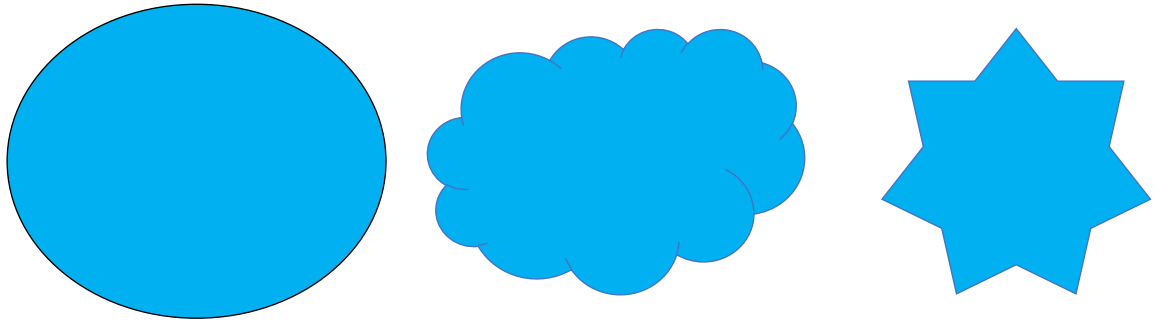
# Radiogenomics applied (Hands-on) - Pyradiomics

Esra Sümer, MSc
Boğaziçi University

**GliMR Training School 2022**
**Artificial Intelligence in Neuro-oncology**

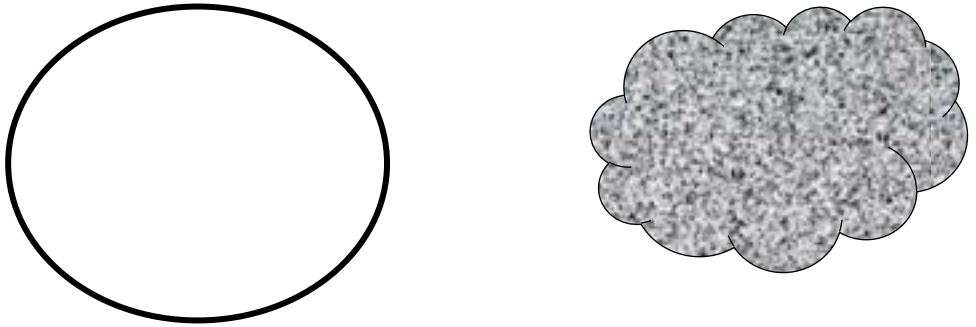
# Introduction

Tumor shape



‘spherical’ ‘lobular’ ‘spiculated’

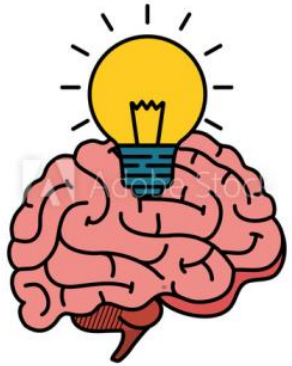
Contrast enhancement



‘homogeneous’ ‘heterogeneous’

*Subjective assessment based on human eye perception*

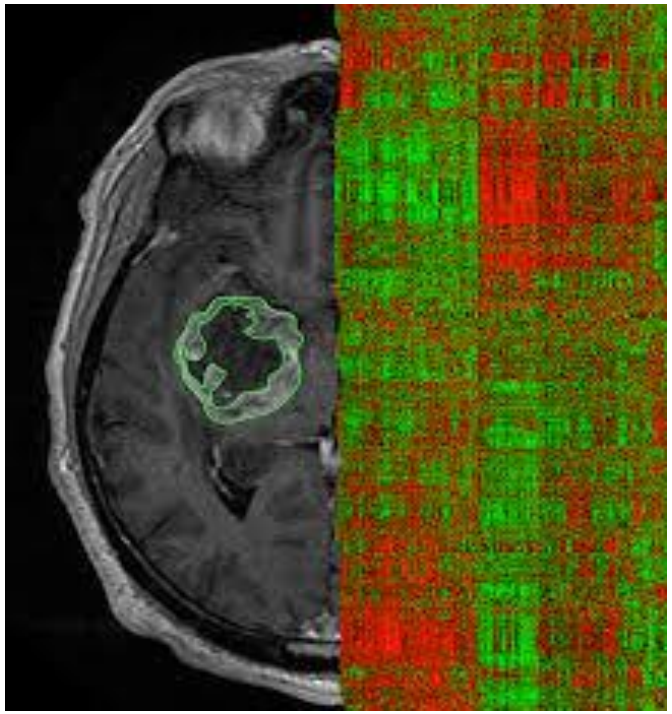
Automated approach
Quantitative patterns



# PyRadiomics

*"Transforming medical images into minable high dimensional data"*

- Harvard School of Medicine
- Open source Python package[1,2,3]



~

- Tumor imaging characteristics
- Intratumoral heterogeneity
- Genetic characteristics
- Tumor phenotype
- Treatment outcome

