PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

Leveraging Unlabeled Data for Glioma Molecular Subtype and Survival Prediction

Nicholas Nuechterlein

Paul G. Allen School of Computer Science & Engineering
The University of Washington

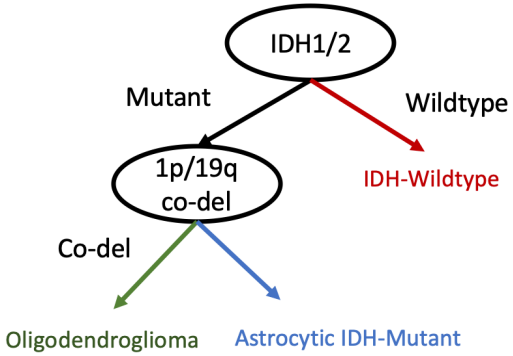
Overview

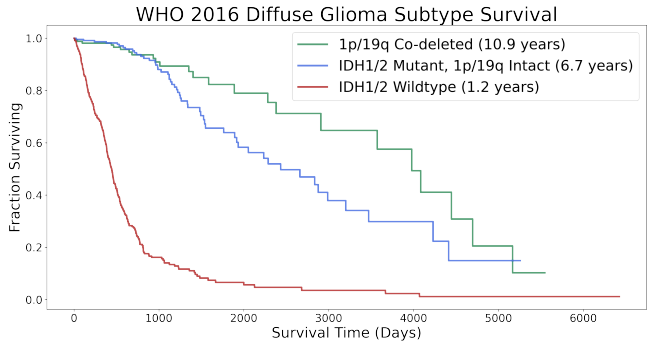
- Setting
  - Genomic biomarkers define brain cancer subtypes which stratify patient survival
  - However, obtaining biomarker status requires invasive surgery
  - Imaging is a non-invasive method that has been used to infer biomarker status
  - Unlabeled imaging data (MRI) is relatively abundant and underutilized
- Questions
  - Can we leverage unlabeled imaging data to better predict these biomarkers (subtypes)?
  - Can we incorporate genomic data, when available, in these models?
  - Subtypes are surrogates for survival—can we predict survival directly?
- Datasets
  - 2018 Brain Tumor Segmentation (BraTS) Challenge: magnetic resonance imaging (MRI)
  - The Cancer Genome Atlas (TCGA): somatic copy number alteration (SCNA)

Glioma

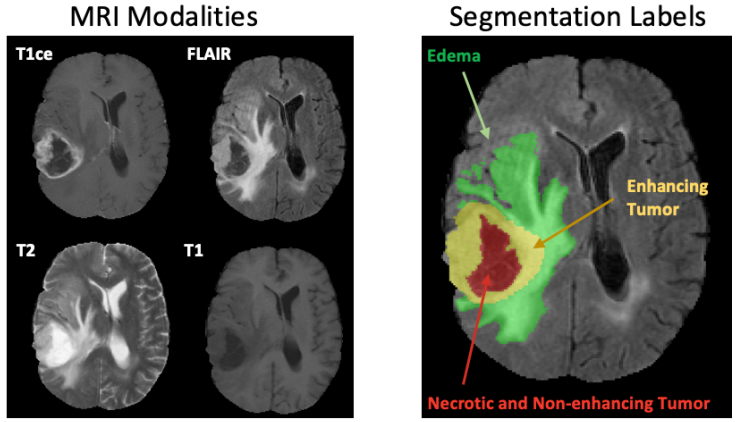
- 80% of primary malignant brain tumors in adults
- Biomarkers: IDH1/2 mutation, 1p/19q co-deletion

WHO 2016 molecular-based glioma classification

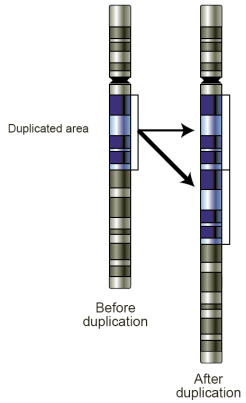


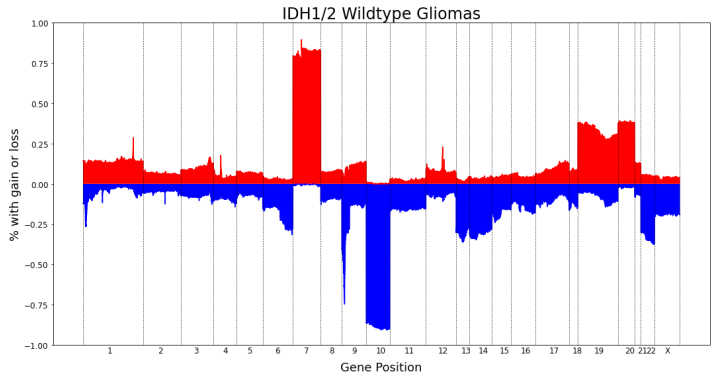


Magnetic resonance imaging (MRI)



