Radiogenomic Classification of Brain Tumor Using MRI Sequences

CS5402 Spring 2024

Team members:

Ganesh, Hasibur, Mizanur, Sazedur *(order is insignificant)*

Team MediMiners

April. 30, 2024

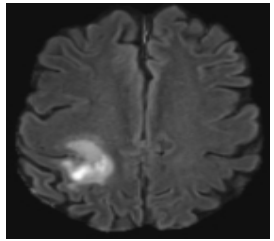
Contents

Motivation

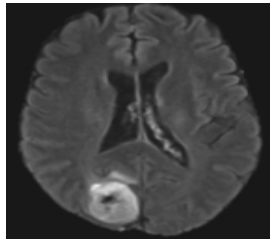
Genetic analysis of cancer typically requires invasive surgery and considerable time for results, leading to delays in treatment decisions and required additional surgeries.

Problem Statement

- **Motivation:** Currently, genetic analysis of cancer requires surgery to extract a tissue sample. Then it can take several weeks to determine the genetic characterization of the tumor. Depending upon the results and type of initial therapy chosen, a subsequent surgery may be necessary. If an accurate method to predict the genetics of the cancer through imaging (i.e., radiogenomics) alone could be developed, this would potentially minimize the number of surgeries and refine the type of therapy required.
- The presence of a specific genetic sequence in the tumor known as MGMT promoter methylation has been shown to be a favorable prognostic factor and a strong predictor of responsiveness to chemotherapy.
- **Design an efficient model to predict the presence of MGMT promoter methylateon in the glioblastoma tumor using MRI (magnetic resonance imaging) scans.**



Methylated



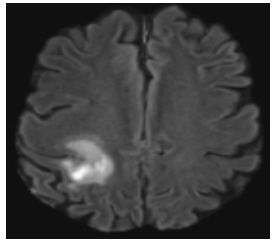
Unmethylated

Challenges

- Different patterns for the same class due to heterogeneous **sources** and data collection **techniques**.
- Model tasked with predicting MGMT value based on MRI sequences, whereas in reality, MGMT value is determined by molecular testing on biopsy.
- How the related papers address the challenges?

Problem Statement

- **Objective**: Utilize radiogenomics to analyze MRI scans for predictive markers of MGMT promoter methylation, enabling non-invasive and rapid assessment of tumor genetics.

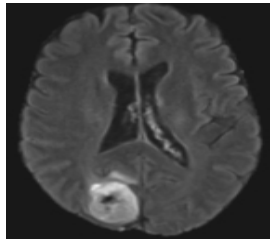


Methylated

- **Advantages**:

✓ Non-invasive: Eliminates the need for surgery to obtain tissue samples.

✓ Time-efficient: Provides rapid genetic characterization, expediting treatment planning.

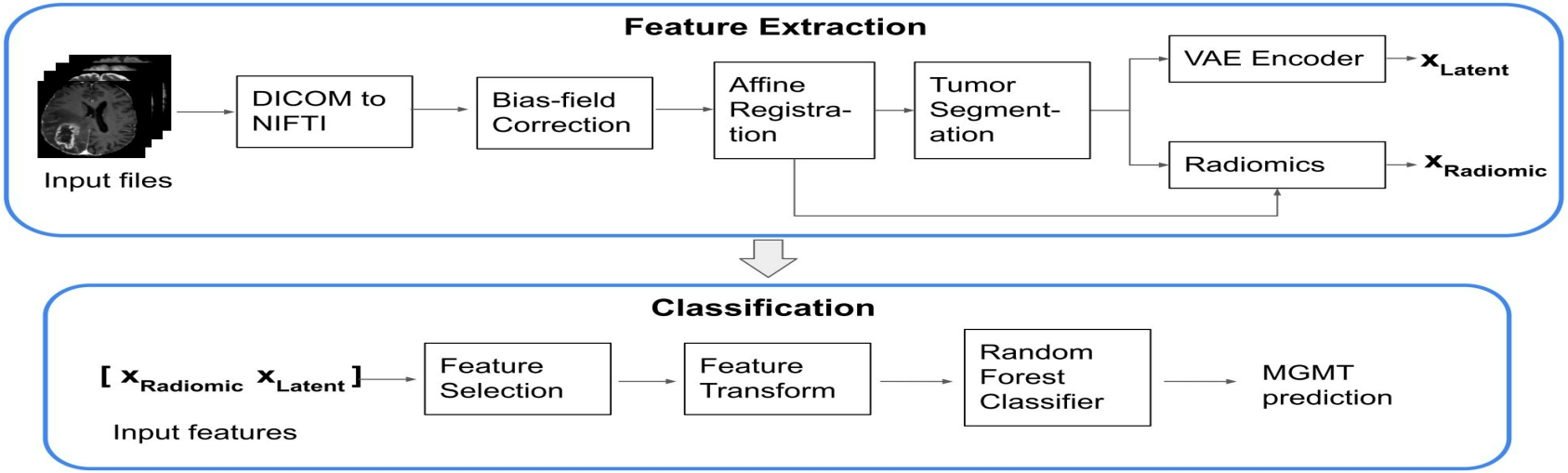


Unmethylated

Challenges

- Different patterns for the same class due to heterogeneous **sources** and data collection **techniques**.
- Data Processing
- Model tasked with predicting MGMT value based on MRI sequences, whereas in reality, MGMT value is determined by molecular testing on biopsy.
- Resource constraint: Need high computational power