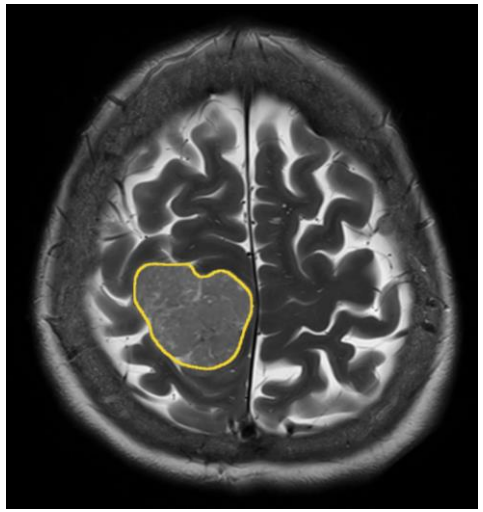
Radiogenomics

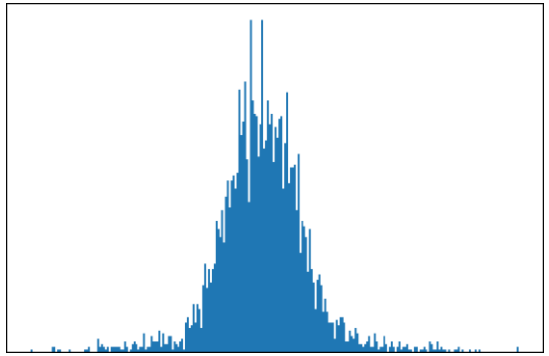
✓ Supporting the clinical decision making,
✓ Personal treatment,
✓ Prediction of prognosis,
✓ …

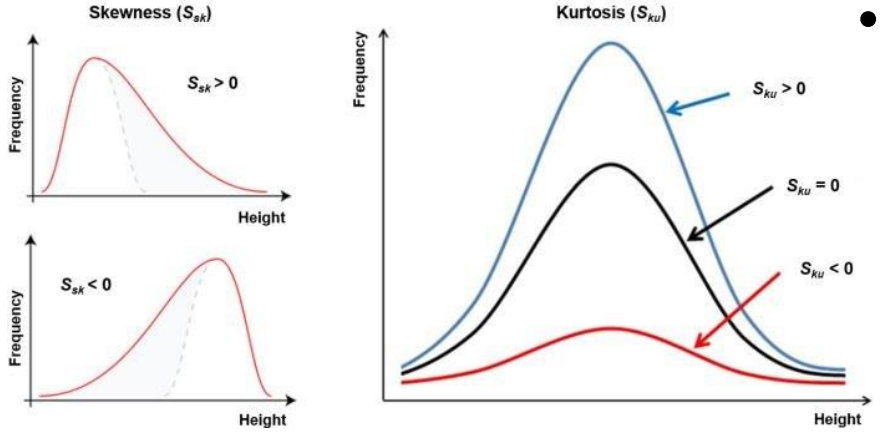

[1] *Van Griethuysen JJM et al. Cancer Res. 2017;77(21):e104-e107. doi:10.1158/0008-5472.CAN-17-0339*
[2] *https://pyradiomics.readthedocs.io/en/latest/features.html*
[3] *https://github.com/AIM-Harvard/pyradiomics*
[4] *Stanford Medicine, https://med.stanford.edu/gevaertlab/ReviewRadiomicsBrain.html, 2022.*

