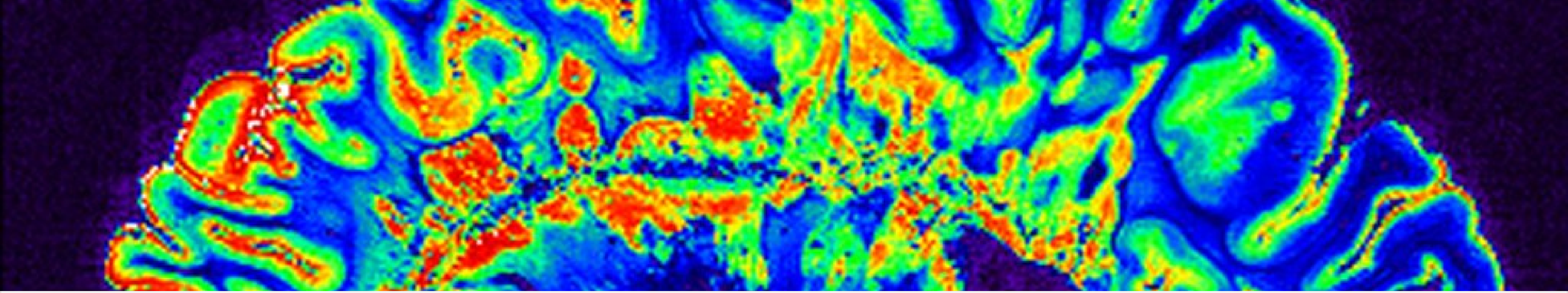


Diagnostic Aid System for Diffuse Gliomas using Deep Learning: Determination of IDH Status through Magnetic Resonance Imaging

