

# MLRF Lecture 06

J. Chazalon & N. Boutry, LRDE/EPITA, 2021

# Introduction to practice session 6

Lecture 06 part 04

# Using classification to segment images

## Until now

1 image  $\rightarrow$  many vectors (instance recognition)

1 image  $\rightarrow$  1 vector (image retrieval, image classification)

## Today / next practice session :

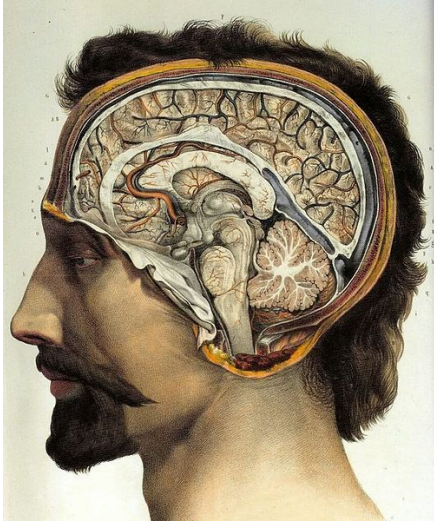
1 pixel  $\rightarrow$  1 vector (pixel classification, image semantic segmentation)

# Brain Anatomy and Imaging



# To investigate brain malfunction, two options:

Open human (then dispose)



Public domain: Bourguery, 1831-1854 via [Wikimedia Commons](#)

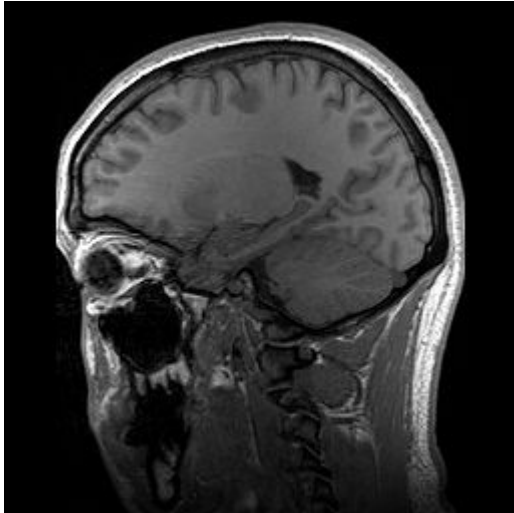
Use clever tools (like MRI)



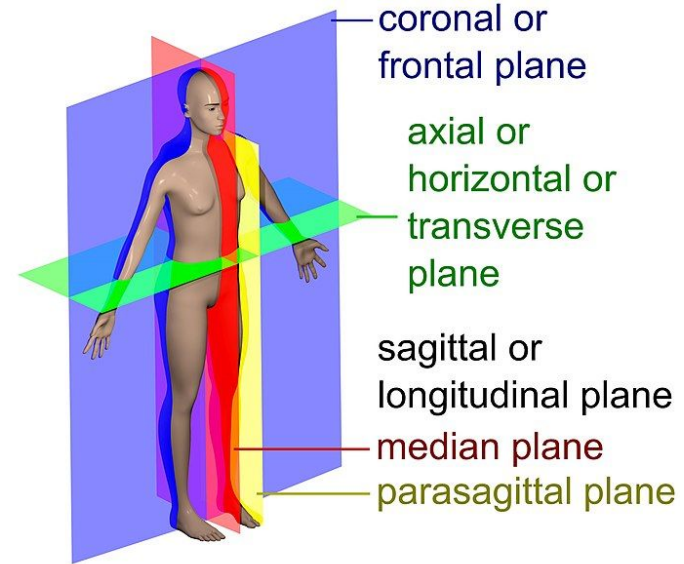
CC BY SA 4.0: Ptrump16 via [Wikimedia Commons](#)

# Magnetic Resonance Imaging (MRI)

Produces beautiful images...



...captured in standard orientations.



# Everything you always wanted to know about MRI

*Certain atomic nuclei are able to absorb radio frequency energy when placed in an external magnetic field; the resultant evolving spin polarization can induce an RF signal in a radio frequency coil and thereby be detected.*

— Hoult, D.I.; Bahkar, B. (1998). "NMR Signal Reception: Virtual Photons and Coherent Spontaneous Emission". *Concepts in Magnetic Resonance*. **9** (5): 277–297.