# Radiomic Features

*First order intensity (histogram-based) features*



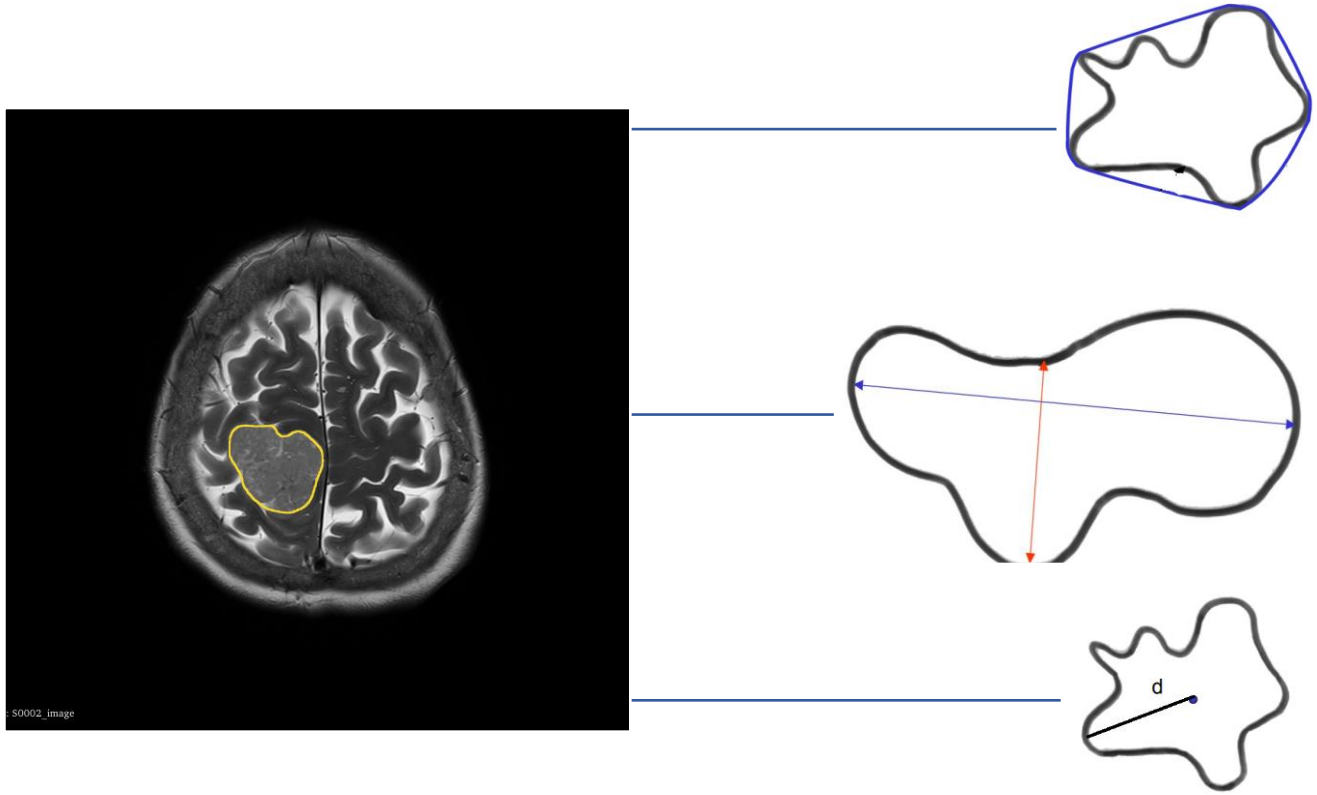





- Energy
- Skewness
- Entropy
- Median
- 15. ve 75. percentiles
- Kurtosis
- …

# Radiomic Features

*2D Shape Features*




Perimeter

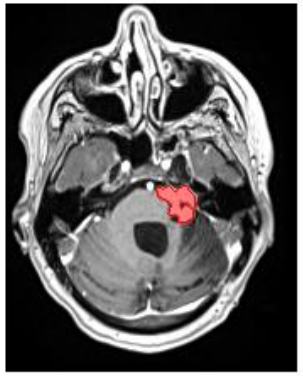
First and second principle axes

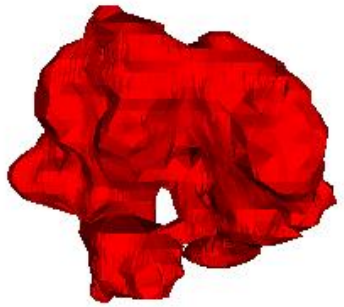
Maximum 2D radius

…

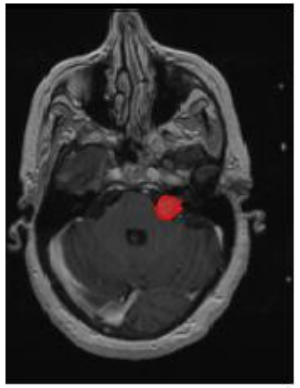
# Radiomic Features

*3D Shape Features*

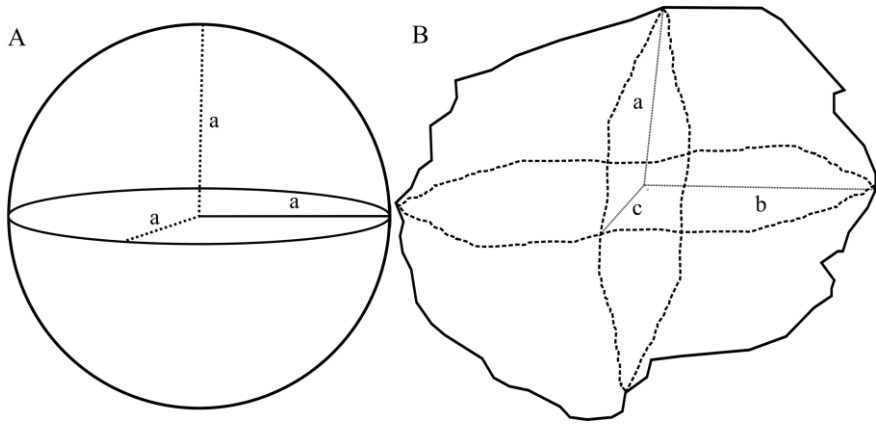




- Sphericity: 0.5
- Spherical disproportion: 1.99
- Elongation: 0.12
- Flatness: 0.67



