

Automated classification of glioma treatment response: benchmarking via modality selection and coarse tumor detection

Ana Matoso^{1,2*}, Catarina Passarinho¹, Marta P. Loureiro¹, José Maria Moreira², Pedro Vilela³, Patrícia Figueiredo¹, Rita G Nunes¹

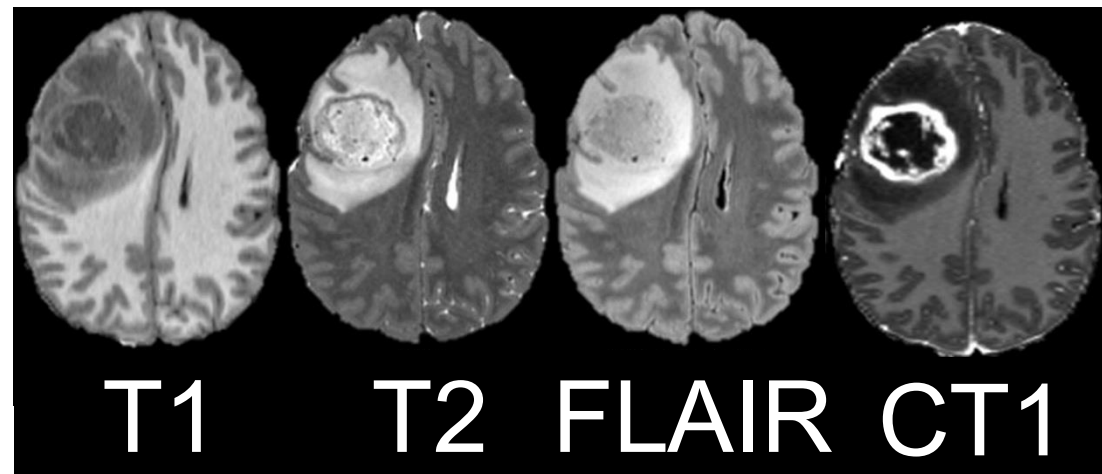
¹Institute for Systems and Robotics – Lisboa and Department of Bioengineering, Instituto Superior Técnico, University of Lisbon, Portugal;

²Learning Health, Hospital da Luz, Luz Saúde, Lisbon, Portugal; ³Imaging Department, Hospital da Luz, Luz Saúde, Lisbon, Portugal;

