

# Multi-sequence learning for multiple sclerosis lesion segmentation in spinal cord MRI

Ricky Walsh<sup>1</sup>, Malo Gaubert<sup>1,2</sup>, Cédric Meurée<sup>1</sup>, Burhan Rashid Hussein<sup>1</sup>, Anne Kerbrat<sup>1,3</sup>, Romain Casey<sup>4</sup>, Benoit Combès<sup>1,†</sup>, Francesca Galassi<sup>1,†</sup>

<sup>1</sup>Univ Rennes, Inria, CNRS, Inserm, IRISA UMR 6074, Empenn, Rennes, France

<sup>2</sup>Department of Neuroradiology, Rennes University Hospital, Rennes, France

<sup>3</sup>Department of Neurology, Rennes University Hospital, Rennes, France

<sup>4</sup>Univ Lyon, Université Claude Bernard Lyon 1, Hospices Civils de Lyon, Fondation EDMUS, OFSEP, Centre de Recherche en Neurosciences de Lyon, Lyon, France

## Introduction

Spinal cord lesions in Multiple Sclerosis (MS) have *prognostic value*.

Detecting spinal MS lesions is a *challenging task* for clinicians.

*Multiple imaging protocols* are used, and not every sequence is available for each patient.

How can we use all available data for training and inference despite **missing sequences**?

1. Pre-processing pipeline to deal with varying fields of view and inter-sequence misalignments.
2. Analysis of improvement of multi-sequence training for single-sequence inference.
3. Latent feature augmentation for single-sequence training.

## Method

**Data** Multiple acquisitions from 247 subjects. Sagittal T2w (247 subjects), sagittal STIR (137), axial T2\*w (119), T1w (63), axial T2w (52), MP2RAGE (36).

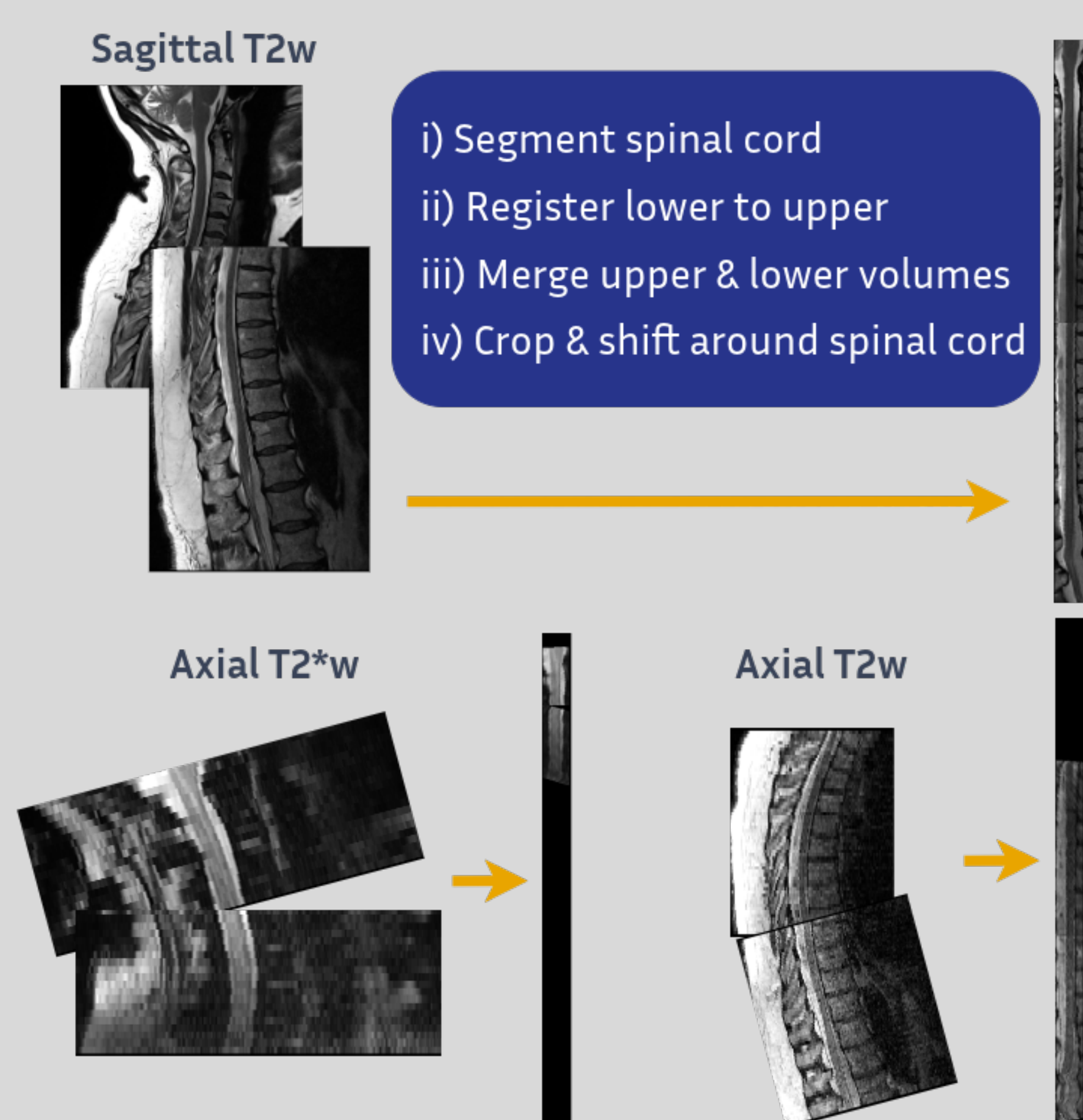
**Preprocess** Pre-process all sequences for patient to common space. Crop and centre around spinal cord [2]. No non-linear registration to avoid introducing artifacts.