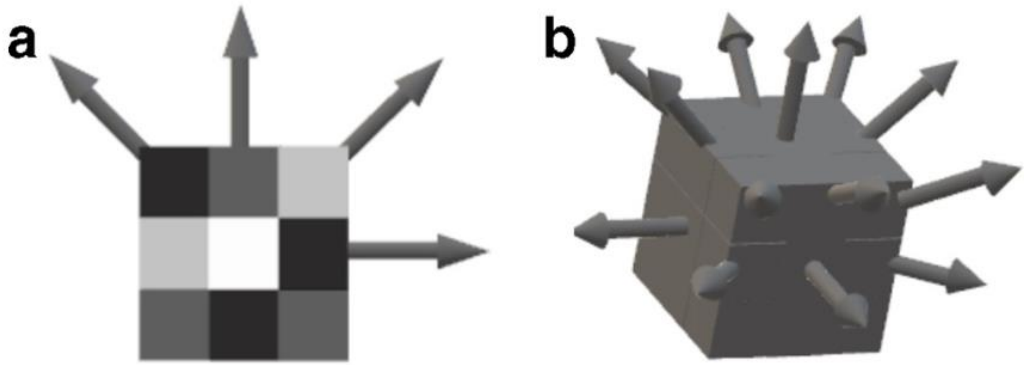
- Sphericity: 0.87
- Spherical disproportion: 1.15
- Elongation: 0.65
- Flatness: 0.83



# Radiomic Features

*Texture Features*

- GLCM
- GLRM
- GLSZM
- NGTDM
- GLDM




2D 3D

GLCM ($\theta$=0)

| | | | |
|---|---|---|---|
| 0 | 3 | 2 | 1 |
| 2 | 1 | 3 | 0 |
| 3 | 2 | 0 | 1 |
| 1 | 0 | 3 | 2 |

2

3

| | | | |
|---|---|---|---|
| 0 | 1 | 0 | 2 |
| 1 | 0 | 0 | 1 |
| 1 | 2 | 0 | 0 |
| 1 | 0 | 3 | 0 |

GLRM ($\theta$=0)

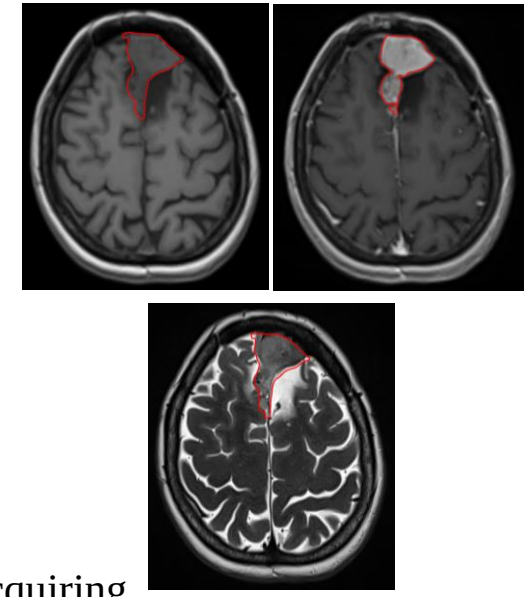
| | | | |
|---|---|---|---|
| 0 | 1 | 2 | 3 |
| 0 | 2 | 3 | 3 |
| 2 | 1 | 1 | 1 |
| 3 | 0 | 3 | 0 |

Runs 1-4

| | | | |
|---|---|---|---|
| 4 | 0 | 0 | 0 |
| 0 | 1 | 1 | 0 |
| 3 | 0 | 0 | 0 |
| 3 | 1 | 0 | 0 |

# Workflow of Radiomics

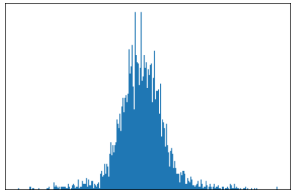
Benign/malignant?

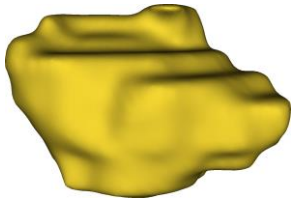


Acquiring medical images

Segmentation

Pre-processing (resampling, normalization, etc.)

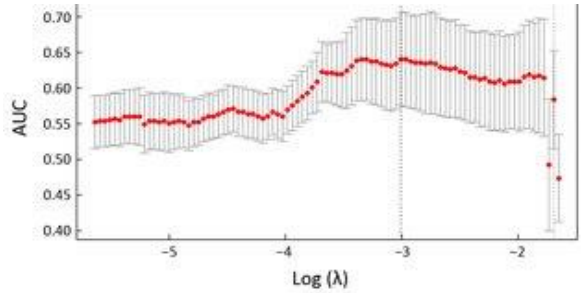




| | | | |
|---|---|---|---|
| 1 | 2 | 3 | 3 |
| 1 | 3 | 5 | 4 |
| 1 | 4 | 2 | 5 |
| 5 | 1 | 2 | 2 |