Sergio Aizcorbe Pardo

Supervisor: Fernando Díaz de María

Collaborators: Juan Adan Guzmán-De-Villoria and Pilar Fernández

Project Design

1. Problem Approach

- Understanding Gliomas
- Deep Learning for MRI

2. Data Processing

- Data Sources
- Data Standardization

3. Modeling

- Convolutional Neural Networks (CNNs)
- Attention Models

4. Evaluation

- Performance Metrics
- Model Comparison

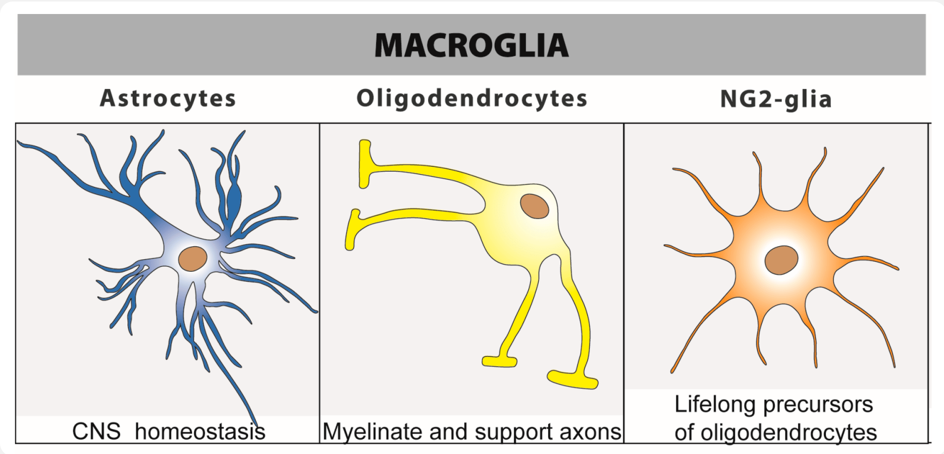
5. Conclusions

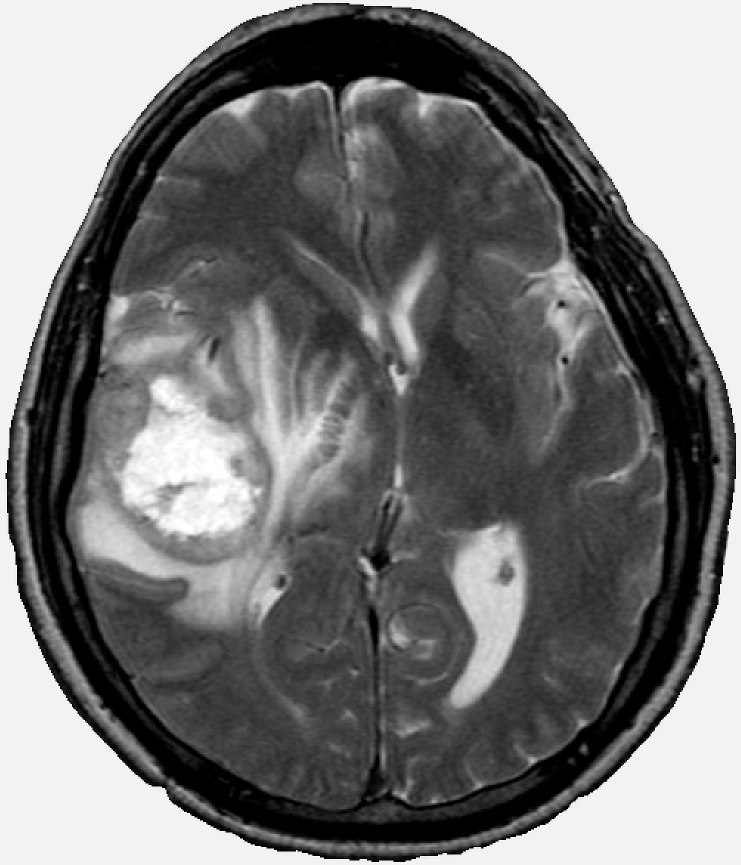
- Summary of Findings
- Challenges & Improvements

Introduction

Diffuse Gliomas - What are they?

- Most common type of brain tumor in adults (ADG)
- 80% are malignant
- They originate from glial cells





Glioma Characterization

Grades

1 — 2 — 3 — 4

Histological Subtype

	Origin cell	Grades
Astrocytoma	Astrocytes	2, 3, 4
Oligondendroglioma	Oligodendrocytes	2, 3, 4
Glioblastoma	-	4

