

# Glioma Classification and Biopsy Guidance with Multimodal MRI and Deep Learning

Nhat Le

## 1. Topic and Goals

Glioma is a common type of brain tumor with average survival time ranging from 12-18 months to 7 years depending on the subtypes. Accurate classification and localization of tumor regions are needed for suitable treatment plans to increase the chance of survival. In recent years, deep learning has shown excellent performance in automated image-based diagnostic tools. Therefore, we seek to develop a deep learning pipeline using multimodal MR images, including a special sequence developed in Johns Hopkins called "Amide proton transfer-weighted" (APTw), to perform glioma classification tasks including tumor grading and treatment assessment and highlight salient regions. We will further analyze the performance of using APTw images to confirm that they add valuable information to the automated pipeline.

## 2. Team Members and Mentors

- Student: Nhat Le
- Mentors: Prof. Mathias Unberath, Prof. Shanshan Jiang

## 3. Relevance/Importance

Currently, the clinical standard for glioma diagnosis relies on MR images and subsequent biopsies, but standard MR sequences are not sufficiently tissue-specific enough to guide treatment decisions. On the other hand, APTw MRI has shown clinical values in tasks such as glioma grading and assessment of tumor recurrence. However, interpreting APTw images for such tasks requires radiologists to have additional training and expertise, which affects the rate of adoption in clinical settings. A fully automated pipeline using deep learning that could provide accurate classifications and explainable decisions with less human input will be valuable for a clinical decision support system.