Feature extraction

Feature elimination



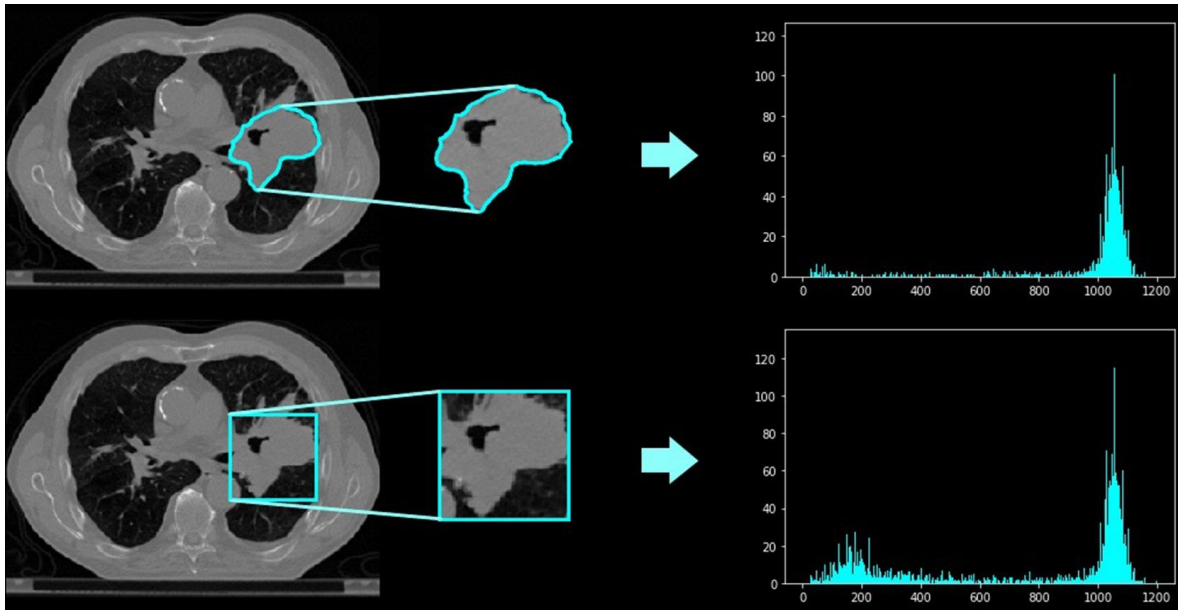

Wu, L. et al., *Chinese J Cancer Res*. 2018;30(4):396-405. doi:10.21147/j.issn.1000-9604.2018.04.02

