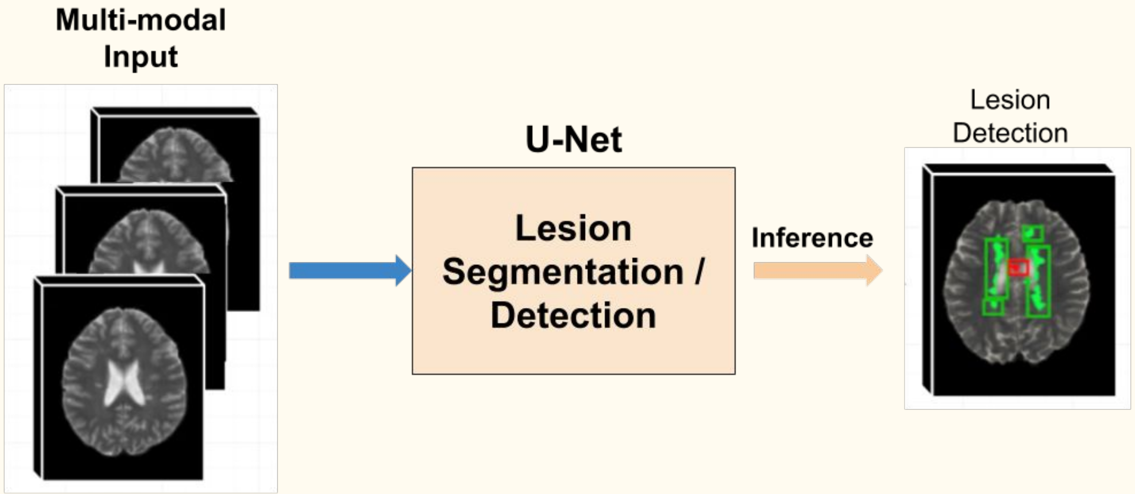
(a) T1w (b) T2w (c) FLAIR (d) PDW (e) Expert labels

MS T2 Lesion Segmentation/Detection Pipeline

- **Evaluation Metric**
  - Accurate detection of MS T2 lesion is of interest
  - Segmentation converted into lesion detection with connected component analysis
  - Lesions divided into 3 categories based on size.
    - Small (3-10 voxels) --- 40% of total lesions are small
    - Medium (11-50 voxels)
    - Large (50+ voxels)
  - Receiver operating characteristic (ROC) curves for each lesion size and for all lesions combined
    - Area under the curve (AUC) of ROC curve
    - True Positive Rate (TPR) at False detection rate (FDR) of 0.2

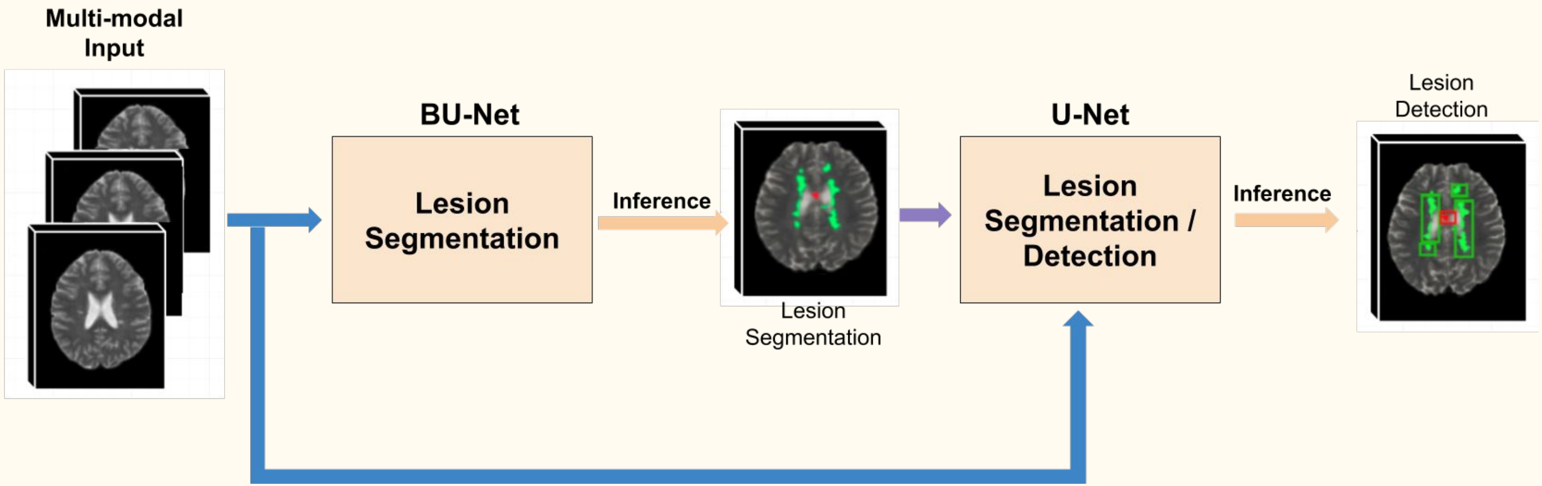
MS T2 Lesion Segmentation/Detection Pipeline

