Deep Learning in Radiomics

Wistan Marchadour – Ph.D. Student

2023.01.13

1 – Standard Radiomics

Definition

- **Before**, medical images are only complementary information for clinicians
- « Radiomics » calling appeared in **2010**
- **Extraction of quantitative features** from medical imaging data
- **Statistical information** for modeling in oncology :
    - Tumor / organ caracteristics for diagnosis
    - Clinical outcome (response to therapy, survival)
- **Target** → Get all relevant information we can
- **1st Idea** → «Handcrafted» features

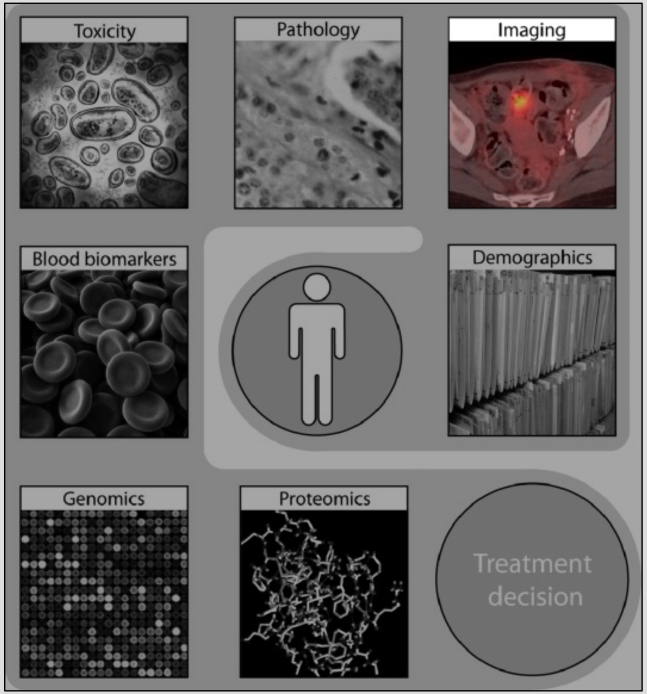
Data types used in Diagnostic & Prognosis

Lambin, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012
Gillies, et al. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016

1 – Standard Radiomics

Definition

- **Before**, medical images are only complementary information for clinicians
- « Radiomics » calling appeared in **2010**
- **Extraction of quantitative features** from medical imaging data
- **Statistical information** for modeling in oncology :
  - Tumor / organ caracteristics for diagnosis
  - Clinical outcome (response to therapy, survival)
- **Target** → Get all relevant information we can
- **1st Idea** → «Handcrafted» features



Data types used in Diagnostic & Prognosis

Lambin, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012
Gillies, et al. Radiomics: Images Are More than Pictures, They Are Data. Radiology 2016

1 – Standard Radiomics

Definition

Tumours are **heterogeneous entities** (genetic, cellular, tissular)

- **Hypothesis 1** : information in Images (macro scale) reflect at least some caracteristics in Smaller Scales
- **Hypothesis 2** : Images contain more information « than meets the eye »

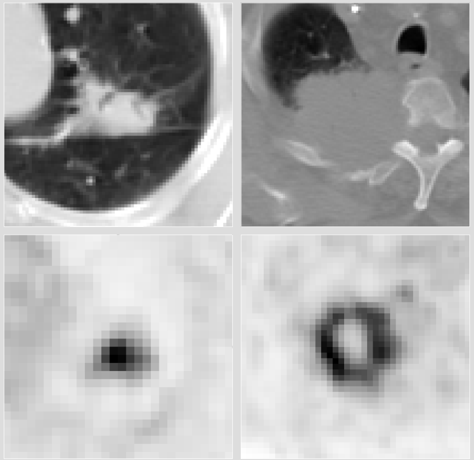
Gerlinger, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012
Leijenaar. Radiomics: Images are more than meets the eye. Thesis, University of Maastricht

1 – Standard Radiomics

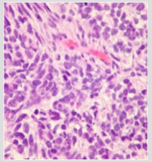
Definition

Tumours are **heterogeneous entities** (genetic, cellular, tissular)

- **Hypothesis 1** : information in Images (macro scale) reflect at least some caracteristics in Smaller Scales
- **Hypothesis 2** : Images contain more information « than meets the eye »