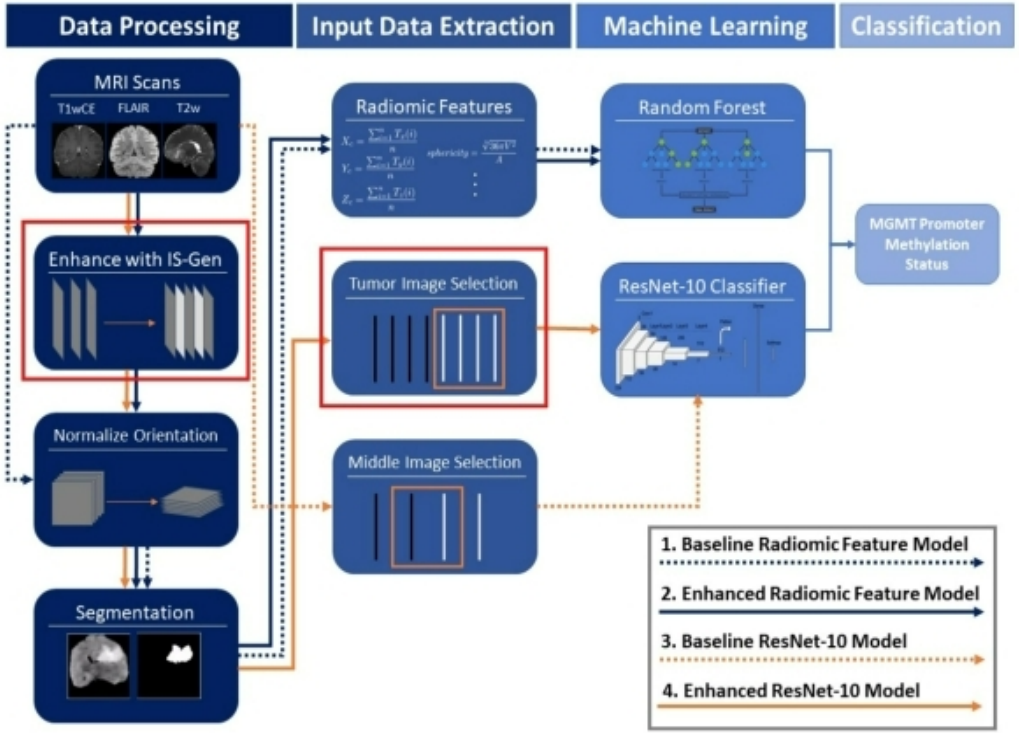
Related Work



AUC Score : 0.598

MICCAI Brain lesion 2021 Workshop[3]

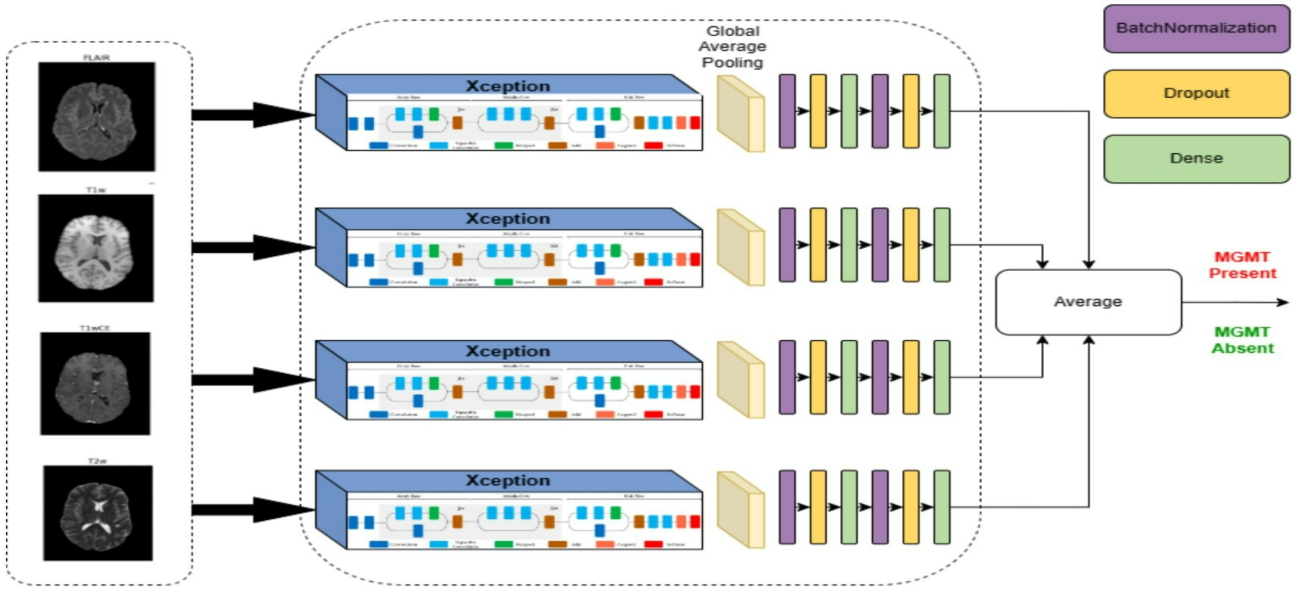
Related Work



AUC Score : 0.5458

(ICTAI). IEEE, 2022 [1]

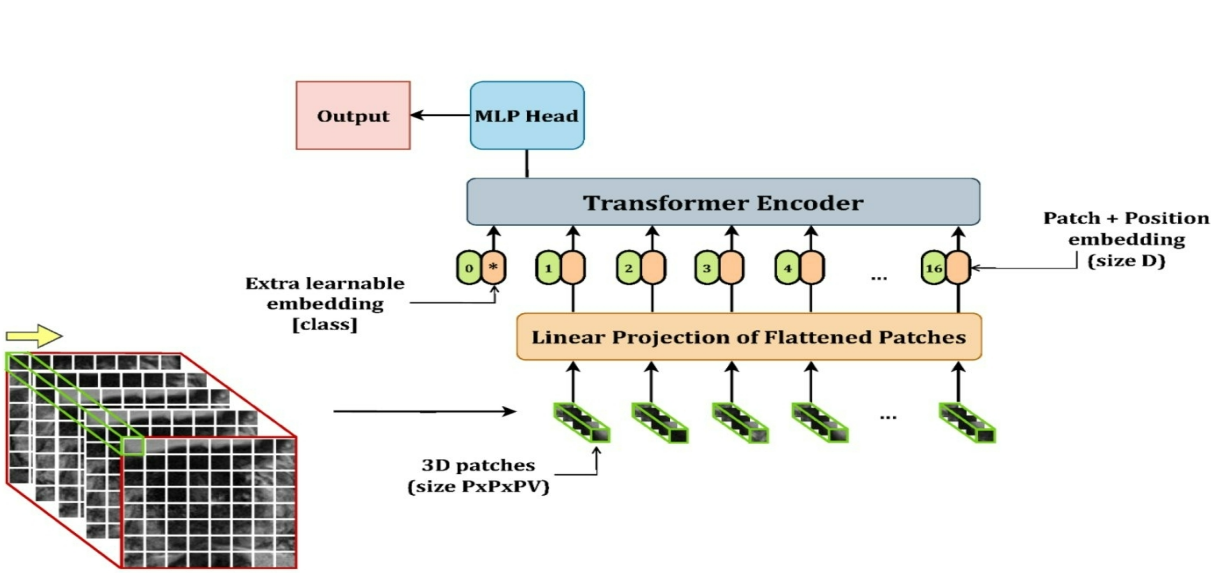
Related Work

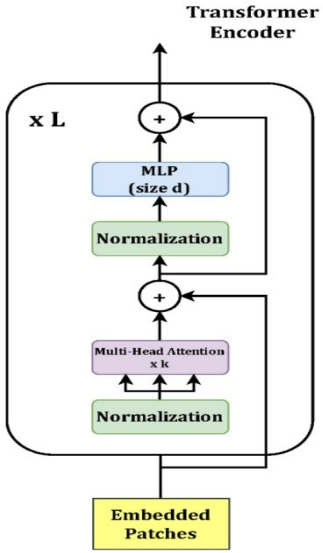


AUC Score : 0.61745

Mohamed et. Al. [4]

