Soutenance de stage

Deep Learning for low-back pain diagnosis in lumbar MRI

Simon Queric

Supervised by Julien Cohen-Adad

From May to September 2024 at Polytechnique Montréal

Télécom Paris & ENS Paris Saclay

Table of contents

Introduction

Introduction

My internship was about designing Deep Learning methods for MRI image analysis and automated diagnosis in lumbar MRI.

Contribution : Propose a solution for the Radiological Society of North America (RSNA) 2024 Kaggle data challenge.

RSNA 2024 data challenge

Data description

In the context of the RSNA 2024 Lumbar Spine Degenerative Classification Kaggle data challenge, we were provided three MRI sequences per subjects, more specifically Sagittal T1, Sagittal T2 and Axial T2 MRI.

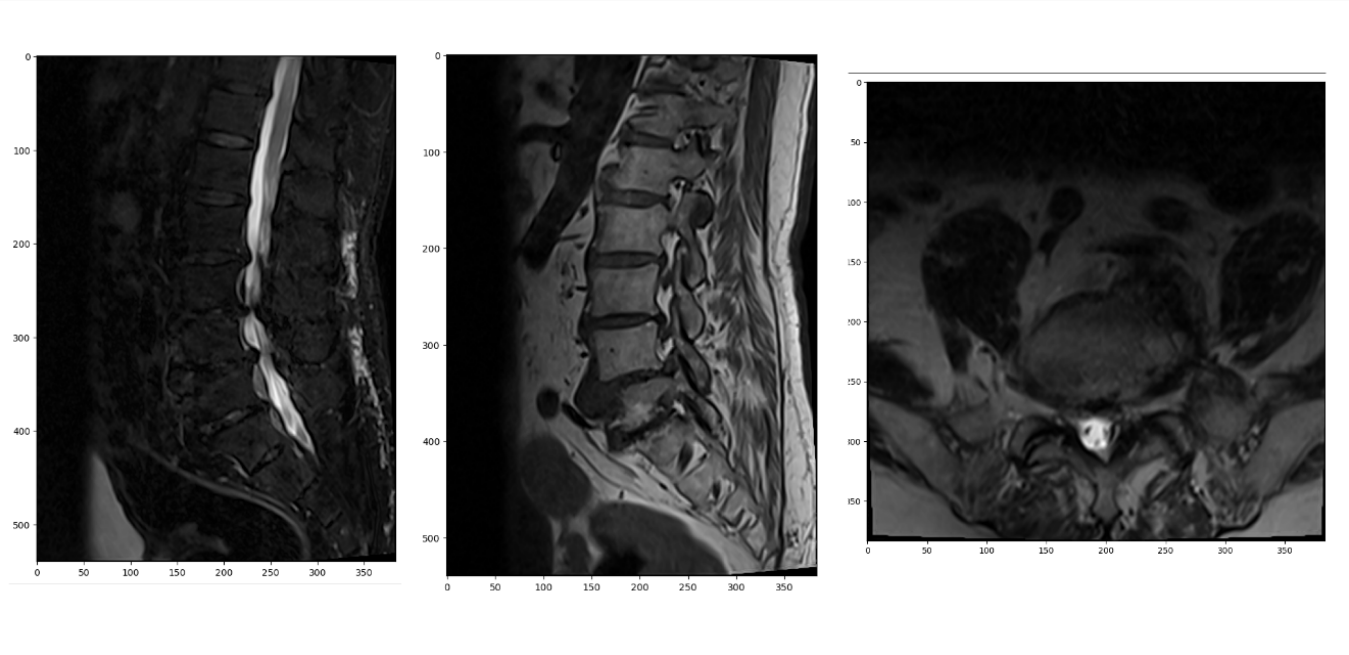

Figure 1: Example of data for one subject. From Left to right : Sagittal T2/STIR, Sagittal T1 and Axial T2 slices.

Goal of the challenge

The goal of this competition is to create models that can be used to aid in the detection and classification of degenerative spine conditions using lumbar spine MR images.

More specifically, this is a multi-classification task.

5 conditions and 3 labels (*Normal/Mild*, *Moderate*, *Severe*) for 5 levels, namely L1/L2, L2/L3, L3/L4, L4/L5 and L5/S1.

The performance criterion used by organizers is the weighted cross-entropy loss :