- **Baseline-1**
  - No Task-1 Network (BU-Net)



MS T2 Lesion Segmentation/Detection Pipeline

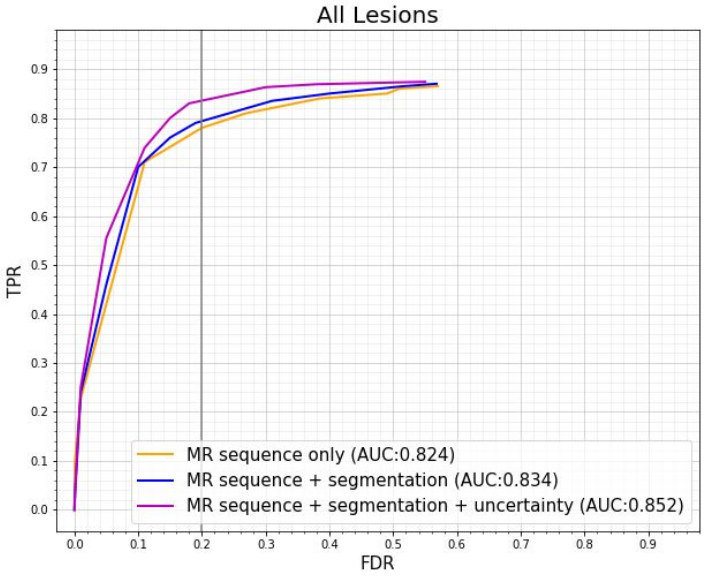
- **Baseline-2**
  - Only inference from Task-1 Network (BU-Net) is propagated to Task-2 Network (3D U-Net)



McGill
Centre for
Intelligent Machines

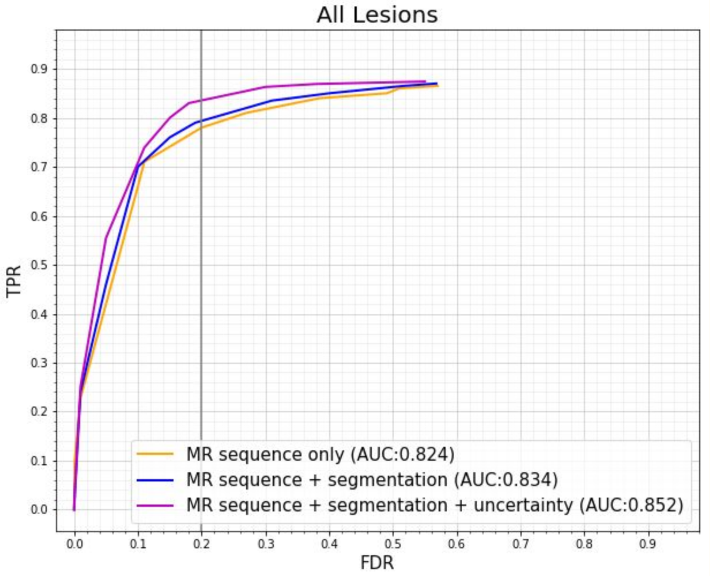
MS T2 Lesion Segmentation/Detection Pipeline