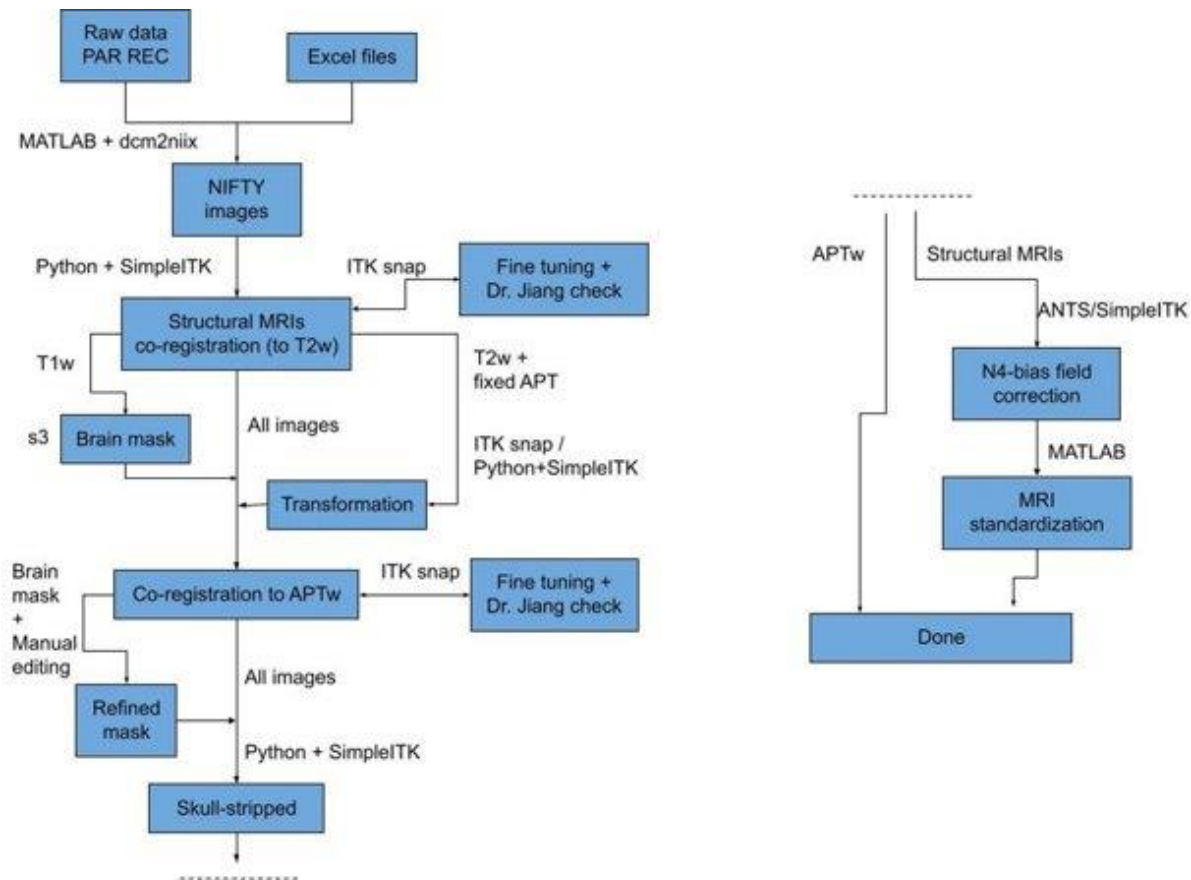
## 4. Technical Approach

The classification tasks we focused on are:

- Low-grade tumors vs. High-grade tumor for newly-diagnosed patients grading
- Recurrent tumors vs. Treatment effects for post-treatment assessment

### a. Data preprocessing

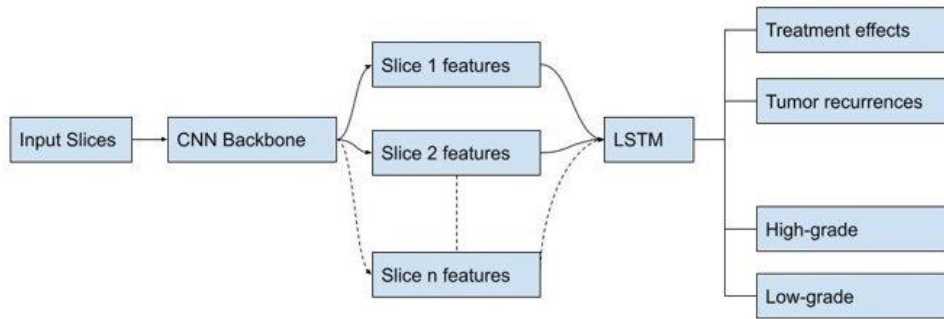
All scans in the dataset will be preprocessed to ensure high input quality for deep learning models with the following steps performed sequentially: co-registration, skull-stripping, N4-bias correction, and MRI standardization. The detailed workflow is described in the graph below:



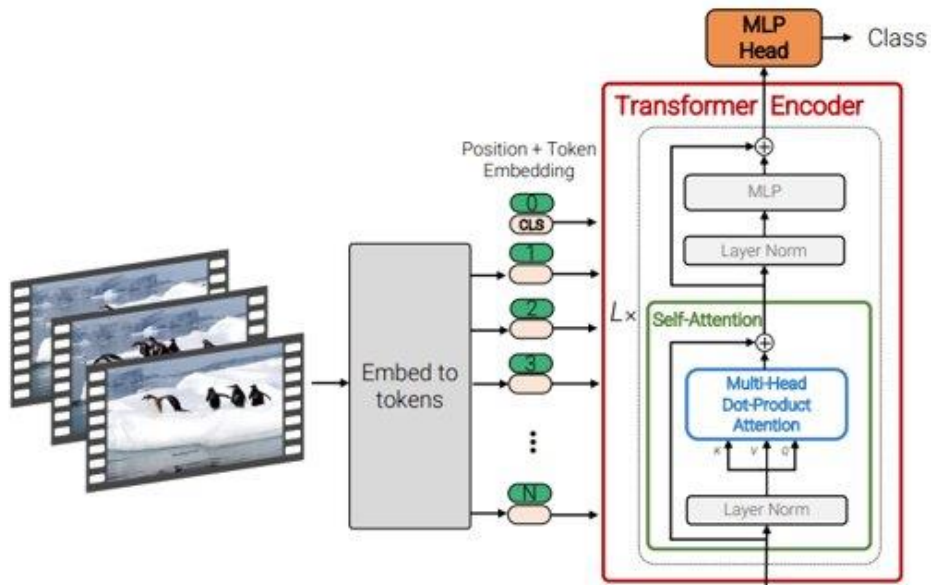
### b. Deep learning architecture

We will treat the 3D MRI scans as a sequence of 2D slices and use existing video classification algorithms for baseline models.

- Baseline 1: CNN-LSTM. This is an architecture commonly used in similar works in literature [4]. The CNN (convolutional neural network) module extracts feature from 2D images while LSTM (long-short term memory) acts as feature aggregator gathering information across all slices to perform classification tasks.

