Supplementary Material

Nicholas Nuechterlein, Beibin Li, Mehmet Saygin Seyfioğlu, Sachin Mehta, Patrick J Cimino, and Linda Shapiro

University of Washington

Note: A longer version of this work has been submitted to the International Conference on Pattern Recognition (ICPR2020).

1 Training details

For consistency, we use the same hyperparameters across all experiments. We use a batch size of 4 and the ADAM optimizer with a learning rate of 0.0005. We use a dynamic learning rate scheduler that decreases the learning rate by a factor of 10 if the validation loss does not decrease by 0.0001 over 10 epochs. We deliberately use minimal prepossessing and data augmentation. We min-max normalize each MR sample. For both 4-channel and 1-channel input, we zero 20% of voxels with probability 0.8. For 4-channel input, we set a random channel to zero with probability 0.5. All cropped tumors are resized to $64 \times 64 \times 64$. When using the entire MR images, we crop the MR volumes down to the brain edges and resize the volumes to $144 \times 144 \times 144$. All models are trained using PyTorch on either an NVIDIA 1080 or 1080 Ti GPU. We report the average maximum ROC-AUC score for each model trained for 50 epochs over 10 trials. We mitigate overfitting by evaluating our MTL models exactly once on the validation set. All PCA dimensions were chosen empirically.

2 Data details

2.1 Magnetic Resonance (MR) Imaging

We download multi-modal MR data for 542 patients from the 2018 BraTS challenge, 285 of which are from the challenge's training set, and 67 and 192 are from the challenge's validation and testing set, respectively [1–3]. Each patient is composed of four MR modalities: pre- (T1) and post-contrast (T1ce) T1-weighted sequences along with the T2-weighted (T2) and T2 Fluid-Attenuated Inversion Recovery (FLAIR) volumes pictured in Figure 1. Each modality is a volume of dimension $240 \times 240 \times 155$. All MR volumes are skull stripped, co-registered, and resampled to 1mm^3 isotropic resolution [1].

Of these 542 patients, 235 are among The Cancer Imaging Archive (TCIA) low-grade glioma and glioblastoma cohorts and have survival data and genomic data in The Cancer Genome Archive (TCGA) [4]. Of these 235 patients, 160 are part of the BraTS 2018 training set, and 75 are part of either the BraTS 2018 validation or testing set. The 160 patients in the training set have ground-truth 4-class segmentation maps that delineate three tumor compartments and

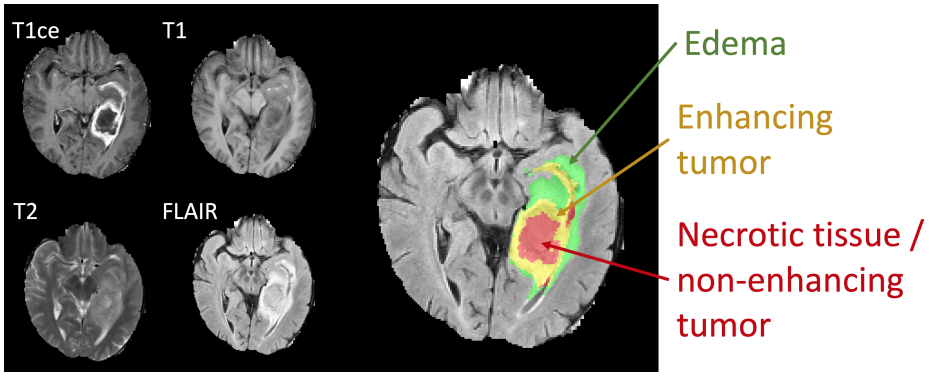

Fig. 1. Tumor compartments: enhancing tumor (yellow); peritumoral edema (green); necrotic tissue and non-enhancing tumor (red).

a background class. The three tumor compartments are listed below and shown in Figure 1.

