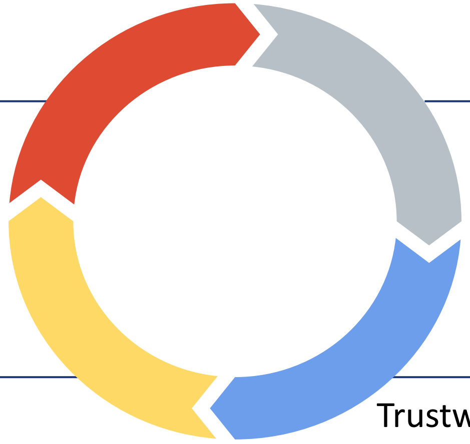
Exploiting XAI maps to improve MS lesion segmentation and detection in MRI



Explainability

Insight on internal actions and/or features of the DL-model.

Interpretability

Links the DL-model to the human observer knowledge.

Uncertainty

Quantification of the reliability of a model's prediction.

Integration

Trustworthy and transparent decision support for patient assessment.

Federico Spagnolo[1,2,3,4], Nataliia Molchanova[4,5], Mario Ocampo-Pineda[1,2,3], Lester Melie-Garcia[1,2,3], Meritxell Bach Cuadra[5,6], Cristina Granziera[1,2,3], Vincent Andrearczyk[4], and Adrien Depeursinge[4,7]

[1]Translational Imaging in Neurology (ThINk) Basel, Department of Medicine and Biomedical Engineering, University Hospital Basel and University of Basel, Basel, Switzerland
[2]Department of Neurology, University Hospital Basel, Basel, Switzerland
[3]Research Center for Clinical Neuroimmunology and Neuroscience Basel (RC2NB), University Hospital Basel and University of Basel, Basel, Switzerland
[4]MedGIFT, Institute of Informatics, School of Management, HES-SO Valais-Wallis University of Applied Sciences and Arts Western Switzerland, Sierre, Switzerland
[5]CIBM Center for Biomedical Imaging, Lausanne, Switzerland
[6]Radiology Department, Lausanne University Hospital (CHUV) and University of Lausanne, Lausanne, Switzerland
[7]Nuclear Medicine and Molecular Imaging Department, Lausanne University Hospital (CHUV), Lausanne, Switzerland