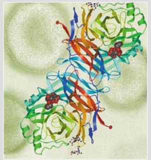


CT (top) and PET (bottom) scans of Tumors

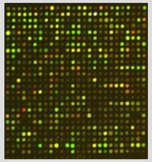


Histology

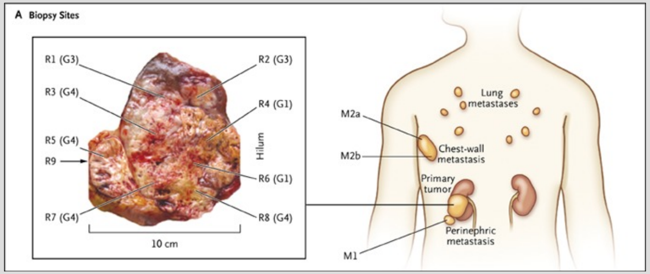
Biomarkers



Proteomics



Genomics



Non-imaging representation of Tumors

Gerlinger, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 2012

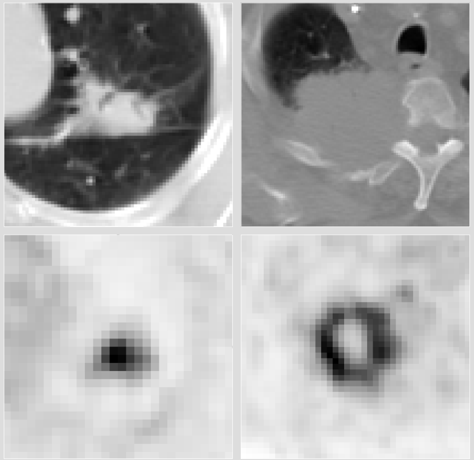
Leijenaar. Radiomics: Images are more than meets the eye. Thesis, University of Maastricht

1 – Standard Radiomics

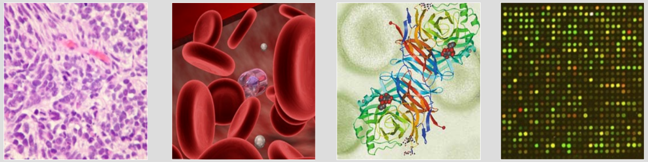
Definition

Tumours are **heterogeneous entities** (genetic, cellular, tissular)

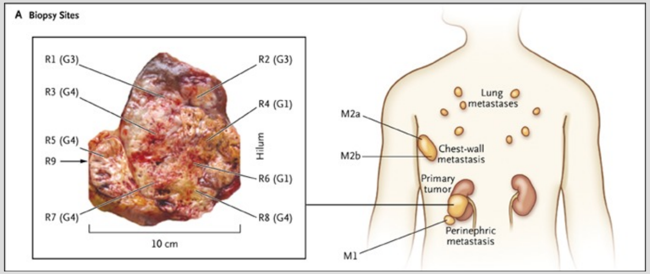
- **Hypothesis 1** : information in Images (macro scale) reflect at least some caracteristics in Smaller Scales
- **Hypothesis 2** : Images contain more information « than meets the eye »



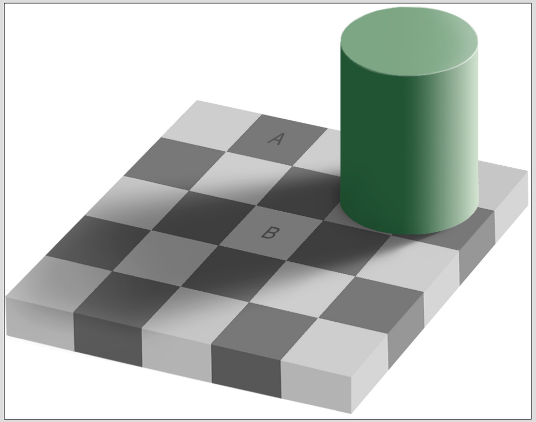
CT (top) and PET (bottom) scans of Tumors



Histology *Biomarkers* *Proteomics* *Genomics*



Non-imaging representation of Tumors

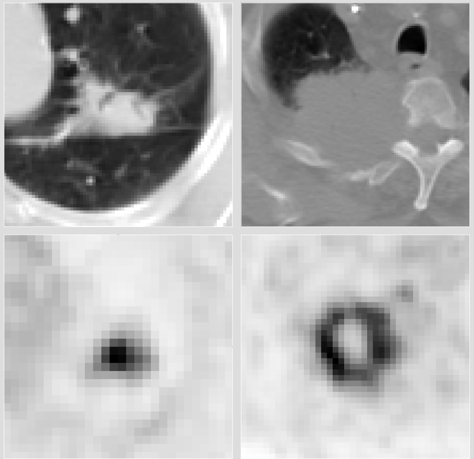


1 – Standard Radiomics

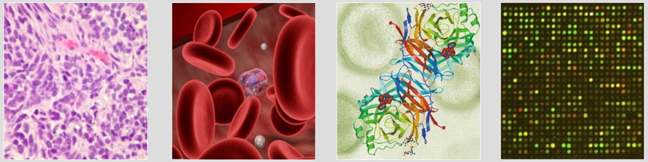
Definition

Tumours are **heterogeneous entities** (genetic, cellular, tissular)

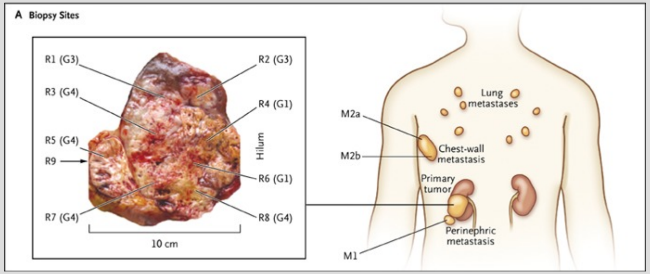
- **Hypothesis 1** : information in Images (macro scale) reflect at least some caracteristics in Smaller Scales
- **Hypothesis 2** : Images contain more information « than meets the eye »