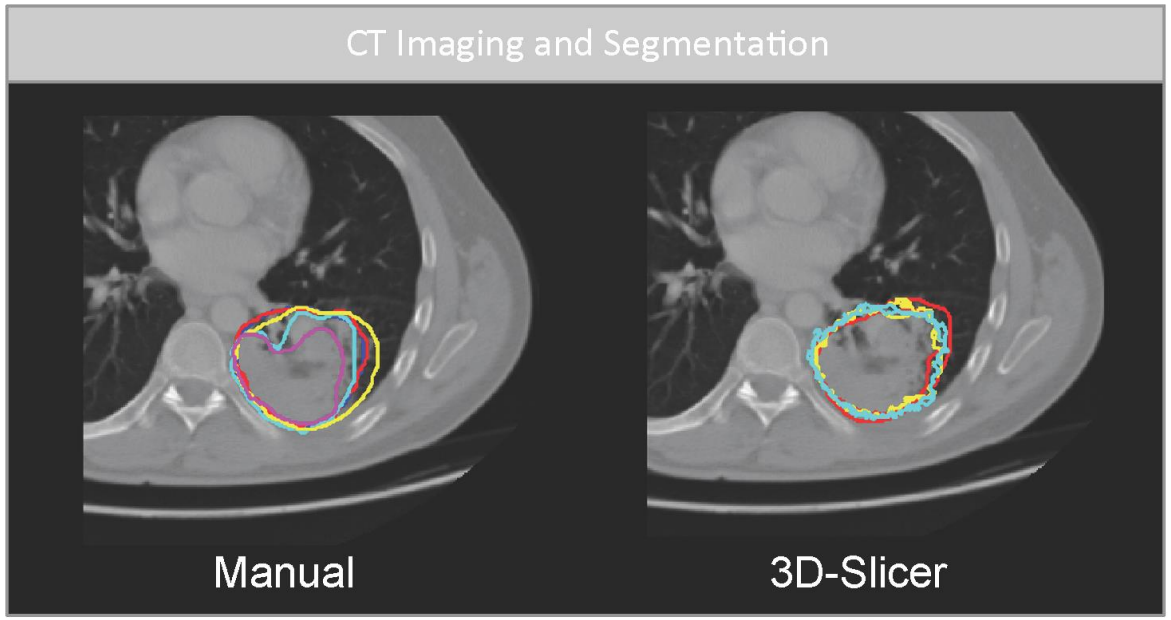
Statistical tests,

Model training and evaluation

Benign ☺

# 1- Segmentation



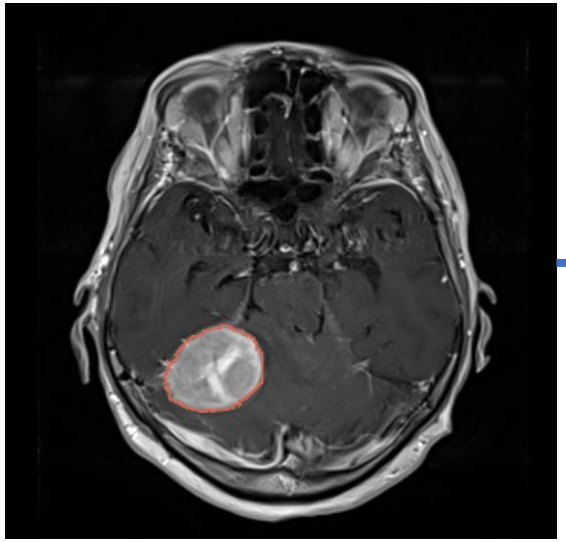

*Parmar C et al. PLoS One. 2014;9(7):1-8. doi:10.1371/journal.pone.0102107*

Manual segmentation vs. bounding box

Manual segmentation vs. 3D Slicer assisted (3 different operators)

- Manual, semi-automatic, and automatic
- Manual methods are time consuming and operator dependent
- Stability and reproducibility are low in manual methods
- 3D Slicer assisted segmentations have higher stability and reproducibility

# 2- Image Pre-processing and Feature Extraction



Filters

- Laplace+Gaussian
- Wavelet,
- Logarithmic,
- Exponential, etc.

Normalization,
Resampling,
Bin width etc.