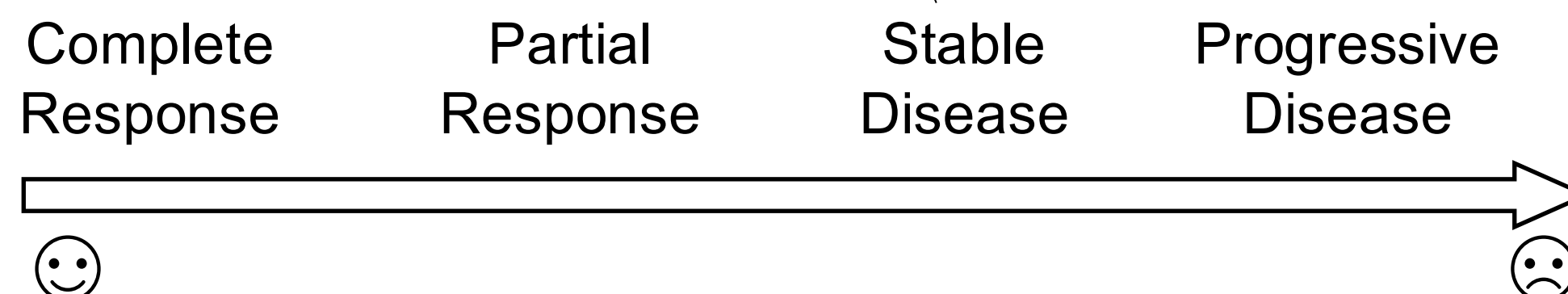
*anamatoso@tecnico.ulisboa.pt

Introduction

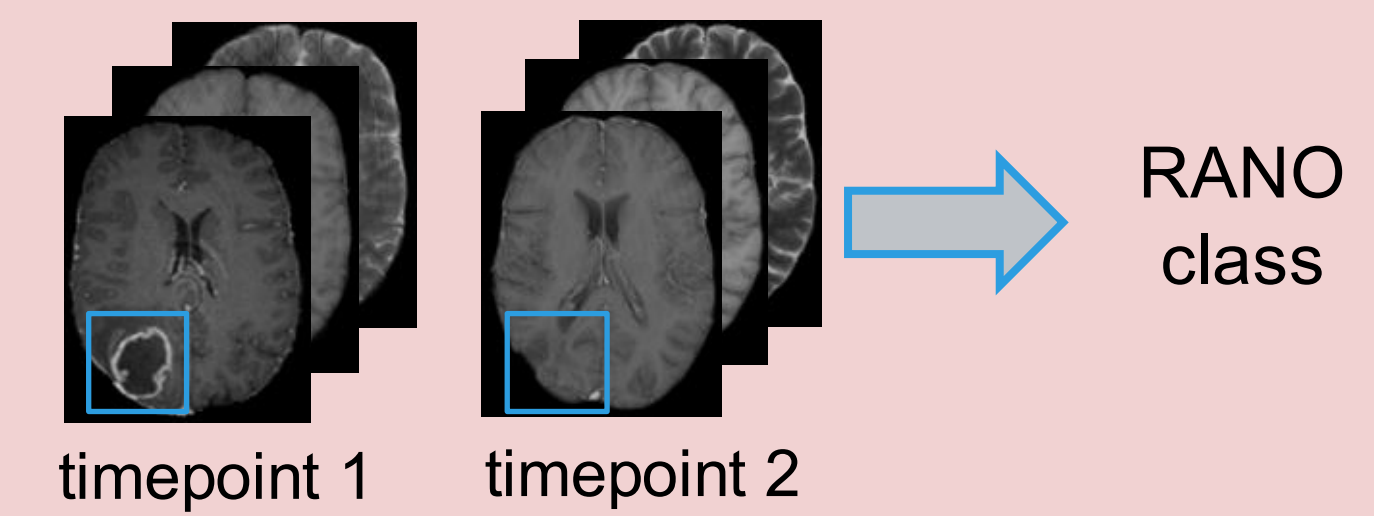
- Glioblastoma is the most malignant type of brain tumor¹
- Frequent MRI scans are done to assess **treatment response**²



