Brain Tumor Grade Classification in MR images using Deep Learning

Eleftheria Chatzitheodoridou

Supervisor: Anders Eklund

External Supervisors: Iulian Emil Tampu

Neda Haj-Hosseini

LINKÖPING UNIVERSITY

Overview

- Motivation
- Background
- Aim
- Data Overview
- Research Questions
- Pre-processing pipeline
- Results
- Conclusion

Motivation

Brain Tumor

- High morbidity
- Difficult to diagnose

Diagnostic tools

- Biopsy
- MR images (radiology)

Biopsy

- Invasive
- Riskful
- Expensive

Computer-Aided Diagnosis (CAD)

- Deep Learning implemented on MR imaging

Deep Learning

- Automated feature extraction
- Fast inference

LINKÖPING UNIVERSITY

Background

Gliomas: primary brain tumor, classified into 4 grades (WHO 2021)

- G1 → Benign tumor
- G2
- G3
- G4

Low-Grade Glioma

High-Grade Glioma

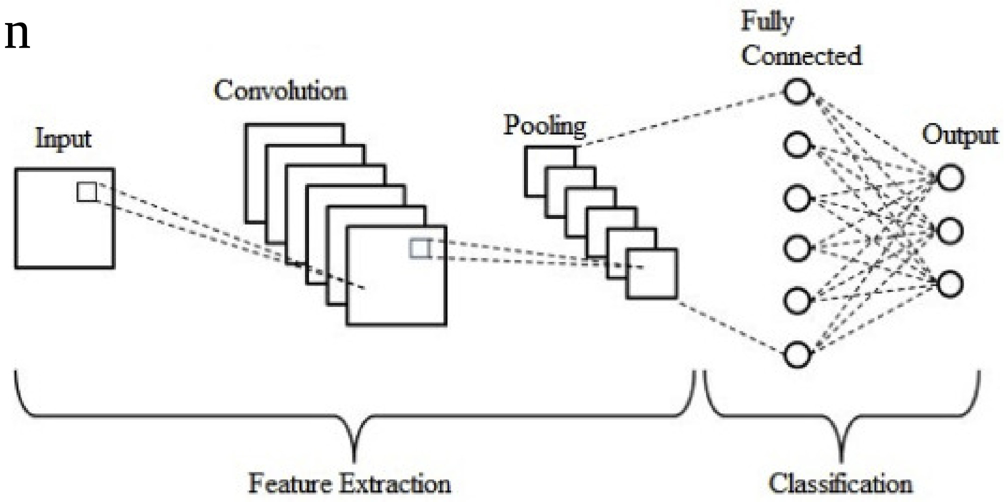
Malignancy

Magnetic Resonance Imaging: uses non-ionizing radiation, offers high-spatial resolution images of different tissue contrast with varying repetition times by means of powerful magnets

Background

DL algorithms:

- inspired by structure + function of human brain
- provide state-of-the-art results in image classification
- can perform automatic feature extraction



Convolutional Neural Networks (CNNs):

- subclass of DL, used with great success in analysis of images
- require minimal preprocessing + little prior knowledge
- can achieve great levels of abstraction by stacking many layers

Image source: https://www.sciencedirect.com/science/article/pii/S2352914820305621

Aim

To classify the grade of brain tumor in MR images of different modalities from adult patients using deep learning

Image source: https://www.philips.se/healthcare/solutions/magnetic-resonance

Data Overview

Source: The Cancer Genome Atlas (TCGA) ➡ G2, G3

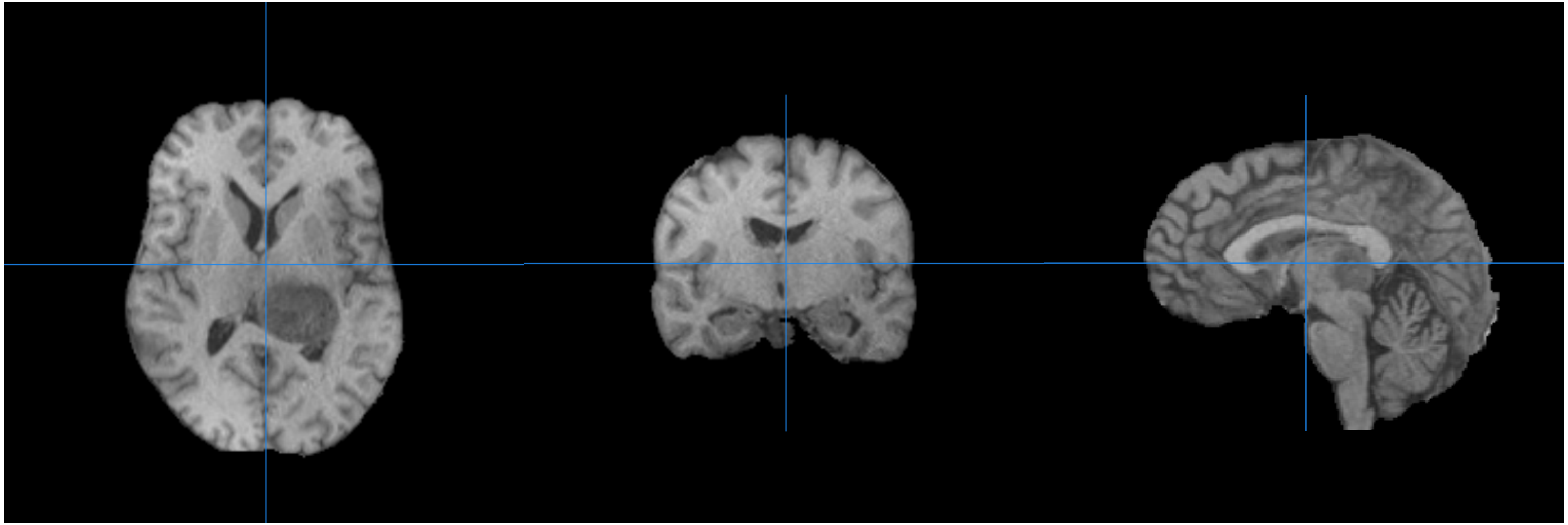
Brain Tumor Segmentation (BraTS 2020) ➡ G4

- 142 subjects (47 × G4, 45 × G2, 50 × G3)
- All images are: anonymized, registered to T1-weighted image, skull-stripped, bias-field corrected
- 3D volumes of 240 × 240 × 155 voxels with 1mm^3 isotropic resolution
- 155 slices for each subject
- Each subject has images of 4 MRI modalities (T1-weighted, T1-weighted post-contrast, T2-weighted, Fluid Attenuated Inversion Recovery (FLAIR))

TCGA dataset: https://portal.gdc.cancer.gov

BraTS 2020 dataset: https://arxiv.org/abs/1811.02629

Data Example – Grade 4 T1-weighted

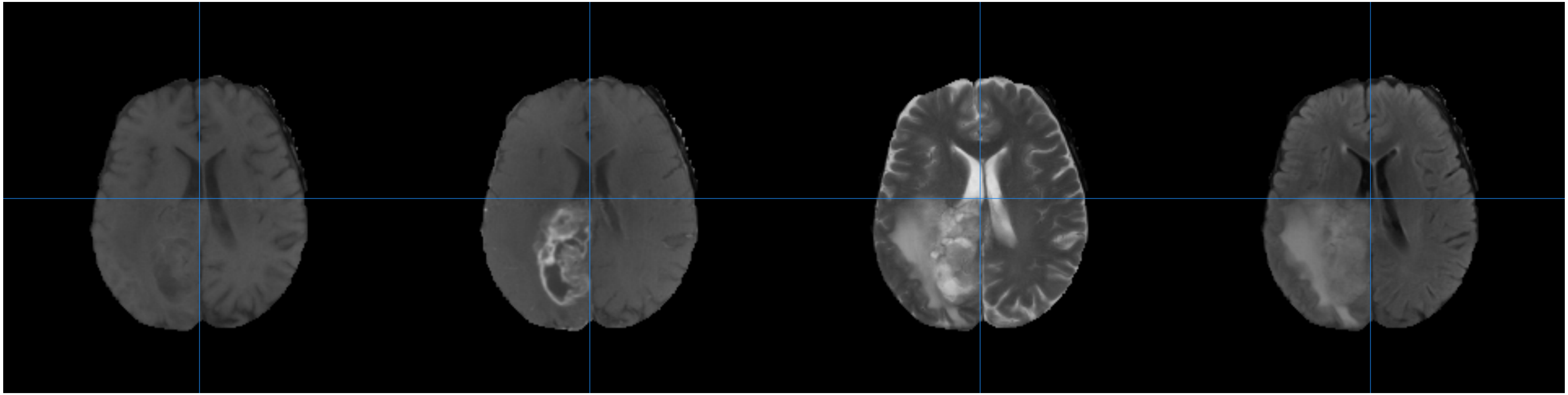


Axial view
(x-y plane)

Coronal view
(x-z plane)