Biological Markers

– IDH Status

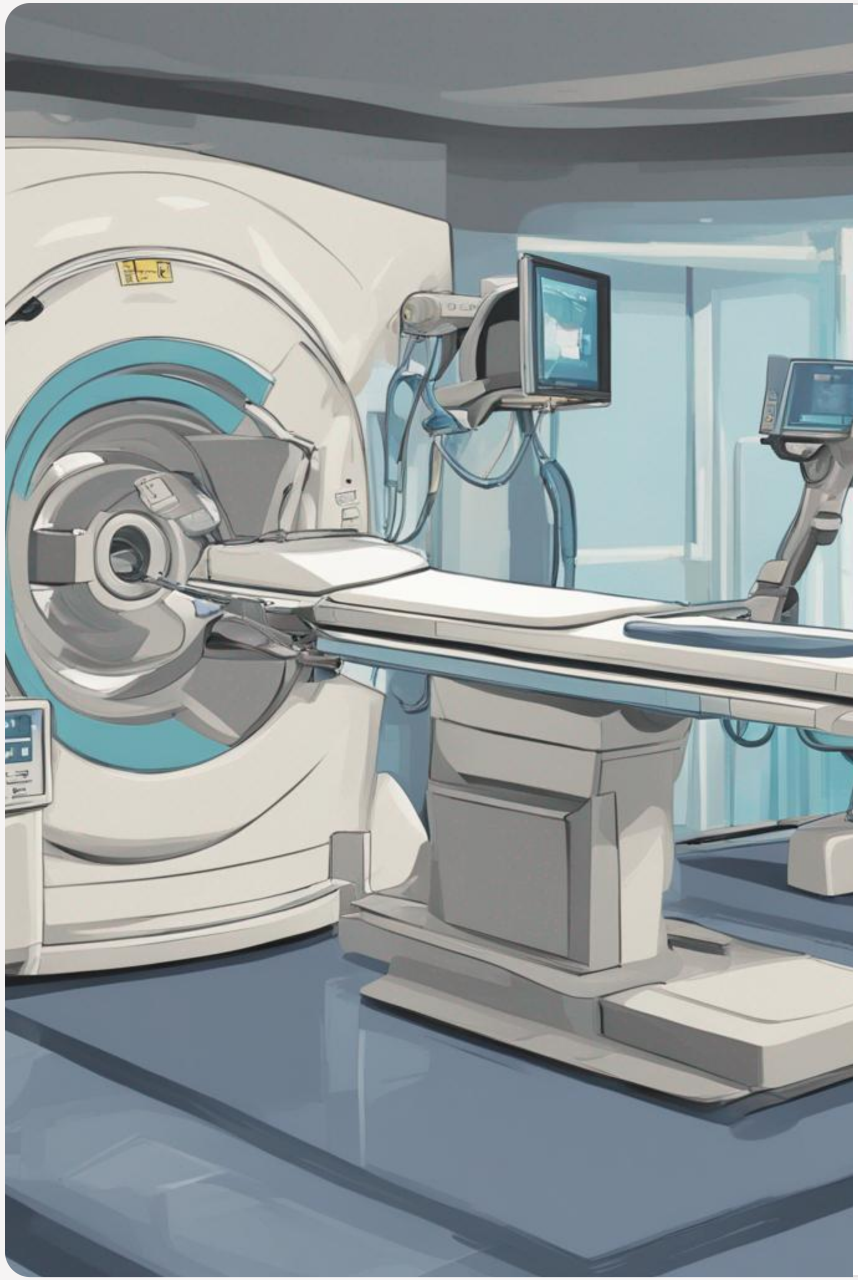
- IDH Positive → Mutant 👍
- IDH Negative → Wildtype 👎

– 1p/19q codeletion

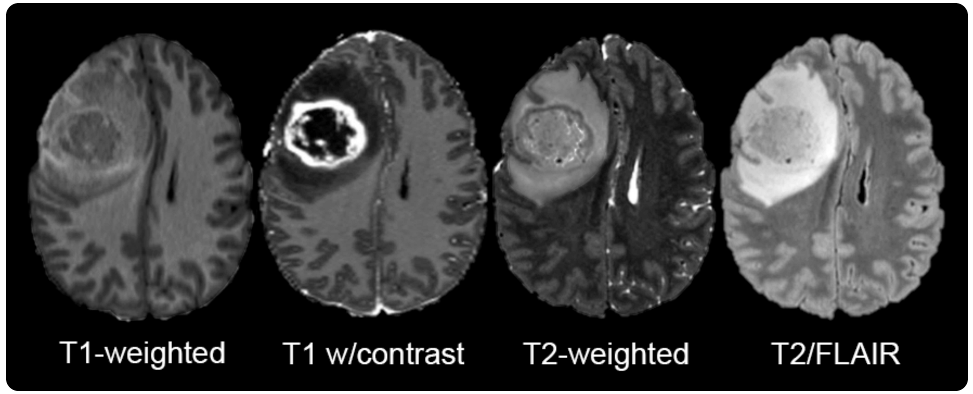
- Codeleted
- Non-codeleted

– ATRX, TERT, MGMT, TP53...

⚠️ **OPERATING RISKS AND SAMPLING ERRORS**



Types of MRI Sequences



	T1	T1CE	T2	FLAIR
Bright	Fat, hemorrhage	Blood-barrier areas (tumors)	Fluid, certain tumors	Lesions
Dark	CSF	Normal tissue	Bone, air, calcifications	CSF

- DWI, SWI, HARDI, ASL...