1. Introduction
2. Methods
3. Results
4. Discussion
5. Conclusion

MS lesion segmentation

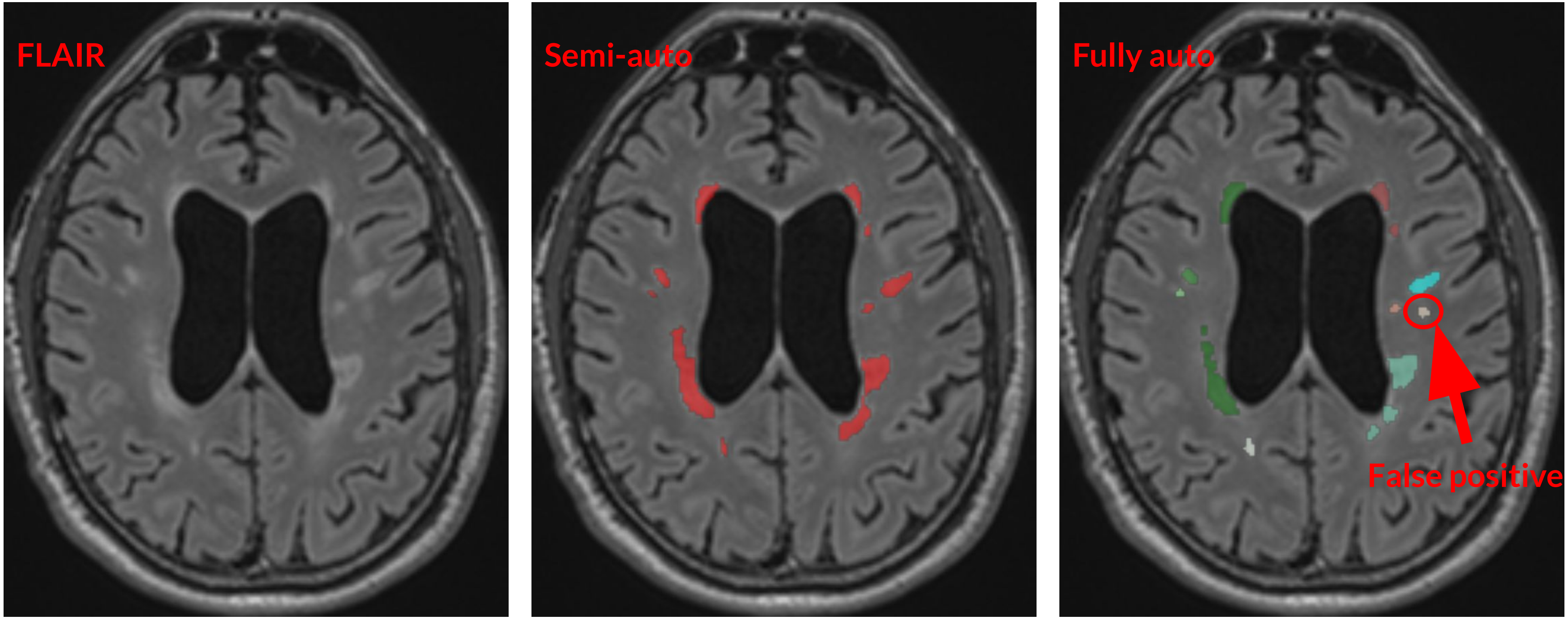


Fig.2 Axial slice of three MR images from a same MS patient, same visit.

Motivation

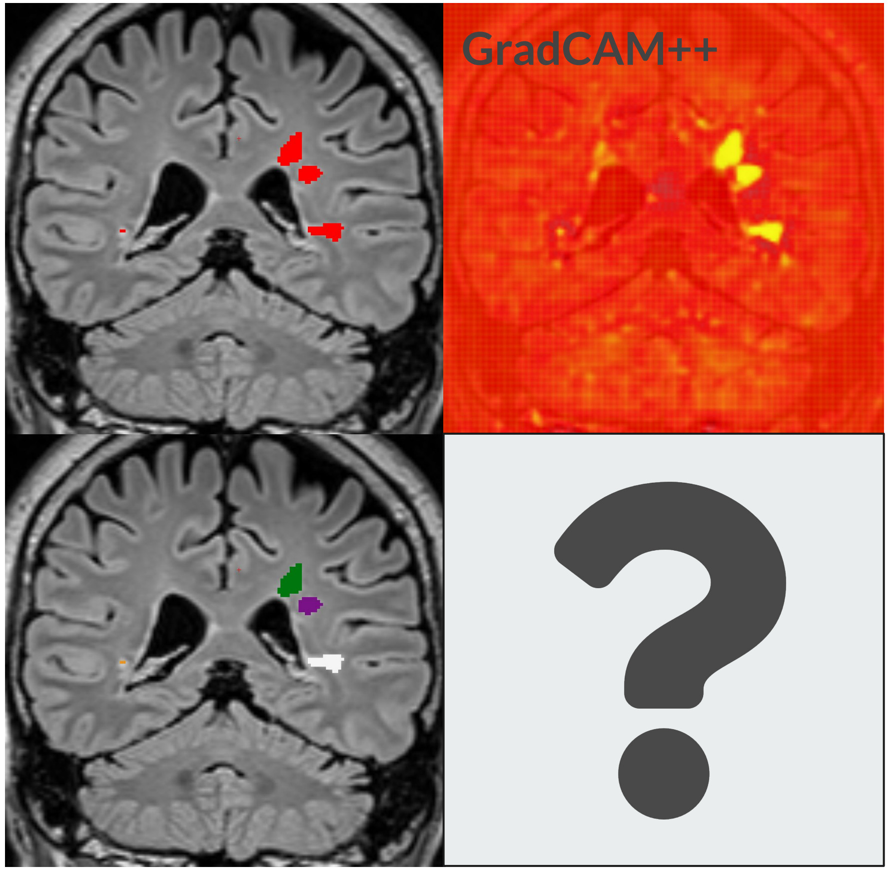


Fig.3 Output of a semantic segmentation network showing several instances of the considered class (top left). An XAI method (top right) applied to all the spatial predictions.

- Can we explain the segmentation of a **lesion of interest**?

If we have a **lesion-specific XAI**[1]…

- How to exploit it?

[1]Spagnolo, F., Molchanova, N., Schaer, R., Bach Cuadra, M., Ocampo-Pineda, M., Melie-Garcia, L., Granziera, C., Andrearczyk, V., Depeursinge, A.: Instance-level quantitative saliency in multiple sclerosis lesion segmentation. arXiv (2024). https://doi.org/10.48550/ARXIV.2406.09335.