Response Assessment in Neuro-Oncology (RANO) criteria³

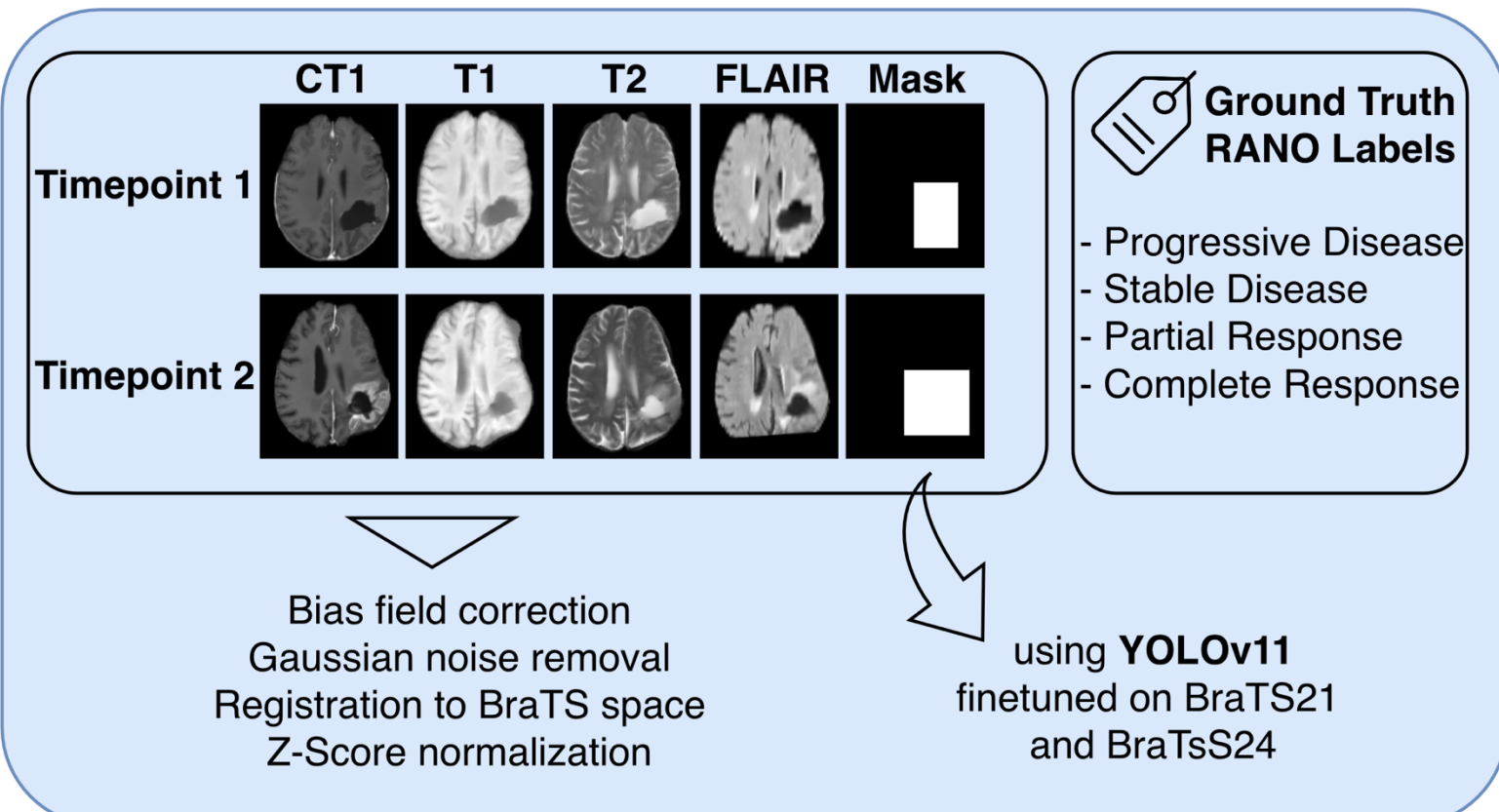


GOAL: Test different strategies for RANO criteria classification using MRI



Methods

Data Preprocessing

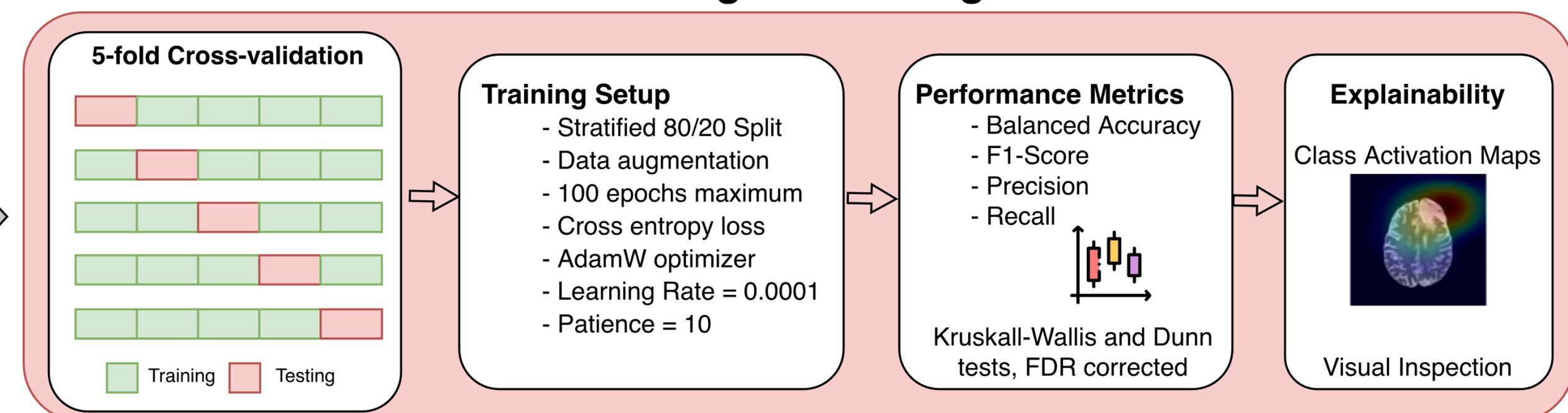


Data Selection

Modality Combinations	Number of timepoints	Classes Distribution			
CT1+T1+T2+FLAIR	337	228	66	20	23
T1+T2+FLAIR	338	229	66	20	23
T1+FLAIR	338	229	66	20	23
CT1+FLAIR	344	235	66	20	23