Related Work – ViT 3D





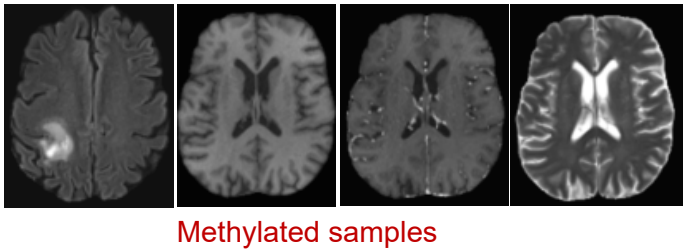
AUC Score : 0.6015

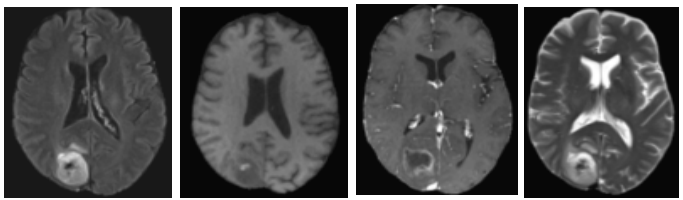
Mohamed et. Al. [4]

Task and Dataset [2]

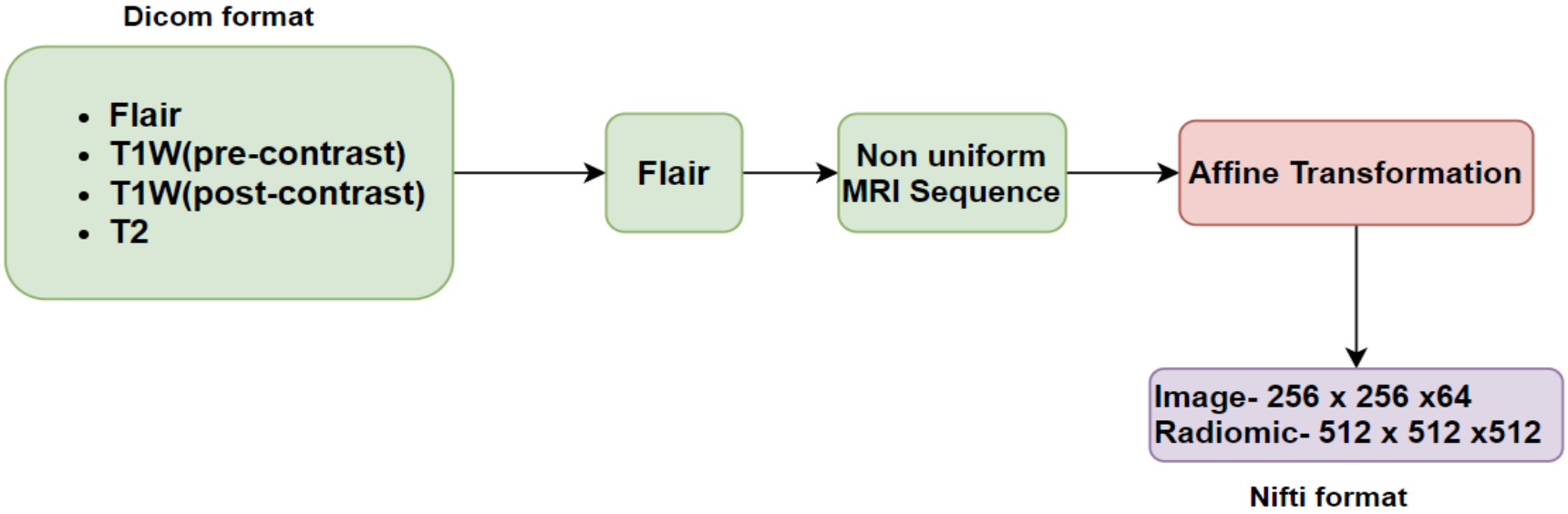
- **Predict the genetic subtype** of glioblastoma using MRI (magnetic resonance imaging) scans to train and test our model
  to **detect** for the presence of **MGMT promoter methylation**.
- Total **672 Patients** Sample(**train-585, test-87***(no label for them yet)*)

- **X**
  - *Fluid Attenuated Inversion Recovery (FLAIR)*
  - *T1-weighted pre-contrast (T1w)*
  - *T1-weighted post-contrast (T1Gd/T1wCE)*
  - *T2-weighted (T2)*
- **Y**
  - Methylated (1): 306
  - Unmethylated (0): 277


Methylated samples


Unmethylated samples

Data Processing - DICOM to NIFTI



Data Processing (Contd..)

Challenge:

- Each MRI scan case contains a varying number of image sequences.
- Utilizing all scans may confuse the model to learn spatial dependence of brain pixels

Solution 1(Regular Interval):

- Calculate sequence indices based on the number of images per scan
- Stack corresponding sequences to construct a 3D image (**256 x 256 x 64**)

Issue:

- Random stacking doesn't ensure consistent brain portion representation.
- Model struggles to learn crucial tumor patterns and learns irrelevant spatial patterns.