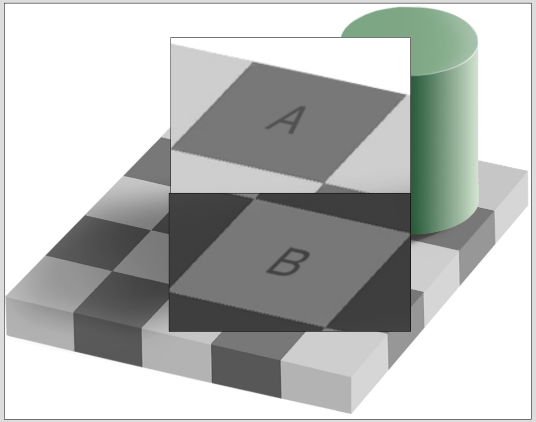
CT (top) and PET (bottom) scans of Tumors



Histology *Biomarkers* *Proteomics* *Genomics*



Non-imaging representation of Tumors



Statistical analysis extract more information than visual assessment

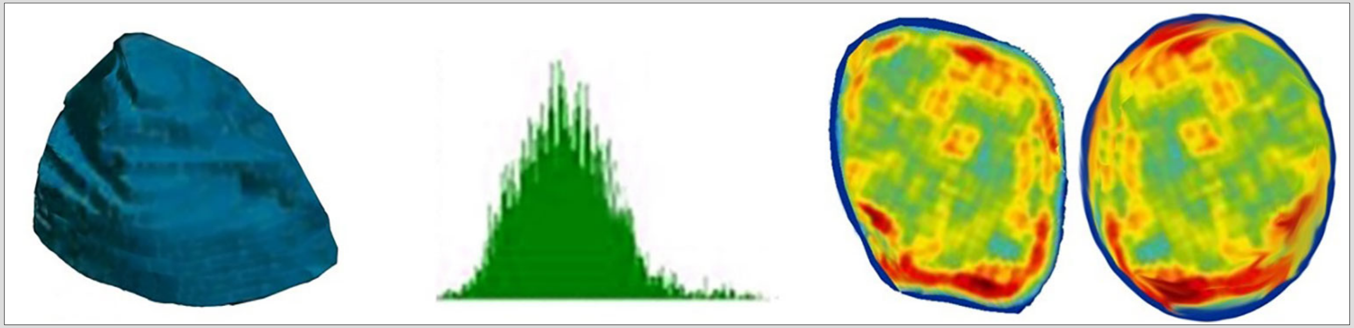
1 – Standard Radiomics

Definition : Features

«Handcrafted / engineered» features, i.e., designed by human experts decades ago

Most often used :

- Shape descriptors
- Intensity histogram / statistics
- 2nd or higher order textures



Sphericity of the Tumor *Histogram analysis* *Co-occurence entropy*

1 – Standard Radiomics

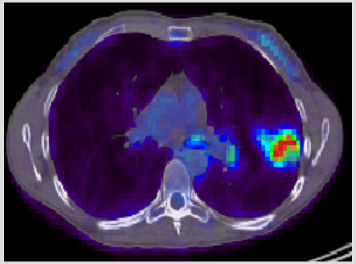
Workflow Process

Study Design
(application, data…)

1 – Standard Radiomics

Workflow Process

Study Design
(application, data…)

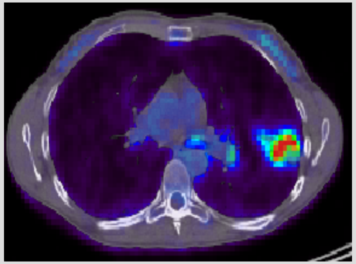


Images
Acquisition / Collection

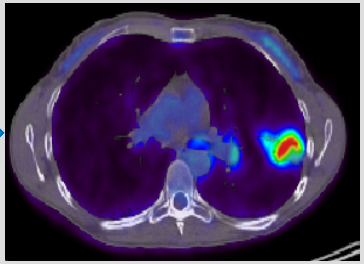
1 – Standard Radiomics

Workflow Process

Study Design
(application, data…)



Images
Acquisition / Collection

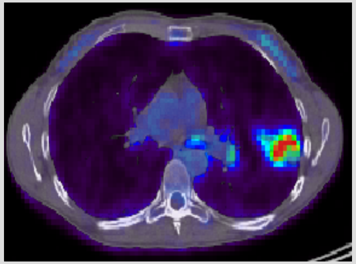


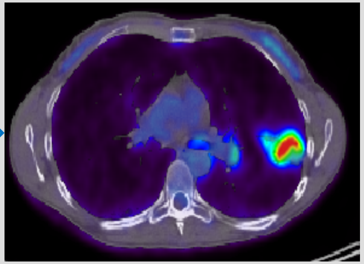
Pre-processing
(denoising, registration…)

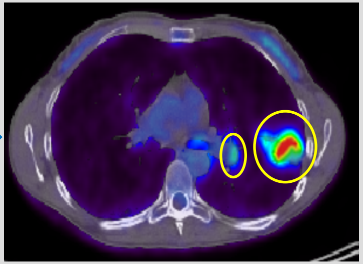
1 – Standard Radiomics

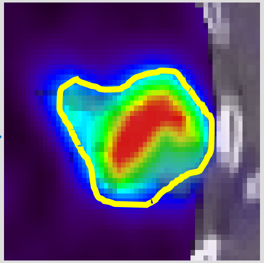
Workflow Process

Study Design
(application, data…)