Motivation

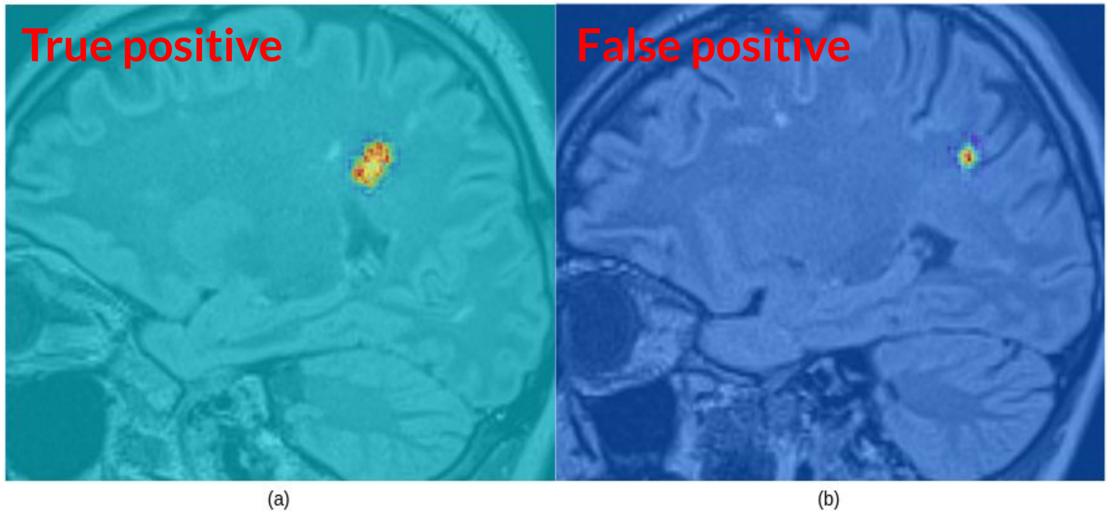


Fig.5 Instance-level saliency overlay on FLAIR, for a true positive case (a) and a false positive case (b).

	PRECISION	RECALL
@thr=.5	0.6265	0.7945
@thr=.8	0.6338	0.7848
@thr=1	0.6419	0.7778
@thr=1.5	0.7013	0.6983

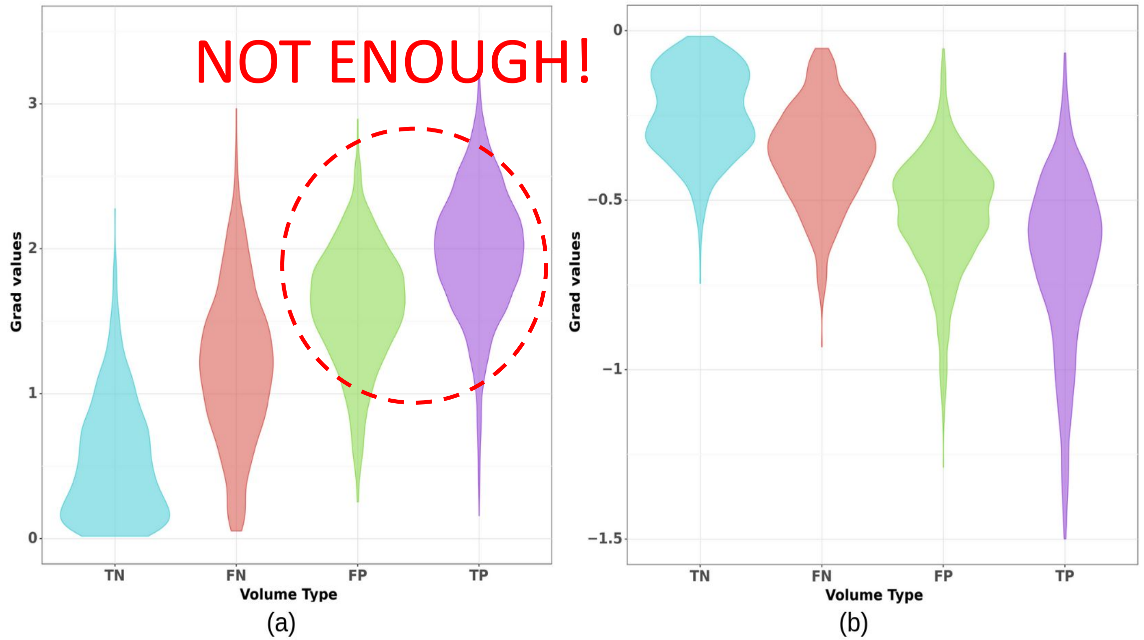


Fig.4 Maximum (a) and minimum (b) distributions in XAI maps for true positive, false positive, false negative, and true negative volumes[1].