**Method** One encoder per sequence and concatenate features at bottleneck. **Mean Imputation:** When a *sequence is missing*, replace its features with the mean across the available sequences [3].

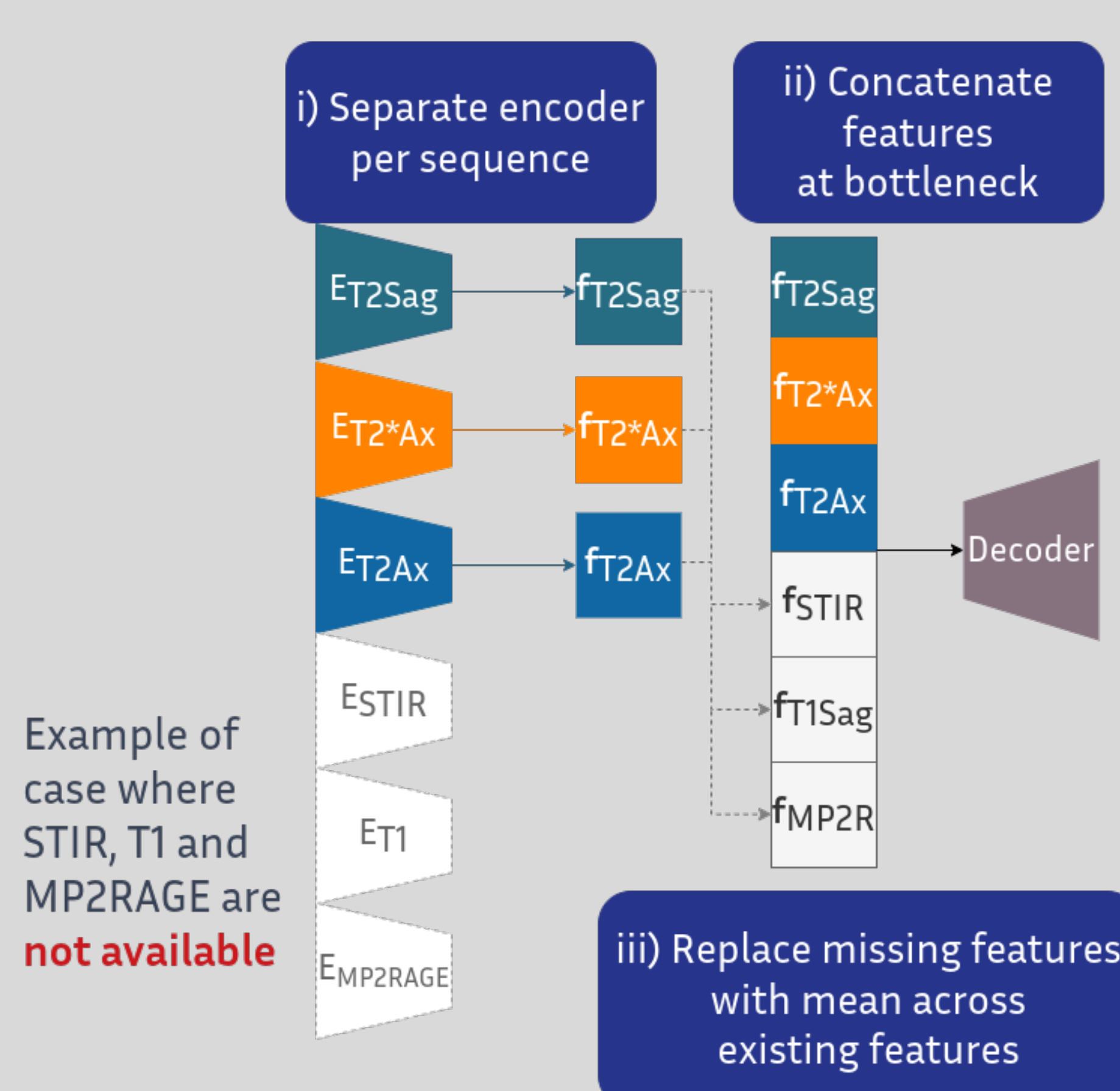
**Feature Augmentation:** Apply noise at the bottleneck while training single-sequence model.

