FOUNDATION MODELS for GLIOMAS HISTOPATHOLOGY IMAGE CLASSIFICATION

Lucas RINCÓN DE LA ROSA[1], POLA network and Agustí ALENTORN[1].
[1]*Huillard-Sanson, Genetics and Development of Brain Tumors, Paris Brain Institute CNRS UMR 7225/INSERM U 1127/Sorbonne Université, Paris, France*

INTRODUCTION

Diffusely infiltrating gliomas in adults are currently classified into WHO grades I–IV, reflecting varying malignancy levels. Recent advances have led to a refined classification of diffusely infiltrating gliomas in adults into three groups: IDH-mutant, 1p/19q codeleted tumors (best prognosis), IDH-mutant, 1p/19q non-codeleted tumors (intermediate prognosis), and IDH wild-type tumors (poor prognosis).

This study aims to use a deep learning (DL) image classification model on histopathology images from various cohorts to predict glioma grade, IDH status, and 1p/19q codeletion.

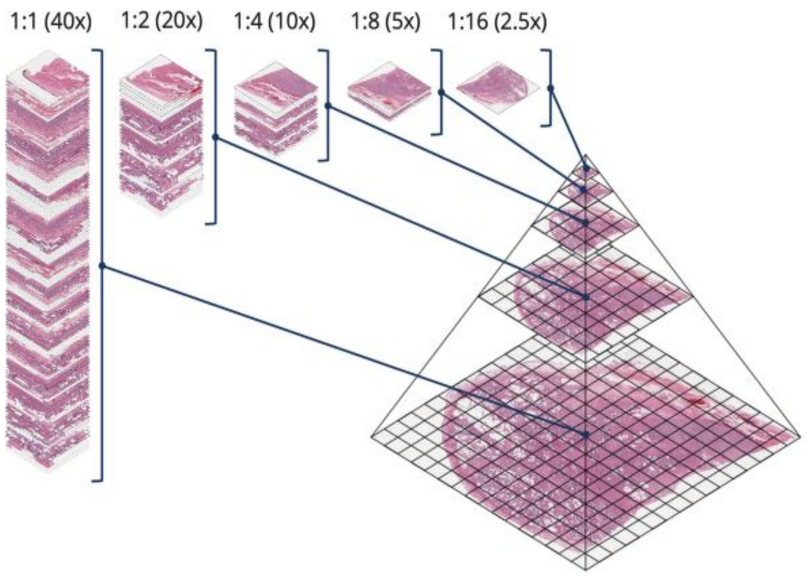

Figure 1: Whole Slide Image (WSI) structure used for the histopathology images.

METHODS

WSIs are high-resolution images that can store gigabytes of data, making them significantly larger than standard images The general structure is represented in **Figure 1**. To facilitate processing by neural networks (NNs), WSIs are divided into smaller units called patches. Pre-segmentation is performed to exclude irrelevant patches, reducing unnecessary computation. This processing method is outlined in **Figure 2**.

Figure 3 illustrates the outline of our experiments, where WSIs from two public datasets, TCGA (5) and eBrains (6), are used for training and validation of the model, while the external validation from POLA cohort is used for testing. The summary of the datasets used is collected in **Table 1**.

For the classification model, we use a well-known foundation model called clustering-constrained-attention multiple-instance learning (CLAM) (1) with pre-trained weights for feature extraction (2).

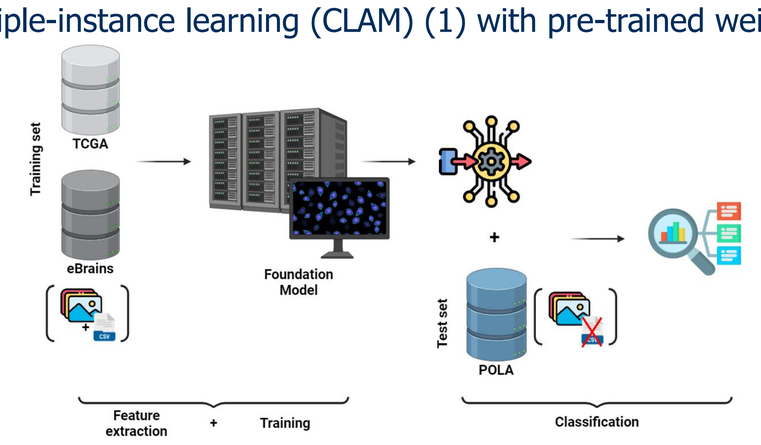

Figure 3: Architecture outline of the training and testing process of our experiments.

Use	Cohort	Total
Training + Validation	TCGA	833
	eBrains	805
Test	POLA	269

Table 1: Resume of the WSI datasets used for training and testing the classification model.