Neural Networks

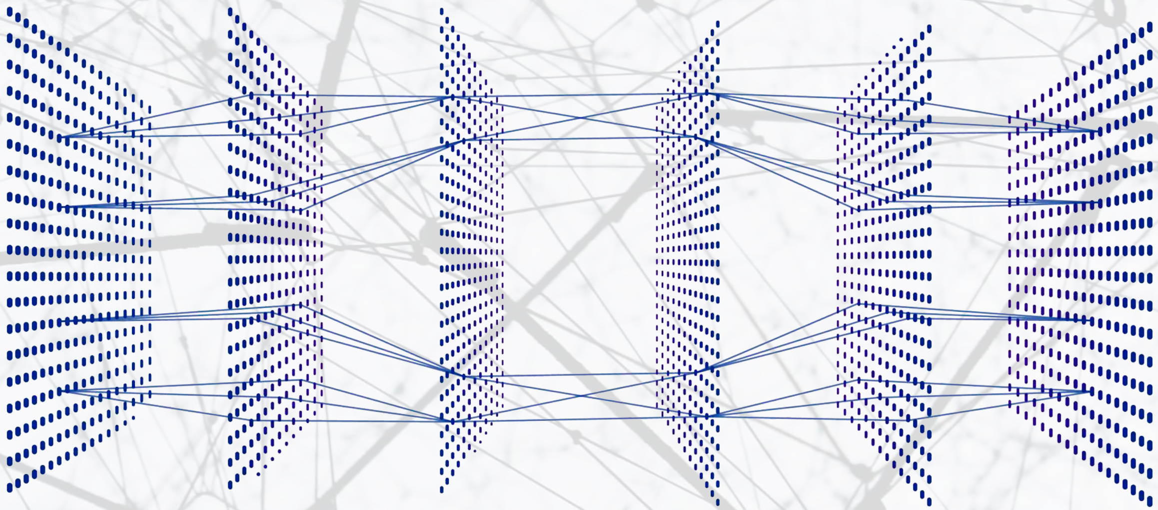
- Brain inspired structure
- Interconnected nodes
- Weights adjusted dynamically

✅ Extract **features** automatically

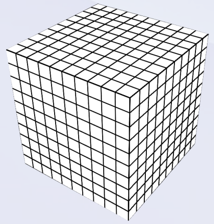
✅ Good at **generalization**

❌ **Time** and computational cost

❌ Lack of **interpretability**



Neural Networks for Computer Vision



Convolutional Neural Networks (CNNs)

- Uses convolutional layers
- Specialized for local patterns
- Less data-dependant for good performance
- Widely-adopted in medical imaging

Attention Models

- Weights input data importance
- Focuses on specific image regions
- Requires large amounts of data for optimization
- Emerging promise in imaging analysis

Data gathering and pre-processing

Data sources

TCIA Database → 501 patients

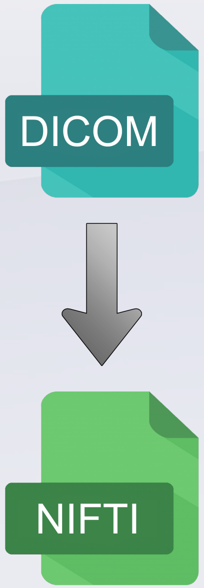
- **MRI Scanners** → T1, T1CE, T2, FLAIR, SWI, HARDI, ASL...
- **Molecular Data** → IDH Status, 1p/19q codeletion, MGMT
- **Additional Information** → Sex, age, tumor grade, histological subytpe, OS
- **Scanner Magnetic Field Strength:** 3.0T

HGUGM Database → 42 patients

- **MRI Scanners** → T1, T1CE, T2, FLAIR
- **Molecular Data** → IDH Status, 1p/19q codeletion, MGMT
- **Additional Information** → Sex, age, tumor grade, histological subytpe
- **Scanner Magnetic Field Strength:** 1.5T

Data Processing I

1. Data Conversion



2. Spatial Normalization

- **Image Resampling**
  - Size → 240×240×155
  - Windowed Sinc Interpolation
- **Isotropic Resampling**
  - 1 mm. isotropic resolution
- **Image Reorientation**
  - RAI (Right, Anterior, Inferior)