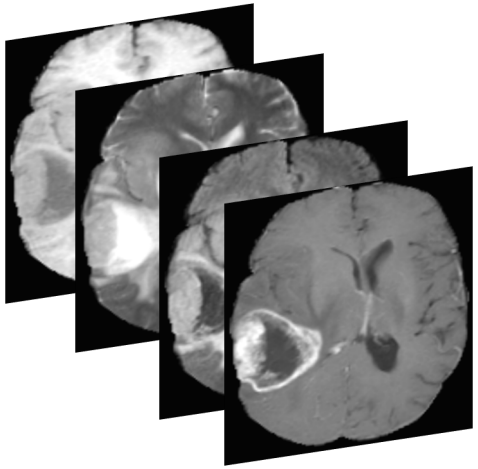
Somatic copy number alteration (SCNA)



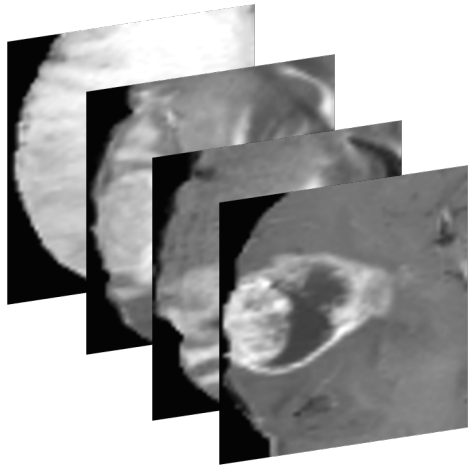


Menze *et al.*, IEEE transactions on medical imaging 2014, Bakas *et al.*, Scientific Data 2017, Bakas *et al.*, arXiv 2018

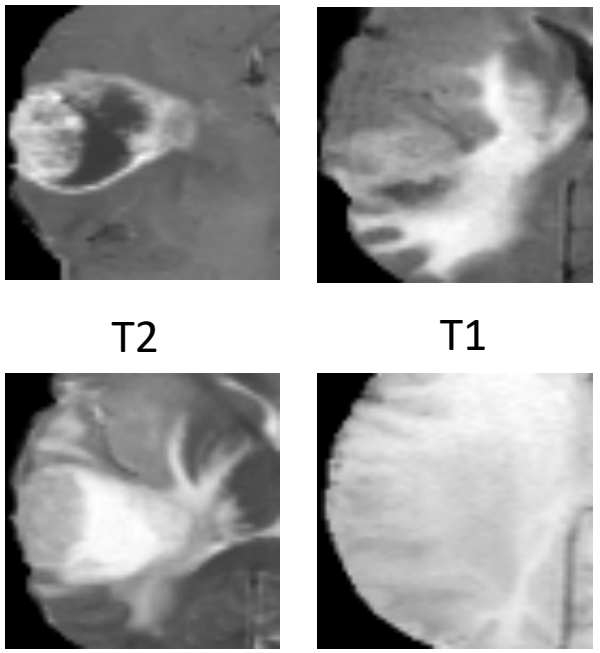
MRI inputs



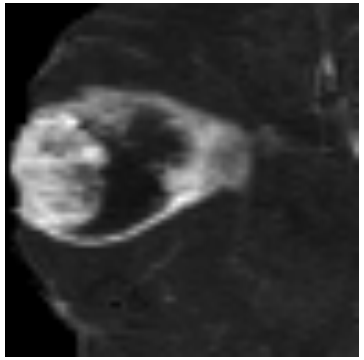
All Modalities
(4 x 144 x 144 x 144)



All Modalities (cropped)
(4 x 64 x 64 x 64)



Single Modalities (cropped)
(64 x 64 x 64)