- **Can we improve this trade off?**

[1]Spagnolo, F., Molchanova, N., Schaer, R., Bach Cuadra, M., Ocampo-Pineda, M., Melie-Garcia, L., Granziera, C., Andrearczyk, V., Depeursinge, A.: Instance-level quantitative saliency in multiple sclerosis lesion segmentation. arXiv (2024). https://doi.org/10.48550/ARXIV.2406.09335.

1. Introduction
2. **Methods**
3. Results
4. Discussion
5. Conclusion

Network

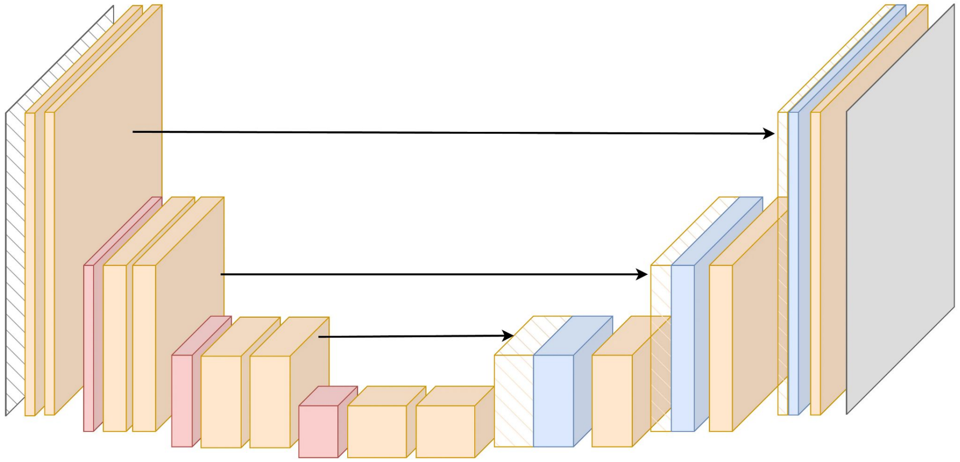
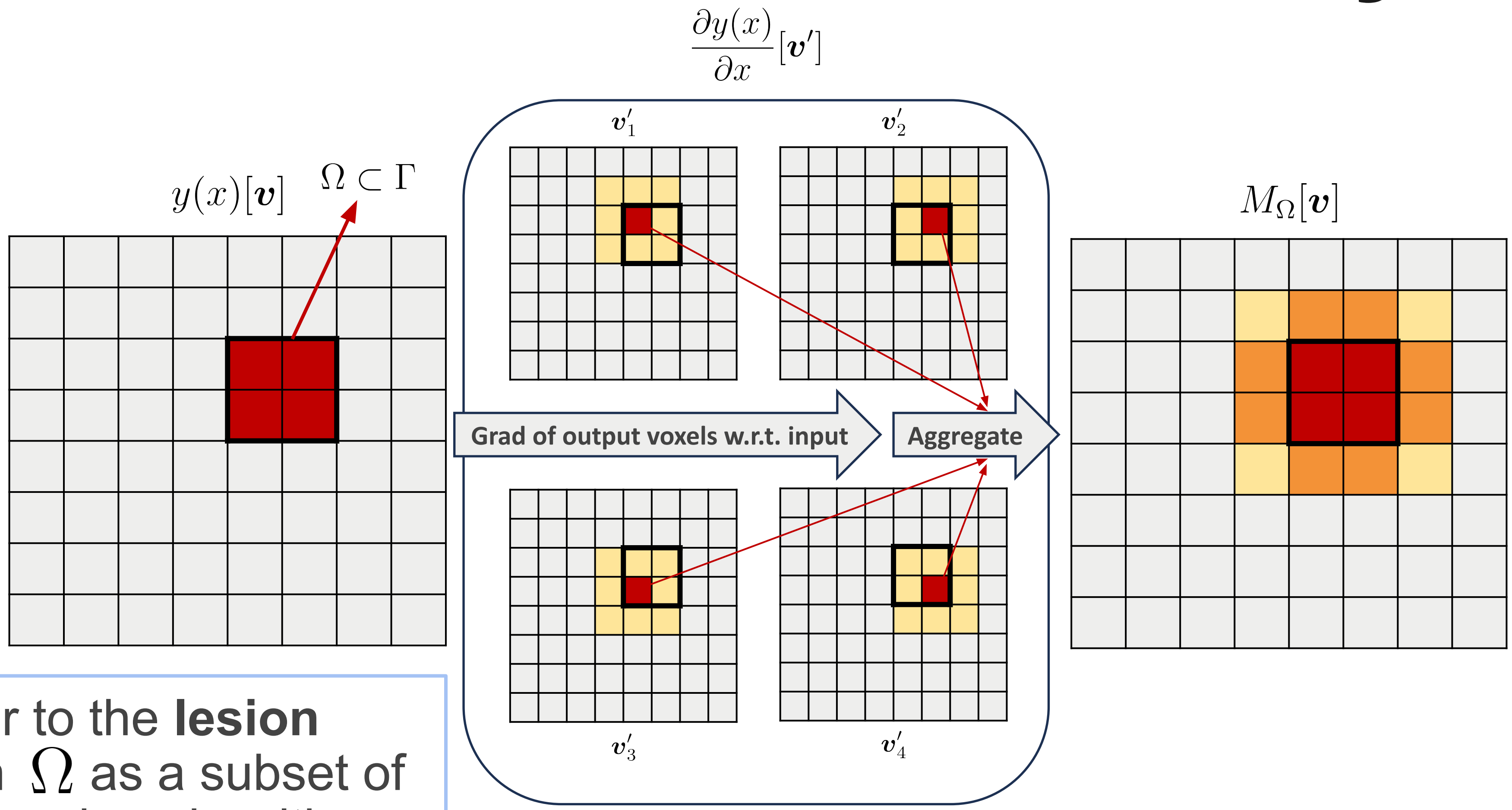