3. BRATS Pre-processing

- **N4 Bias Field Correction**
  - Minimize lighting artifacts
- **Rigid Registration**
  - T1CE as reference (SRI)
- **Skull-stripping**
  - BRATS 2019
- **Tumor Segmentation**
  - DeepMedic (ISLES 2015)

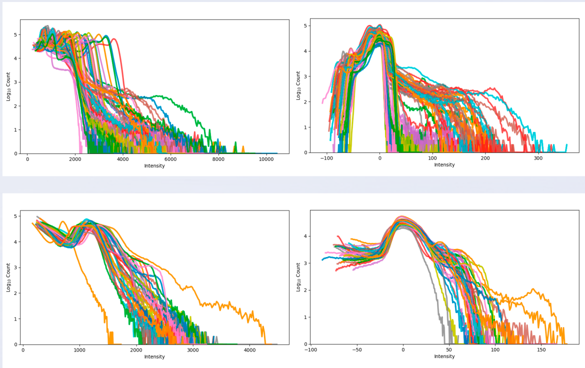
Data Processing II

4. Intensity Matching

- **KMeans Clustering**
  - Compute histograms
  - Group with Kmeans
  - Predict histogram
  - Identify closest histogram match

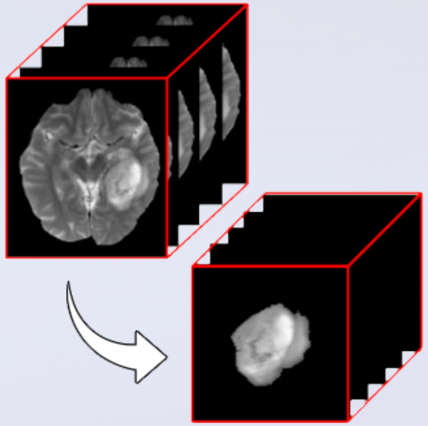
5. Intensity Normalization

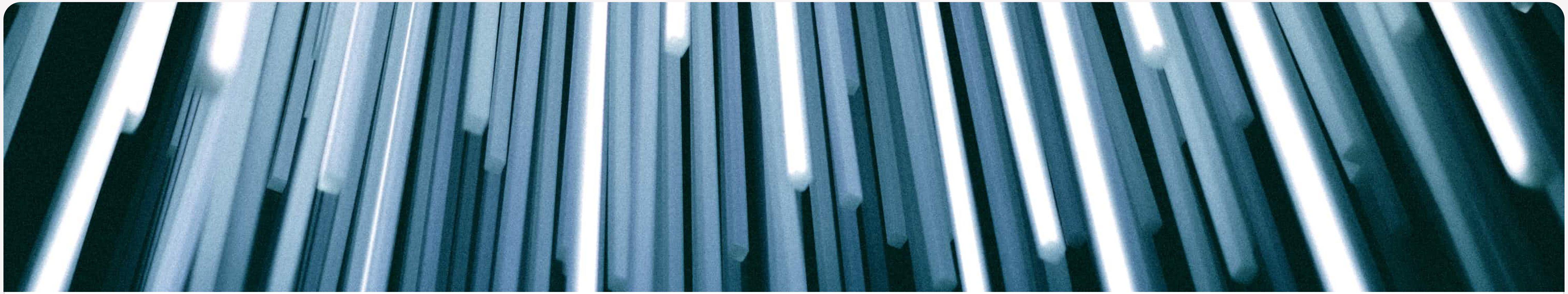
- Voxel values **standarization**
- Adjust **contrast range**
- **WhiteStripe** method



6. Tumor Cropping

- **Tumor region isolation**
  - Shape → 128×128×64





Training set		Validation set		Testing set	
318 patients		107 patients		109 patients	
295 patients	23 patients	99 patients	8 patients	100 patients	9 patients
TCIA	HGUGM	TCIA	HGUGM	TCIA	HGUGM
The Cancer Imaging Archive	Hospital General Universitario Gregorio Marañón	The Cancer Imaging Archive	Hospital General Universitario Gregorio Marañón	The Cancer Imaging Archive	Hospital General Universitario Gregorio Marañón