■ Progressive Disease
 ■ Stable Disease
 ■ Partial Response
 ■ Complete Response

Training and Testing



Lumiere Dataset⁴
638 timepoints of 91 patients
➤ MRI scans
➤ RANO label

YOLOv11 trained on FLAIR images⁵

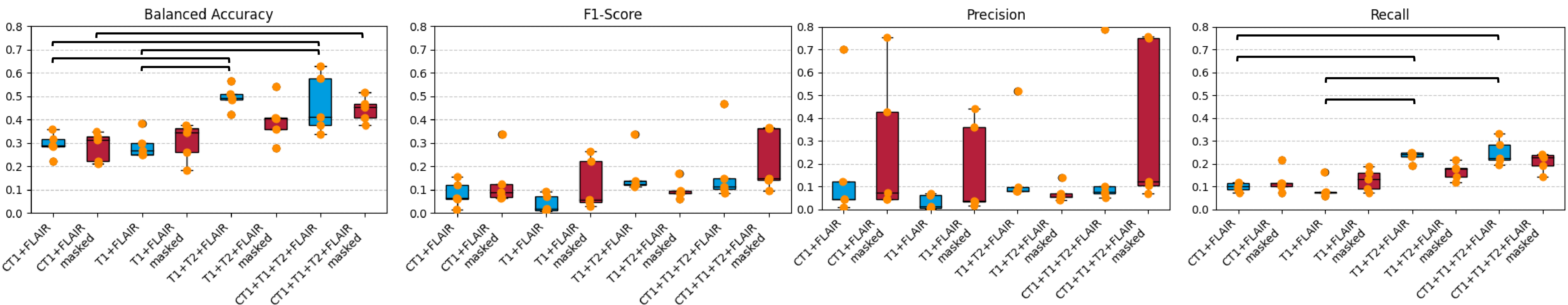
Selection of consecutive timepoints with at least the specified modalities

Using DenseNet264 from **MONAI[†]** and **PyTorch**

Results and Discussion

■ Unmasked ■ Masked

Significant results highlighted with brackets



- Using **all modalities unmasked** yielded the **best performance**: Balanced Accuracy = 62%
- No significant difference between using all modalities and **T1+T2+FLAIR**, which presented **higher median** values of metrics

- **Masking** the tumor region **generally improved classification robustness** though no significant differences were found
- Adding **T2** significantly **improved performance** compared to T1w+FLAIR

Unmasked

Progressive Disease

Classes	PD	SD	PR	CR
Probability (%)	21.3	55.5	3.4	19.8

Stable Disease

Classes	PD	SD	PR	CR
Probability (%)	32.1	24.7	39.7	3.6

Partial Response

Classes	PD	SD	PR	CR
Probability (%)	27.4	21.0	48.2	3.4

Complete Response

Classes	PD	SD	PR	CR
Probability (%)	5.3	10.9	13.3	70.5

Masked

Progressive Disease

Classes	PD	SD	PR	CR
Probability (%)	20.9	48.5	19.0	11.5

Stable Disease

Classes	PD	SD	PR	CR
Probability (%)	21.5	57.0	13.2	8.3

Partial Response

Classes	PD	SD	PR	CR
Probability (%)	9.1	13.1	75.6	2.2

Complete Response

Classes	PD	SD	PR	CR
Probability (%)	0.2	0.9	25.7	73.2

PD=Progressive Disease

SD=Stable Disease

PR=Progressive Response

CR=Complete Response

- In most explainability images the tumor region was highlighted, especially when tumor detection is performed beforehand
- External Validation is needed
- More **advanced modalities**, namely diffusion and perfusion could provide complementary information

Conclusion

- This study **benchmarked** the **automatic classification of RANO criteria** in a longitudinal dataset of glioblastoma
- **Masking** the tumor region tends to **improve performance/robustness** in **most cases**
- Top performance was achieved using **all modalities unmasked** reaching a **Balanced Accuracy of 62%**
- **T1+T2+FLAIR performance was more robust** than using all modalities
- Explainability methods highlighted the steering of the model's activations toward the tumor region

References: 1. Schwartzbaum, et al. Nat Rev Neurol, 2006. 2. Horbinski et al. J Nat Comprehensive Cancer Network, 2023. 3. Ellingson, et al. Neurotherapeutics, 2017. 4. Suter et al. Scientific Data, 2022; 5. Palmeirão, et al. Magn. Reson. Mater. Phys., Biol. Med., 2025