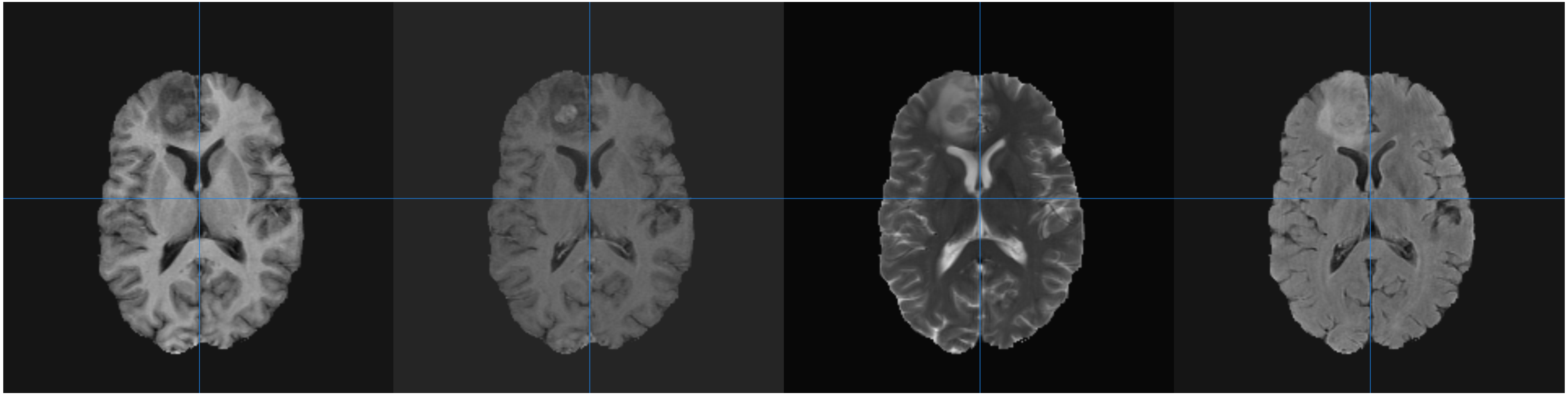
Sagittal view
(y-z plane)

Data Example – Grade 4 (axial view)



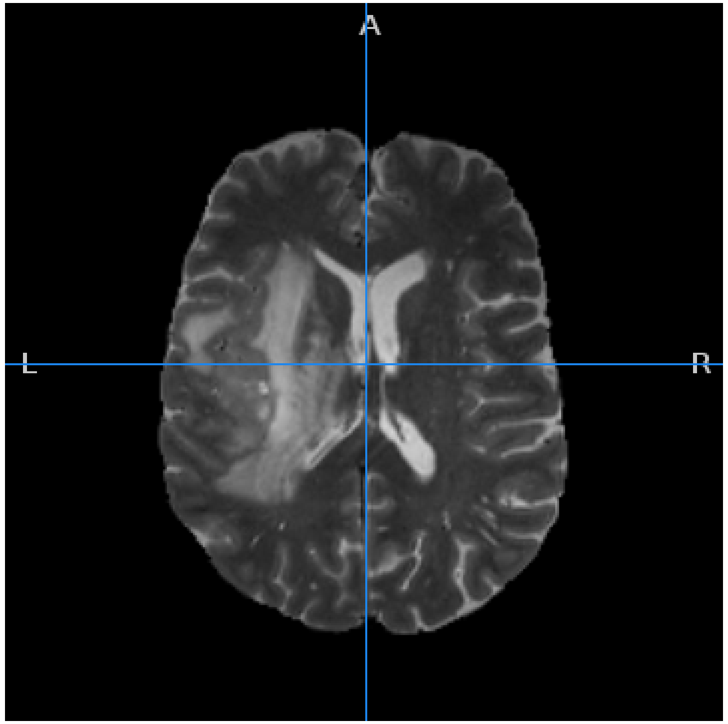
T1-weighted T1-post_contrast T2-weighted FLAIR

Data Example – Grade 2 (axial view)

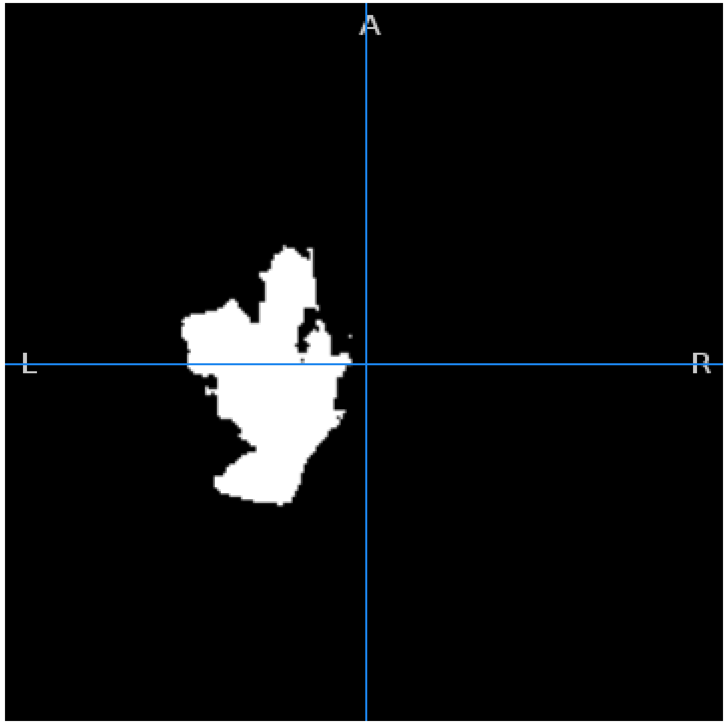


T1-weighted T1-post_contrast T2-weighted FLAIR

Data Example – Grade 4 (axial view)



T2-weighted



Annotated ROI

Research Questions

1. Which of the popular CNN models performs best for brain tumor grade classification on the available dataset?

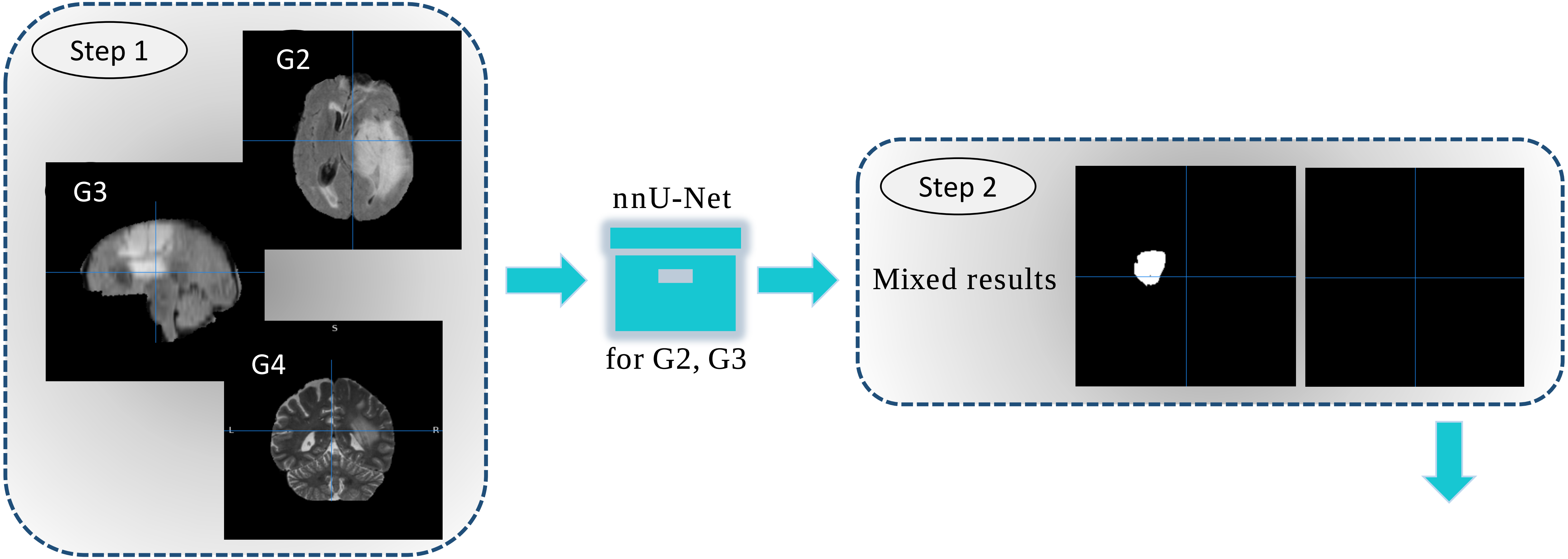
2. Which combination of MRI modalities yields the best results for classification?

3. Is it possible to construct representations of model explainability?

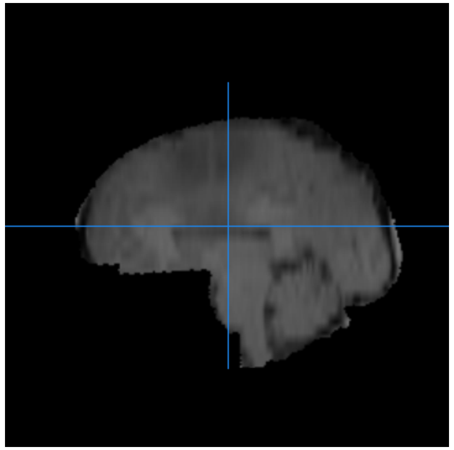
Challenges:

- Little data available ⇨ need for data augmentation
- Images come from 19 different institutions ⇨ need for intensity normalization
- Classification of grades in multiple classes as relatively new concept ⇨ not many experiments conducted, little supporting literature

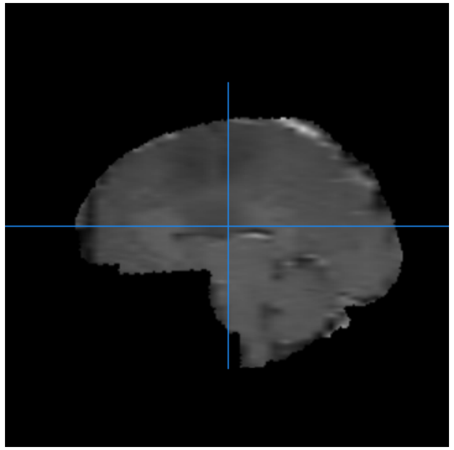
Pre-processing pipeline – Data collection



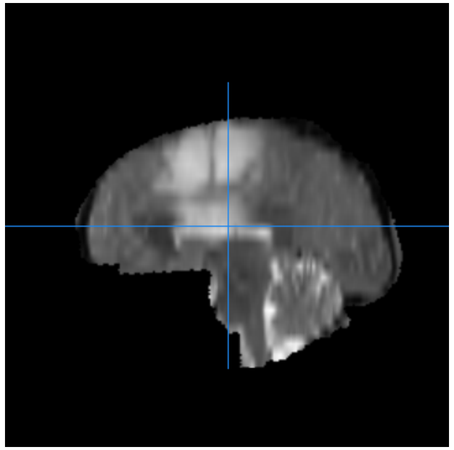
Tumor segmentation with nnU-Net on Grade 3 case - result (1)