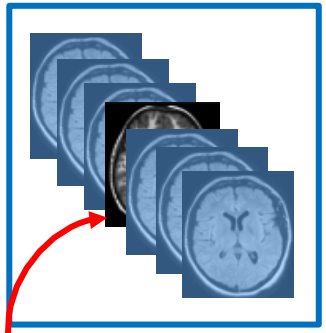


Fig: 3D Stack

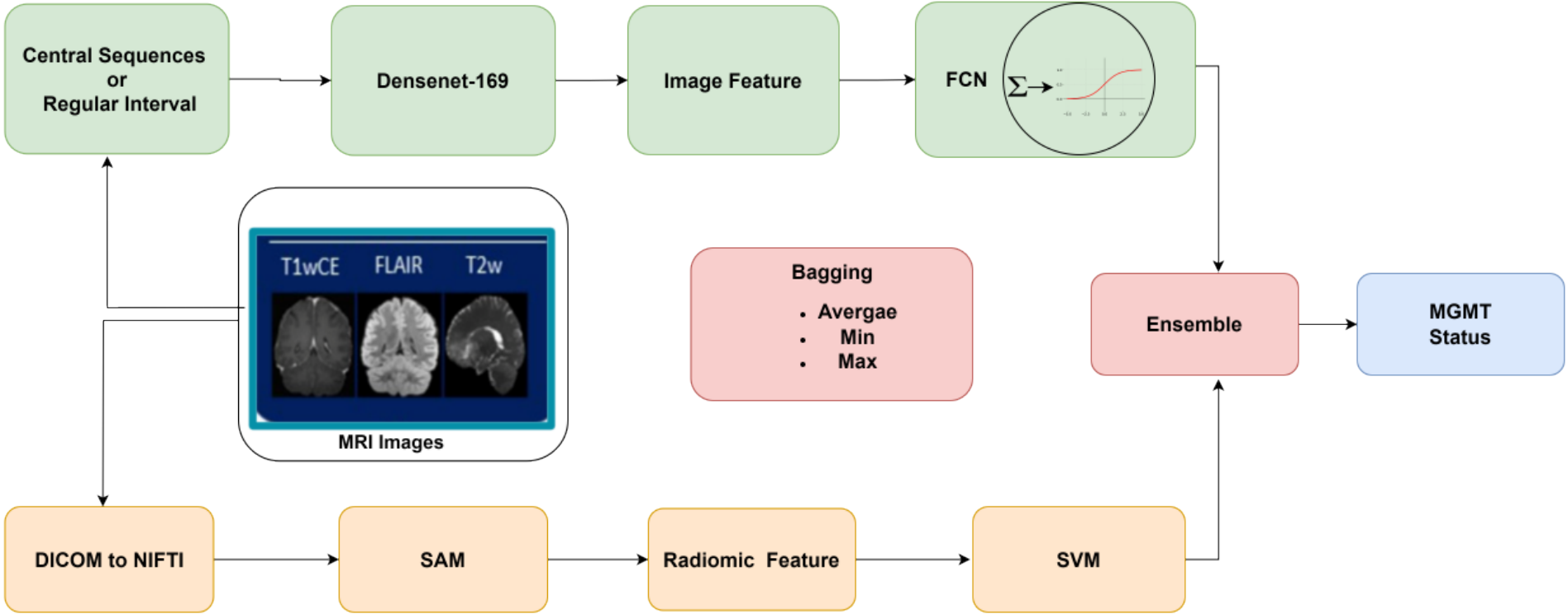
Solution 2(Central Sequences):

- Utilize the image with the largest brain cutaway view as the **central image** for constructing the 3D stack.
- Consistent brain portion representation.

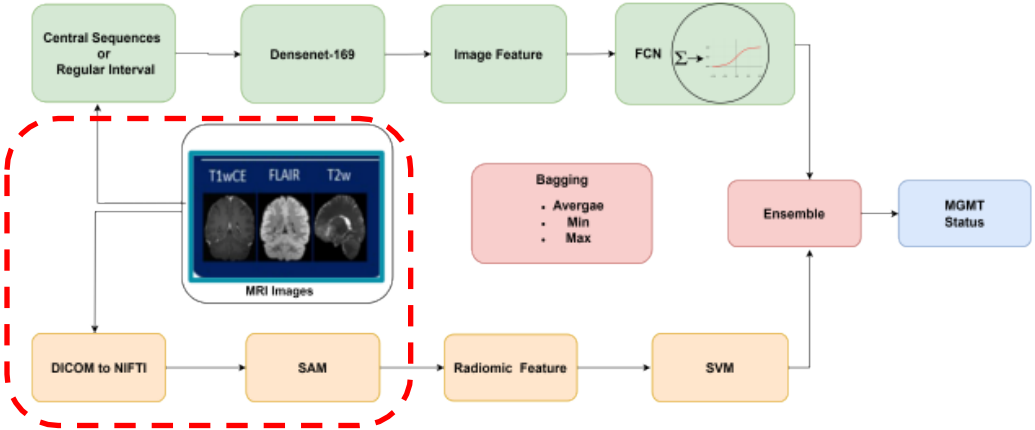
Data Processing (Contd..)

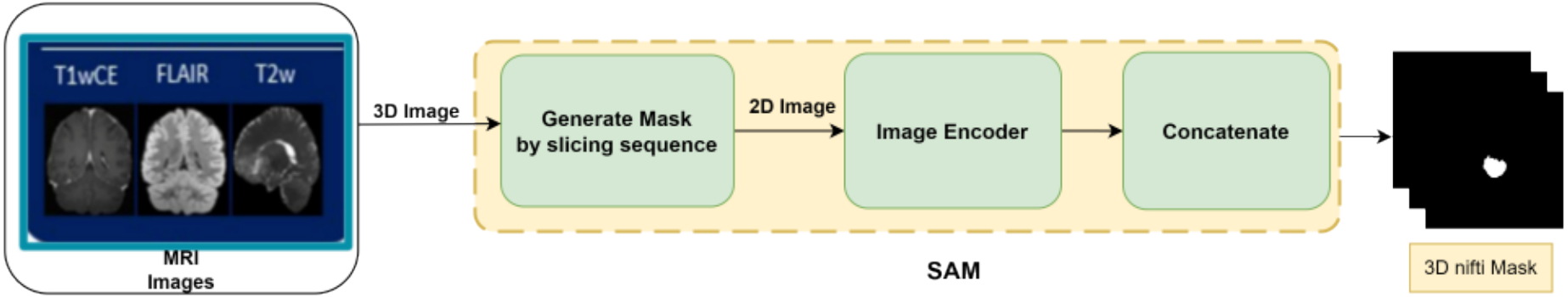
- 585 train samples: 3 of them are corrupted and discarded
- Total dataset split into 80% train and 20% test
- Stratification is used to handle class imbalance problem