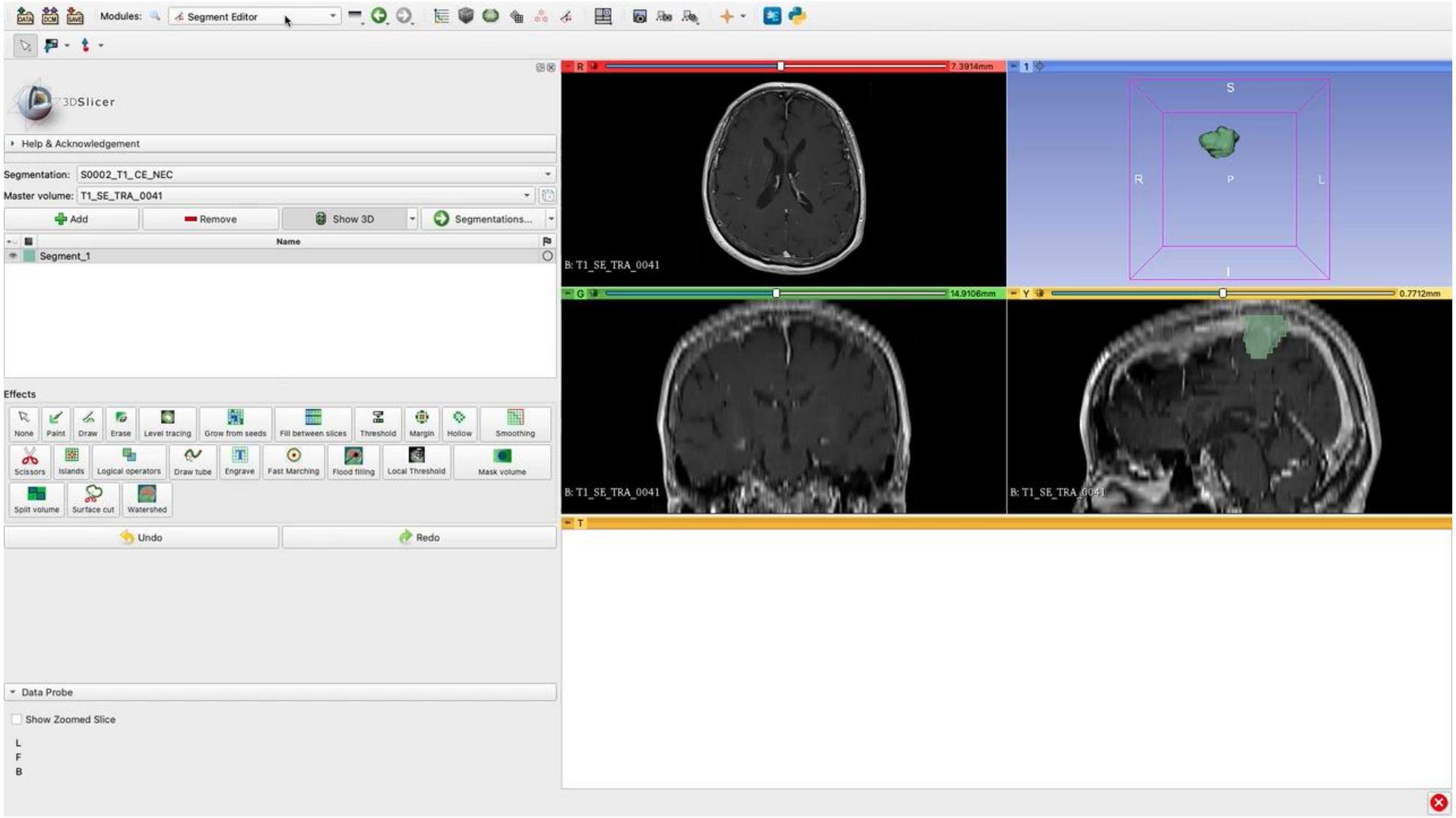
Settings

Feature extraction

Classification of Radiomics Features

- First-order intensity
- Shape (2D/3D)
- Texture

# 2- Feature Extraction with 3D Slicer



# 2- Feature Extraction with Python

| | |
|---|---|
| 1 | ID,Image,Mask,IDH,TERT |
| 2 | G0001,Gliom/nii_gliom.../G0001/segmentations..,1,0 |
| 3 | G0002,Gliom/nii_gliom.../G0002/segmentations..,1,0 |

File directories of image and mask pairs for each patient

Settings .yaml file

Feature extraction

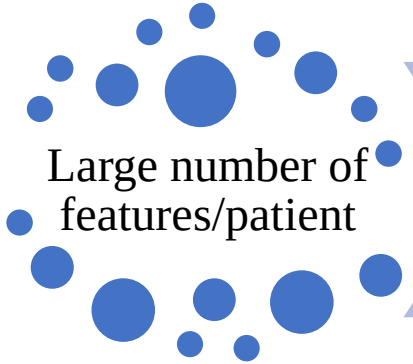
Features

# of patients

**Libraries:**

- Pandas
- SimpleITK
- Radiomics
- Radiomics feature extractor

# 3- Feature Selection



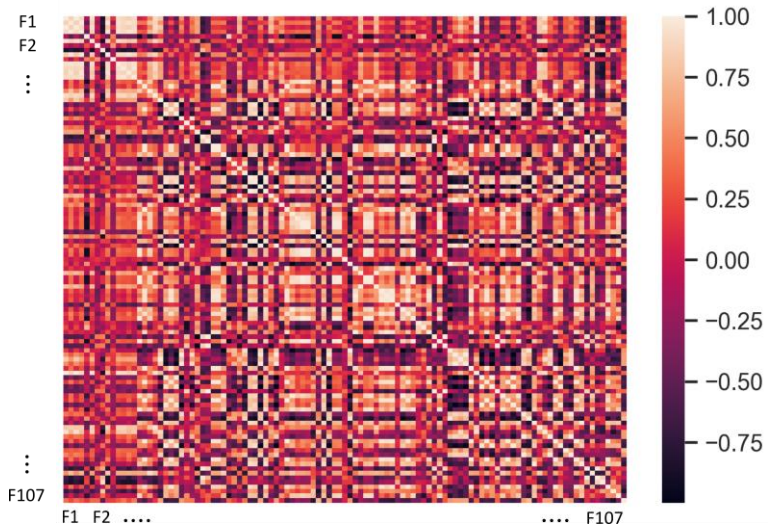
Large number of features/patient

Predictive features

Model training

Feature selection algorithms

- Supervised
- Unsupervised
- Dimension reduction




Supervised Feature Elimination

**Filter**

- Variation threshold
- Correlation: multicollinear features are eliminated
- Interclass correlation coefficient (ICC: 0.75-0.80)
- …

**Wrapper**

- Recursive feature elimination
- Sequential feature selection (forward-backward)
- …