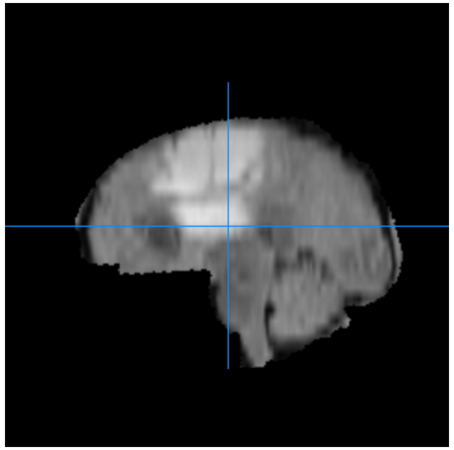
T1



T1Gd

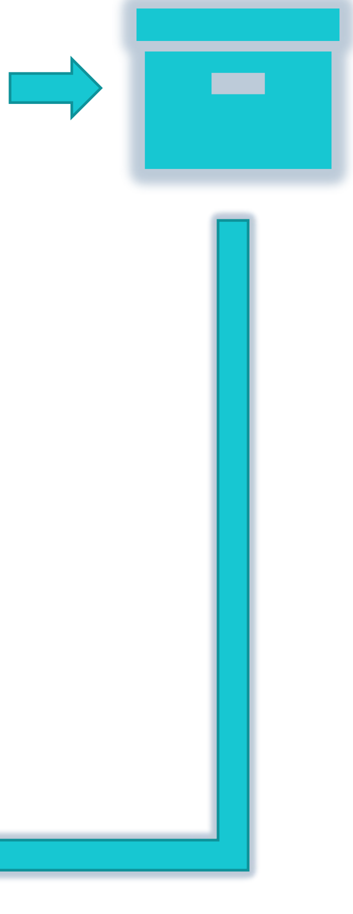


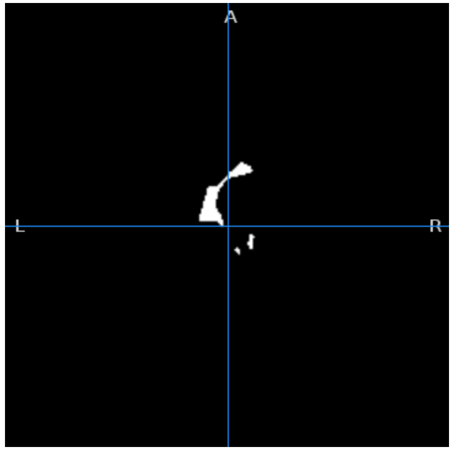
T2



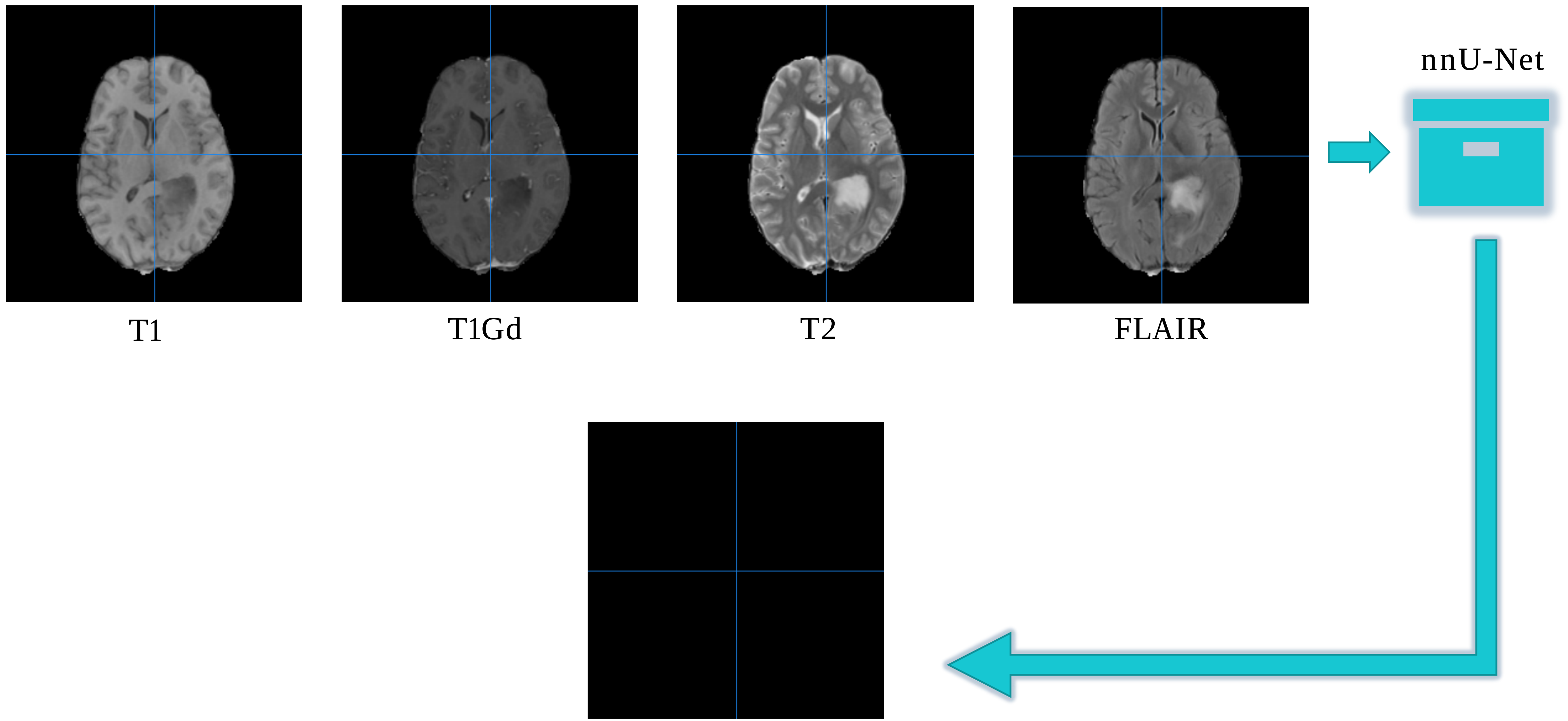
FLAIR

nnU-Net



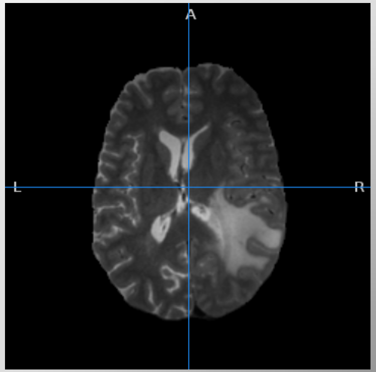


Tumor segmentation with nnU-Net on Grade 3 case – result (2)