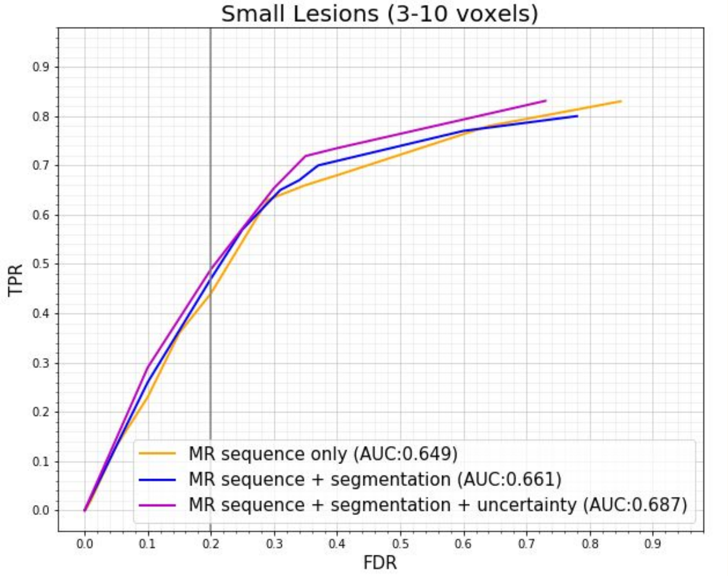
- **Quantitative Results**



MS T2 Lesion Segmentation/Detection Pipeline

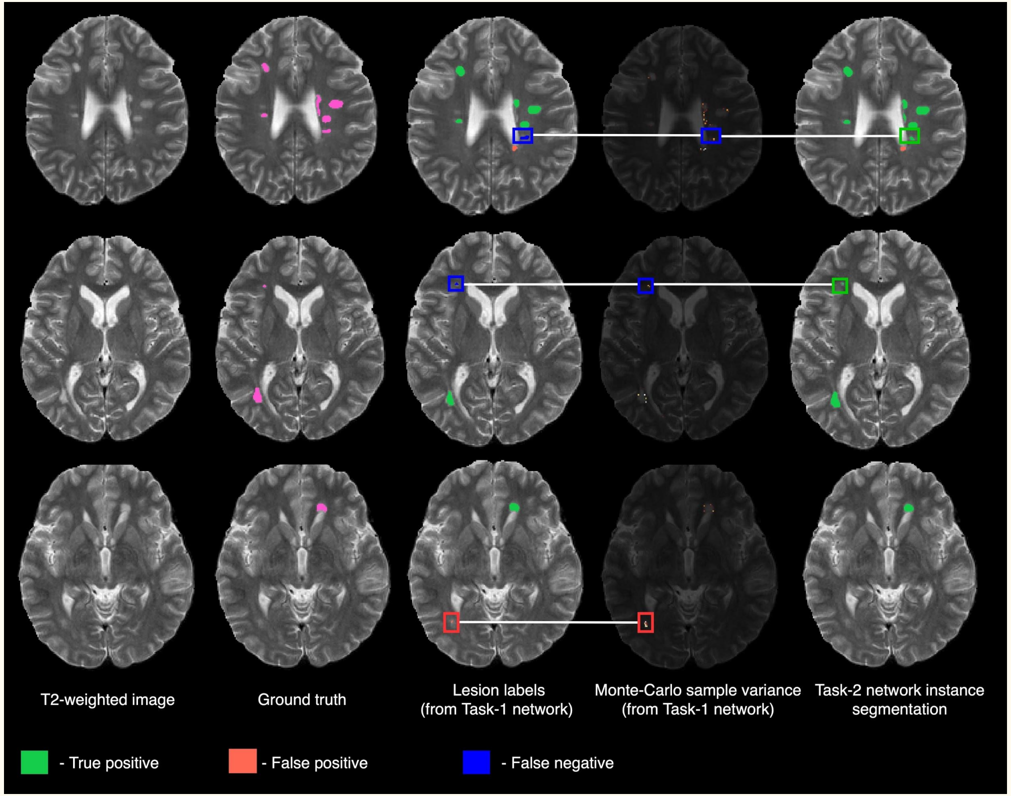
- **Quantitative Results**





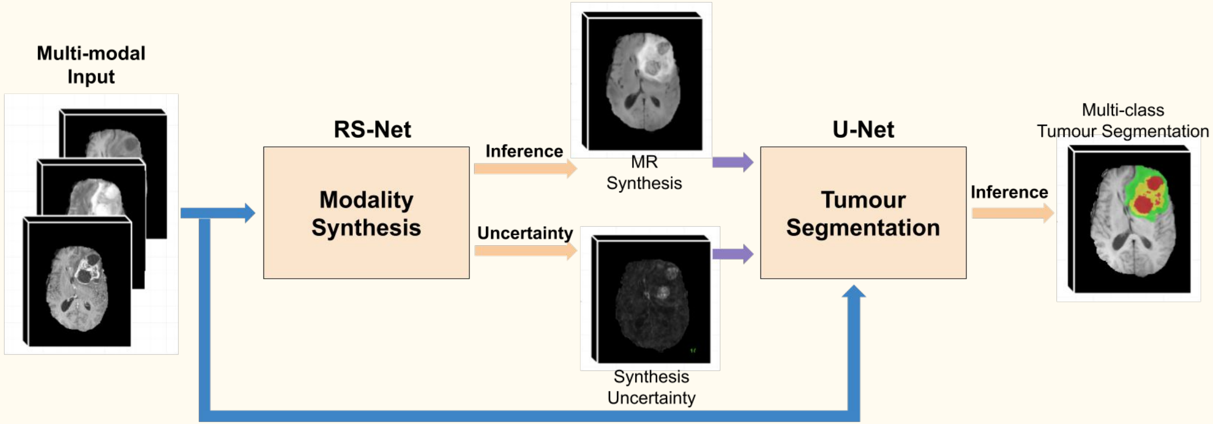
MS T2 Lesion Segmentation/Detection Pipeline

- Qualitative Results



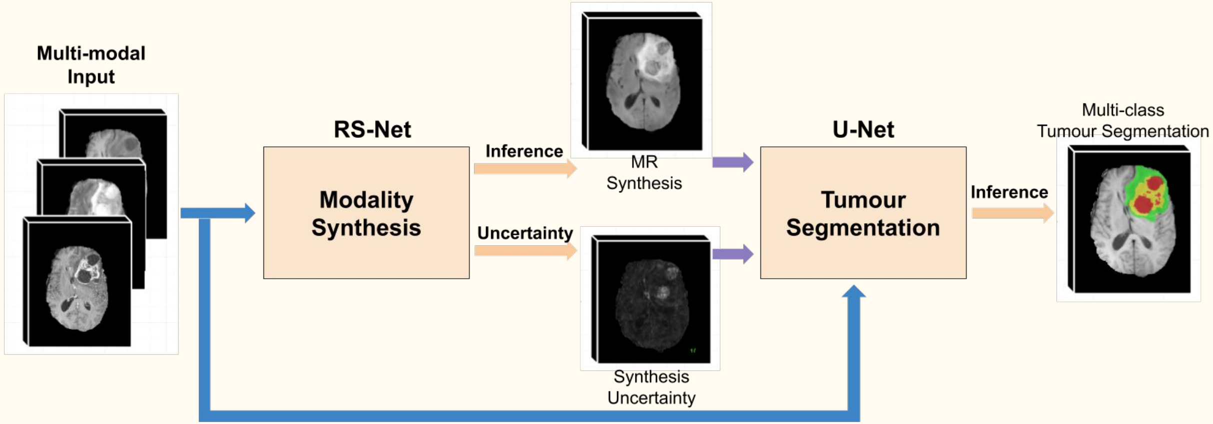
Brain Tumour Segmentation Pipeline

- **Task:** Accurate multi-class tumour segmentation in case of missing modality