Images
Acquisition / Collection

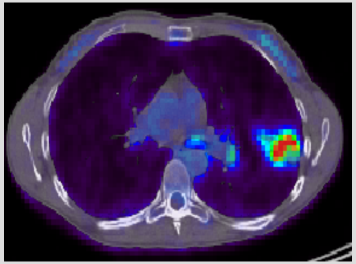
Pre-processing
(denoising, registration…)

Object(s) Detection / Segmentation (ex. tumors)

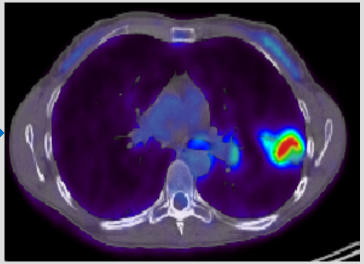
1 – Standard Radiomics

Workflow Process

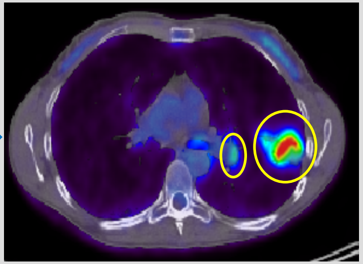
Study Design
(application, data…)

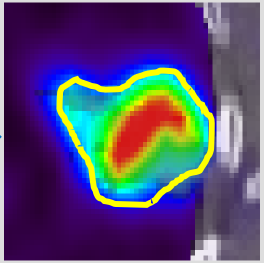


Images
Acquisition / Collection

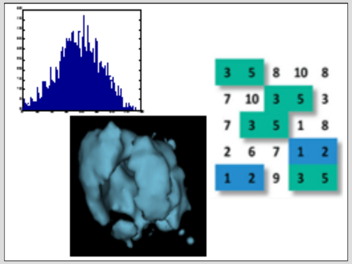


Pre-processing
(denoising, registration…)





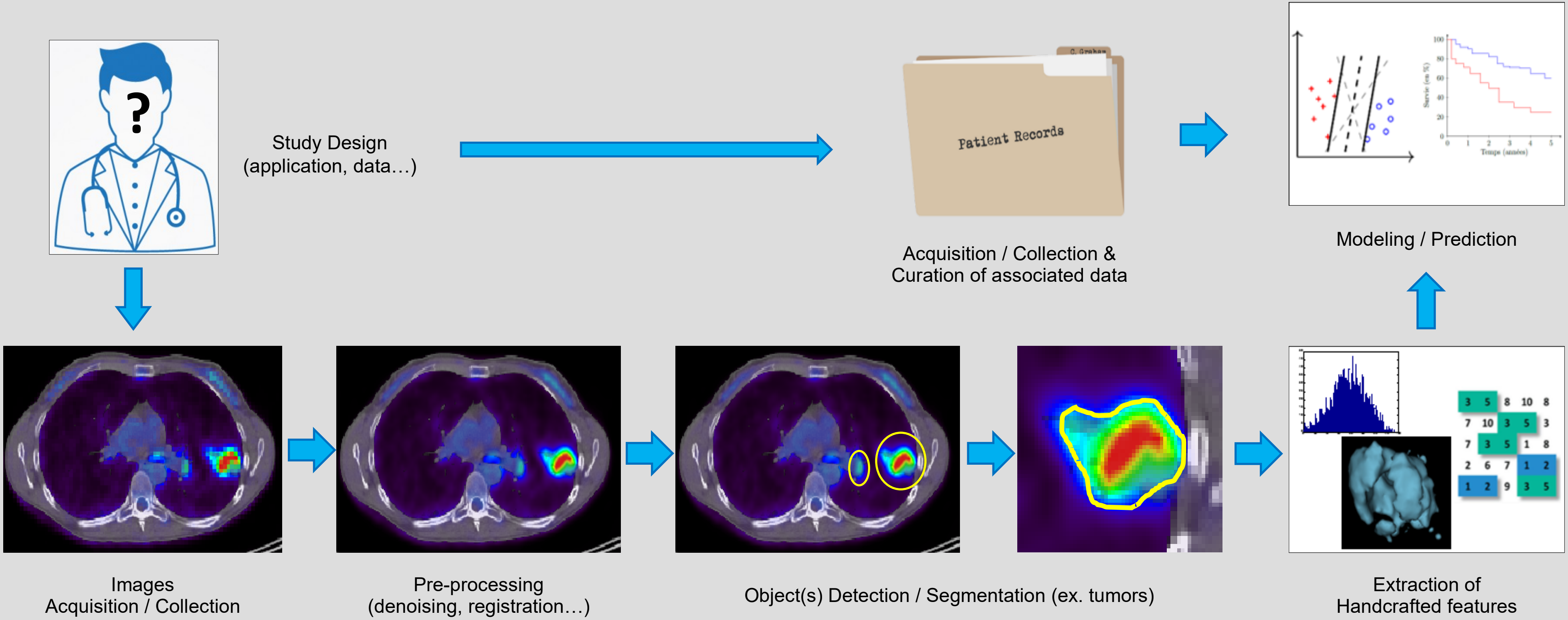
Object(s) Detection / Segmentation (ex. tumors)