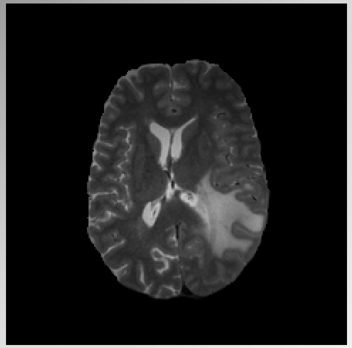
Pre-processing pipeline

Step 5



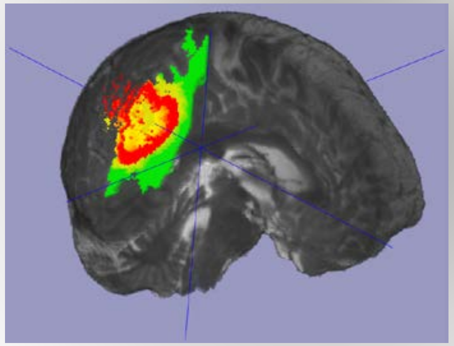
240 × 240

Intensity normalization & resizing



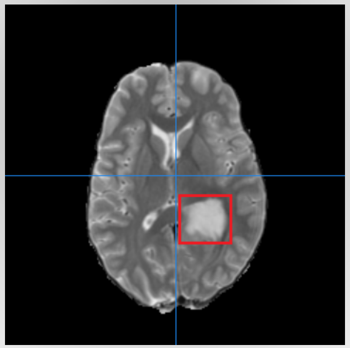
224 × 224

Step 4

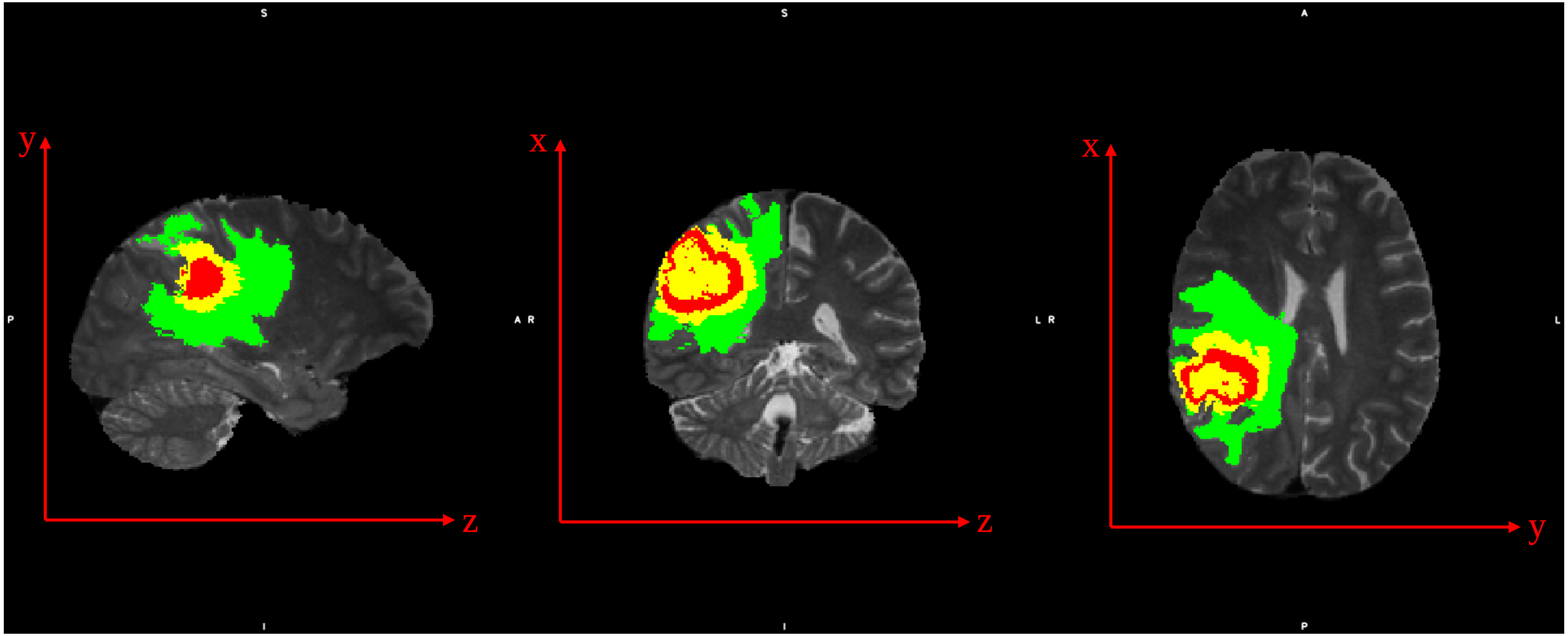


Tumor extraction

Step 3

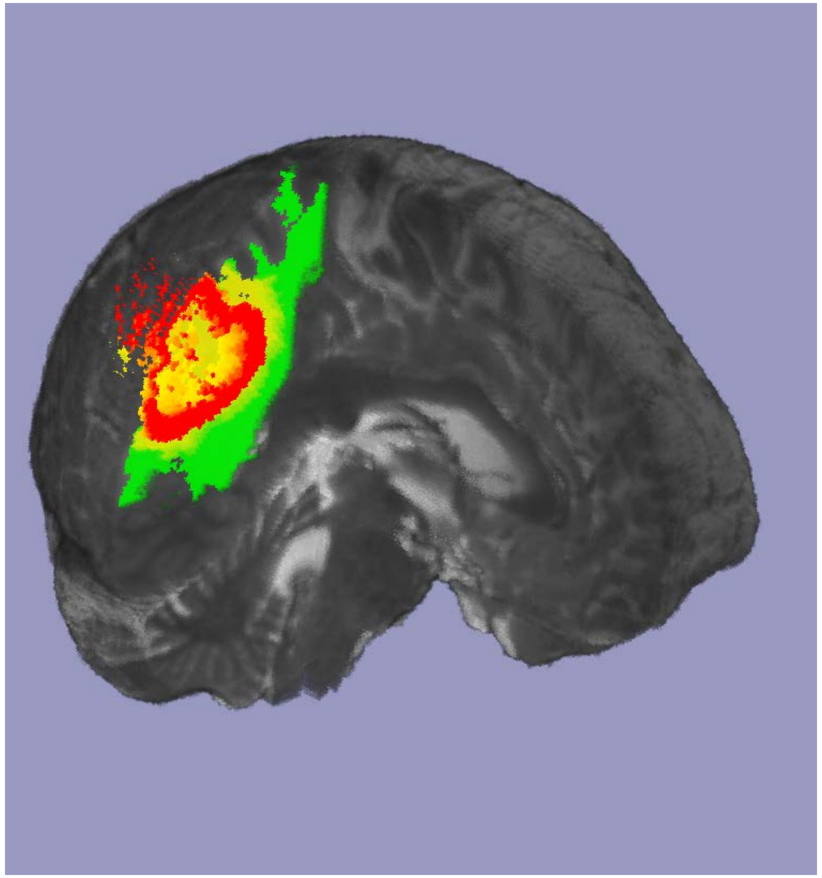


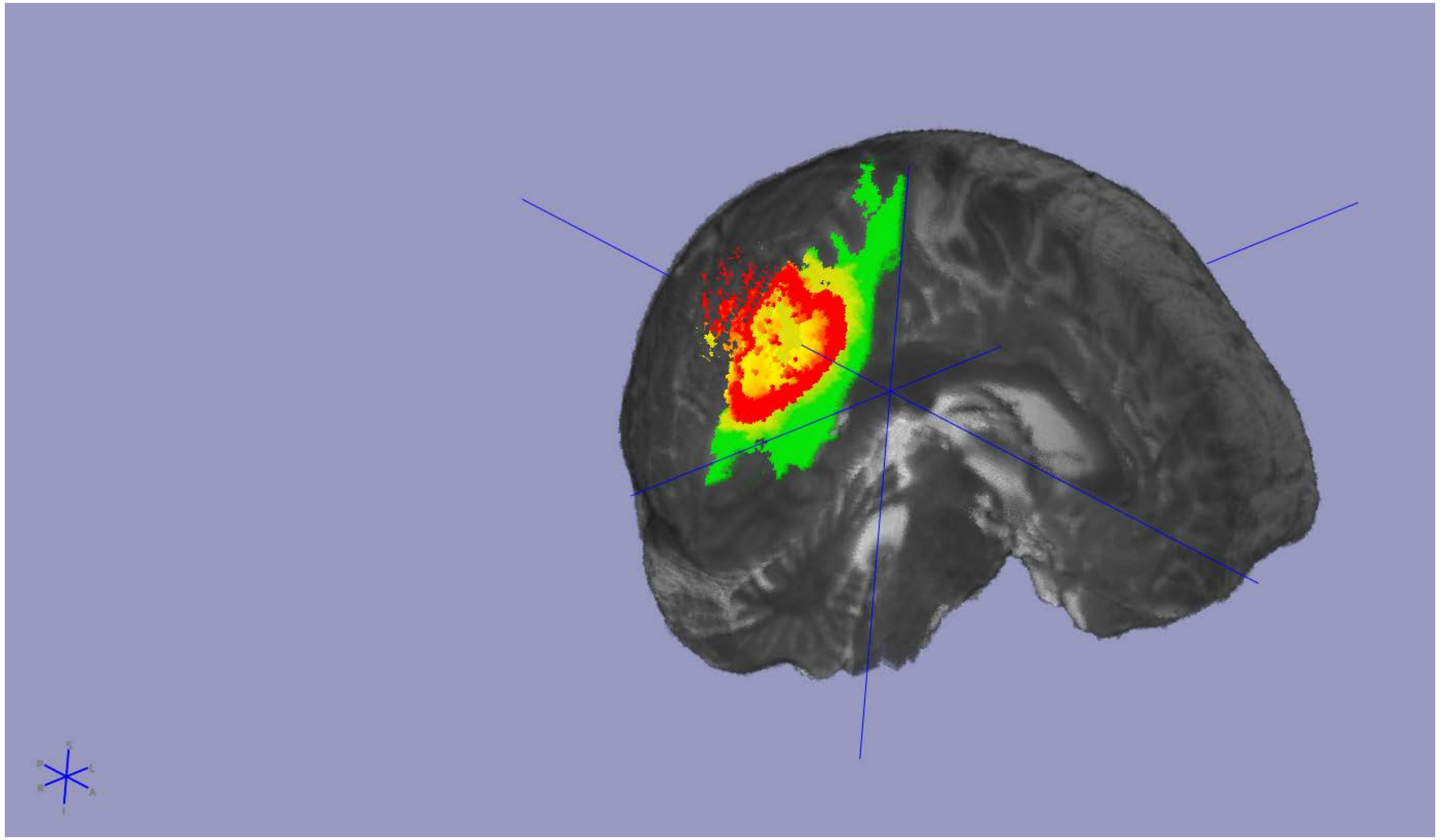
Extraction of tumor slices from 3D image for a Grade 4 case



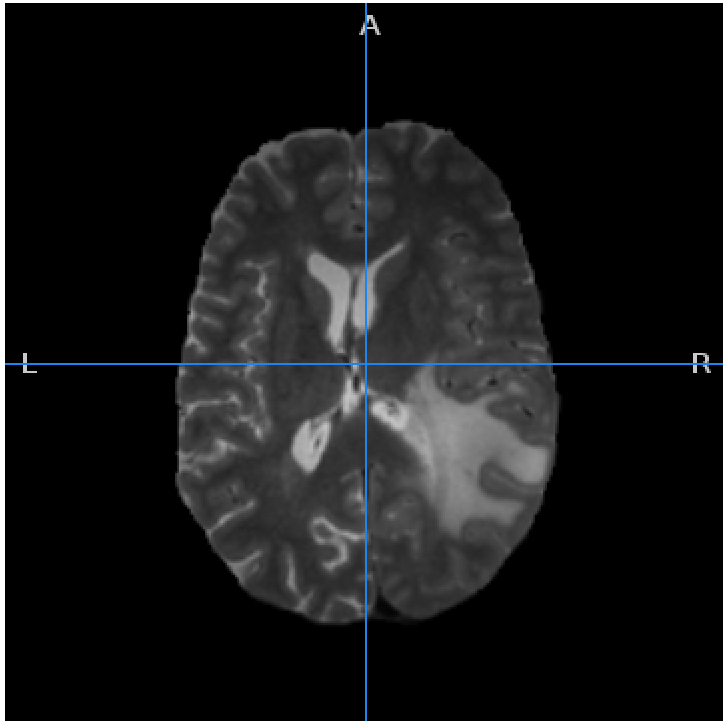
Sagittal view   Coronal view   Axial View