**Frozen Encoder:** Freeze T2Sag baseline encoder and train the rest of the multi-sequence model.

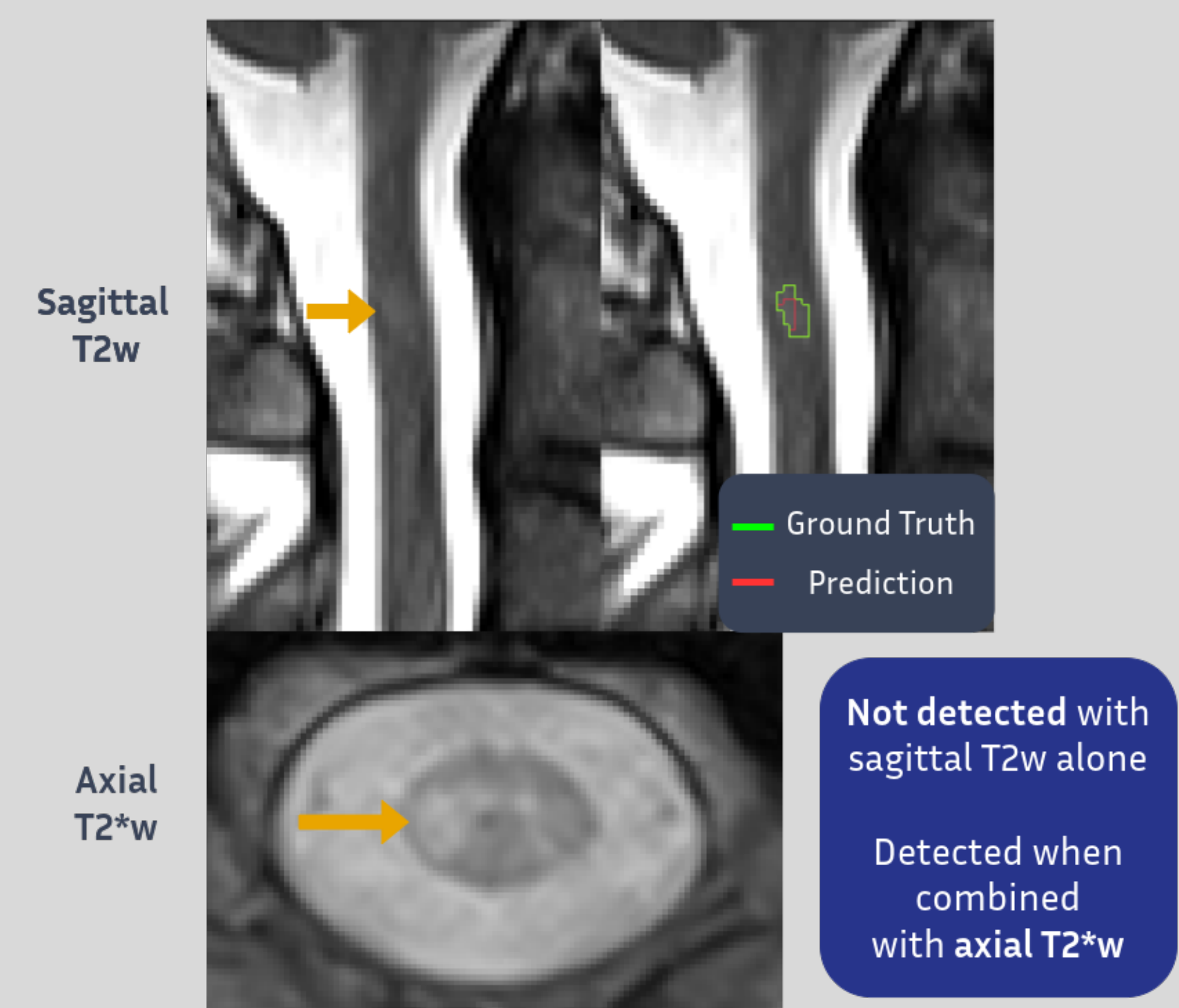
### 1. Preprocessing



### 2. Method



### 3. Example Output



Multi-sequence *training* improves single-sequence *inference*.

## Results

Table 1: Results for two test cohorts,  $\mathcal{D}_A$  and  $\mathcal{D}_C$ .

Significance levels for Wilcoxon test vs. T2Sag Baseline: \* 0.1; \*\* 0.05.

| Model          | Sequences at Inference |       |      |    | Lesion F1       |                 | Dice            |                 |
|----------------|------------------------|-------|------|----|-----------------|-----------------|-----------------|-----------------|
|                | T2Sag                  | T2*Ax | T2Ax | T1 | $\mathcal{D}_A$ | $\mathcal{D}_C$ | $\mathcal{D}_A$ | $\mathcal{D}_C$ |
| T2Sag Baseline | ●                      | ○     | ○    | ○  | 0.593           | 0.583           | 0.452           | 0.469           |
| Mean           | ●                      | ○     | ○    | ○  | 0.641**         | 0.606           | 0.466*          | 0.493**         |
| Imputation     | ●                      | ●     | ○    | ●  | -               | <b>0.622</b>    | -               | <b>0.518**</b>  |
|                | ●                      | ●     | ●    | ○  | <b>0.650**</b>  | -               | <b>0.480**</b>  | -               |

Table 2: Results using only T2Sag at inference (mean  $\pm$  std. dev.).