$$\ell(\widehat{y}, y) = -\sum_{c=1}^{C} w_c y_c \log(\widehat{y}_c)$$

Lumbar anatomy

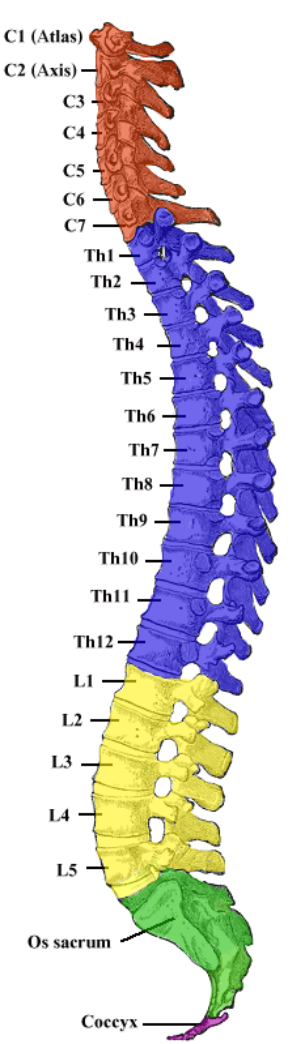

Figure 2: From Wikipedia

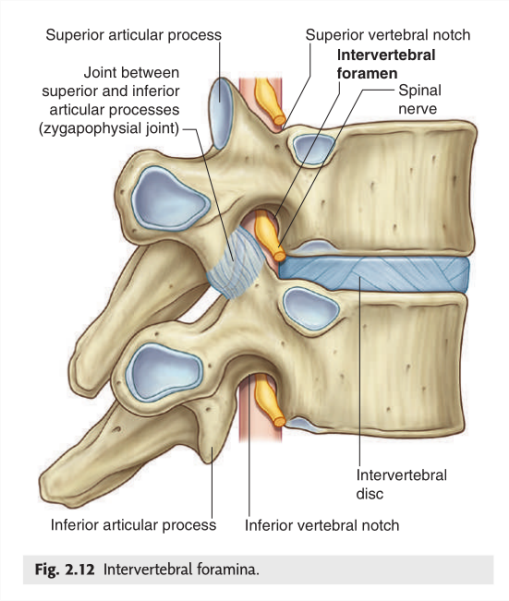

Figure 3: Foramen illustration. From *Gray's anatomy for students*

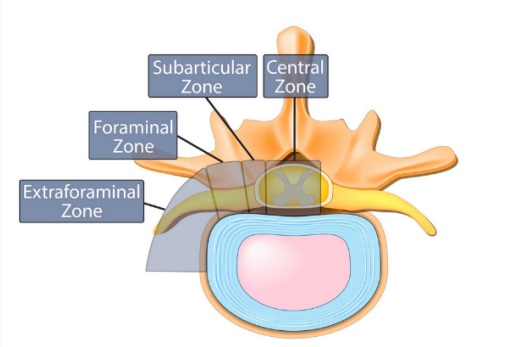

Figure 4: Axial view of lumbar vertebrae. From *Gray's anatomy for students*

Anatomy and degenerative disease

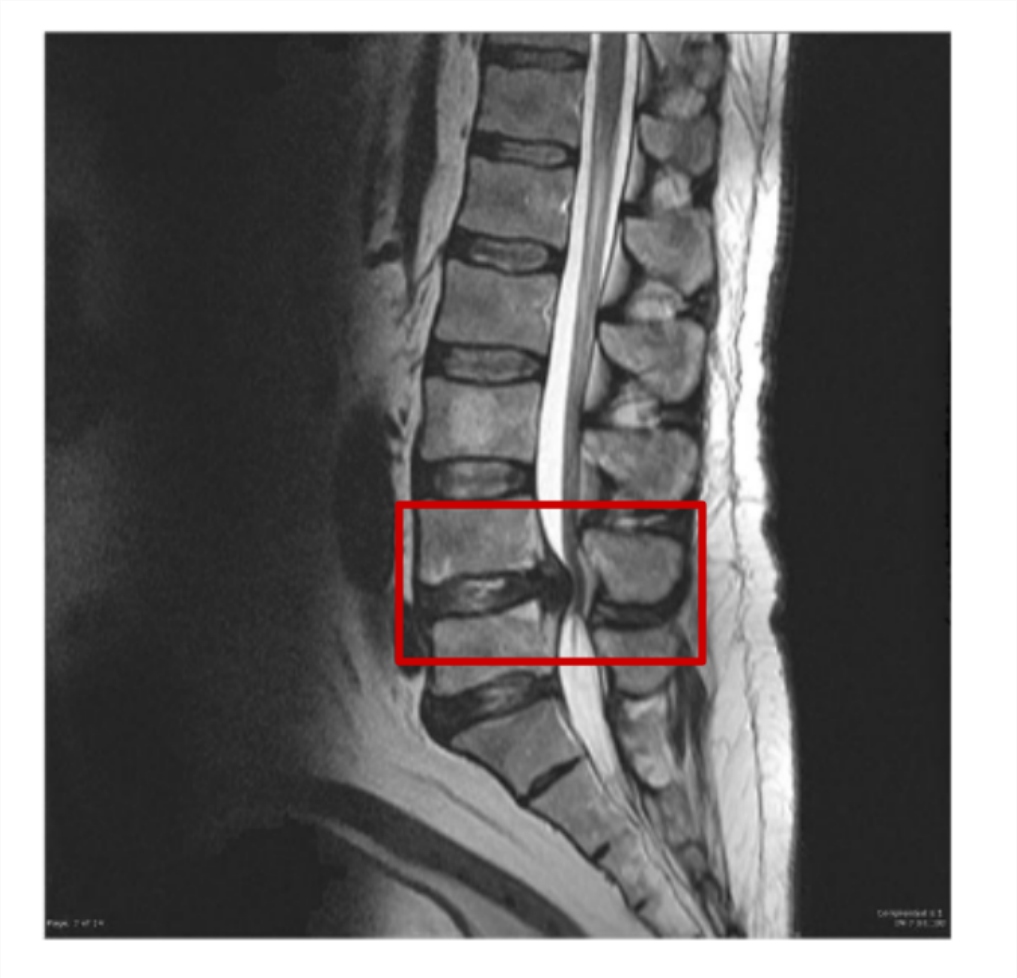