Extraction of tumor slices from 3D image for a Grade 4 case

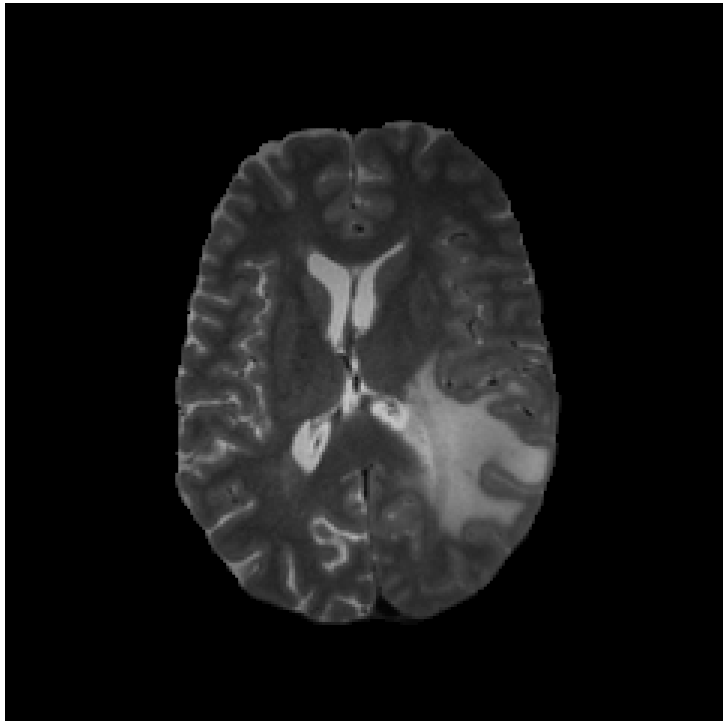




Intensity normalization on Grade 4 case, T2-w, axial view



Original image



Normalized 2D slice

Pre-processed 2D image dataset

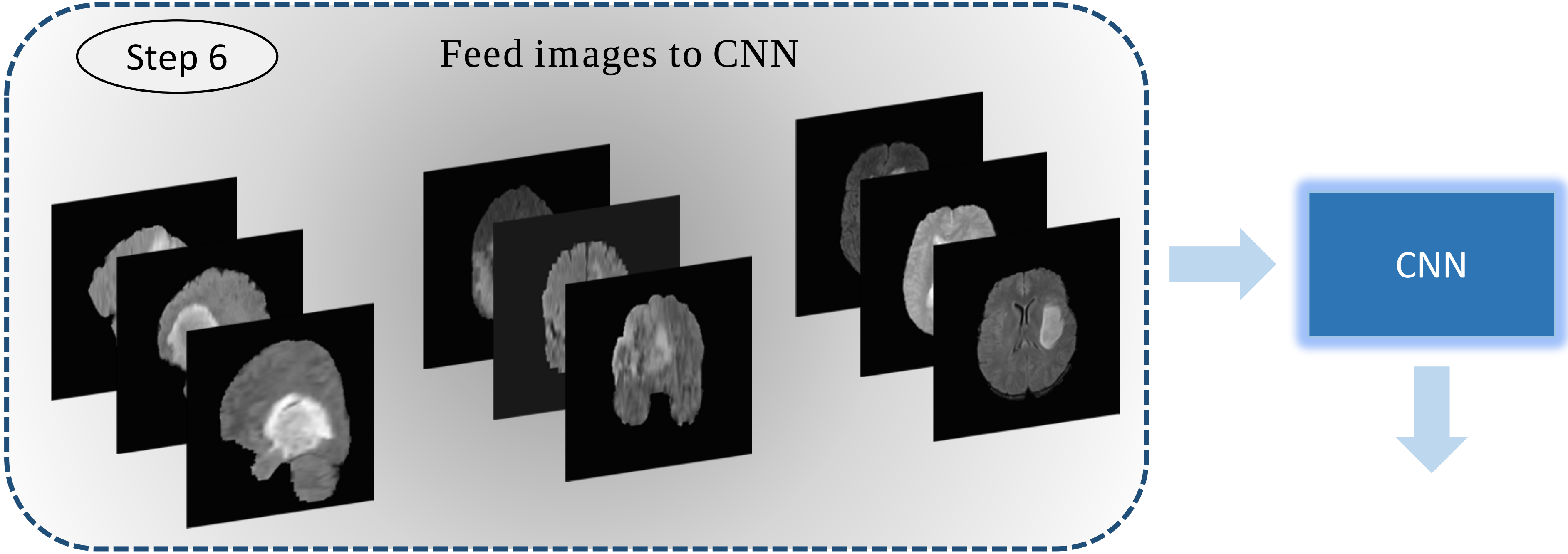
Grade	Frontal/coronal slices	Sagittal slices	Transversal/Axial slices
G2	8,131	5,415	6,591
G3	9,959	7,910	7,978
G4	15,613	11,470	12,485
Total	**33,703**	**24,795**	**27,054**

Contains tumor for all 4 modalities (before partition)

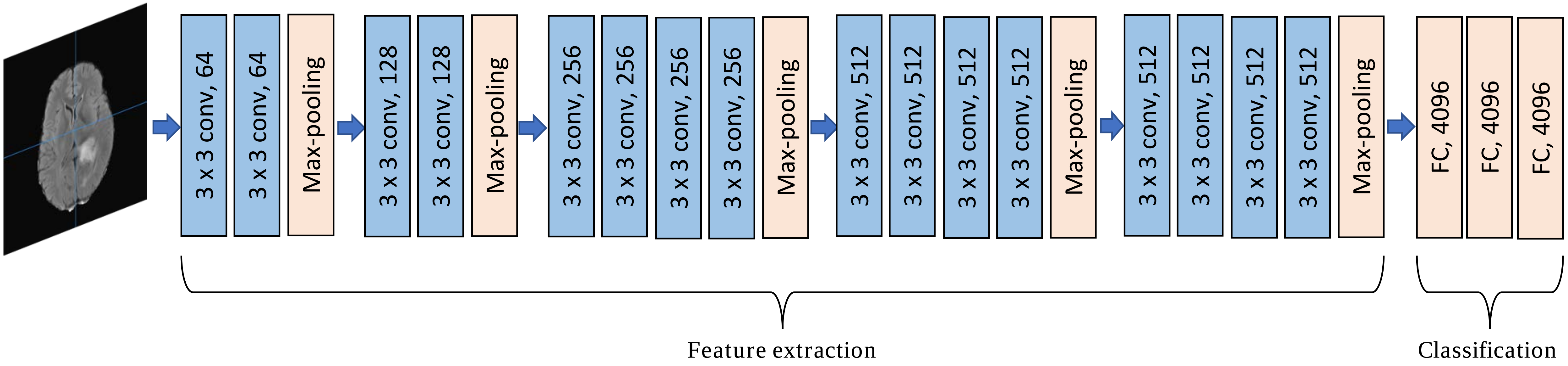
Grade	Train	Validation	Test
G2	1,364	393	103
G3	1,661	293	173
G4	2,034	609	166
Total	**5,059**	**1,295**	**442**

Contains T2, T1ce, FLAIR images on the sagittal plane

Pre-processing pipeline – CNN Implementation

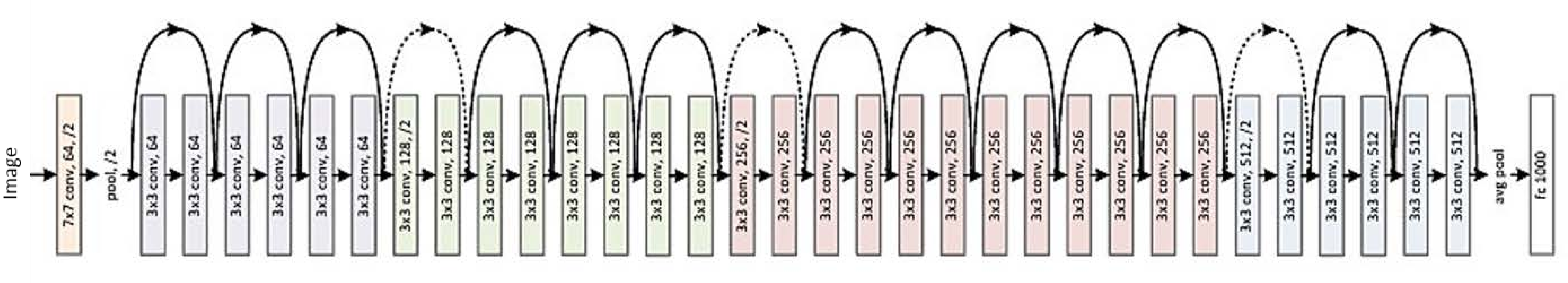


Pre-processing pipeline – Implementation of VGG-19



- Data augmentation: Increased brightness (usually within [0.5, 1.25] range)
- Dropout for regularization in Dense nodes ⇨ reduced overfitting
- Early stopping after 15 epochs, lr = 1e-05 with ReduceLROnPlateau callback

Pre-processing pipeline – Implementation of ResNet50



- Data augmentation: Increased brightness (usually within [0.5, 1.25] range)
  Horizontal+vertical flip
  45° rotation with fill_mode = 'constant'
- Dropout for regularization in Dense nodes → reduced overfitting
- Early stopping after 15 epochs, lr = 1e-05 with ReduceLROnPlateau callback