Extraction of
Handcrafted features

1 – Standard Radiomics

Workflow Process



Hatt, et al. Data are also images. J Nucl Med 2019

1 – Standard Radiomics

Limitations & Objectives

1st **Reference** to Radiomics in **2010**, yet **no transfer to clinical practice :**

- Lack of standards (hardly reproducible)
- No automation of process (cannot inspect large datasets manually)
- Too much heterogeneity of data between centers (no harmonization)
- Complex modeling (many available algorithms, with too much hyper-parameters)
- Trust & acceptance issues (Explainability / Interpretability)

Ideal Radiomics Process :

- Can support medical images from any center (Robust)
- Fully Automatized
- Standardized Radiomics features (Reusable)
- Has a strong Clinical Value (comparison / evaluation)
- Is Interpretable for end-users (clinicians trust)

→ **Why not use AI ?**

Hatt, et al. Data are also images. J Nucl Med 2019
Papadimitroulas, et al. Artificial intelligence: Deep learning in oncological radiomics and challenges of interpretability and data harmonization. Phys Med. 2021

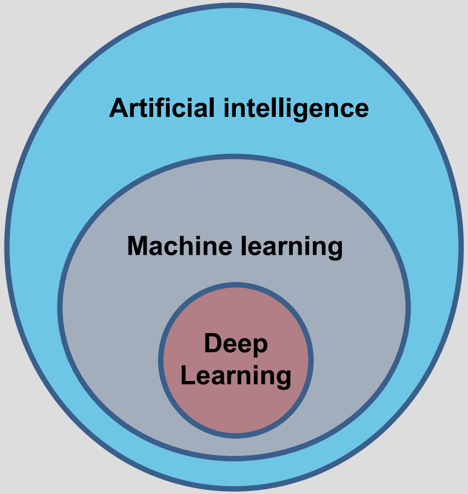
"… a system's ability to interpret external data correctly, to learn from such data,
and to use those learnings to achieve specific goals and tasks through flexible adaptation."

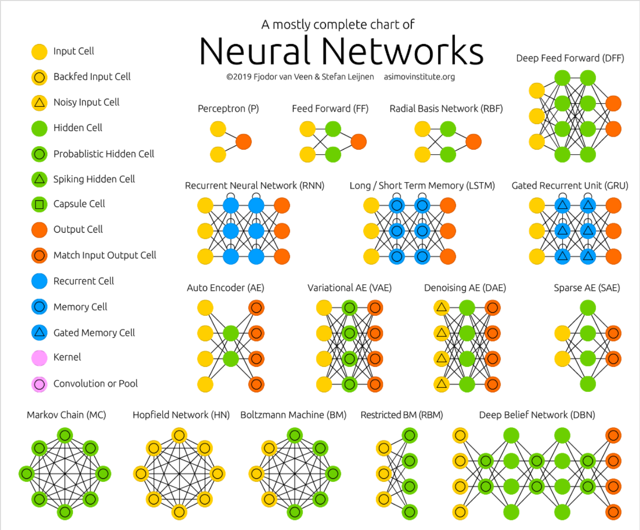
→ **Complex / Adaptive / Reasoning** → **Intelligence**

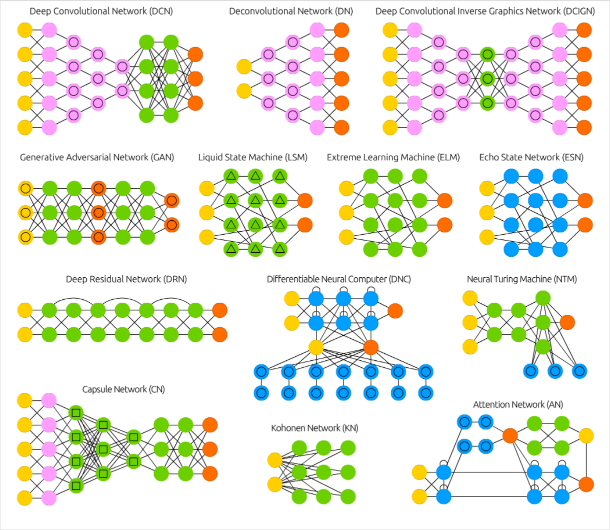
	Expert Systems	Analytical AI	Human-Inspired AI	Humanized AI	Human Beings
Cognitive Intelligence	×	✓	✓	✓	✓
Emotional Intelligence	×	×	✓	✓	✓
Social Intelligence	×	×	×	✓	✓
Artistic Creativity	×	×	×	×	✓
		Supervised Learning, Unsupervised Learning, Reinforcement Learning			

Kaplan & Haenlein. Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. Bus Horiz. 2019

- Use **existing (un)labeled data** to train/learn a model (all knowledge is memorized)
- Apply the model to new data (**inference**)
- Relies on **advanced analysis & statistical** methods