Hydrogen atoms are naturally abundant in humans, particularly in water and fat.

Pulses of radio waves excite the nuclear spin energy transition, and macroscopic polarization that is detected by antennas.

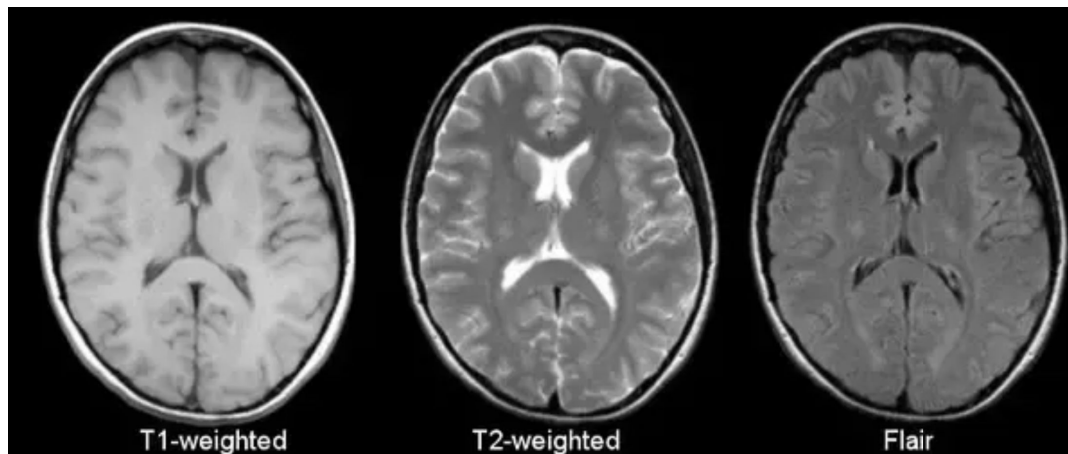
Magnetic field gradients localize the polarization in space. By varying the parameters of the pulse sequence, different contrasts may be generated between tissues based on the relaxation properties of the hydrogen atoms therein.



# What you actually need to know

MRI is a large family of imaging techniques.

They can produce 3D scans of various appearances in order to emphasize some human tissues versus others.



# Types of scan we will consider

1. **T1** – *T1-weighted MRI*: image contrast is based predominantly on the T1 (longitudinal) relaxation time of tissue; tissue with short T1 relaxation time appears brighter (hyperintense).
2. **T2** – *T2-weighted MRI*: image contrast is based predominantly on the T2 (transverse) relaxation time of tissue; tissue with long T2 relaxation time appears brighter (hyperintense).
3. **T1C** – *T1-weighted MRI after administration of contrast media*: many tumors show signal enhancement after administration of contrast agent.
4. **FLAIR** – *Fluid-attenuated inversion-recovery MRI*: bright signal of the CSF (cerebrospinal fluid) is suppressed which allows a better detection of small hyperintense lesions.

*There are dozens of other scan types.*

# BraTS: Brain Tumor Segmentation Competition

# Original segmentation task

Given a 3D scan (skull-stripped, registered) of a patient with T1, T2, T1C and FLAIR modalities, predict a tumor class for each voxel (the patient suffers from a glioma): edema (yellow), non-enhancing solid core (red), necrotic/cystic core (green), enhancing core (blue).