Methodology

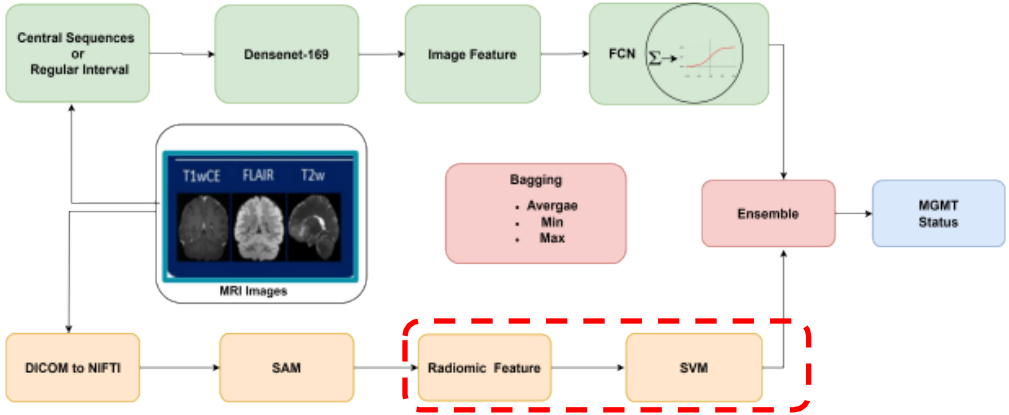


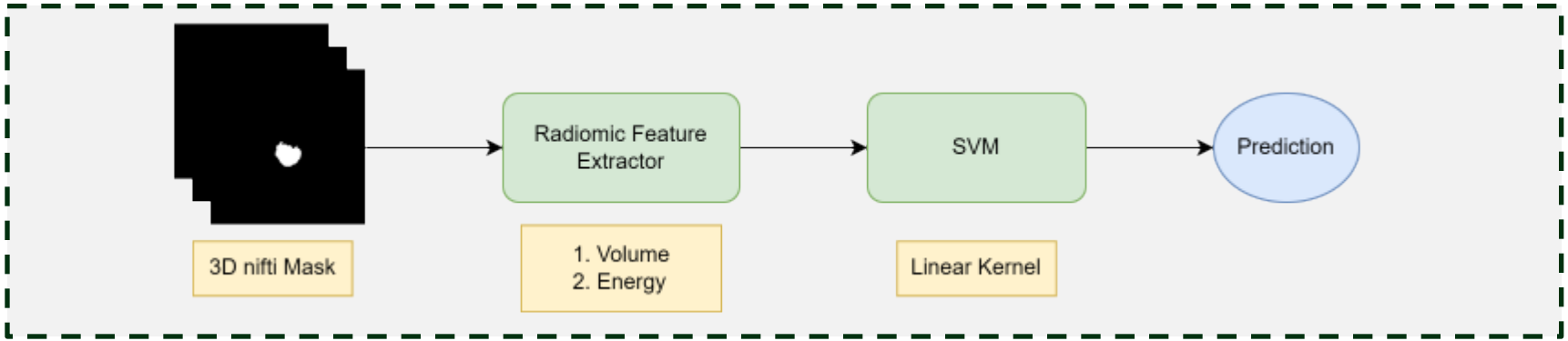
Methodology (Cont.) - Mask Generation



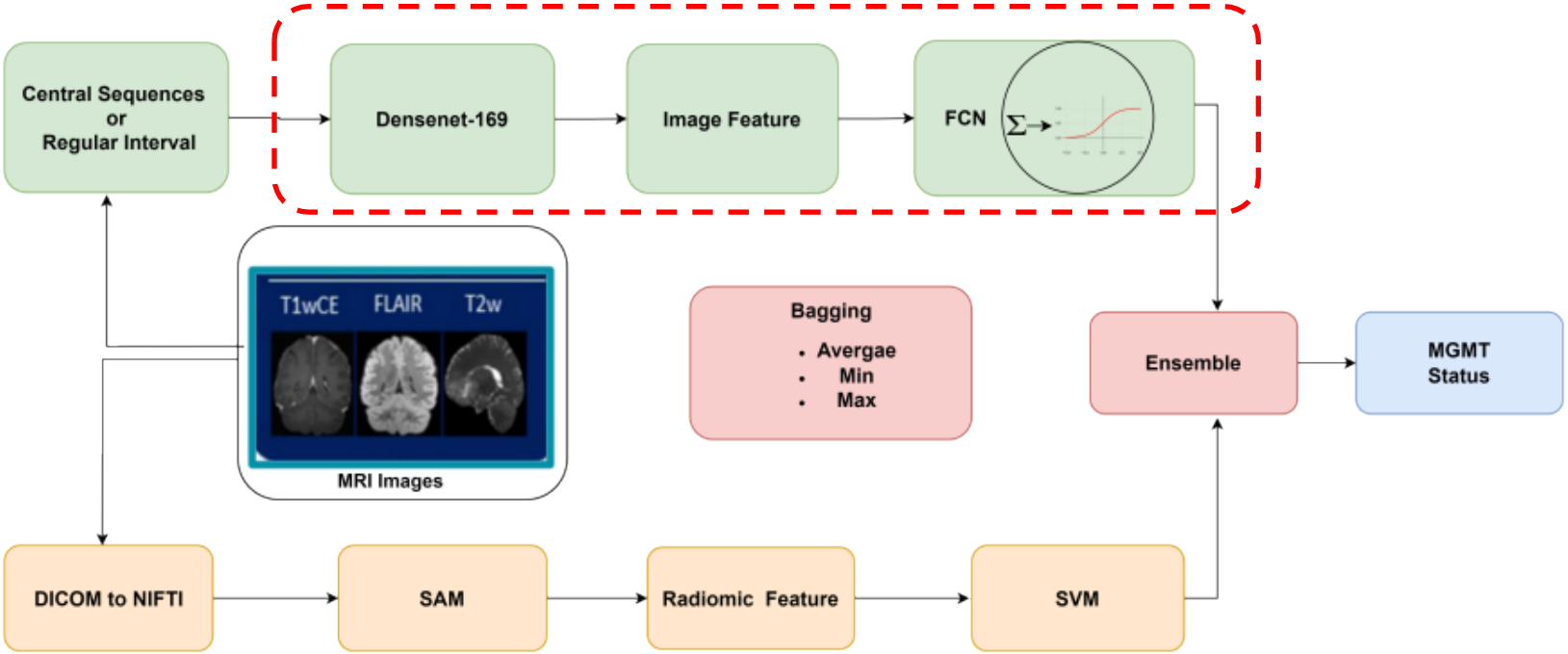


Methodology (Cont.) - Radiomic Feature





Methodology (Cont.) - DenseNet

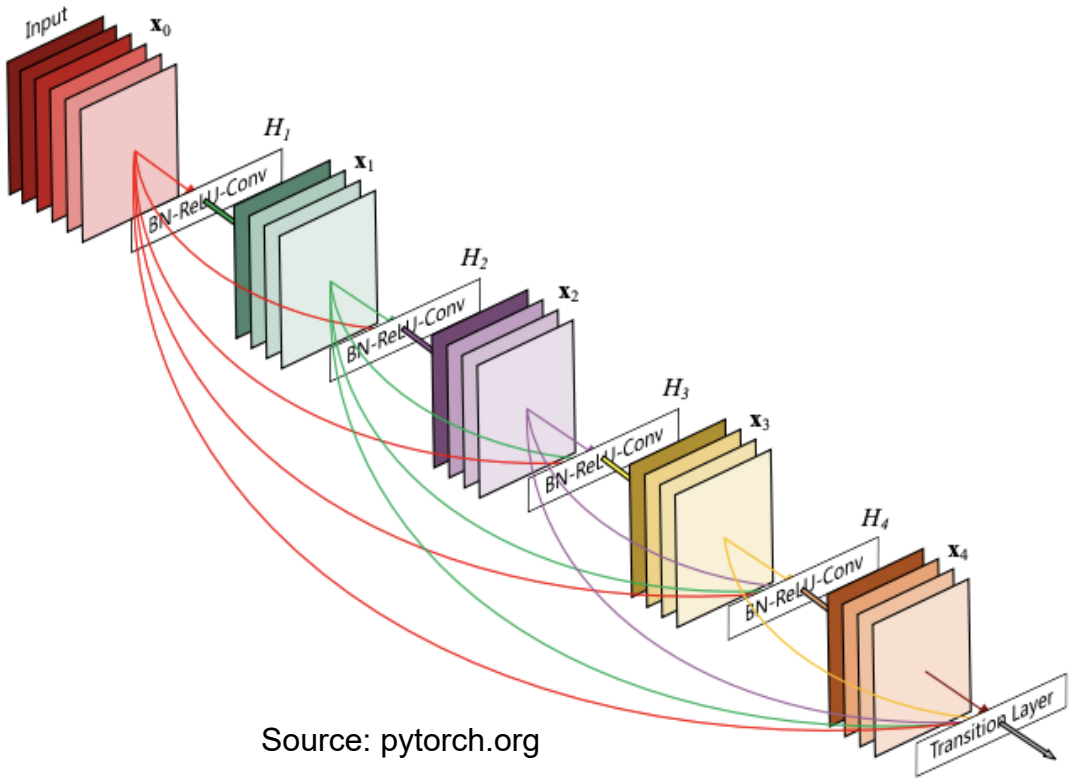


Densenet - Dense Convolutional Network