- Enhancing tumor
- Peritumoral edema
- Necrotic tissue and non-enhancing tumor

2.2 Somatic Copy Number Alteration (SCNA)

For all 235 TCGA/BraTS patients, we downloaded thresholded gene-level somatic copy number alteration (SCNA) data estimated using the R package GISTIC2.0 [5] and gene-level non-silent mutation data (n = 169) from the University of California Santa Cruz cancer browser (https://genome-cancer.ucsc.edu/) to determine IDH1/2 mutational status. Clinical data for these patients were obtained from the Genomic Data Commons Data Portal from the National Institutes of Health (NIH) [6]. Relevant survival fields in the clinical data include overall survival and whether death was observed. We validate all models on the 75 samples in the TCGA that have MR data in the 2018 BraTS validation or testing sets. For IDH1/2 classification, we only use the 59 samples with IDH1/2 labels.

3 Additional Integrated Gradient Figures

We additionally interpret our classification results using integrated gradients, a gradient attribution technique that can visualize a network's per-sample focus, in Figures 2, 3, and 4 [7].

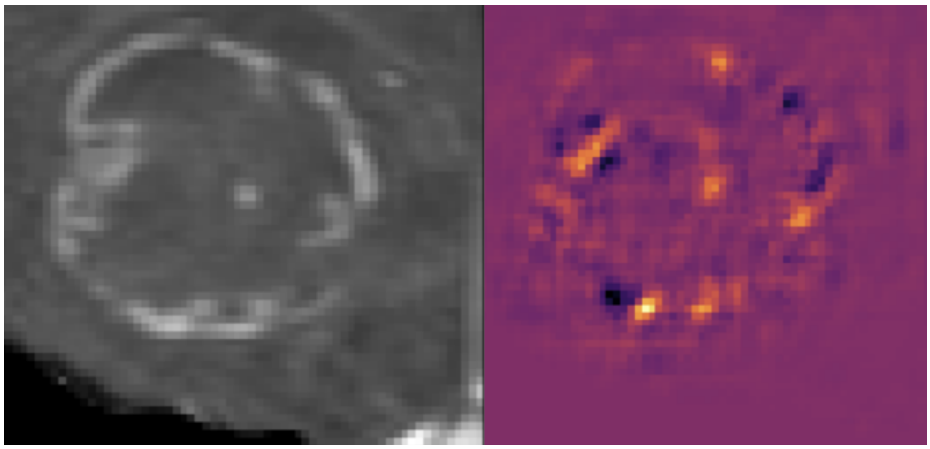

Fig. 2. IDH1/2 mutant misclassified as an IDH1/2 wildtype glioma. This sample is consistent with our conjecture that our models are associating contrast enhancement with IDH1/2 wildtypes and WHO grade IV glioblastomas. As an enhancing IDH1/2 mutant tumor, we would expect that our models would predict it is an IDH1/2 wildtype. Additional evidence is apparent in the integrated gradients plot, where it seems that the model is focused on the clearly visual enhancing ring.