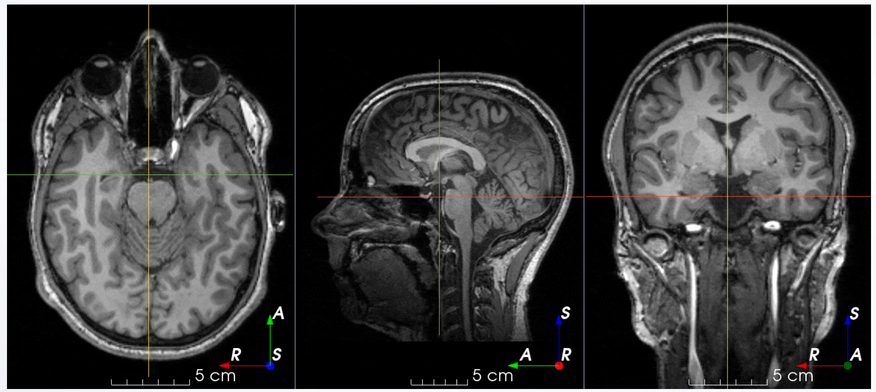
Data Augmentation

Spatial transformations

- Crop foreground
- Random 90º rotation
- Random zoom
- Random affine

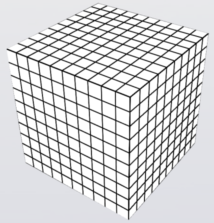
Intensity transformations

- Random scale intensity
- Random adjust contrast
- Random Gaussian noise
- Rand. Gaussian sharpen
- Random shift intensity
- Random K Spike Noise



Modeling

Models



Convolutional Neural Networks (CNNs)

- DenseNet
- HighResNet

Attention Models

- Attention UNET
- ViT Autoencoder
- UNETR

Auxiliary Features

Age → (17 - 94)

- Numerical variable → Normalized

Sex → (M / F)

- Categorical variable → One-hot encoded

Tumor Grade → (2, 3, 4)