Huang. et. al. ICCV 2017 [5]

- Feature Reuse
- No Vanishing Gradient Issue

- Densenet169
- 3D Densenet



Source: pytorch.org

Densenet Implementation – Loss Function and Hyperparameters

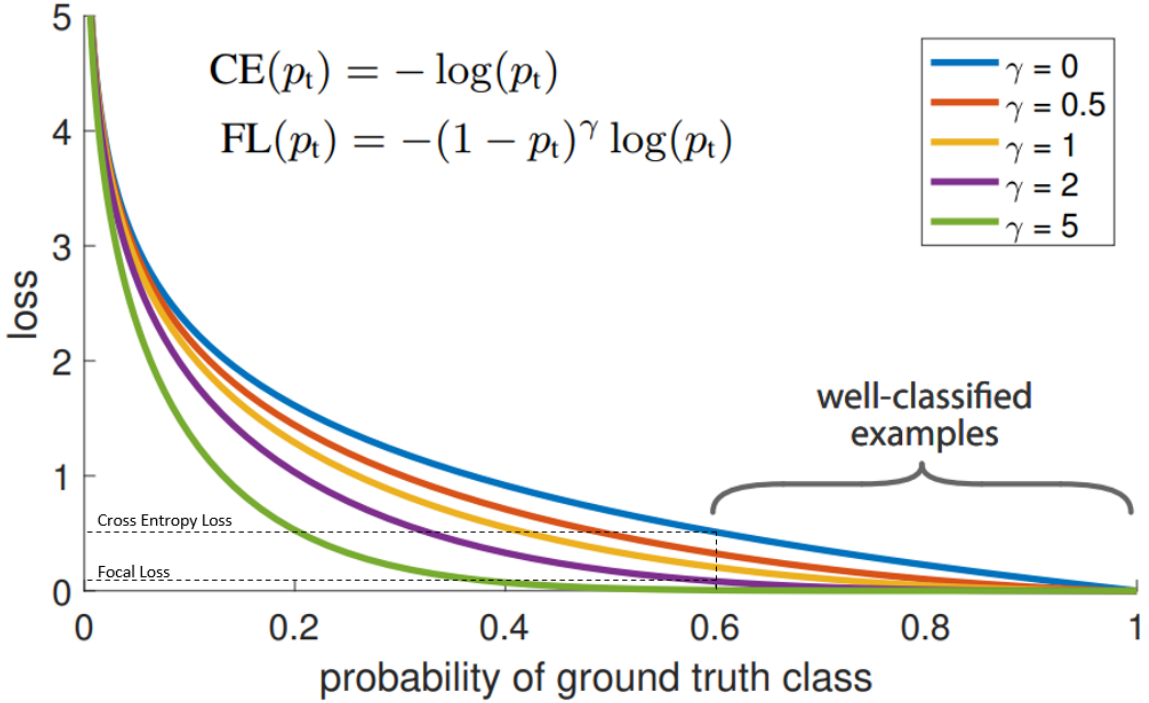


Image Size = 256
Num_Img= 64
Bathch_Size = 1
Learning Rate = 1e-6
Lr_Decay = 0.9
Num_Epochs = 150