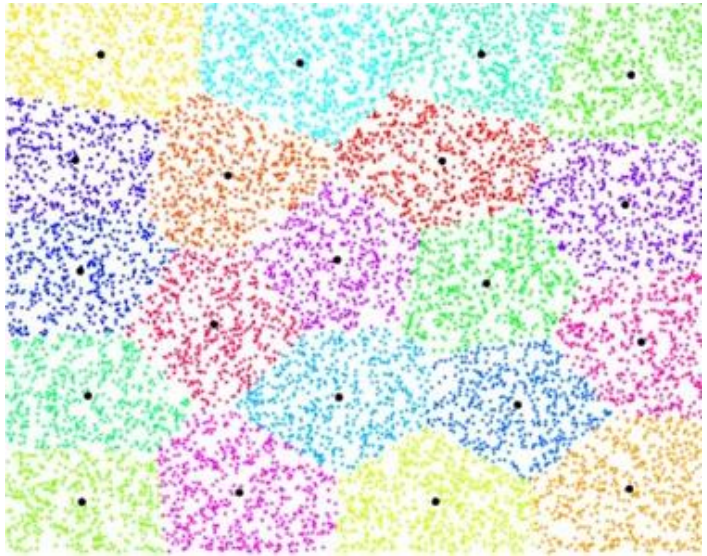
**Embedded**

- Lasso
- Ridge

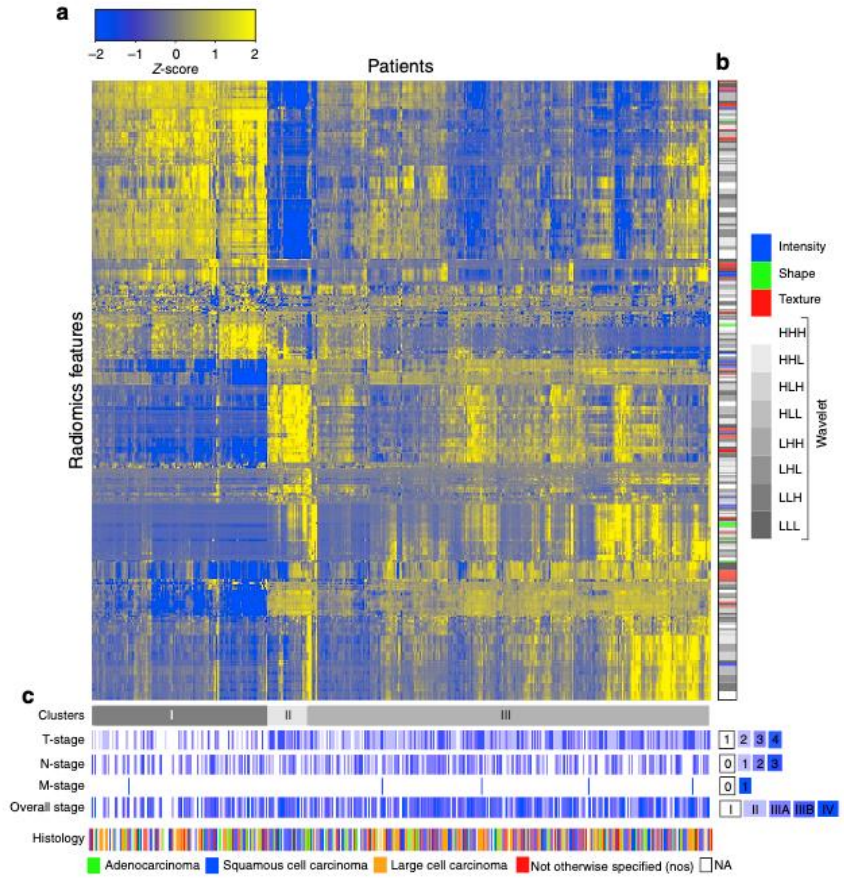
# 3- Feature Selection

Unsupervised

- K-means clustering



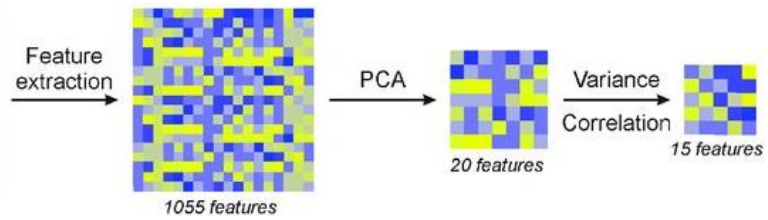
https://medium.com/@serapgur169/k-means-k%C3%BCmeleme-nas%C4%B1l-ya-985498eea343




*Aerts HJWL. et al., Nat Commun. 2014;5. doi:10.1038/ncomms5006*

Dimension reduction

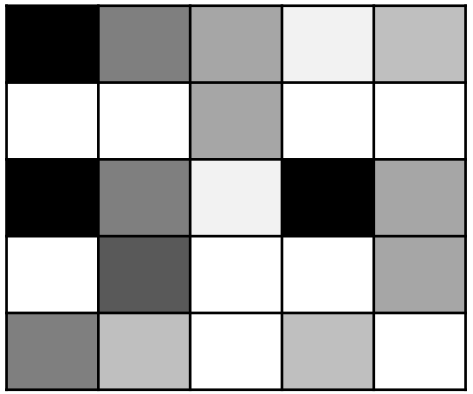
- Linear discriminating analysis (LDA)
- Principle component analysis (PCA)



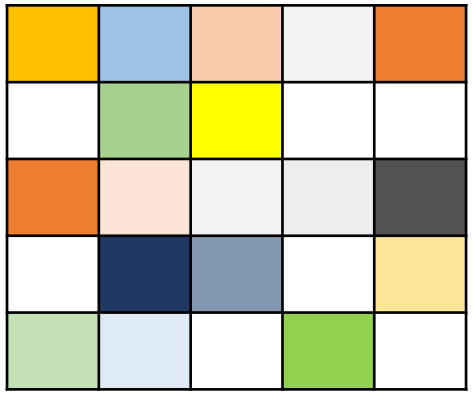

*Coroller TP et al. PLoS One. 2017;12(11):1-15. doi:10.1371/journal.pone.0187908*

# 4- Model Construction and Analysis

Clinical/genetic data



Radiomics features



**1. Statistical tests**

- Mann Whitney U (2 grups),
- Kruskal Wallis (more than 2 groups),
- Tukey Kramer post-hoc test,
- Dunn post-hoc test, etc.

**2. Classification**

- K- nearest neighbor,
- Linear discriminating analysis,
- Logistic regression,
- Support vector machine,
- Random forest classifier, etc.

**3. Survival Analysis**

- Cox proportional hazards model,
- Kaplan-Meier,
- Log-rank test, etc.