- Categorical variable → One-hot encoded

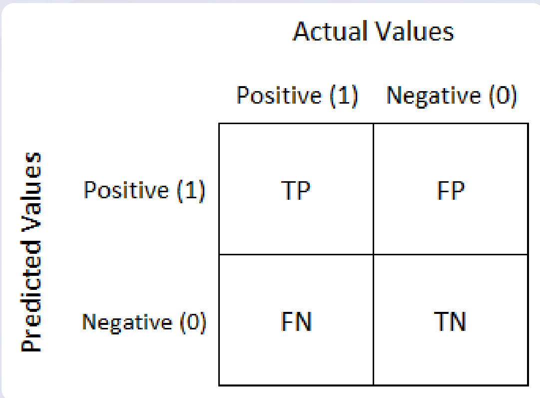
Late Fusion

Information is "fused" in the last layers of the neural network

Evaluation and results

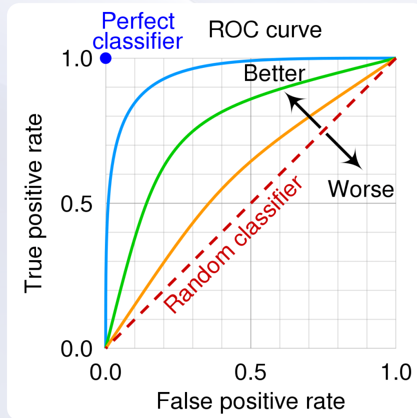
Metrics

Confusion Matrix



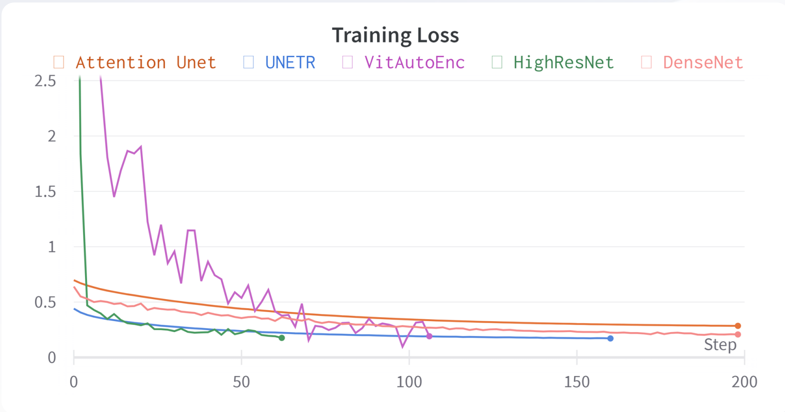
- **Accuracy** → Correct predictions over total
- **Precision** → True positives over predicted positives
- **Recall (TPR)** → True positives over actual positives
- **F1 Score** → Weighted average of precision and recall

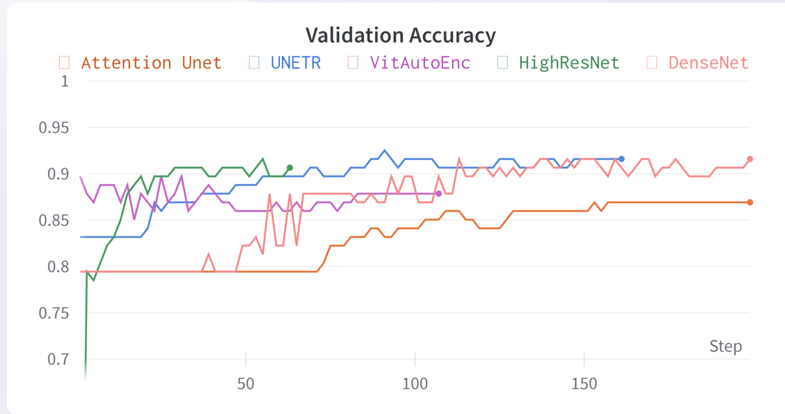
AUC Score



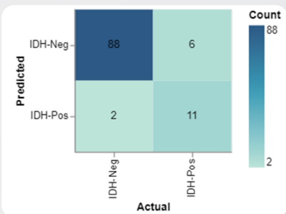
- Evaluates the **performance** of binary classifications models
- **ROC Curve** → TPR against FPR
- ***AUC = 1*** → Perfectly distinguishes classes
- ***AUC <= 0.5*** → No better than random choice

Comparative Analysis I

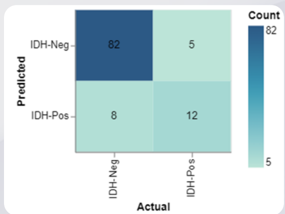




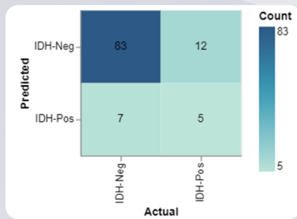
DenseNet



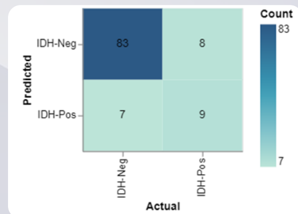
HighResNet



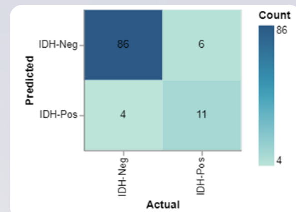
Attention UNET



ViT Autoencoder



UNETR



Comparative Analysis II

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC (%)
DenseNet	**92.52**	**84.62**	64.71	**73.33**	91.18
HighResNet	87.85	60.00	**70.59**	64.86	91.18
Attention UNET	82.24	41.67	29.41	34.48	82.22
ViT Autoencoder	85.98	56.25	52.94	54.55	86.11
UNETR	90.65	73.33	64.71	68.75	**91.24**

DenseNet exhibits the highest accuracy, precision and F1 score percentage.

UNETR is the attention-based model with the best results.

The most common FP samples are **Grade 4 Glioblastomas** from **younger patients**.

The most common FN samples are **Grade 4 Astrocytomas** from **patients aged 40-60 years**.

Conclusions

CNN designs remain robust, even in data-limited circumstances; large amounts of data are necessary for transformer-based models.