Brain Tumour Segmentation Pipeline

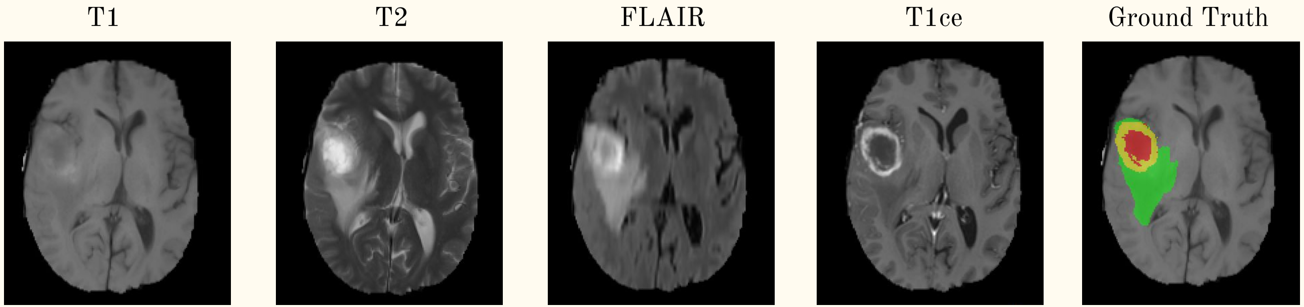
- **Task:** Accurate multi-class tumour segmentation in case of missing modality
- **Task-1 Network:** Regression-Segmentation Network (RS-Net) [10] for modality synthesis
- **Task-2 Network:** 3D U-Net [15] for multi-class brain tumour segmentation
- MC-Dropout [5] to estimate uncertainty in RS-Net



