Using Pathomic Imaging Data to Predict Histological Classifications of Pediatric Medulloblastoma

Children's Hospital of Philadelphia
Center for Data-Driven Discovery in Biomedicine

Phillip Duarte* (SEAS 2027), Andrew Wang* (CAS 2026)
Ariana Familiar*, Ali Nabavizadeh*▲ (Perelman School of Medicine, Neurosurgery)

*Center for Data-Driven Discovery in Biomedicine, Children's Hospital of Philadelphia, Philadelphia PA USA
▲ Department of Radiology, Hospital of the University of Pennsylvania, Philadelphia PA USA

PURM Penn Undergraduate Research Mentoring Program

Background

- Medulloblastoma is one of the most common malignant brain tumors in children
- Diagnosis of medulloblastoma involves an integrated analysis of molecular and histologic characteristics
- Pathology and radiology data collected through standard clinical care of pediatric neuro-oncology patients could offer predictive value in novel characterization of tumor types
- Machine learning methods have promise in harnessing the data to perform predictive forecasting and aid in clinical decision-making

Objective

Implement and evaluate the effectiveness of pathomic and radio-pathomic features in predicting clinically-relevant properties of pediatric medulloblastoma.

Significance

- There exists independent radiomic and pathomic analyses in pediatric neuro-oncology, but integrated radio-pathomic analyses are limited
- Micro-scale pathology data can provide insight into the biological meaning of macro-scale radiomic features, of which little is known
- Previous radio-pathomic analyses for pediatric brain tumors are limited by small sample sizes

Data

- Multi-institutional data was collected from the Children's Brain Tumor Network
- 203 samples of H&E stained whole-slide images
- Histological categories extracted from clinical pathology notes by team member into classic, desmoplastic, large cell/anaplastic (LCA) categories

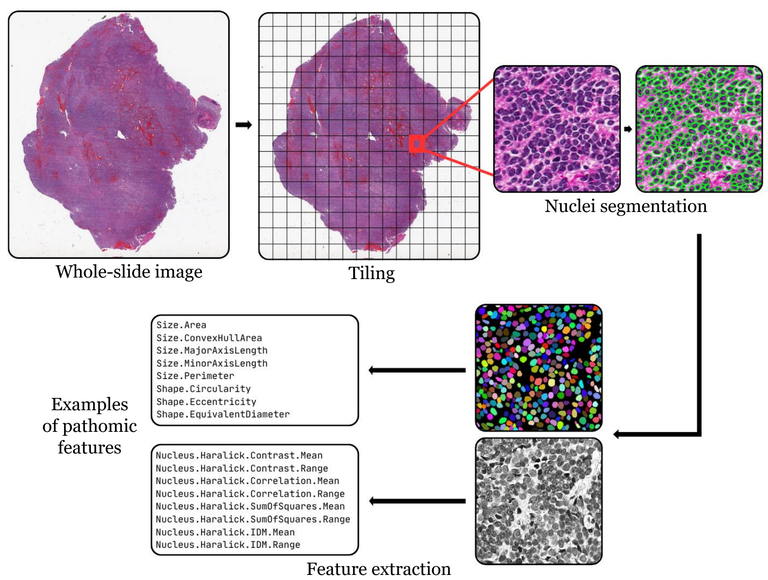
DATA PROVIDED BY CHILDREN'S BRAIN TUMOR NETWORK

References

1. This research was conducted using data and/or samples made available by The Children's Brain Tumor Network
2. Familiar et. al. (2023). Radio-pathomic approaches in pediatric neuro-oncology: Opportunities and challenges, *Neuro-Oncology Advances*.
3. Yan et al., 2020 Radiomic features from multi-parameter MRI combined with clinical parameters predict molecular subgroups in patients with medulloblastoma. *Frontiers in Oncology*.
4. Cooney, T., Lindsay, H., Leary, S., & Wechsler-Reya, R. (2023). Current studies and future directions for medulloblastoma: A review from the pacific pediatric neuro-oncology consortium (PNOC) disease working group. *Neoplasia*, 35, 100861. Chicago

Methods

Pathomic Analysis Workflow

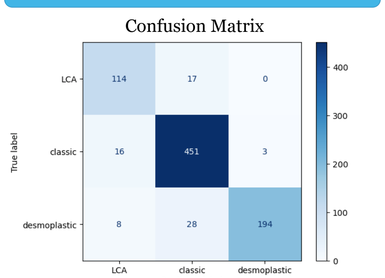


Model Learning

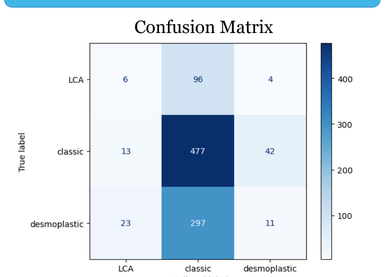
- Compiled a dataframe of median values of nuclei data for each tile
- Tested different resampling methods and number of features selected based on ANOVA of features and histological classifications
- Trained an SVC with grid search for parameter optimization
- Used five-fold cross validation with stratification of histological classes

Results

Under-sampling

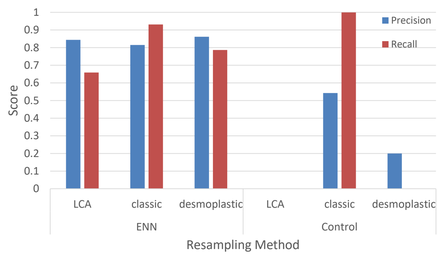


No Under-sampling

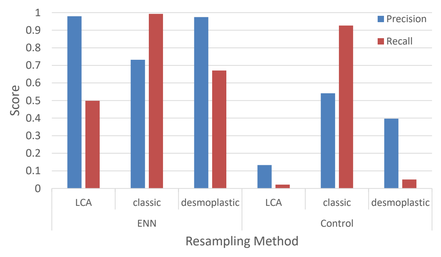


Results - Continued

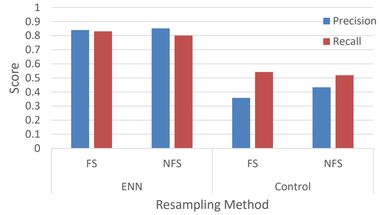
Feature Selection (k=60)



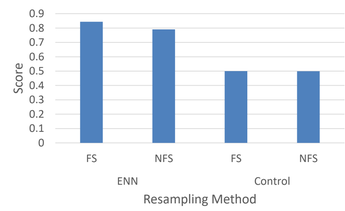
No Feature Selection



Average Precision and Recall of ENN vs Control



Area Under Curve (AUC) of ENN vs Control



Summary of Main Findings

- Under-sampling the entire dataset with feature selection resulted in best model performance (0.84 AUC) compared to the control (0.50 AUC)
- Feature selection resulted in less discrepancy between precision and recall compared to no feature selection but had little effect on AUC

Conclusions

- Under-sampling the entire dataset shows promising results but is limited in practice due to lack of generalizability
- The success of under-sampling the dataset may be due to removal of noisy data

Next Steps

- Remove uninformative tiles from dataset to improve data quality
- Try other classification algorithms to improve model performance
- Train and test model to classify samples into genomic categories
- Incorporate radiomic data to create integrated radio-pathomic model