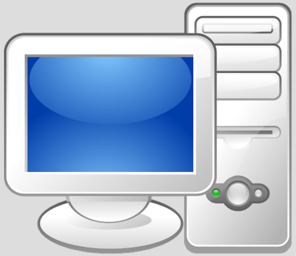
3 – Deep Learning in Radiomics

Advantages / Drawbacks of DL

Calculator

Input data

Designed program



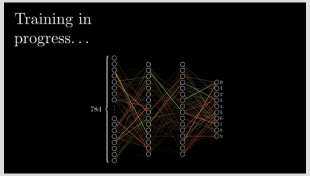
Numeric results

Handcrafted / engineered features

Calculator

Input data

Known outputs (labels)



Trained Program

Features learned from the data

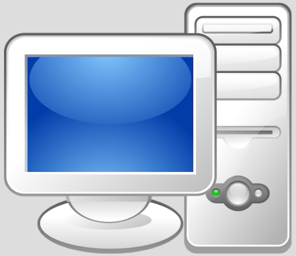
3 – Deep Learning in Radiomics

Advantages / Drawbacks of DL

Calculator

Input data

Designed program



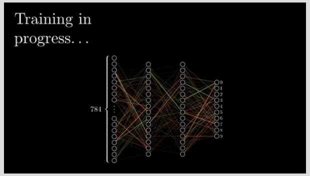
Numeric results

Handcrafted / engineered features

Calculator

Input data

Known outputs (labels)



Trained Program

Features learned from the data

High	**Human Expertise**	Moderate
Low	**Data Amount**	High
Moderate	**Interpretability**	Low

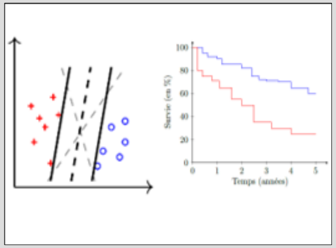
3 – Deep Learning in Radiomics

Data Harmonization

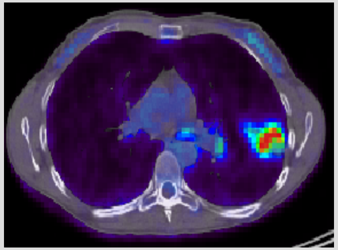
Robustness Objective → framework needs to be portable on any data (without reajusting)

Study Design
(application, data…)

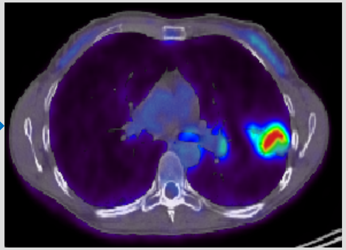
Acquisition / Collection & Curation of associated data



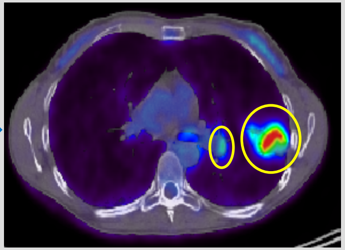
Modeling / Prediction

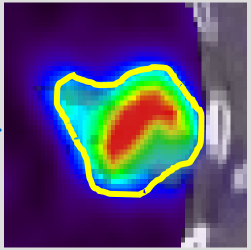


Images
Acquisition / Collection

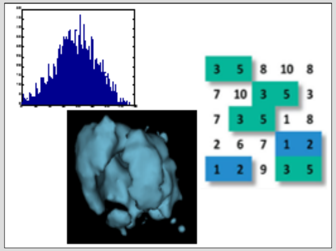


Pre-processing
(denoising, registration…)





Object(s) Detection / Segmentation (ex. tumors)