[10] Mehta et al., “RS-Net: Regression-Segmentation 3D CNN for Synthesis of Full Resolution Missing Brain MRI in the Presence of Tumours.”, SASHIMI 2018.
[5] Gal and Ghahramani, “Dropout as a bayesian approximation: Representing model uncertainty in deep learning.”, ICML 2016.
[15] Cicek et al., “3D U-Net: learning dense volumetric segmentation from sparse annotation.”, MICCAI 2016.

Brain Tumour Segmentation Pipeline

- **Dataset**
  - Brain Tumour Segmentation (BraTS) 2018 [16] challenge dataset
  - Multi-class tumour segmentation ground truth
    - Edema
    - Enhancing Tumour
    - Non-enhancing core
  - Multi-modal MRI (T1, T2, FLAIR, and T1ce)
    - BraTS 2018 Training set to train and validate RS-Net and 3D U-Net (285 patients)
    - BraTS 2018 Validation set (held-out) to test 3D U-Net (66 patients)



[16] S. Bakas, et al.: "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge", arXiv preprint arXiv:1811.02629 (2018)

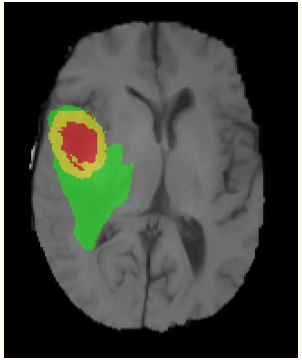
McGill
Centre for
Intelligent Machines

Brain Tumour Segmentation Pipeline

- **Evaluation Metric** [16]
  - Dice scores for three different tumour subtypes:
    - enhancing tumour (DE)
    - whole tumour (DT)
    - tumour core (DC)

$$Dice(G, P) = \frac{2|GP|}{|G| + |P|}$$

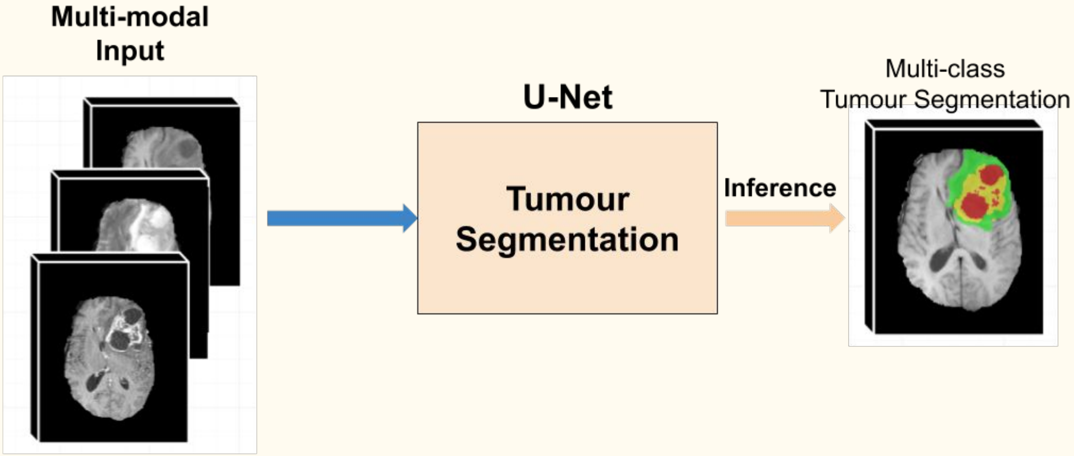
where |G| denotes the number of positive elements in the binary segmentation G and |GP| is the number of shared positive elements by G and P. Dice∈[0, 1]. A higher Dice value indicates a better segmentation performance.