| Method                           | Lesion F1                              | Dice                                   |
|----------------------------------|--|--|
| T2Sag Baseline                   | 0.564 $\pm$ 0.031                      | 0.441 $\pm$ 0.011                      |
| T2Sag – Feature Augmentation     | 0.586 $\pm$ 0.026 **                   | 0.446 $\pm$ 0.014 *                    |
| Mean Imputation – Frozen Encoder | 0.591 $\pm$ 0.022 **                   | 0.448 $\pm$ 0.015 *                    |
| Mean Imputation                  | <b>0.603 <math>\pm</math> 0.025 **</b> | <b>0.457 <math>\pm</math> 0.013 **</b> |

## Conclusion

**Mean Imputation** implicitly regularises both the encoder and decoder.

Latent **feature augmentation** can achieve a similar regularisation effect.

Multiple sequences at inference leads to some modest improvements, especially with axial T2\*w.

Future work: pre-train single sequence models to better capture the sequence-specific information.

[1] R. Walsh et al., MICCAI 2024, <https://hal.science/hal-04668565>

[2] C. Gros et al., NeuroImage 2019, DOI: 10.1016/j.neuroimage.2018.09.081

[3] H. Wang et al., MICCAI 2023, DOI: 10.1007/978-3-031-43901-8\_21

Download the paper here:



### Acknowledgements

This work was supported by the following grants managed by the French Agence Nationale de la Recherche (ANR): ANR-20-THIA-0018, ANR-21-RHUS-0014, ANR-10-COHO-002 OFSEP.