Pre-processing and the choice of model architecture remain critical.

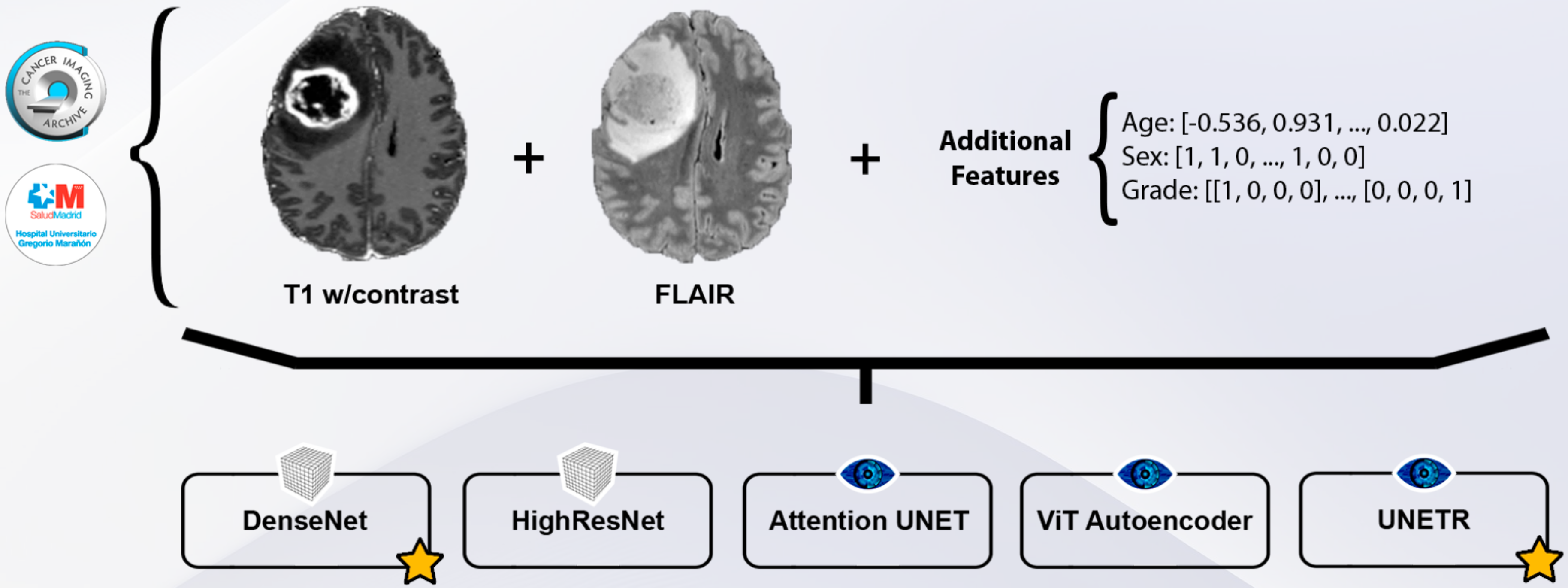
Deep learning has the potential to improve medical diagnosis.

Challenges, limitations and enhancements

	Challenges / Limitations	Potential Enhancements
Data Sources	• Data imbalance • Scanner intensity variations	• Larger datasets • More data augmentation
Pre-Processing	• Interpolation methods • Intensity matching & Normalization	• Exploratory pre-processing • Standard pipelines
Modeling	• MRI sequences • Additional features • Hyperparameter tuning	• More network architectures • Feature Analysis • Cross-Validation

Review

IDH Classification



Thank you.

Questions & Answers

saizk/**GlioScan**

IDH Classification for Gliomas using CNN and Transformers.

1 Contributor · 0 Issues · 0 Stars · 0 Forks

GitHub

GitHub - saizk/GlioScan: IDH Classification for Gliomas using CNN and...

IDH Classification for Gliomas using CNN and Transformers. - GitHub - saizk/GlioScan: IDH Classification for Gliomas using CNN and Transformers.