Figure 5: Example of a spinal stenosis at L4/L5

Anatomy and degenerative disease

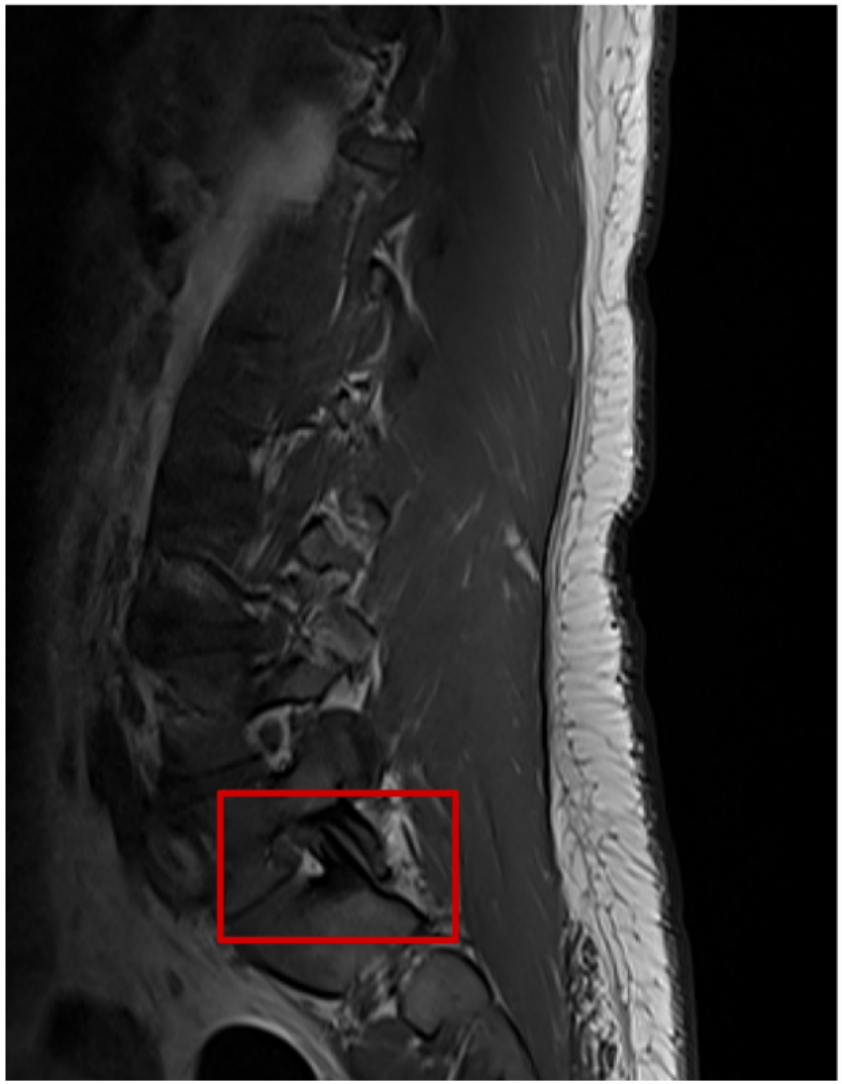

Figure 6: Example of a foraminal narrowing at L5/S1

Anatomy and degenerative disease

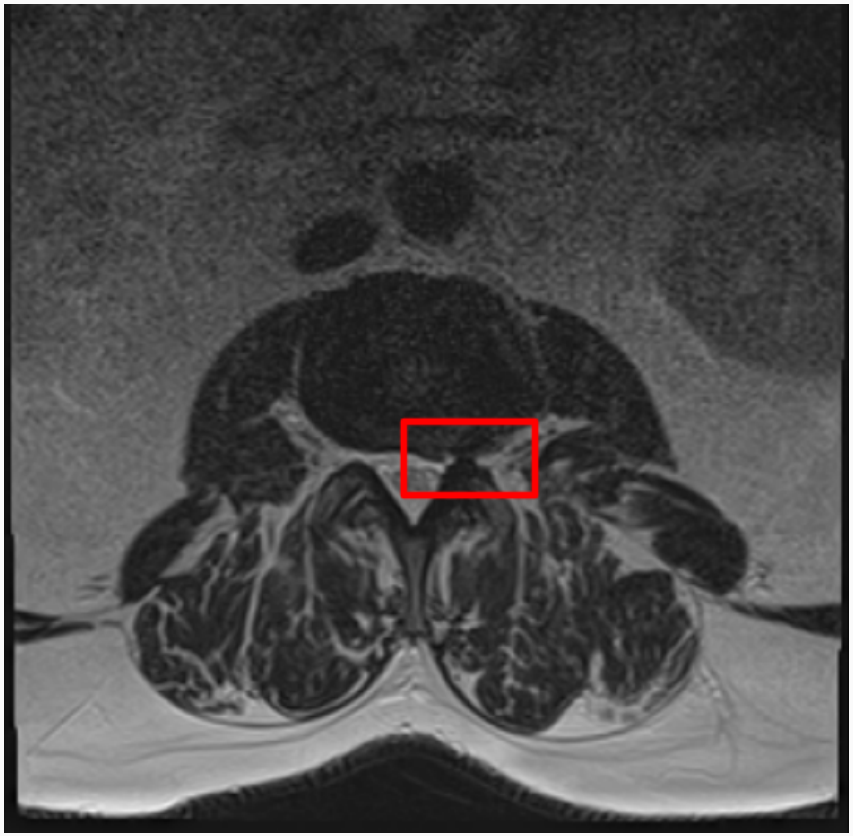

Figure 7: Example of a subarticular stenosis at L4/L5

Unbalanced classification task

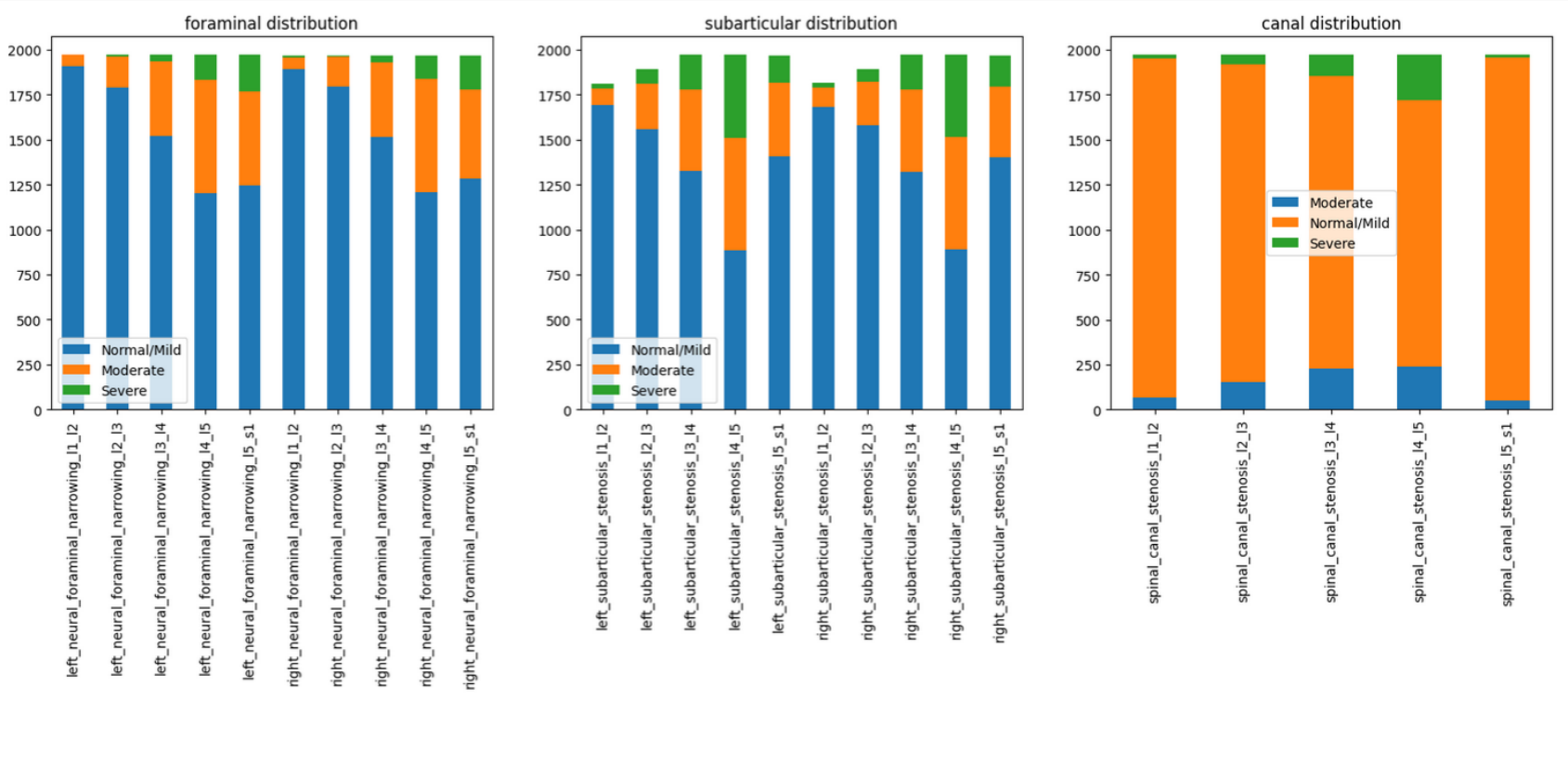

Figure 8: Prévalence des différentes pathologies au sein du jeu de données. Tiré du notebook *Anatomy and Vizualisation* du challenge RSNA 2024

Methods

Litterature review

Classical architectures for this task

- You Only Look at Once (object detection)
- CenterNet: Keypoint Triplets for Object Detection

Approaches in medical imaging

- For IDD detection, a YOLO architecture has already been proposed.
- For cervical degenerative disease, an approach based on anatomical structure has been proposed.