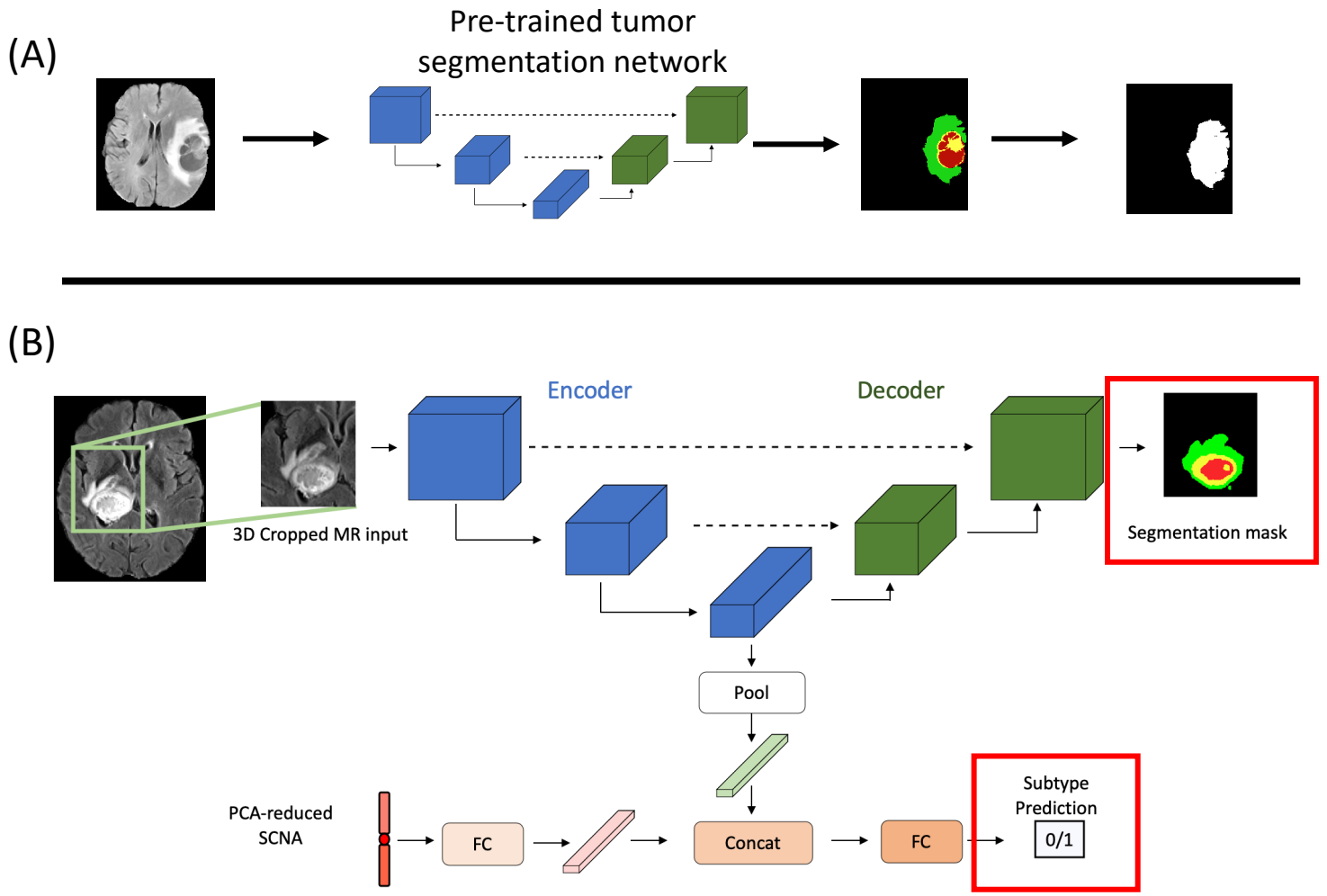
T1ce-T1 Subtraction
Map (64 x 64 x 64)

MTL network, glioma subtype classification

We use tumor segmentation as an auxiliary learning task.

(A) We assign weak segmentation labels to MRI samples without segmentation labels.

(B) We train our MTL model to learn glioma subtype jointly with tumor segmentation. For samples without subtype labels, we only take the segmentation loss.