Fig.6 Architecture of a U-Net.

- 3D U-Net[2], inputs FLAIR and MPRAGE
- 687 MS patients (4023 acquisitions)
- 101 acquisitions as test
- Linear combination of normalized dice[3] and blob loss[4]
- Pre-processing: registration to FLAIR space, bias field correction, z-score intensity normalization
- DSC of 0.60 and nDSC of 0.71

[2]Çiçek, O., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. arXiv.
[3]Raina, V., Molchanova, N., Graziani, M., Malinin, A., Muller, H., Cuadra, M. B., and Gales, M. (2023). Tackling Bias in the Dice Similarity Coefficient: Introducing NDSC for White Matter Lesion Segmentation. In 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), pages 1–5.
[4]F. Kofler, S. Shit, I. Ezhov, L. Fidon, I. Horvath, R. Al-Maskari, H. Li, H. Bhatia, T. Loehr, M. Piraud, A. Erturk, J. Kirschke, J. Peeken, T. Vercauteren, C. Zimmer, B. Wiestler, and B. Menze. blob loss: instance imbalance aware loss functions for semantic segmentation. arXiv, 2022

Instance-level saliency

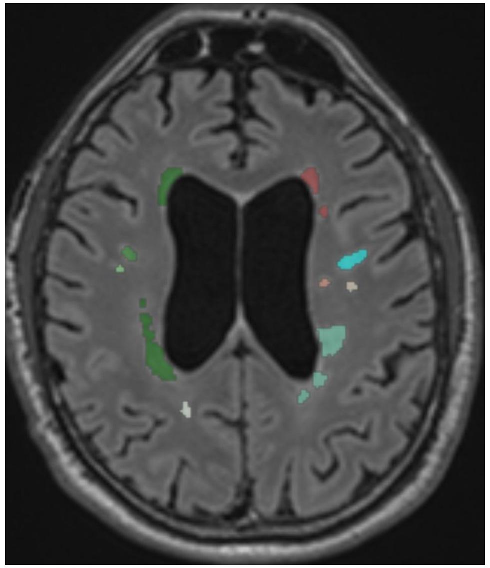


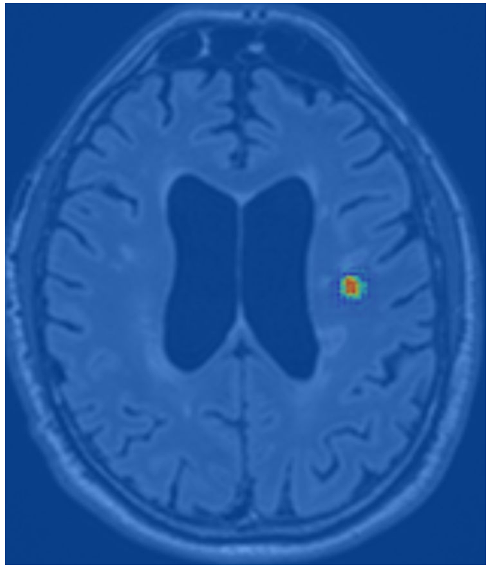
We refer to the **lesion domain** Ω as a subset of the image domain with cardinality $|\Omega|$