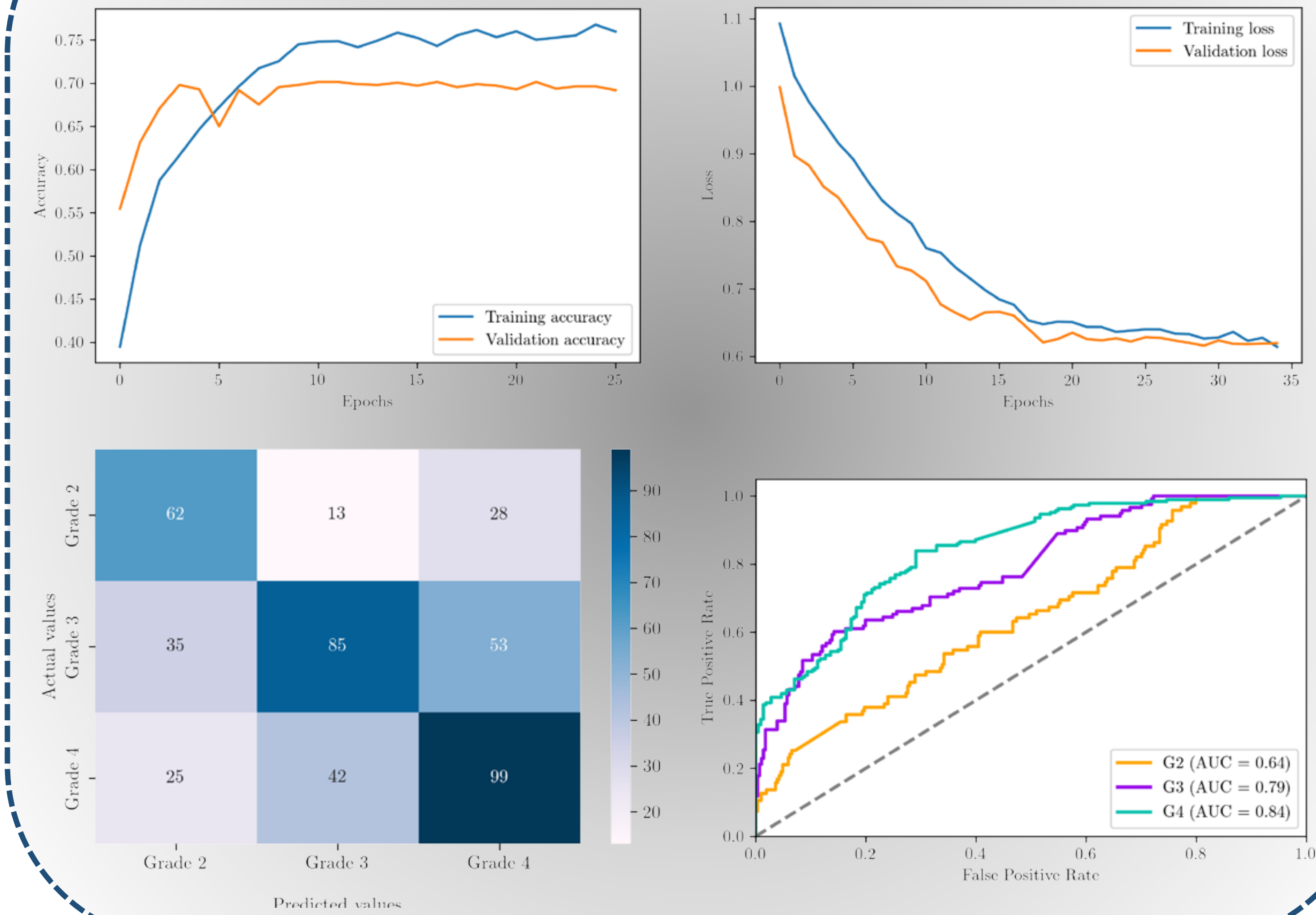
Image source: https://becominghuman.ai/resnet-convolution-neural-network-e10921245d3d

LINKÖPING UNIVERSITY

Pre-processing pipeline

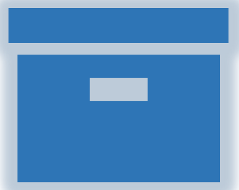
Step 7

Evaluation methods



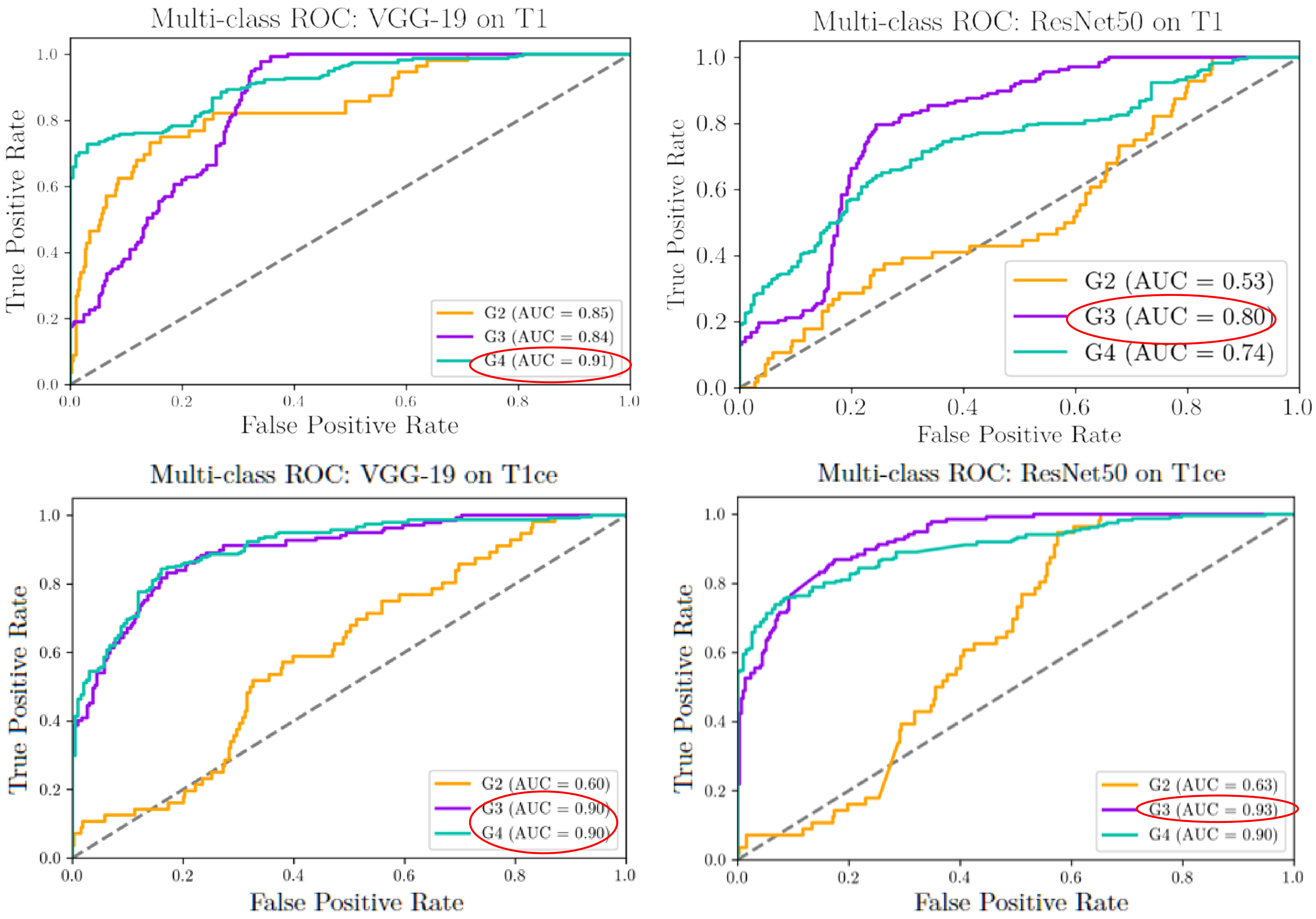
Best Classifier

Grad-CAM



Results (1)

Single modalities

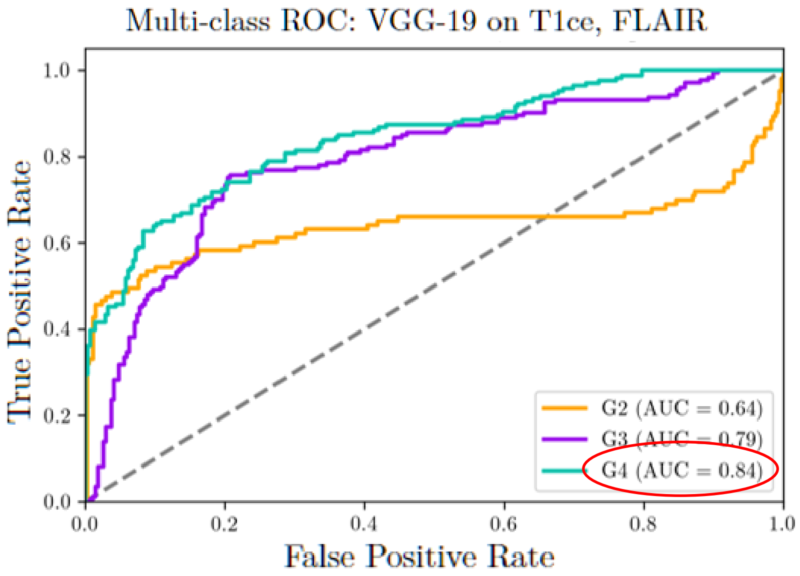


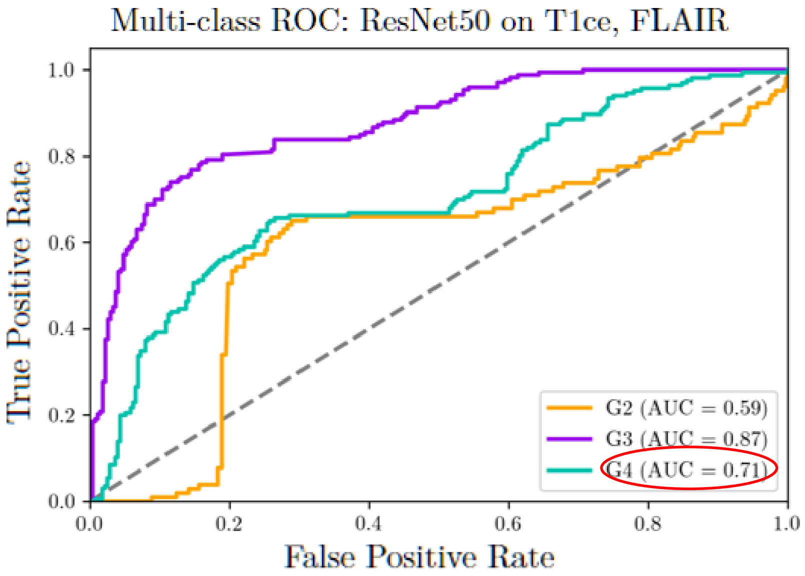
Summary (1)