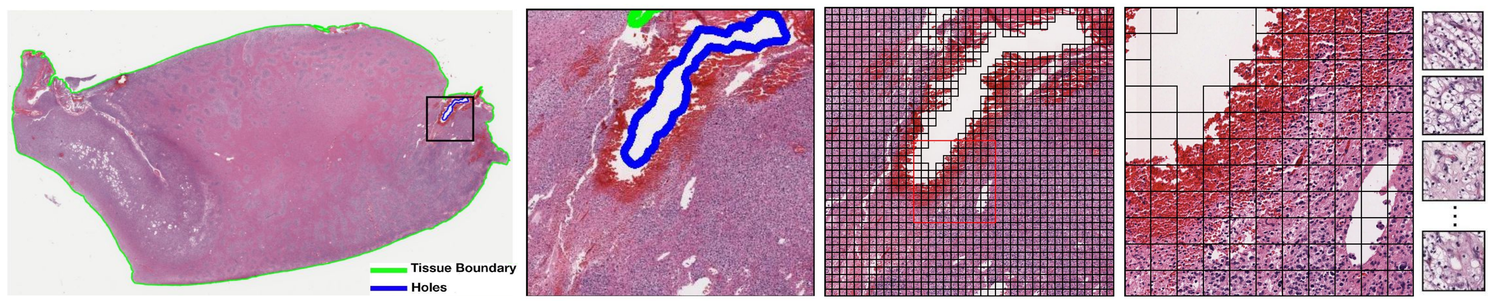

Figure 2: Resume of the WSI preprocessing technique for patch and feature structure for MIL-based models like CLAM.

RESULTS

We performed 10-fold classification training for three cases: grade (only grades II, III, and IV are present in the test cohort), IDH status, and 1p/19q codeletion. Each case was trained and tested separately. In **Figure 4**, we present the test results for IDH status classification across the different folds, showing the ROC curve and AUC as the accuracy metric. The lowest AUC value observed is 0.78, with an average AUC across the folds of 0.82. For the 1p/19q codeletion model, the average AUC across the 10 folds was 0.79, while for grade classification, the obtained AUC was 0.66.

In **Figure 5**, we display one of the original histopathology WSI used for testing along with the corresponding heatmap generated by the neural network weights for the IDH status classification model. The heatmap uses warmer colors to indicate regions of the WSI that the model focused on more heavily in determining the predicted label, and colder colors to represent less relevant regions.

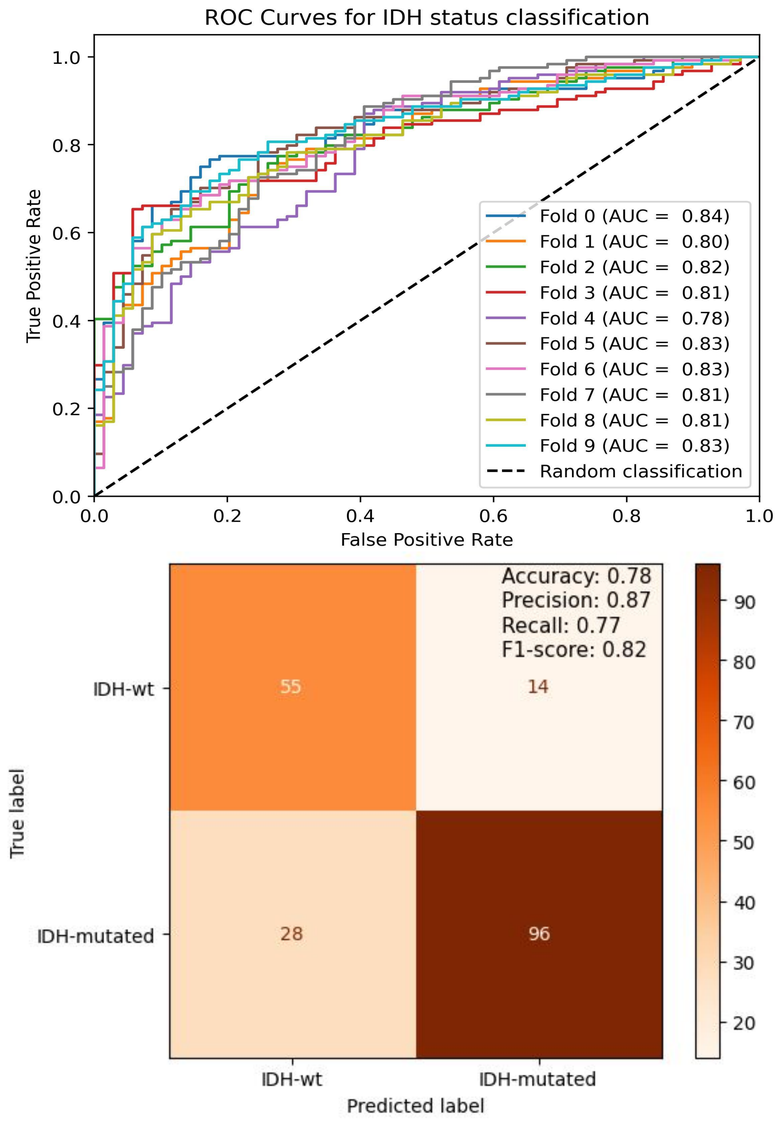

Figure 4: ROC curves for the 10-folds using the POLA dataset for validation along with the AUC values (above) and confusion matrix for fold 1 (below).