[16] S. Bakas, et al.: "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge", arXiv preprint arXiv:1811.02629 (2018)

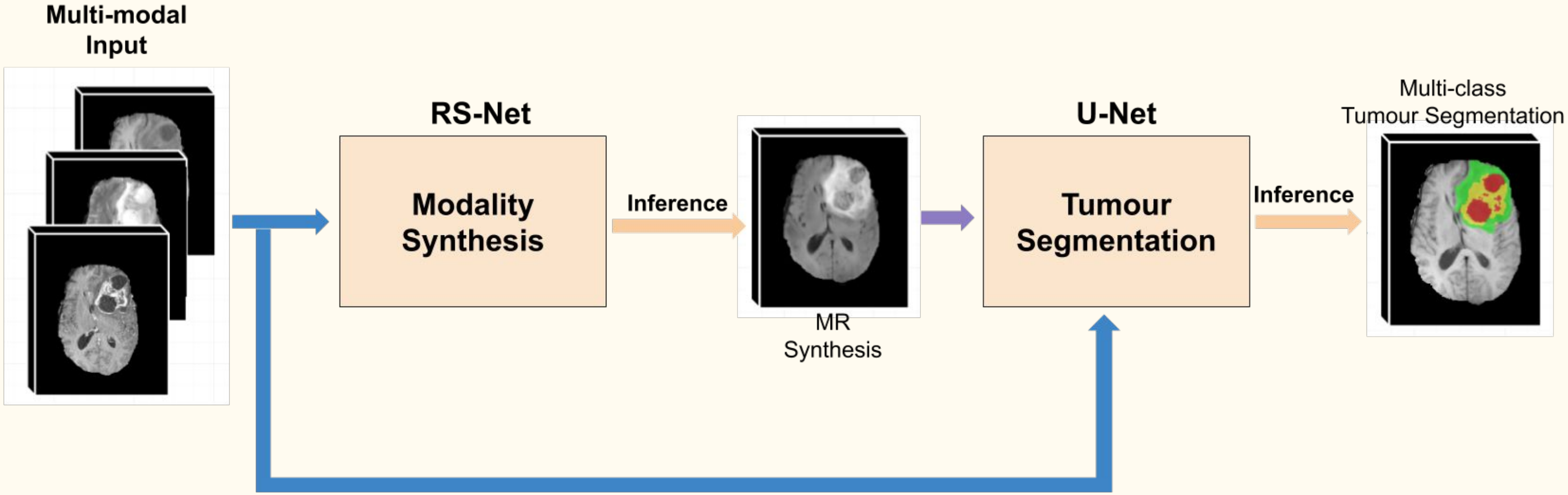
Brain Tumour Segmentation Pipeline

- **Baseline-1**
  - No Task-1 Network (RS-Net): No synthesis of missing modality