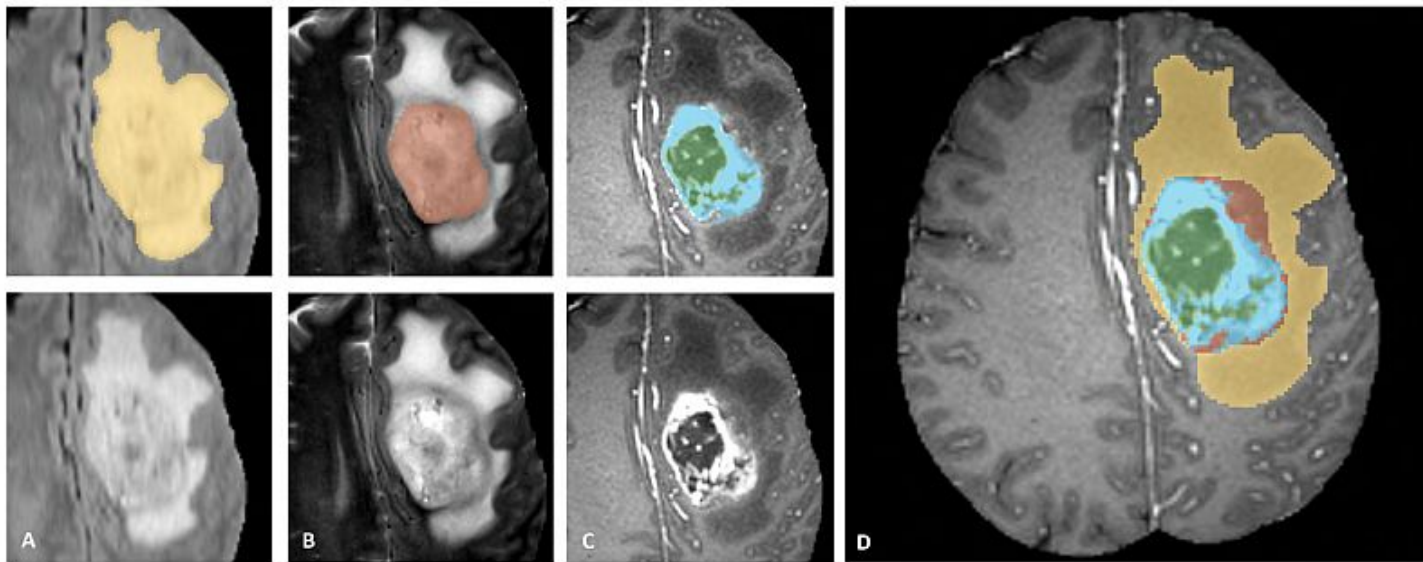


Figure taken from the [BraTS IEEE TMI paper](#).

# Original dataset

The 2018 competition we use the data from originally contains **285 brain scans**.

Each of them

- was acquired by a different patient, potentially by a different team with a different device (hence the intensity values vary a lot)
- has 4 modalities of unequal quality
- have a shape of  $240 \times 240 \times 155$  voxels
- has a manually-annotated ground truth with 4 classes

# Your Mission

# A simplified competition

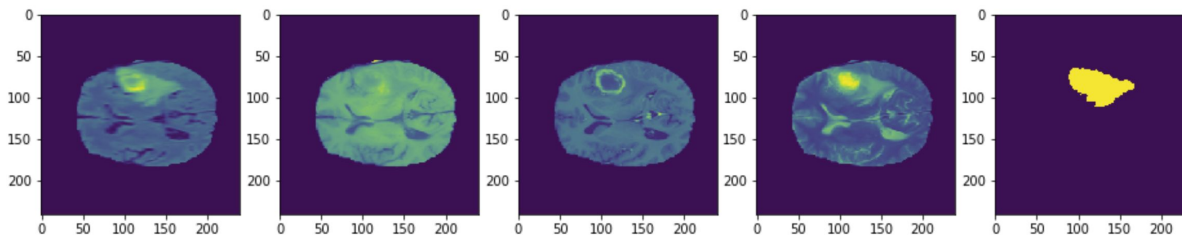
Because dealing with 3D and data normalization would take you much time and pain, we:

1. already performed **data normalization**
2. extracted **2D (axial) slices** that you have to process

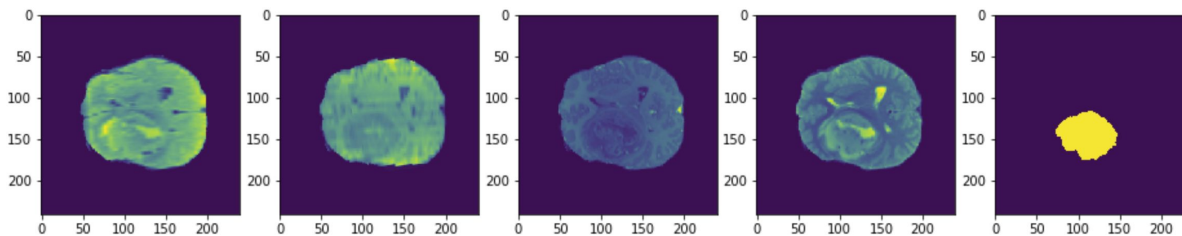
# Actual task

Given a **240\*240 image with 4 modalities** (already normalized),  
predict **for each pixel** whether it belongs to a **tumor or nor**.