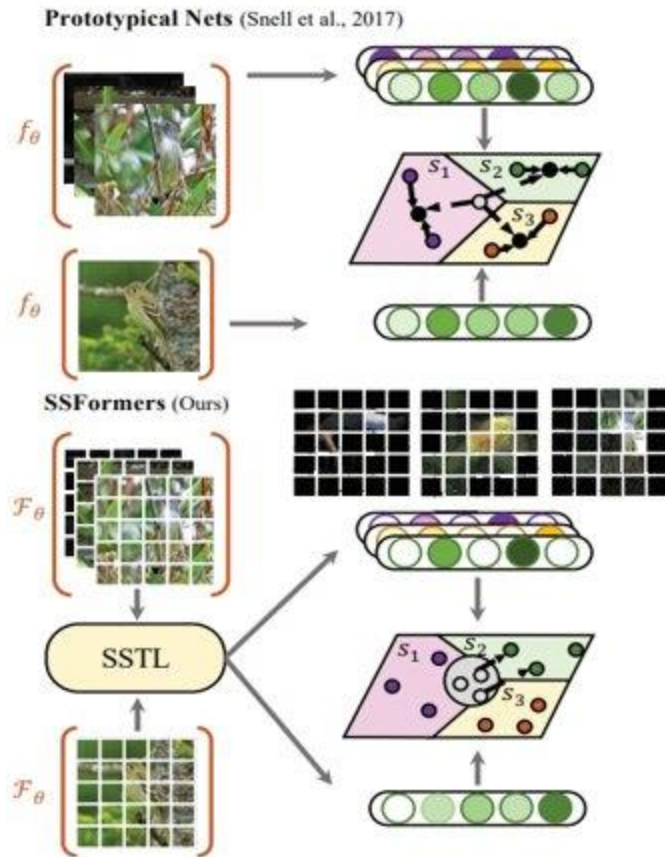


- Baseline 2: Vivit. This is an architecture using pure transformers both as feature extractor and feature aggregator across all slices [3]. Transformer is emerging as the predominant architecture in computer vision with certain advantages over CNN in modeling long dependencies between input sequence elements, which is important in our classification problems.



(Vivit architecture from [3])

We plan to design our improved methods based on the Vivit model. Our direction is to use few-shot classification approach on videos using cross-attention mechanism to highlight similarities between the support set and the query scan with attention scores. This approach will help us deal with the limited size of our dataset as well as provide better explanations for the network output. The figure below from [7] illustrates similar mechanism used in image classification tasks:



### c. Training

The following techniques will be employed for training deep learning models: pretraining, data augmentation, stratified sampling, hyperparameter tuning. These techniques will ensure that the deep learning models can reach optimal performance with the dataset in this project.

### d. Evaluation

The performance of the different deep learning models will be evaluated based on the following metrics: overall accuracy, AUC, and Precision-Recall AUC in case of imbalanced dataset. Salient regions obtained from these models will be compared qualitatively and possibly quantitatively using IoU with pixel-level annotations if available.

The performance of different design features will be evaluated by an ablation study with image modalities and deep learning model components.

## 5. Deliverables

- Minimum deliverable:
  - Validated pipeline for classification tasks with existing deep learning algorithms
  - Analysis on effectiveness of using APTw images

- Documentation of the dataset
- Expected deliverable:
  - Improved methods to increase classification and localization performance
- Maximum deliverable:
  - A precise method validated in clinical context

## 6. Key dates and assigned responsibilities

Project Timeline				
Task	Start	End	Week	Milestone
Literature Review	1/31/2022	3/4/2022	Week 2-6	
Data Preparation	2/7/2022	2/25/2022	Week 3-5	
Plan Proposal	2/7/2022	3/1/2022	Week 3-5	Proposal sent to mentors by 2/25
Baseline 1 evaluation	2/27/2022	3/5/2022	Week 6	
Baseline 2 evaluation	3/6/2022	3/12/2022	Week 7	
Checkpoint presentation + minimum deliverables	3/13/2022	3/19/2022	Week 8	Minimum deliverables by 3/19
Designs for improved method	3/13/2022	3/26/2022	Week 8-9	Initial designs approved by 3/26
Implementation of improved method	3/26/2022	4/2/2022	Week 10	
Evaluation of improved method + expected deliverables	4/3/2022	4/16/2022	Week 11-12	Expected deliverables by 4/16
Additional designs and experiments	4/10/2022	4/23/2022	Week 12-13	Maximum deliverables (if possible) by 4/23
Final report + poster	4/24/2022	4/30/2022	Week 14	
Final presentation	5/1/2022	5/7/2022	Week 15	

The table shows the project timeline with intermediate deadlines and milestones. Further adjustment might be made based on project progress and discussion with mentors.