Radiomics on XAI

SEG — XAI — RAD — LR

DL segmentation	XAI maps generation	Radiomics	Classification
Trained WML segmentation model (on set Tr), probability output maps on test set Te	Generation of instance-level saliency maps (total of 4868)	Extraction of radiomic features from XAI maps of TPs/FPs (dilated masks)	Training (on Tr) and testing (on Te) a logistic regression model to classify TP/FP



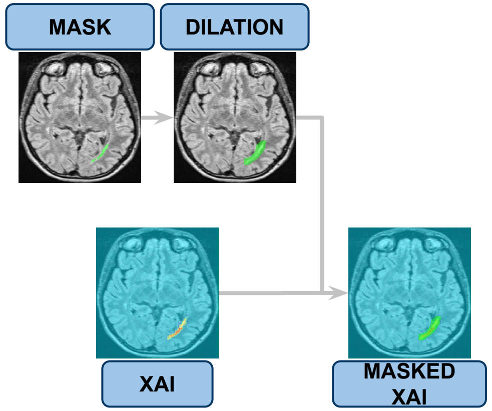


19 first order

74 second order

- Gray Level Co-occurrence
- Gray Level Run Length
- Gray Level Size Zone
- Neighbouring Gray Tone Difference
- Gray Level Dependence

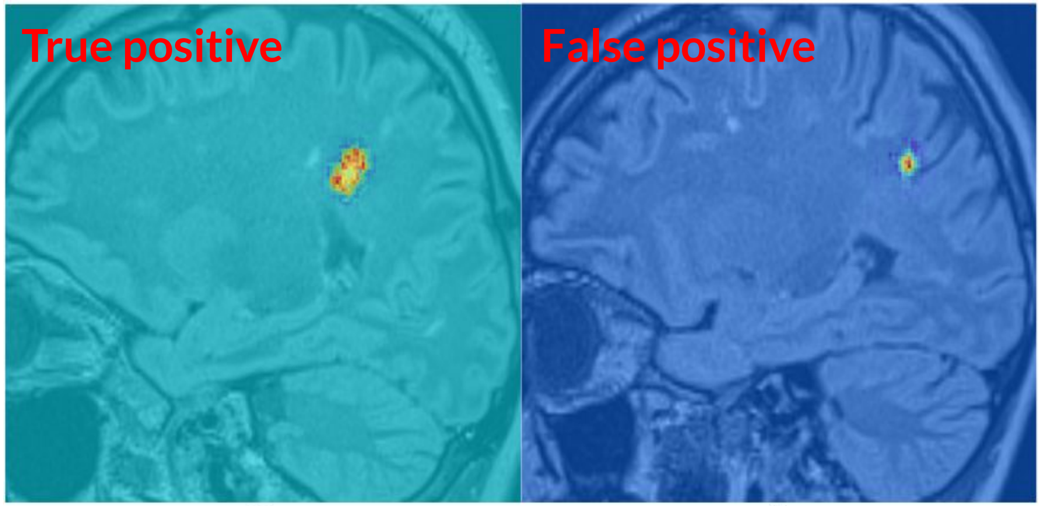
Bootstrap with test set to estimate confidence intervals of the performance



1. Introduction
2. Methods
3. **Results**
4. Discussion
5. Conclusion

Results

	U-Net only	U-Net + saliency [95% CI]
TPs	3050	2732
FPs	1818	763
FNs	789	1107
F1 score	0.7006	0.7450 [0.7358, 0.7547]
PPV	0.6265	0.7817 [0.7679, 0.7962]



(a) (b)

Fig.7 The TP case (a) obtained a score (LR) of 0.9398 for the positive class, while the FP (b) reported 0.0232 and was now classified as TN.