Sample 1



Sample 2



Inputs (modalities)

Expected output



# Actual dataset

## Train set

- 256 normalized slices, one per patient, containing 240x240 images with 4 channels (1 for each modality), float32
- 256 target segmentations, one per patient, containing 240x240 images with 1 channel (indicating tumor or clean region), uint8

## Test set

- 29 normalized slices, one per patient (not in the training set), containing 240x240 images with 4 channels (1 for each modality), float32
- *Ground truth kept secret for grading*

# Suggested Pipeline

# Data preprocessing

 **We already did this step.**

For each 3D sample  $i$ , for each modality  $m$ , we computed the mean  $\mu_{im}$  and the variance  $\sigma_{im}$ , then computed the new value for each voxel  $\mathbf{v}'_{im} = (\mathbf{v}_{im} - \mu_{im}) / \sigma_{im}$ .

Then, we extracted the slices we gave you.

*It means you cannot recover the original data because the normalization was done with data (other slices) we do not provide you with.*

# Choose and train a classifier

There are several suggestions in the reference notebook: SVM, neural network, etc.

Input = 1 vector of 4 components for each pixels

Output = 1 for tumor, 0 for “not tumor”

 **Do not use background (“black”) pixels for training, they would ruin your classification.**

 **Deep nets can work but they are harder to train well.**

 **And don't use deep nets, we'll play with them next semester.**

# Validate your training

Create and use a validation set extracted from the full training set.

To not train on the samples it contains.

`sklearn.model_selection.train_test_split` may be your friend.

Check visually results from both train and val sets!

# Interpret your results

```
y_pred = svcclassifier.predict(x_test)
```

```
from sklearn.metrics import classification_report, confusion_matrix  
print(confusion_matrix(y_test,y_pred))  
print(classification_report(y_test,y_pred))
```

```
[[2976   22]  
 [ 273 255]]
```

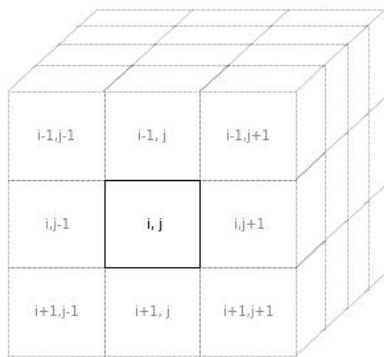
	precision	recall	f1-score	support
0.0	0.92	0.99	0.95	2998
1.0	0.92	0.48	0.63	528
accuracy			0.92	3526
macro avg	0.92	0.74	0.79	3526
weighted avg	0.92	0.92	0.90	3526

# Add some context to each pixel

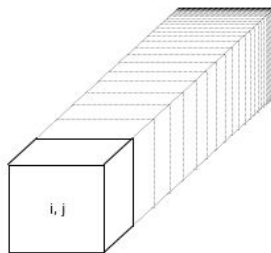
You can get better results by looking at the neighborhood of a pixel to classify it better: train with vectors of size  $N \times M$  instead of  $1 \times M$ .

$N$  = number of pixels in the neighborhood

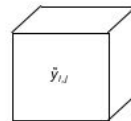
$M$  = number of modalities



From  $x$  seen through  $3 \times 3$  windows with 4 modalities...



... we build  $x_{win}$  where each pixel has  $3 \times 3 \times 4$  features...



... then we predict  $y$  for the pixel at the center of the window.

# Fighting underfitting and overfitting

**You do not have much data to train on.**

If you pick a classifier which is too simple, you may underfit: you will get low and similar scores both on the train and test sets.

Choosing another classifier may be a good idea here.

You may also easily overfit your classifier, especially if you use one with a large capacity: you will get excellent scores on the train set, and bad ones on the test set.

Regularization may be necessary.



# Post processing

We suggest in the notebook to “clean up” the results by removing very small isolated pixels marked as tumor.

You may have many other ideas here.

# Going Further

# Many options

- Data augmentation to increase train set
- Larger / better neighborhood for each pixel
- Better ANN structure than the one suggested in the notebook
- Change the representation space? (Fourier, wavelets...)
- As the tumors under consideration may not have “holes”, improve the post-processing
- ~~— Super heavy classifiers (UNet, Gradient Boosted Trees...)~~
- ...