Principle of the method

Approach in 2 steps

1. Extraction of region of insterest (L1/L2, etc...)
2. Classification network for grading severity

1st attempt : U-Net based approach for ROI extraction

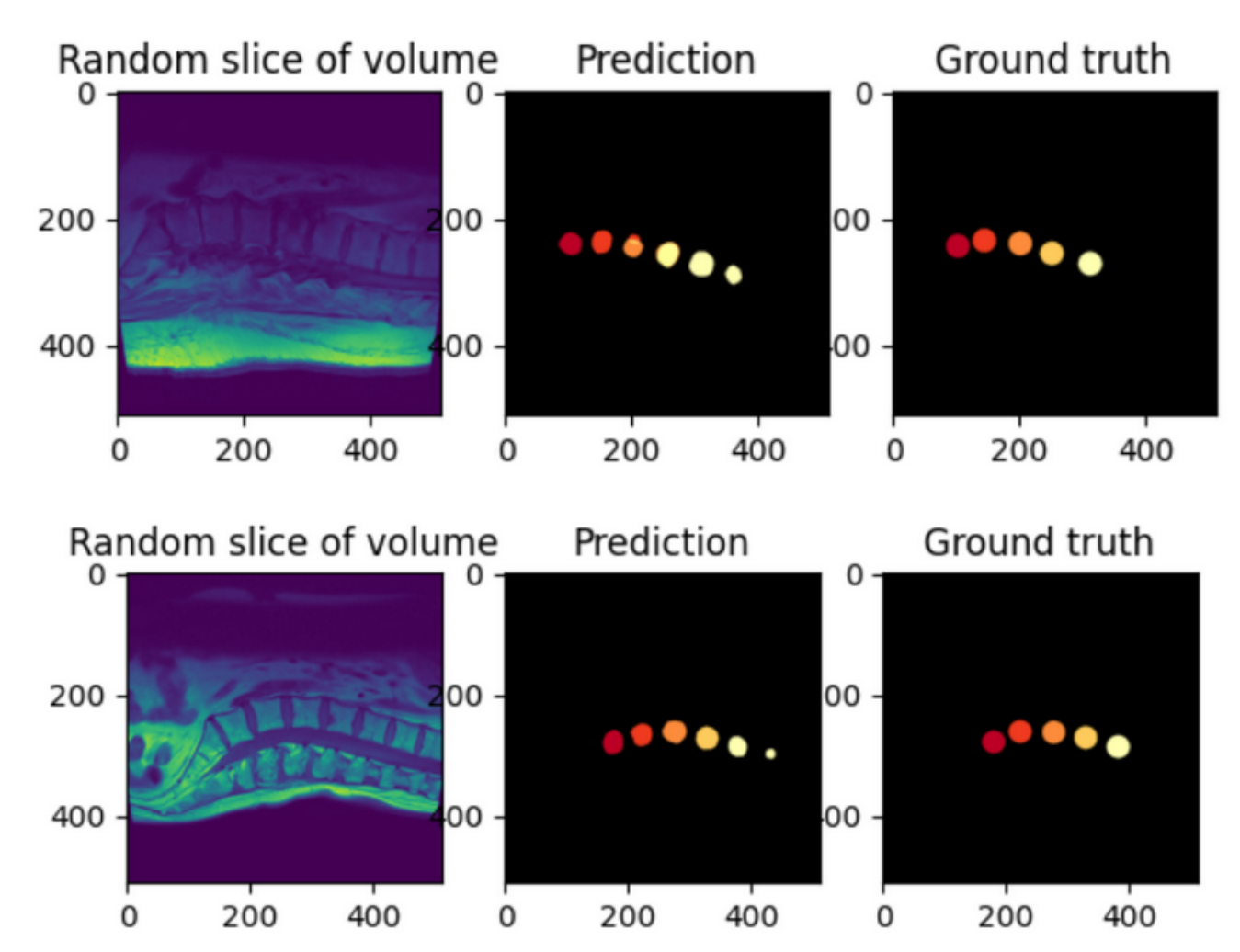

Figure 9: 1st approach for ROI extraction with a U-Net model.

Using TotalSpineSeg for ROI extraction

TotalSpineSeg is a Deep Learning model which labels anatomic parts in MRI T1 and T2, in different FOV (cervical, thoaric and lumbar)

Figure 10: TotalSpineSeg output

ROI patch extraction

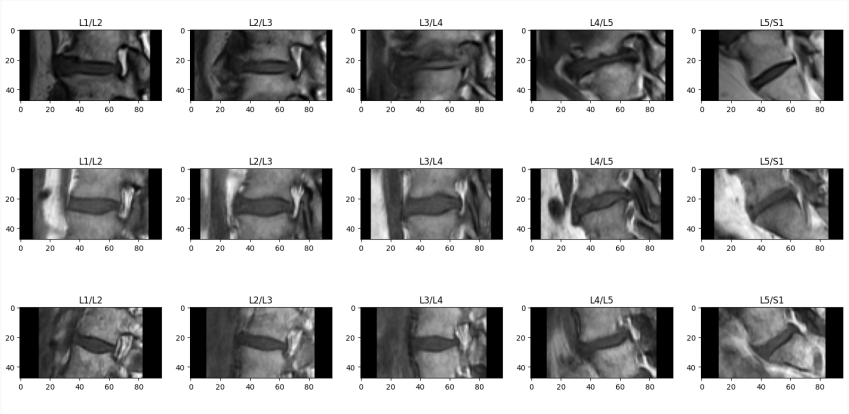

Figure 11: ROI extraction centered around intervertebral foramen.

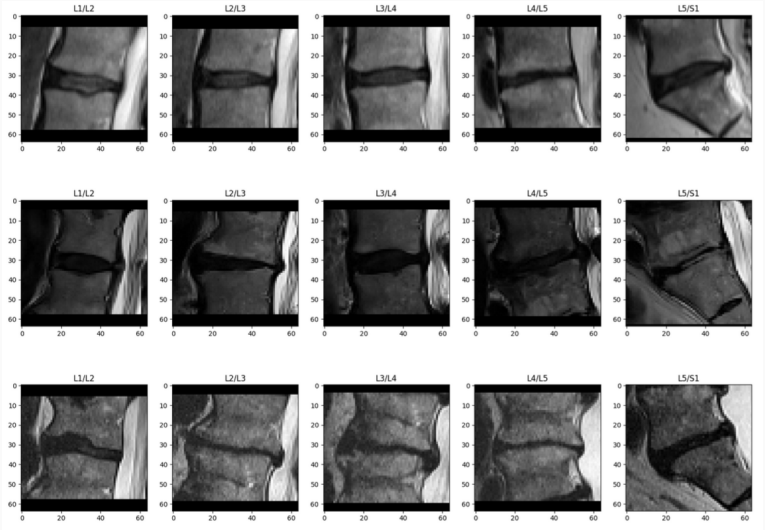

Figure 12: ROI extraction centered around intervertebral discs.