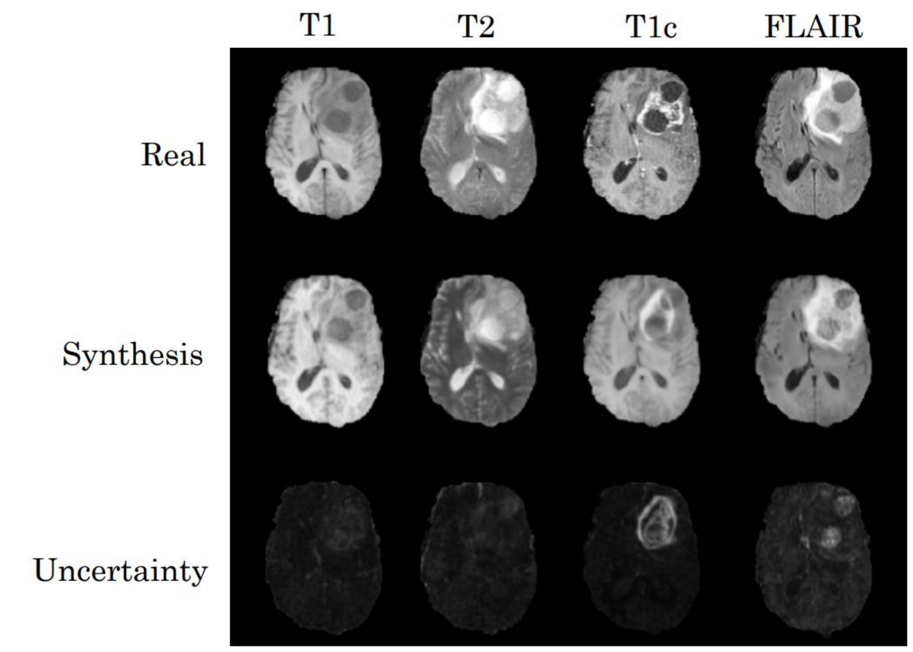
Brain Tumour Segmentation Pipeline

- **Baseline-2**
    - Only inference from Task-1 Network (RS-Net) is propagated to Task-2 Network (3D U-Net)



Brain Tumour Segmentation Pipeline

- **Uncertainty in Synthesis (RS-Net)** [10]



[10] Mehta et al., "RS-Net: Regression-Segmentation 3D CNN for Synthesis of Full Resolution Missing Brain MRI in the Presence of Tumours.", SASHIMI 2018.

Brain Tumour Segmentation Pipeline

- **Quantitative Results**

	T1ce synthesis		
	DT	DC	DE
real(3) sequences	87.17	50.25	26.89
real(3)+synthesized sequences	86.72	52.80	27.35
real(3)+synthesized+uncertainty	**88.20**	**57.29***	**32.86***

(*) indicates statistically significant ($p \leq 0.05$) differences between second and third row.

McGill Centre for Intelligent Machines

McGill
Centre for
Intelligent Machines

Brain Tumour Segmentation Pipeline

- **Quantitative Results**

				FLAIR synthesis		
				DT	DC	DE
real(3) sequences				83.27	73.91	71.07
real(3)+synthesized sequences				84.56	76.72	72.89
real(3)+synthesized+uncertainty				**85.84***	**79.25***	**74.51***

(*) indicates statistically significant ($p \leq 0.05$) differences between second and third row.

Brain Tumour Segmentation Pipeline

- **Quantitative Results**

	T1ce synthesis			FLAIR synthesis		
	DT	DC	DE	DT	DC	DE
real(3) sequences	87.17	50.25	26.89	83.27	73.91	71.07
real(3)+synthesized sequences	86.72	52.80	27.35	84.56	76.72	72.89
real(3)+synthesized+uncertainty	**88.20**	**57.29***	**32.86***	**85.84***	**79.25***	**74.51***

(*) indicates statistically significant ($p \leq 0.05$) differences between second and third row.