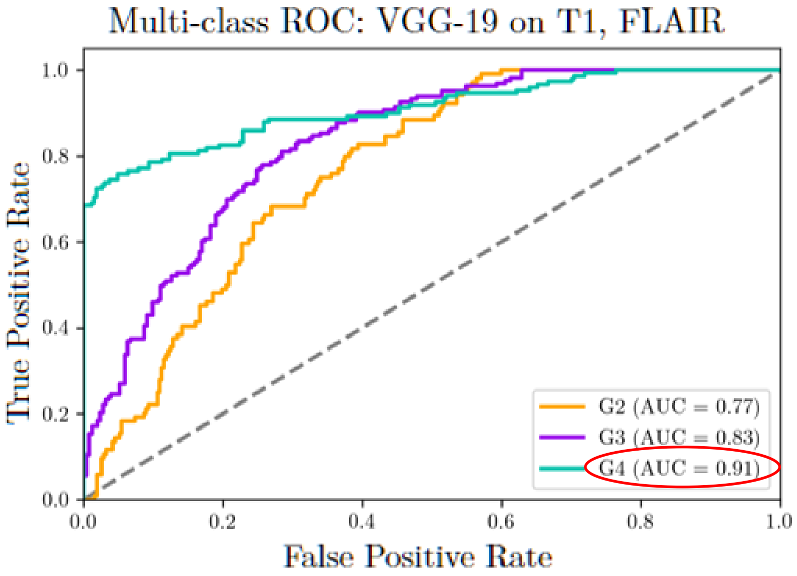
Model	MR Modality	Grade	AUC	Test accuracy
VGG-19	T1	G2	0.85	77.86%
		G3	0.84	
		G4	0.91	
	T1ce	G2	0.60	75.00%
		G3	0.90	
		G4	0.90	
ResNet50	T1	G2	0.53	71.35%
		G3	0.80	
		G4	0.74	
	T1ce	G2	0.63	77.34%
		G3	0.93	
		G4	0.90	

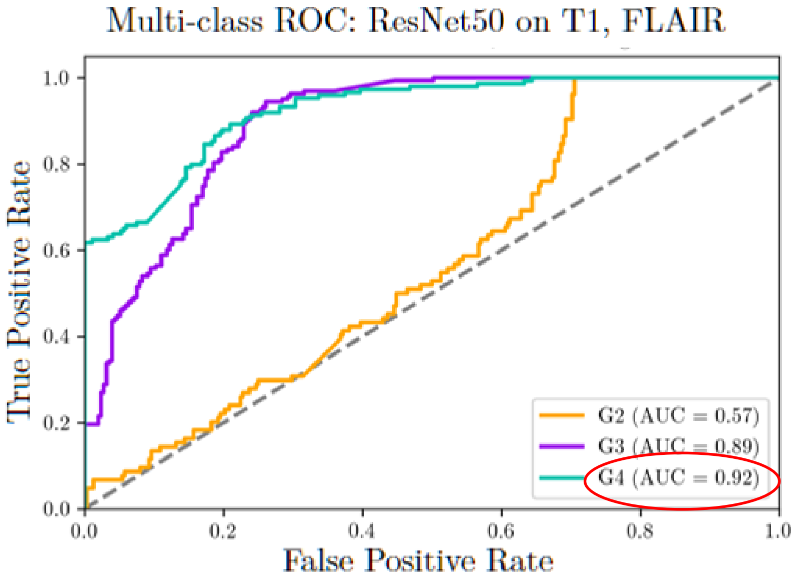
Results (2)

Combinations of 2 modalities







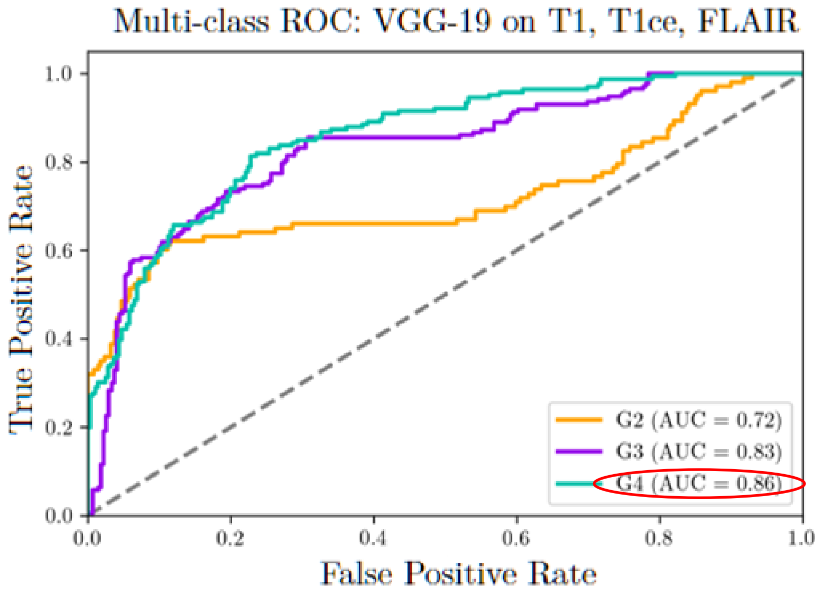


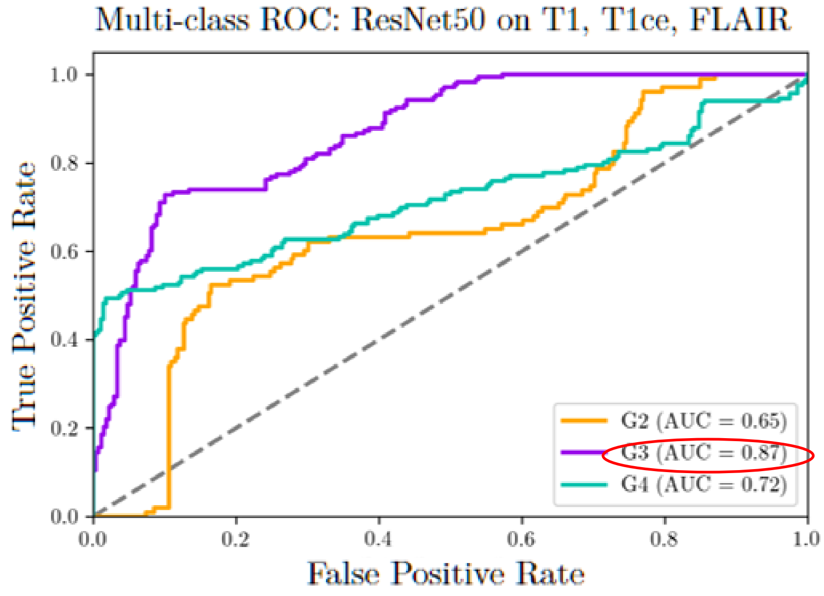
Summary (2)

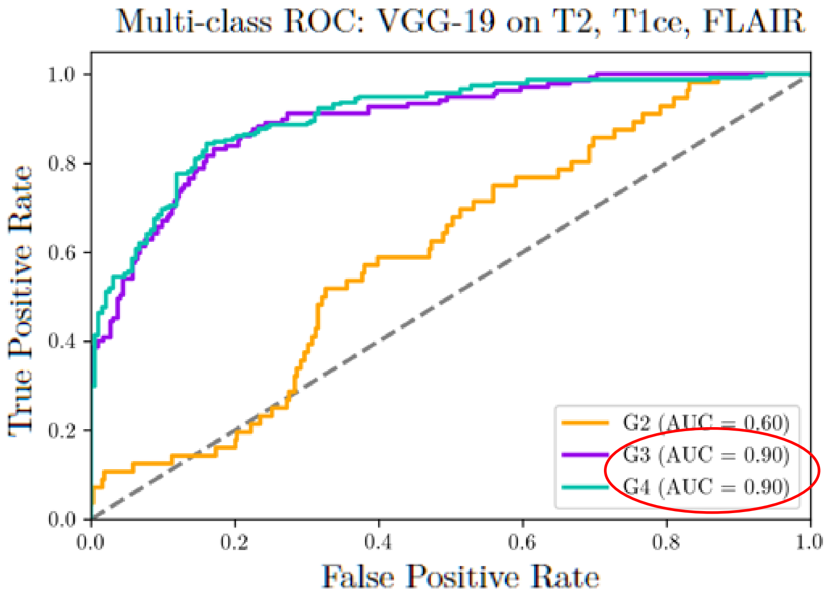
Model	MR Modality	Grade	AUC	Test accuracy
VGG-19	T1, FLAIR	G2	0.77	71.09%
		G3	0.83	
		G4	0.91	
	T1ce, FLAIR	G2	0.64	74.48%
		G3	0.79	
		G4	0.84	
ResNet50	T1, FLAIR	G2	0.57	67.71%
		G3	0.89	
		G4	0.92	
	T1ce, FLAIR	G2	0.59	63.54%
		G3	0.87	
		G4	0.71	