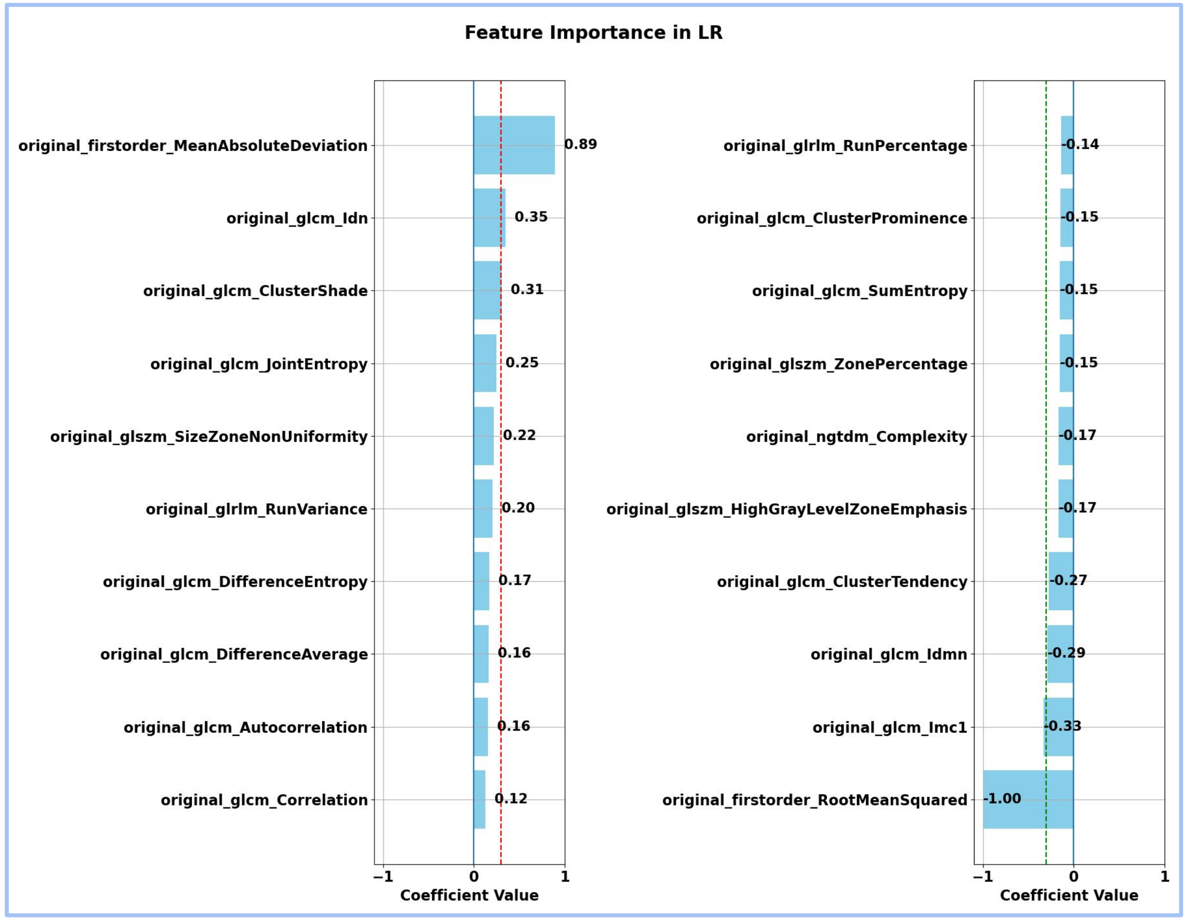
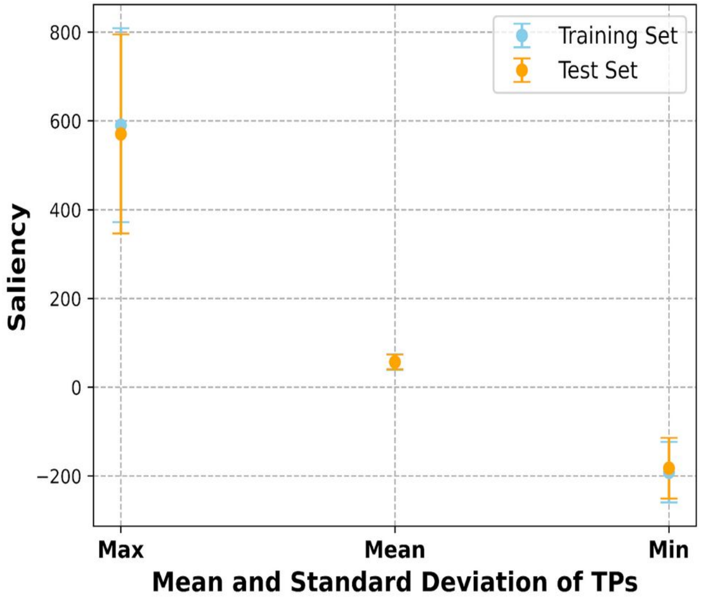