Registration of Axial images

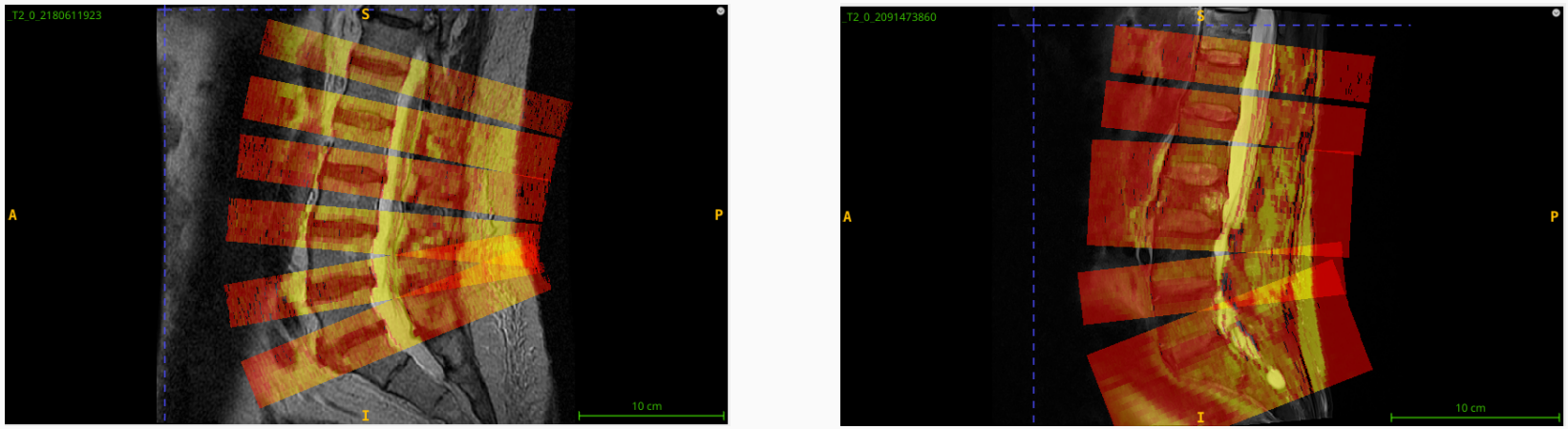

Figure 13: Illustration de la projection des volumes axiaux dans l'espace Sagittal. Les volumes axiaux sont visibles en couleurs orangées.

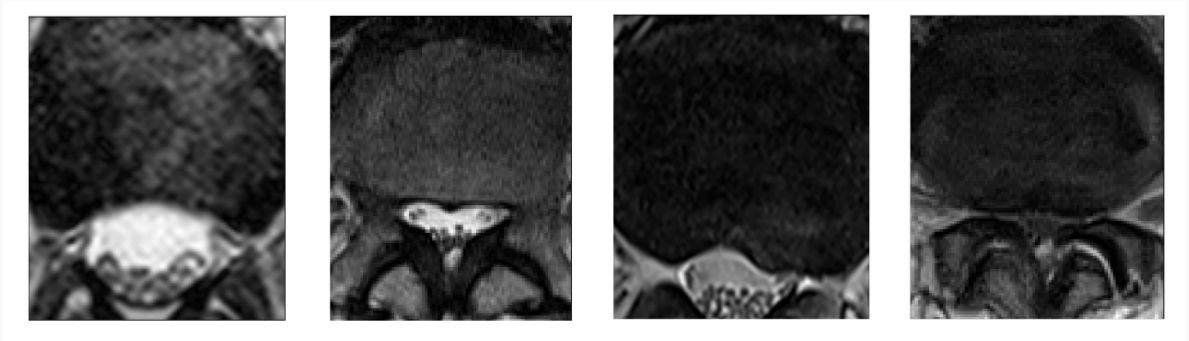

Figure 14: Extraction de patchs en coupe axiale centré sur les disques intervertébraux.