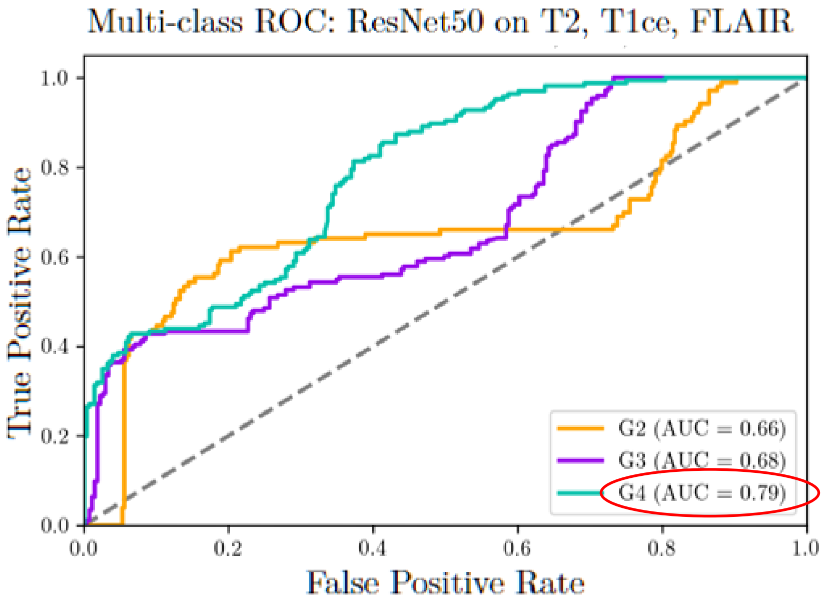
Results (3)

Combinations of 3 modalities





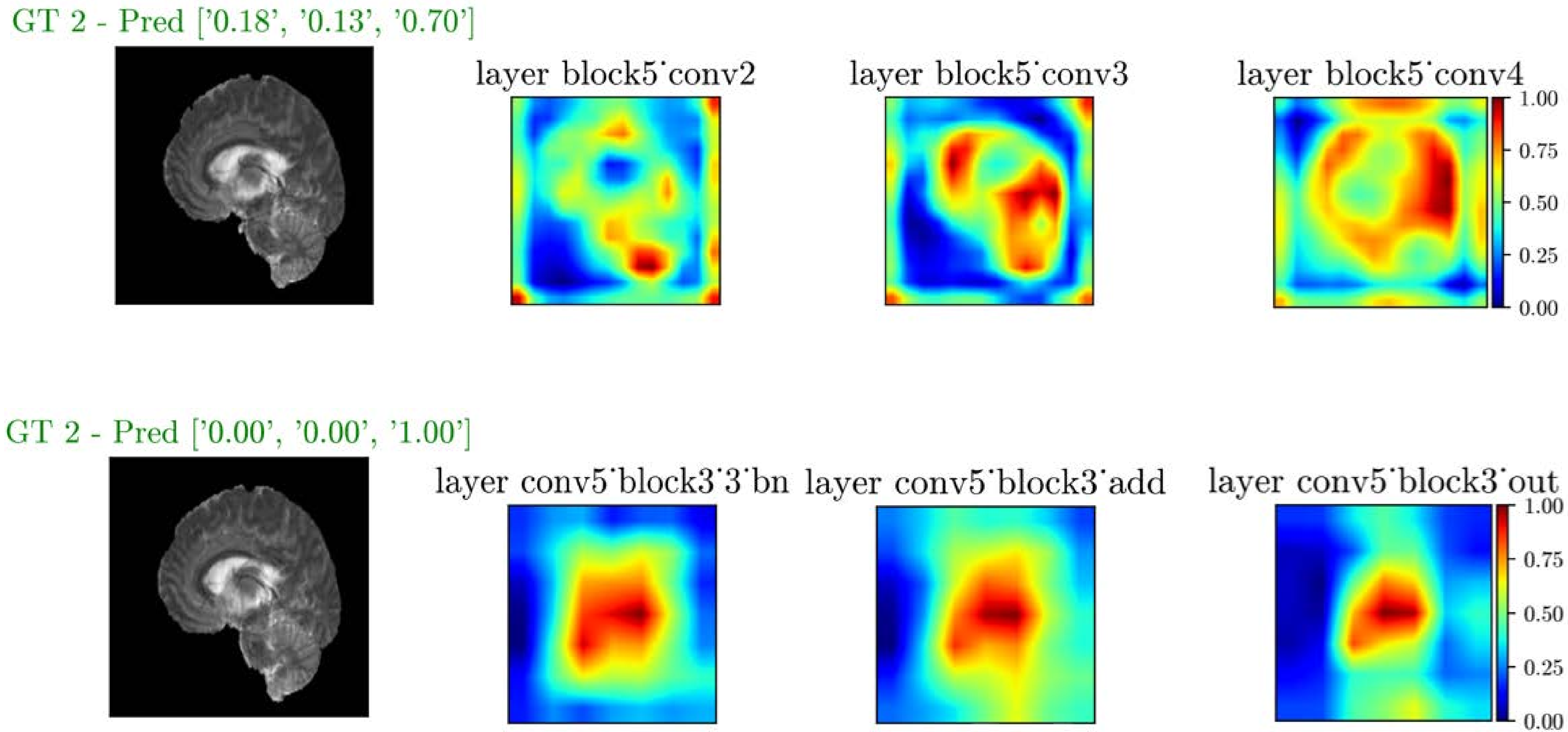




Summary (3)

Model	MR Modality	Grade	AUC	Test accuracy
VGG-19	T1, T1ce, FLAIR	G2	0.72	73.70%
		G3	0.83	
		G4	0.86	
	T2, T1ce, FLAIR	G2	0.60	75.78%
		G3	0.90	
		G4	0.90	
ResNet50	T1, T1ce, FLAIR	G2	0.65	72.40%
		G3	0.87	
		G4	0.72	
	T2, T1ce, FLAIR	G2	0.66	61.46%
		G3	0.68	
		G4	0.89	

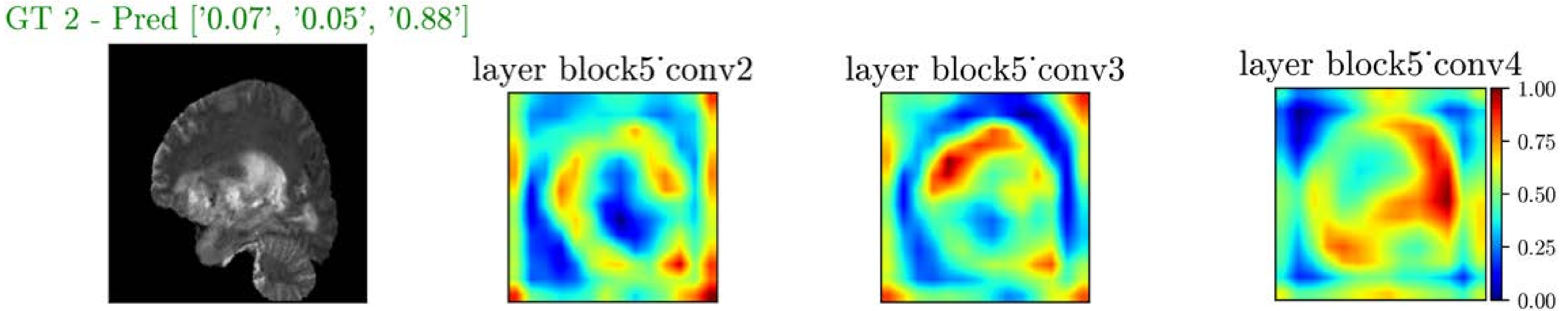
Results – Grad-CAM on T2 images (coarse localization maps)



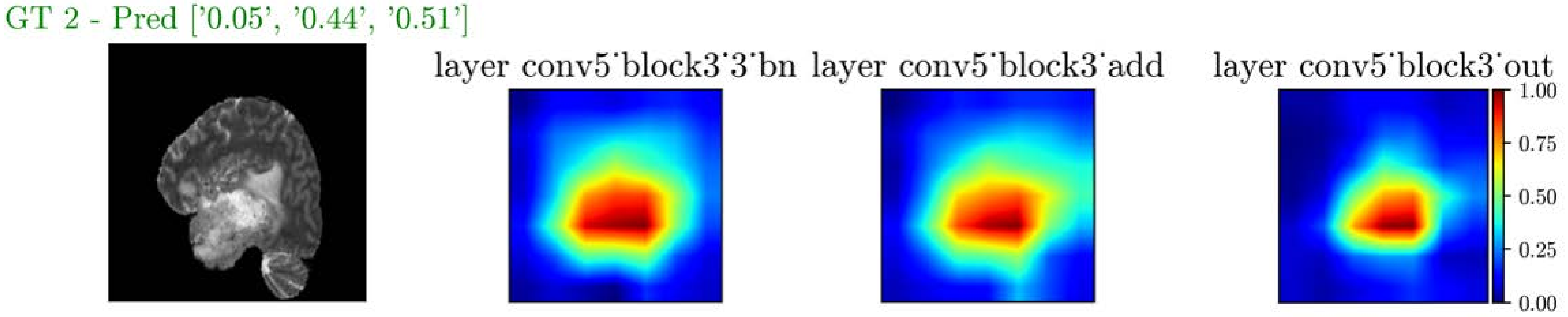
VGG-19

ResNet50

Results – Grad-CAM on T2, T1ce, FLAIR images (coarse localization maps)



VGG-19



ResNet50

LINKÖPING
UNIVERSITY

Conclusion

1. **Which of the popular CNN models performs best for brain tumor grade classification on the available dataset?**

   VGG-19 outperforms ResNet50 in most cases.

2. Which combination of MRI modalities yields the most optimal results for classification?

T1ce, FLAIR and T2, T1ce, FLAIR.

3. Is it possible to construct representations of model explainability?

Yes, in terms of visualizing the regions of the image used for the model's decision.

Thank you!

LINKÖPING
UNIVERSITY