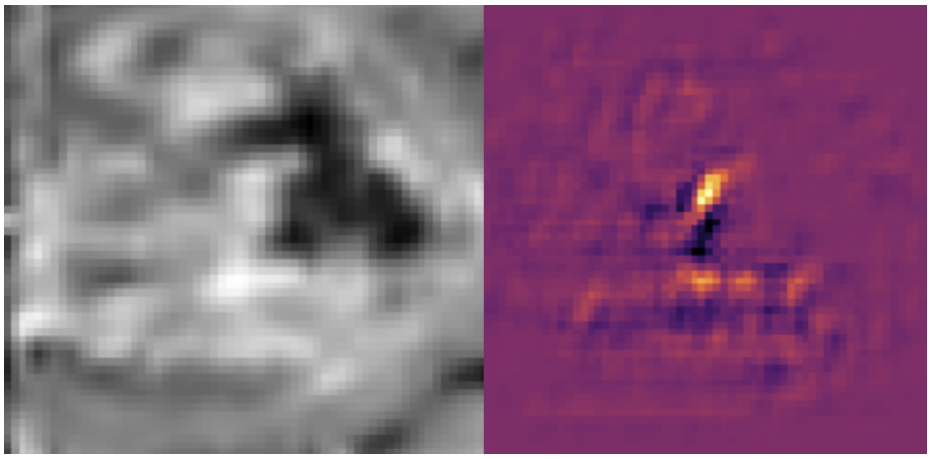

Fig. 3. Correctly classified IDH1/2 wildtype. Emphasis is placed on the borders of the enhancing and necrotic regions, exactly the area which describes the tumor enhancing ring. Again, this is consistent with our conjecture that our models predict tumors with enhancement to be IDH1/2 wildtype.

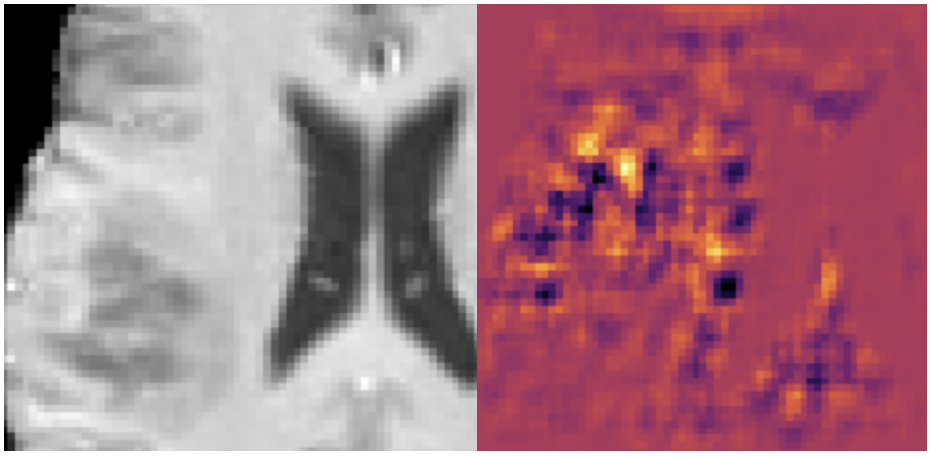

Fig. 4. IDH1/2 wildtype glioma incorrectly classified as an IDH1/2 mutant. The tumor is extremely large though the extent of the peritumoral edema is not apparent on the T1ce modality. Moreover, not only is the tumor relatively small on the cropped image compared to most other tumor, this IDH1/2 tumor has no visual enhancement, a characteristics we have observed our models associate with IDH1/2 mutant tumors. This tumor is labeled as a 'Mixed glioma' in the TCGA, but does not have an IDH1/2 mutation.

4 Additional Notes

We do not give direct comparisons to previously reported results, because our goal is not to outperform others. Instead, we endeavor to test the advantage of using unlabeled MR over supervised CNNs and do not want optimizations, such as network architecture, to obscure the benefit of our strategy. We also do not formally report our segmentation results, because segmentation is performed only to increase classification and survival prediction performance, not to compete with dedicated segmentation models.

References

1. Bjoern H Menze et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10), 2014.
2. Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, et al. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. *Scientific data*, 4, 2017.
3. Spyridon Bakas et al. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge. *arXiv preprint arXiv:1811.02629*, 2018.
4. Spyridon Bakas, Hamed Akbari, et al. Segmentation labels and radiomic features for the pre-operative scans of the tcga-gbm collection. the cancer imaging archive. *Nat Sci Data*, 4:170117, 2017.
5. Craig H Mermel, Steven E Schumacher, Barbara Hill, et al. Gistic2. 0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome biology*, 12(4), 2011.
6. Robert L Grossman, Allison P Heath, Vincent Ferretti, et al. Toward a shared vision for cancer genomic data. *New England Journal of Medicine*, 375(12), 2016.
7. Mukund Sundararajan et al. Axiomatic attribution for deep networks. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR. org, 2017.