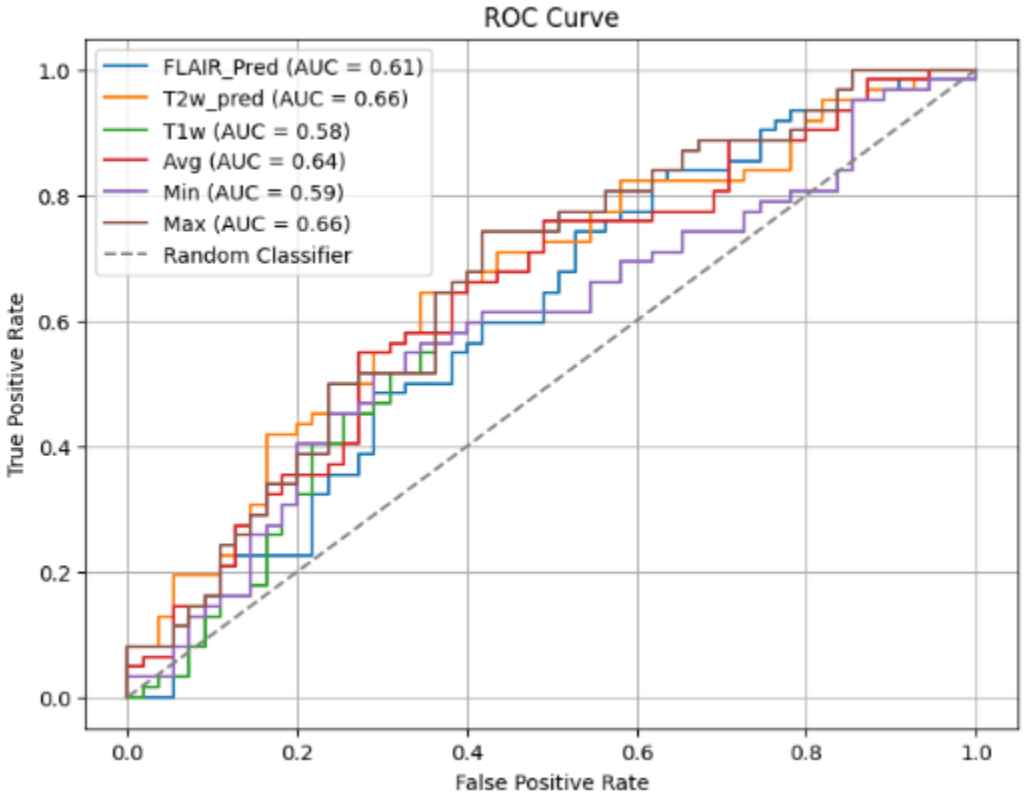
Focal Loss:
Gamma = 2

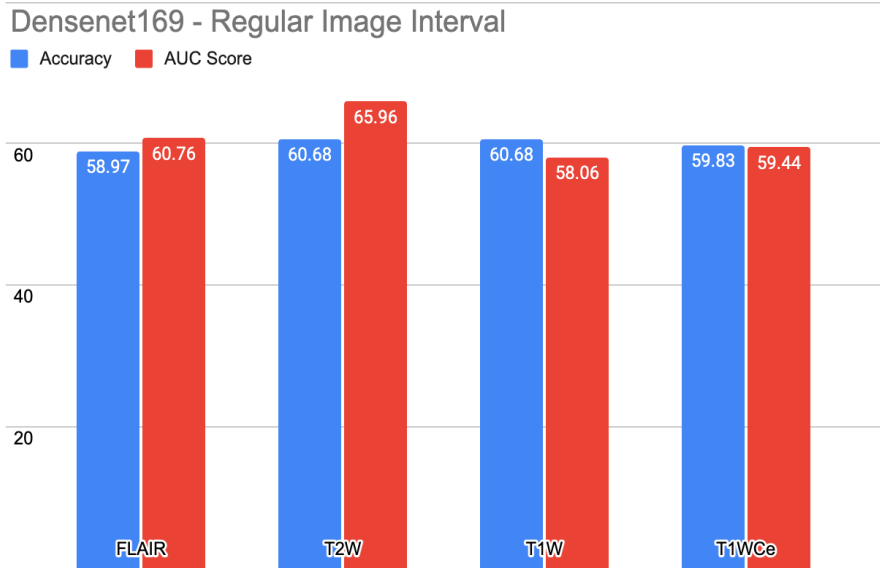
Lin. et. al. ICCV 2017 [7]

Implementation Details

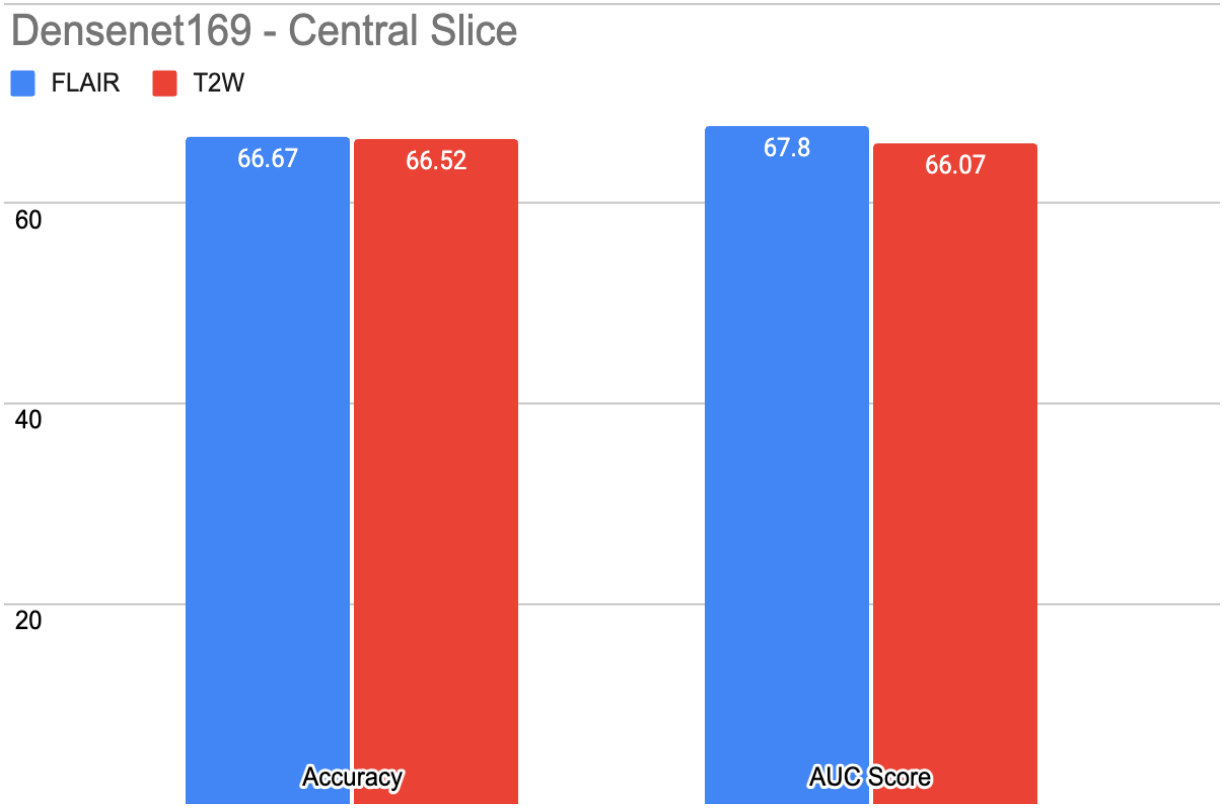
- Total dataset split into 80% train and 20% test
- 585 train images: 3 of them are corrupted and discarded