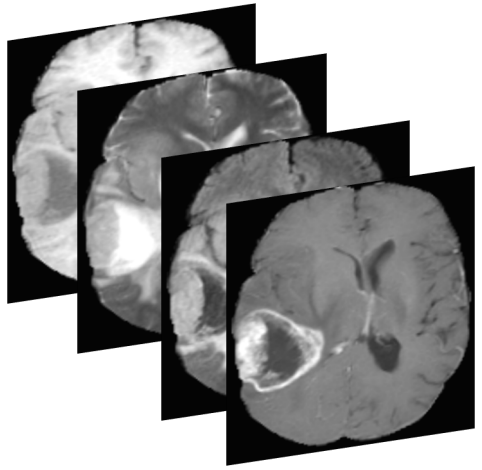
Nuechterlein, *et al.*, International MICCAI Brainlesion Workshop 2018

Subtype prediction results

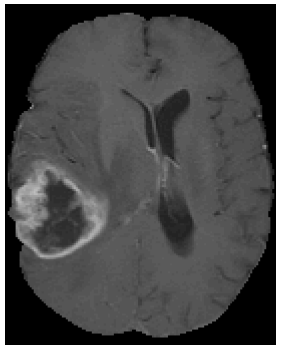
- MTL offers biggest boost to 4-channel, whole brain input
- 1p/19q co-deletion prediction is significantly improved
- T1ce-based modalities are the best predictors of IDH1/2 mutations; these models focus on tumor enhancement

Input Modalities	IDH1/2 Mutation (AUC)		1p/19q Co-deletion (AUC)	
	CNN	MTL (MR)	CNN	MTL (MR)
All (Whole Brain)	0.669	**0.846**	0.605	**0.813**
All (Cropped)	0.872	**0.894**	0.744	**0.871**
T1ce (Cropped)	**0.893**	0.884	0.772	**0.819**
FLAIR (Cropped)	**0.778**	0.690	0.755	**0.818**
T1 (Cropped)	0.731	**0.738**	0.727	**0.757**
T2 (Cropped)	**0.778**	0.732	0.740	**0.755**
T1ce-T1 (Cropped)	**0.895**	0.861	**0.764**	0.742