Overview of the method

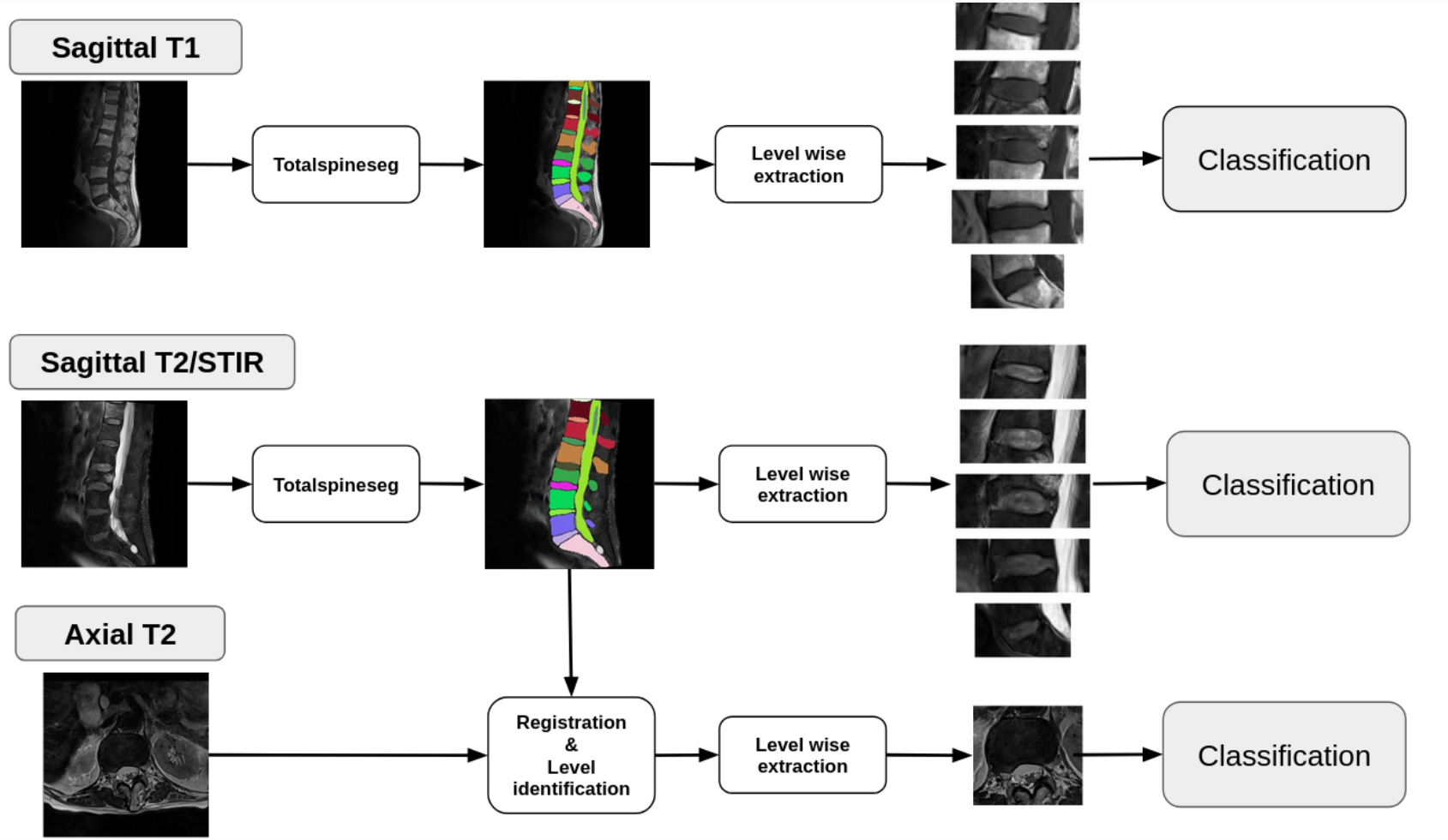

Figure 15: Overview of the method

Results

Confusion matrix for Spinal Canal Stenosis model

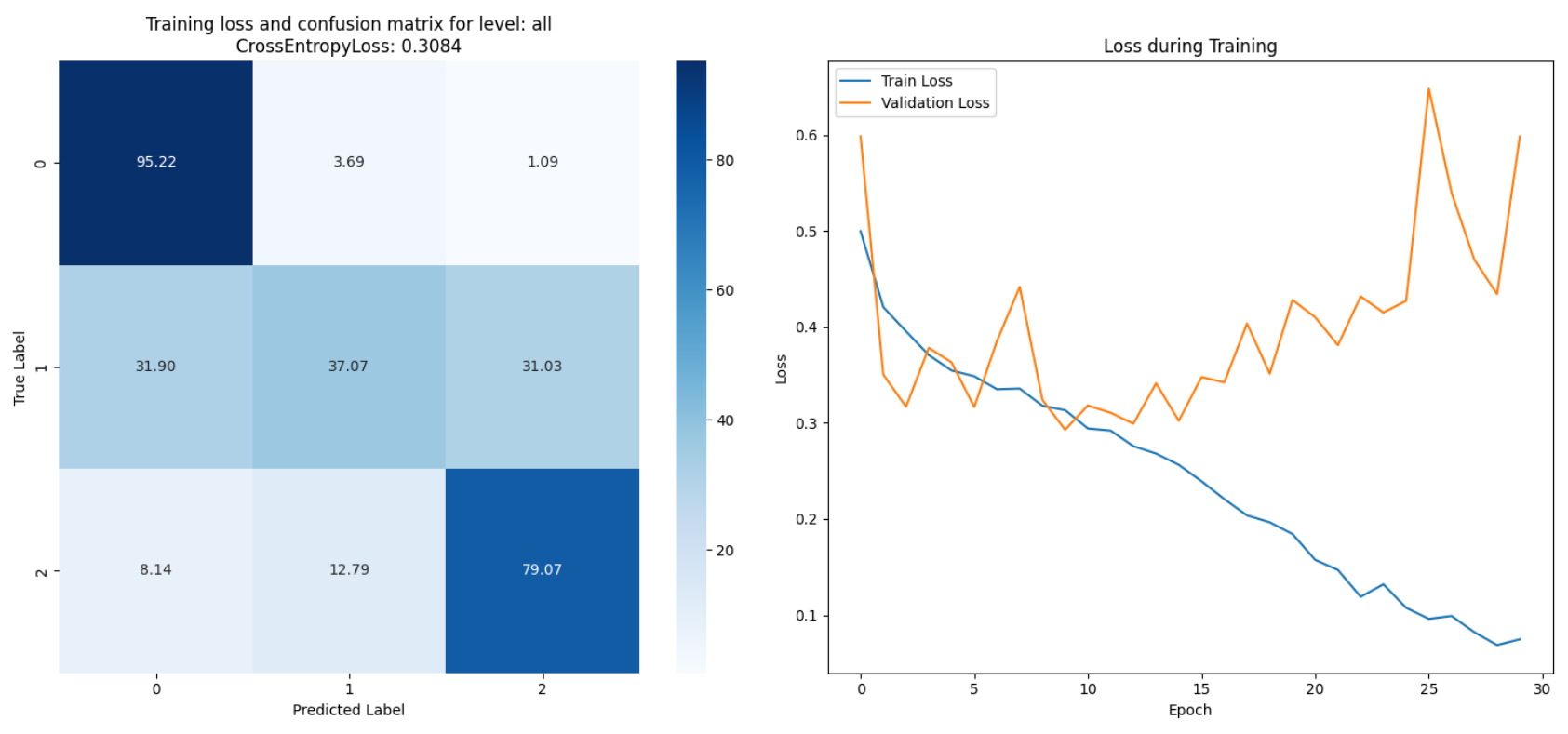

Figure 16: Confusion matrix for spinal canal stenosis. Classification with axial T2 patches.

Confusion matrix for Subarticular Stenosis model

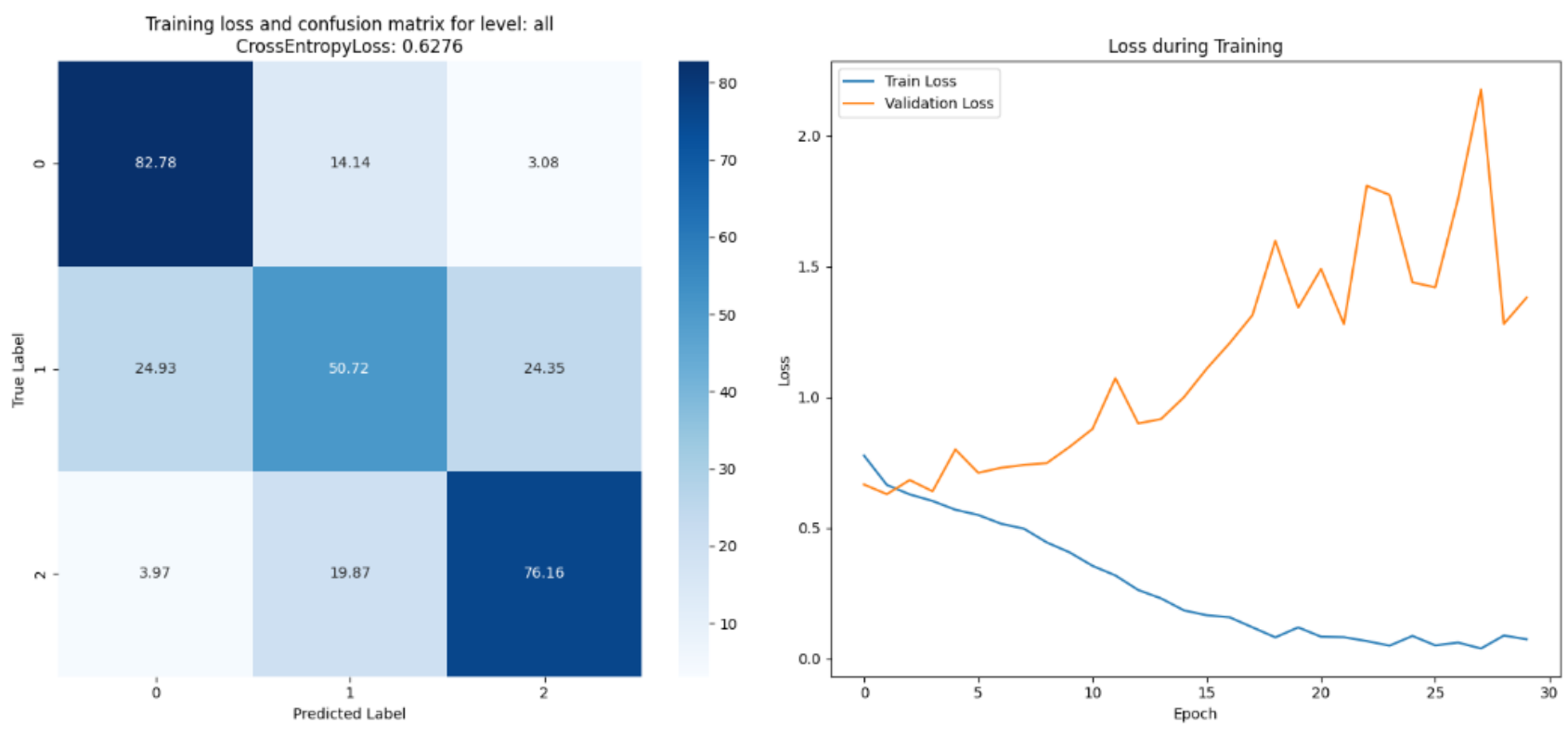

Figure 17: Confusion matrix for subarticular stenosis. Classification with axial T2 patches.

Analysis of the results, limitations and future work

Pros and cons

Pros

- Efficient medically relevent context extraction with TotalSpineSeg
- Discrimnate well Severe against Normal
- Clearly better results than a random prediction (for SAS and SCS, *balanced accuracies* : 69.88% and 67.45%)
- Strong overfitting

Cons & room for improvement

- The information of all sequences is not used
- Experimental choices for hyperparameters
- No rigorous ablation study
- The method fails for moderate samples

Acknowledgements

Acknowledgements

I would like to thank Julien Cohen-Adad for this research opportunity, and also Nathan Molinier and Nilser Laines Medina for their supervision.

Many thanks to my colleagues, Abel Salmona and Thomas Dagonneau who contributed to the final Kaggle submission.