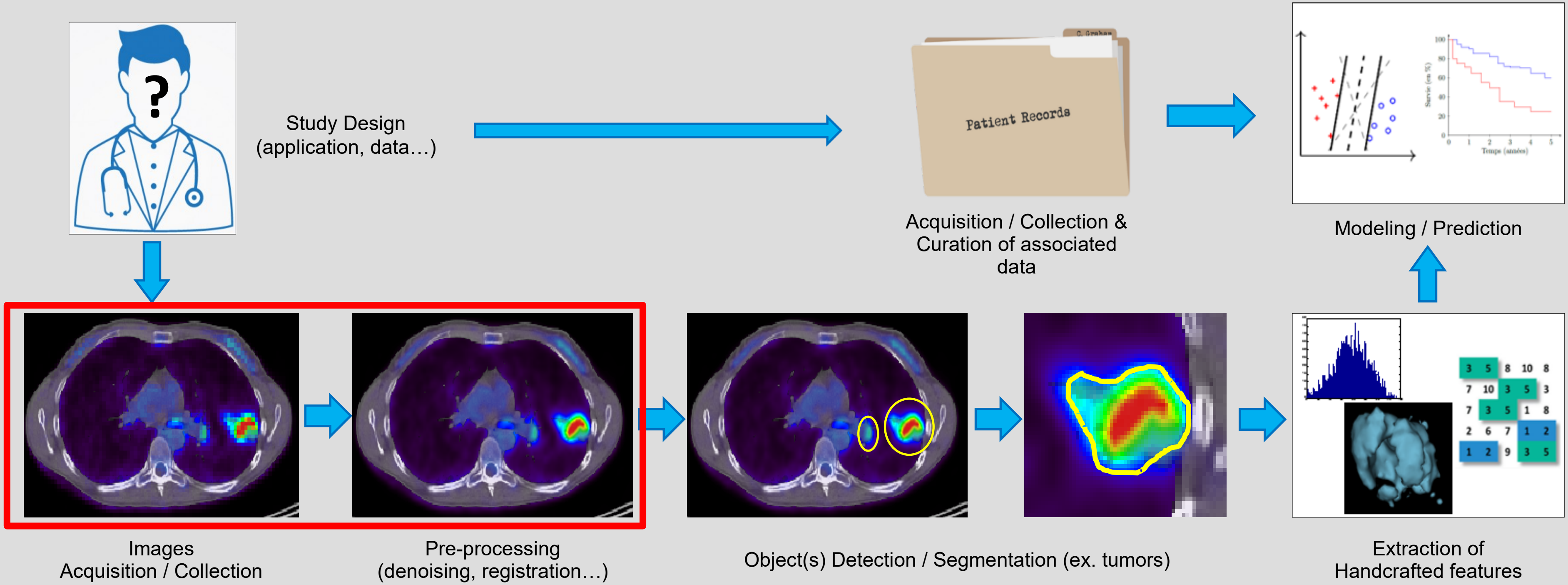
Extraction of
Handcrafted features

Hatt, et al. Data are also images. J Nucl Med 2019

3 – Deep Learning in Radiomics

Data Harmonization

Robustness Objective → framework needs to be portable on any data (without reajusting)



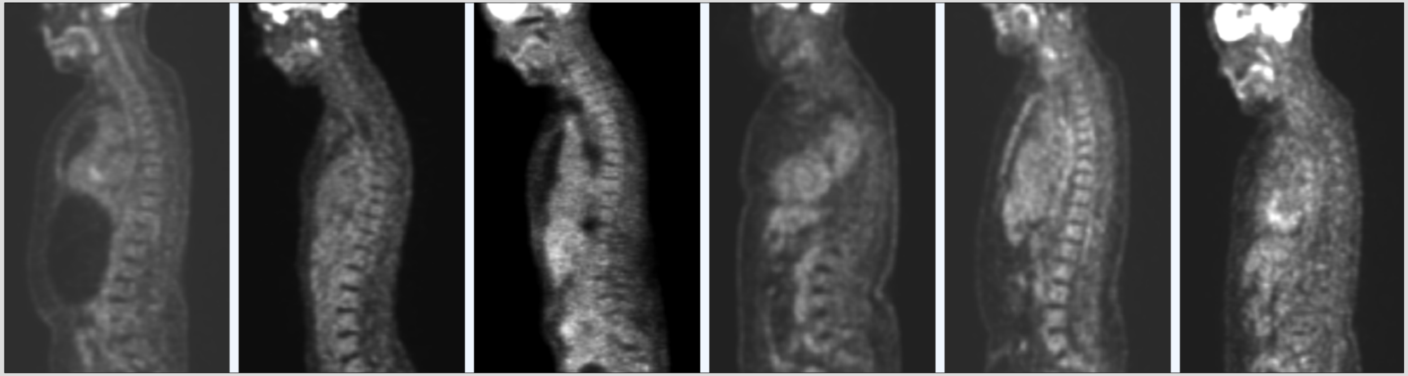
Hatt, et al. Data are also images. J Nucl Med 2019

3 – Deep Learning in Radiomics

Data Harmonization

Many hardware brands & products for medical imaging
Different acquisition parameters for different imaging targets