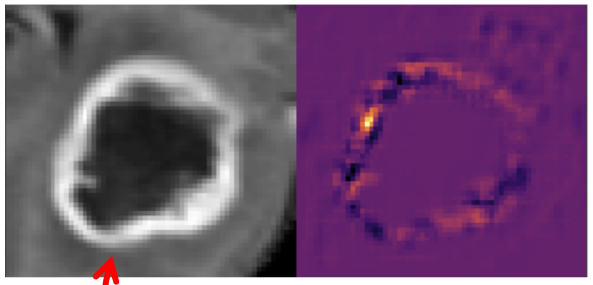
4-channel, whole brain input



T1ce modality



Integrated gradients visualization of a IDH1/2 wildtype prediction



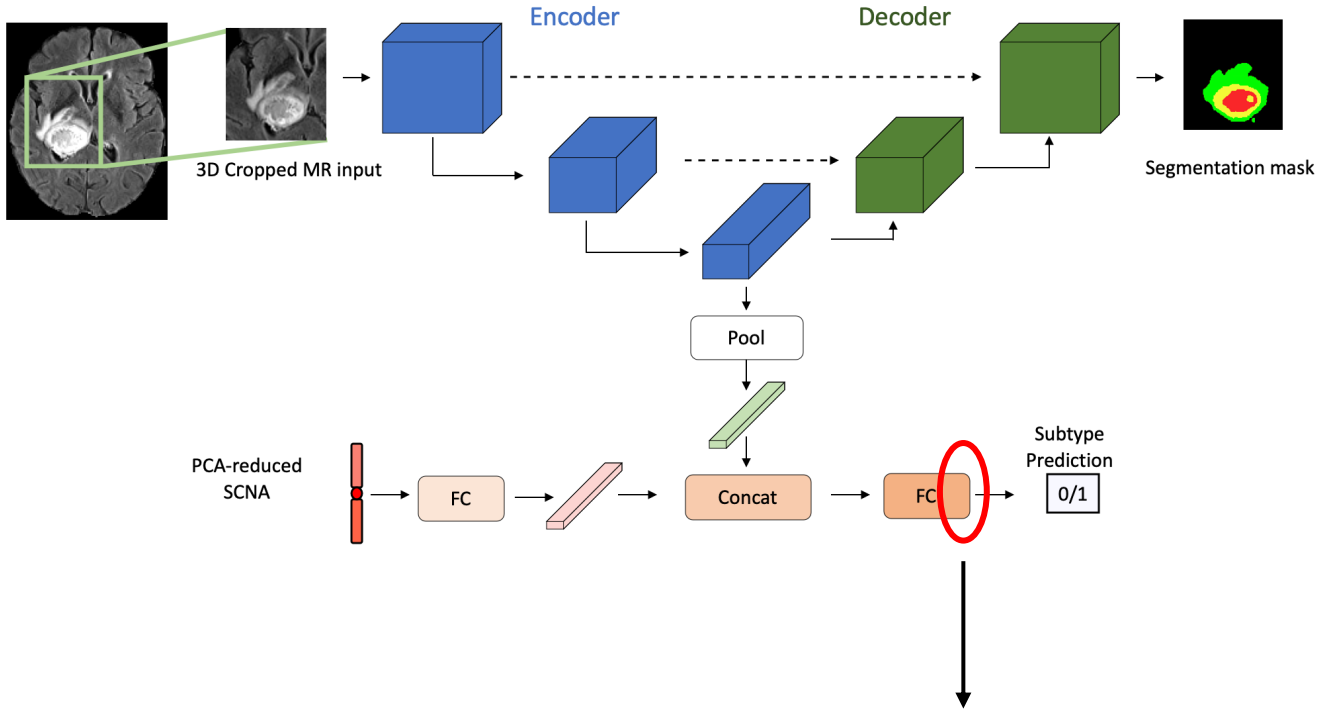
Ring enhancement

MTL network, survival prediction

1. We pretrain our MTL models on the IDH1/2 mutation classification task (both using only MRI and MRI + SCNA).

2. We use the last-layer embeddings of the classification branch in these MTL models to train linear Cox proportional-hazards models.

Note: We do not learn survival concurrently with subtype and tumor segmentation, because the CPH loss function requires large batch sizes that far exceed GPU memory given the size of 3D MRI data.



Survival prediction results

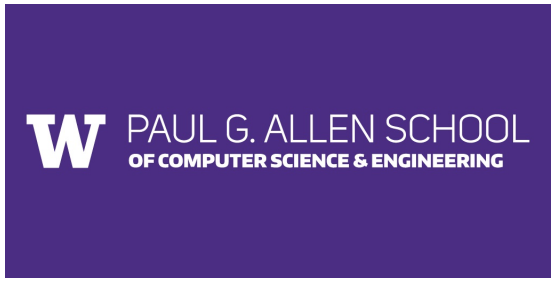
Input Modalities	All Subtypes (C-index)		1p/19q Co-deleted (C-index)		IDH1/2 Mutant, 1p/19q Intact (C-index)		IDH1/2 Wildtype (C-index)	
	MR	MR + SCNA	MR	MR + SCNA	MR	MR + SCNA	MR	MR + SCNA
All (Whole Brain)	0.587	**0.723**	0.714	**1.000**	0.606	**0.727**	0.487	**0.521**
All (Cropped)	0.697	**0.732**	0.607	**0.964**	**0.742**	0.712	0.540	**0.548**
T1ce (Cropped)	0.719	**0.735**	**0.821**	0.786	0.576	**0.742**	**0.644**	0.571
FLAIR (Cropped)	0.565	**0.731**	0.607	**0.964**	0.636	**0.712**	0.527	**0.540**
T1 (Cropped)	0.645	**0.728**	0.643	**0.893**	0.606	**0.636**	0.535	**0.544**
T2 (Cropped)	0.690	**0.718**	**0.679**	0.500	**0.803**	0.697	0.562	**0.563**
T1ce-T1 (Cropped)	0.707	**0.723**	0.821	**0.857**	**0.803**	0.682	0.523	**0.552**
SCNA (PCA = 5)	0.724		0.929		0.667		0.512	

- C-index measures how well model orders survival
- MRI + SCNA is better than MRI or SCNA on their own in most categories
- Models predict survival best in patients with 1p/19q co-deletions
- Survival prediction is difficult for patients with IDH1/2 wildtype tumors

Conclusion

- Unlabeled MRI data can be used to improve glioma subtype predictions, especially those defined by 1p/19q co-deletions
- MRI and SCNA mostly improve survival prediction beyond MRI or SCNA on their own
- MTL models associate tumor enhancement with IDH1/2 wildtype tumors

Acknowledgments



The University of Washington

- **Linda Shapiro**
- **Beibin Li**
- **Mehmet Saygin Seyfioglu**
- **Sachin Mehta**

UW Medicine
UNIVERSITY OF WASHINGTON
MEDICAL CENTER

University of Washington Medical Center

- **PJ Cimino**

Fred Hutch

- Eric Holland
- Sonali Arora

This material is based upon work supported by the National Science Foundation under Grant DGE-1762114.