Fig.8 Normalized radiomic features showing the highest importance (top 10).

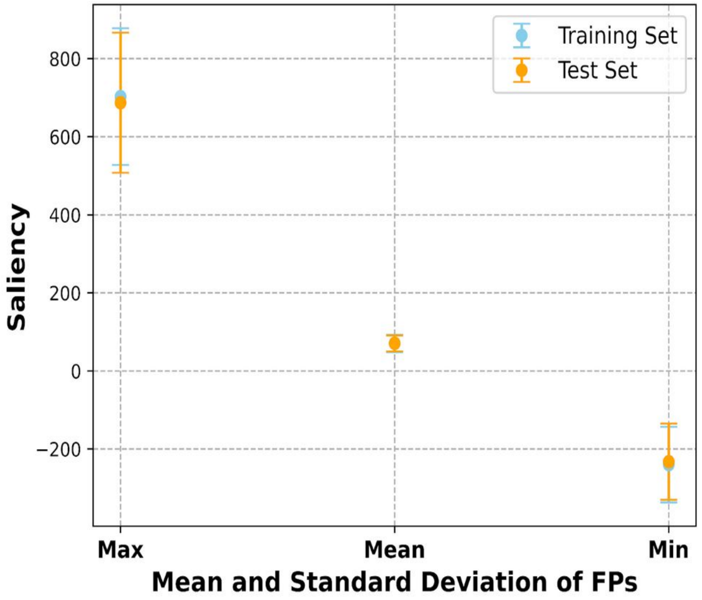
1. Introduction
2. Methods
3. Results
4. **Discussion**
5. Conclusion

Discussion

- Maximum, minimum and mean values of XAI in the training and test set were compared, to exclude domain shift



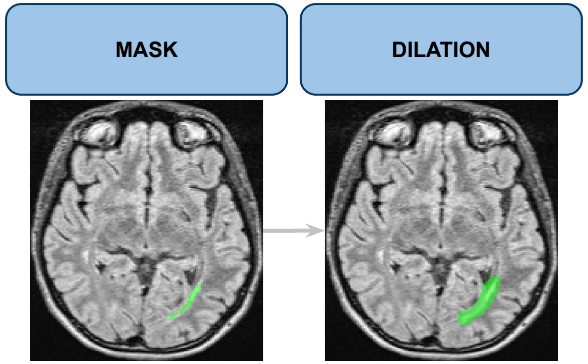
(a)



(b)

Discussion

- Mean absolute deviation (MAD) strong positive: more intensity variability around mean in true positive examples.
- Square root of the mean (RMS) strong negative: false positives present more outliers?



Open questions:

1. How many features are enough?
2. Explore shape features?
3. Location of refined lesions?
4. Apply to different domains?
5. Refine false negatives? Use uncertainty estimation?

1. Introduction
2. Methods
3. Results
4. Discussion
5. **Conclusion**

Conclusion

- **Instance-level XAI** (for segmentation task) can impact model performance and clinical practice
- **Radiomic features** on XAI can improve detection performance (F1 score) with a simple linear model
- First order features (RMS and MAD) seem to separate FP from TP the most

A SPECIAL THANKS TO THE TEAM!


N. Molchanova


B. Spahr


F. Spagnolo


J. Najm


M. Wynen


V. Andrearczyk


M. Graziani


P.J. Lu


A. Cagol


P. Macias Gordaliza


M. Bach Cuadra


A. Depeursinge


H. Muller


C. Granziera


L. Melie-Garcia

CIBM Center for Biomedical Imaging

medGIFT

THINK Translational Imaging in Neurology

HASLERSTIFTUNG

UNIVERSITY OF CAMBRIDGE

Isomorphic Laboratories


M. Gales


V. Raina


A. Malinin

EPFL ECAL LAB


C. Evans


D. Ribes