This is currently a one-person project with all responsibilities assigned to Nhat Le.

## 7. Dependencies

Dependency	Type	Status	Needed by	Alternative	Notes
GPU Server (Dr. Jiang)	Hardware	Available from Dr. Jiang's lab (access acquired)	2/27	Backup GPUs available (access acquired)	Required for deep learning
Data annotation (Dr. Jiang)	Annotation	In progress (Expected 2/26 for scan-level)	2/27	Training on subset of data	Limitations on pretraining and design choices Required scan-level labels as minimum
IRB & HIPAA (Dr. Zhou)	Training	Completed	2/12	N/A	Training and study approval required for new members
DL tools (CUDA, PyTorch, etc.)	Software	Acquired Open-source	2/27	Other DL frameworks	

a. GPU server:

Access to GPU server is needed to train deep learning models. Dr. Jiang's lab provides a main GPU server with 2xNvidia RTX 3090 and a backup GPU server with 2xNvidia RTX 2080 Ti. These GPU servers support remote access through JHU VPN. Training on publicly available resources such as Google Colab is not allowed as the dataset for this project is currently private.

b. Data annotations:

There are 3 levels of annotation for the dataset including scan-level, slice-level, and pixel-level annotations. All level of annotations is in progress and to be provided by medical imaging experts in Dr. Jiang's lab. Complete scan-level annotation is required for this project, while the availability of slice-level and pixel-level ones will affect the choices on pretraining and network design. The slice-level and pixel-level annotations are time-consuming and not expected to be complete by the end of this project.

c. IRB & HIPAA

Team members need to complete the required training per School of Medicine policy. They also need to be added to relevant studies by Dr. Zhou before getting access to the data in case they work on image preprocessing tasks involving personally identifiable information.

d. Deep learning tools

The deep learning pipeline will be built based on open-source or publicly available tools and framework to ensure reproducibility of results.

## 8. Management Plan

Team members are expected to report progress to mentors in weekly lab meetings. One-on-one meetings with mentors are also available upon request for further discussion on the the project. All related code should be organized in a GitHub repository for documentation and

future usage. All results and notes should be kept in a OneDrive folder and updated on the project wiki page.

## 9. Reading List

- Background on APT imaging and relevant works on glioma classification with deep learning

[2] Zhou, J., Heo, H. Y., Knutsson, L., van Zijl, P. C., & Jiang, S. (2019). APT-weighted MRI: Techniques, current neuro applications, and challenging issues. *Journal of Magnetic Resonance Imaging*, 50(2), 347-364.

[4] Lee, J., Wang, N., Turk, S., Mohammed, S., Lobo, R., Kim, J., ... & Rao, A. (2020). Discriminating pseudoprogression and true progression in diffuse infiltrating glioma using multi-parametric MRI data through deep learning. *Scientific reports*, 10(1), 1-10.

- Relevant deep learning papers for transformer architecture and few-shot learning

[1] Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2021). Transformers in vision: A survey. *ACM Computing Surveys (CSUR)*.

[3] Arnab, A., Dehghani, M., Heigold, G., Sun, C., Lučić, M., & Schmid, C. (2021). Vivit: A video vision transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 6836-6846).

[5] Liu, L., Hamilton, W., Long, G., Jiang, J., & Larochelle, H. (2020). A universal representation transformer layer for few-shot image classification. *arXiv preprint arXiv:2006.11702*.

[6] Hou, R., Chang, H., Ma, B., Shan, S., & Chen, X. (2019). Cross attention network for few-shot classification. *Advances in Neural Information Processing Systems*, 32.

[7] Chen, H., Li, H., Li, Y., & Chen, C. (2021). Sparse spatial transformers for few-shot learning. *arXiv preprint arXiv:2109.12932*.