Experimental Results – Densenet169 (Regular Interval)

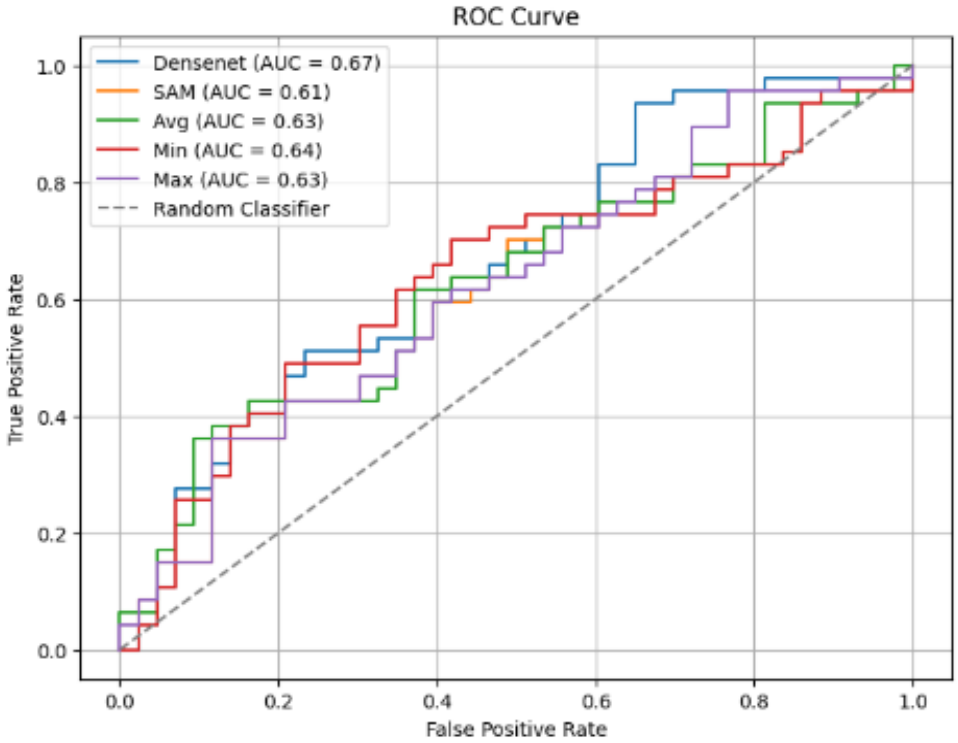


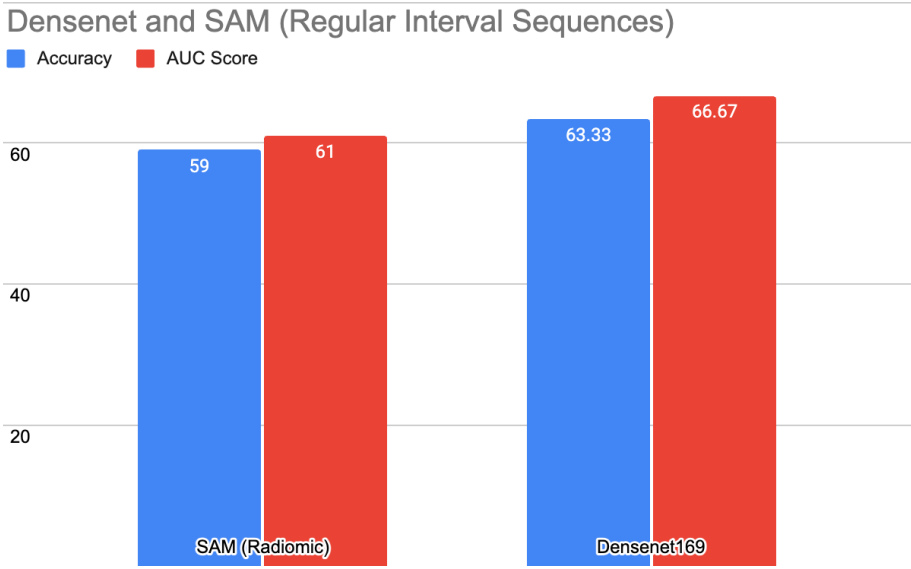


Experimental Results – Densenet169 (Central Sequences)



Experimental Results – Densenet169 and SAM – Regular Interval





Conclusion

- Applied data mining techniques to process medical data.
- Time efficient , non-invasive solution of getting MGMT status.
- Successful application of Densenet169 and SAM model.
- Ensemble method to get better results.
- Intensive experimental analysis.

Future works

- As SAM is not trained on the tumor data, we plan to train the SAM model on the Brain Image and segment mask pair. We have also planned to use the MedSAM[8]. The MedSAM model is developed on a large-scale medical image dataset containing 1,570,263 image-mask pairs
- Use all MRI sequence(T1w, T1wcE, T2w) other than the FLAIR in the current pipeline.
- Comparing with SOTA by collecting test labels.
- Apply other ensemble algorithms (such as AdaBoost) rather than Bagging ensemble only.

References

[1] S. Das, "Optimizing Prediction of MGMT Promoter Methylation from MRI Scans using Adversarial Learning," in 2022 IEEE 34th*International Conference on Tools with Artificial Intelligence (ICTAI), Macao, China*, 2022, pp. 1047-1054.

[2] https://www.kaggle.com/c/rsna-miccai-brain-tumor-radiogenomic-classification

[3] Pálsson, Sveinn, Stefano Cerri, and Koen Van Leemput. "Prediction of MGMT methylation status of glioblastoma using radiomics and latent space shape features." International MICCAI Brain lesion Workshop. Cham: Springer International Publishing, 2021.

[4] Mohamed, Amr, et al. "Brain Tumor Radiogenomic Classification." *arXiv preprint arXiv:2401.09471* (2024).

[5] Huang, Gao, et al. "Densely connected convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.*

[6] Kirillov, Alexander, et al. "Segment anything." *Proceedings of the IEEE/CVF International Conference on Computer Vision.* 2023.

[7] Lin, Tsung-Yi, et al. "Focal loss for dense object detection." *Proceedings of the IEEE international conference on computer vision.* 2017.

[8] Ma, Jun, et al. "Segment anything in medical images." Nature Communications 15.1 (2024): 654.

Thank You

For Your Attention!

Any Questions