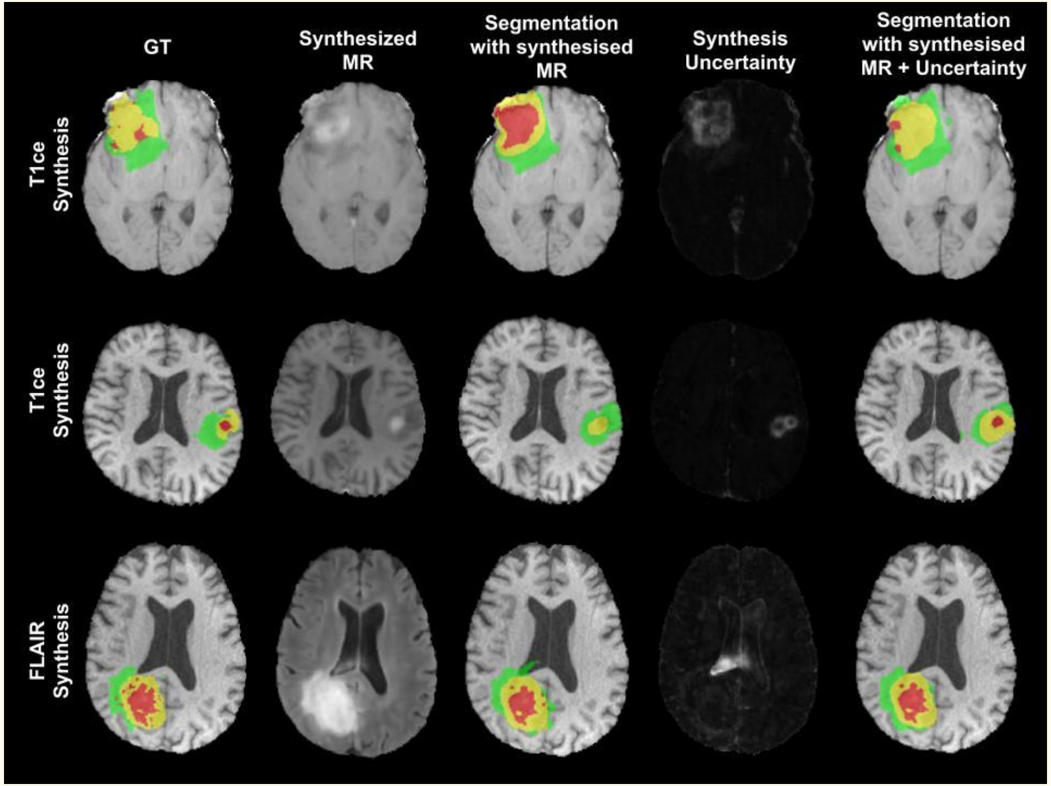
Brain Tumour Segmentation Pipeline

- Qualitative Results

Enhancing Tumour

Edema

Non Enhancing Core



Conclusion

- Proposed **a general deep learning framework** for the propagation of uncertainty across a sequence of inference tasks within a medical image analysis pipeline for improved inference
- **Evaluation on two different contexts** of MS T2 lesion segmentation/detection and Brain Tumour segmentation
- **2-10% improvement** for both tasks on their respected **quantitative** measures
- **Clearly visible qualitative improvement**
- **Future work** will explore how to properly develop **a complete end-to-end system** that includes uncertainty propagation across the inference modules

Thank you

McGill
Centre for
Intelligent Machines

NeuroRx
RESEARCH

NSERC
CRSNG

SYNAPTIVE
MEDICAL

INTERNATIONAL
PROGRESSIVE MS ALLIANCE
CONNECT TO END PROGRESSIVE MS

MS T2 Lesion Segmentation/Detection Pipeline

- **Implementation Details**
  - **Task-1 Network:** BU-Net [4]
  - **Task-1 Network uncertainty:** Variance of 10 MC samples [5]
  - **Task-2 Network:** 3D U-Net [15]
    - 3 resolution U-Net
    - Linear Upsampling
    - Leaky-ReLU non-linear activation [16]
    - Group Normalization [17]
    - Equally weighted Sorensen-Dice loss [18] and binary cross-entropy loss
  - 18 connected component to convert segmentation to detection

[4] Nair et al., “Exploring uncertainty measures in deep networks for multiple sclerosis lesion detection and segmentation.”, Medical Image Analysis 2019.
[5] Gal and Ghahramani, “Dropout as a bayesian approximation: Representing model uncertainty in deep learning.”, ICML 2016.
[15] Cicek et al., “3D U-Net: learning dense volumetric segmentation from sparse annotation.”, MICCAI 2016.
[16] Maas et al., “Rectifier nonlinearities improve neural network acoustic models.”, ICML 2013.
[17] Wu and He., “Group normalization.”, In ECCV 2018.
[18] Milletari et al., “V-net: Fully convolutional neural networks for volumetric medical image segmentation.”, 3DV 2016

McGill
Centre for
Intelligent Machines

Brain Tumour Segmentation

- **Implementation Details**
  - **Task-1 Network:** RS-Net [10]
  - **Task-1 Network uncertainty:** Variance of 20 MC samples [5]
  - **Task-2 Network:** 3D U-Net [15]
    - 4 resolution U-Net
    - Deconvolution [19]
    - ReLU non-linear activation [20]
    - Instance Normalization [21]
    - Weighted categorical cross-entropy loss

[10] Mehta et al., “RS-Net: Regression-Segmentation 3D CNN for Synthesis of Full Resolution Missing Brain MRI in the Presence of Tumours.”, SASHIMI 2018
[5] Gal and Ghahramani, “Dropout as a bayesian approximation: Representing model uncertainty in deep learning.”, ICML 2016.
[15] Cicek et al., “3D U-Net: learning dense volumetric segmentation from sparse annotation.”, MICCAI 2016.
[19] Zeiler et al., “Deconvolutional networks.”, CVPR 2010
[20] Glorot et al., “Deep sparse rectifier neural networks”, AISTATS 2011
[21] Ulyanov et al., “Instance normalization: The missing ingredient for fast stylization.”, arXiv preprint arXiv:1607.08022.