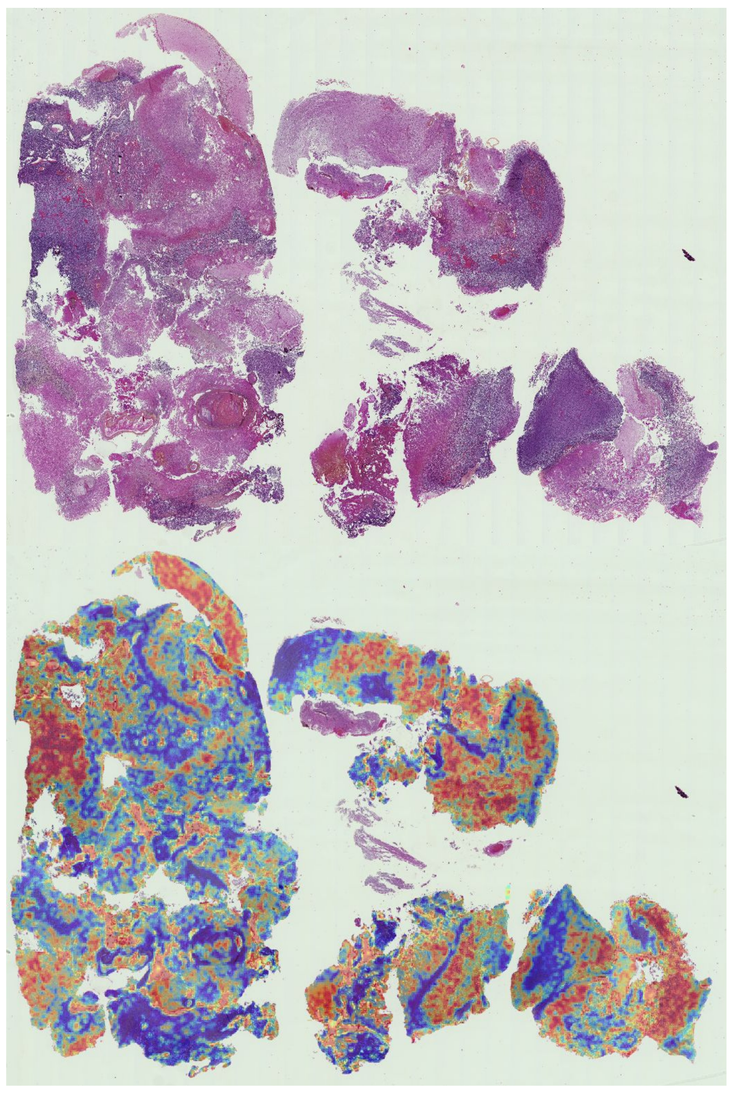

Figure 5: WSI from POLA dataset (above) along with heatmap (below) for the IDH classification model

CONCLUSIONS

MIL-based classification models show high accuracy when applied to histopathology images. While the IDH-status and 1p/19q codeletion models performed well, the grade classification model was less effective. This may be due to annotation inconsistencies and the grouping of WHO grade II and III gliomas, since TCGA and eBrains are classified using the WHO 21. A binary classification approach (grade II-III vs. grade IV) could give better results.

With an AUC above 0.75, we can conclude that the model's accuracy is solid, aligning well with other state-of-the-art methods (3).

FUTURE WORK

- Comprehensive foundation model integrating various data types like different imaging modalities.
- Test classifications for specific glioma groups, such as:
  - Grade IV IDH-wildtype Gliomas.
  - Grade III IDH-mutated Oligodendrogliomas
- Perform survival prediction analysis.
- More detailed intratumoral heterogeneity analysis.
- Ablation study to investigate correlations between different glioma groups.
- Highlight ROI extracted from the model's heatmaps.

REFERENCES

(1) Lu, M.Y., Williamson, D.F.K., Chen, T.Y. et al. Data-efficient and weakly supervised computational pathology on whole-slide images. Nat Biomed Eng 5, 555–570 (2021).
(2) Chen, R.J., Ding, T., Lu, M.Y. et al. Towards a general-purpose foundation model for computational pathology. Nat Med 30, 850–862 (2024).
(3) Wang, X., Zhao, J., Marostica, E. et al. A pathology foundation model for cancer diagnosis and prognosis prediction. Nature (2024).
(4) Katherine J Hewitt, K.J., Löffler, C.M.L. et al. Direct image to subtype prediction for brain tumors using deep learning, Neuro-Oncology Advances, Volume 5, Issue 1, January-December 2023, vdad139
(5) https://www.cancer.gov/ccg/research/genome-sequencing/tcga
(6) Roetzer-Pejrimovsky, T., Moser, AC., Atli, B. et al. The Digital Brain Tumour Atlas, an open histopathology resource. Sci Data 9, 55 (2022).