→ Frameworks cannot adapt directly to this heterogeneity



1 2

3 – Deep Learning in Radiomics

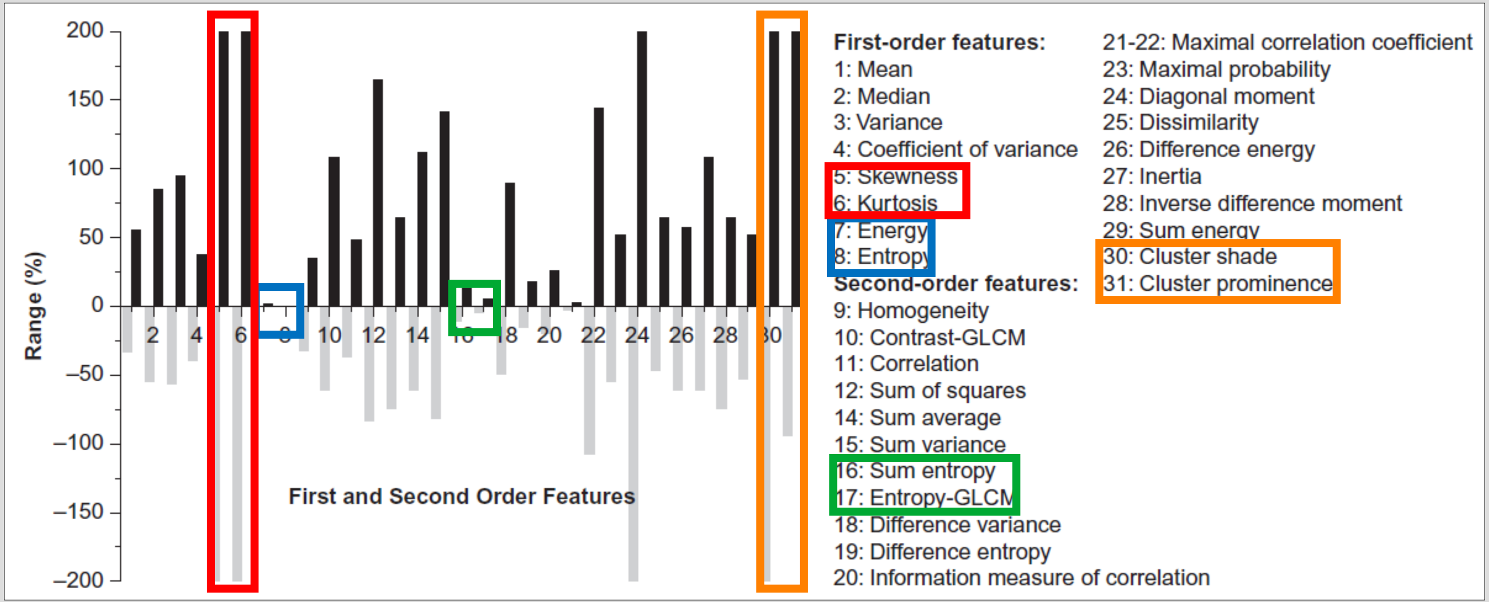
Data Harmonization

First-order features:
1: Mean
2: Median
3: Variance
4: Coefficient of variance
5: Skewness
6: Kurtosis
7: Energy
8: Entropy

Second-order features:
9: Homogeneity
10: Contrast-GLCM
11: Correlation
12: Sum of squares
14: Sum average
15: Sum variance
16: Sum entropy
17: Entropy-GLCM
18: Difference variance
19: Difference entropy
20: Information measure of correlation
21-22: Maximal correlation coefficient
23: Maximal probability
24: Diagonal moment
25: Dissimilarity
26: Difference energy
27: Inertia
28: Inverse difference moment
29: Sum energy
30: Cluster shade
31: Cluster prominence

3 – Deep Learning in Radiomics

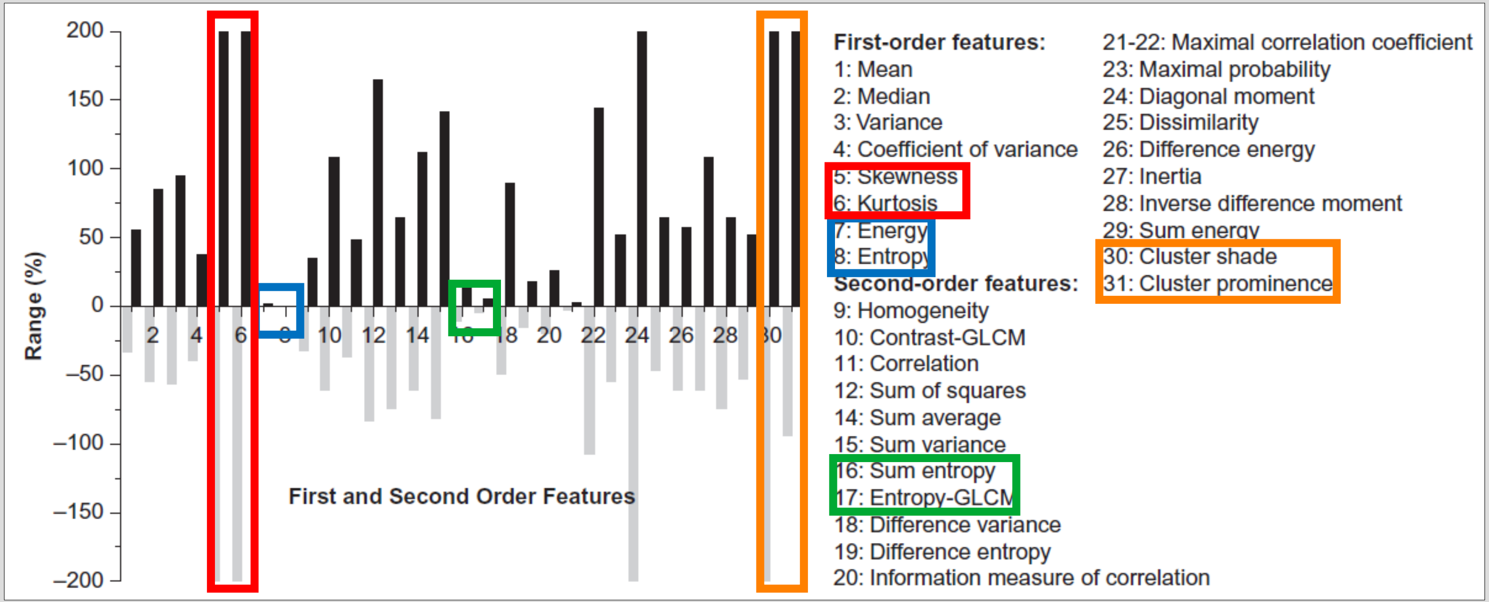
Data Harmonization



Galavis, et al. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. Acta Oncol 2010

3 – Deep Learning in Radiomics

Data Harmonization

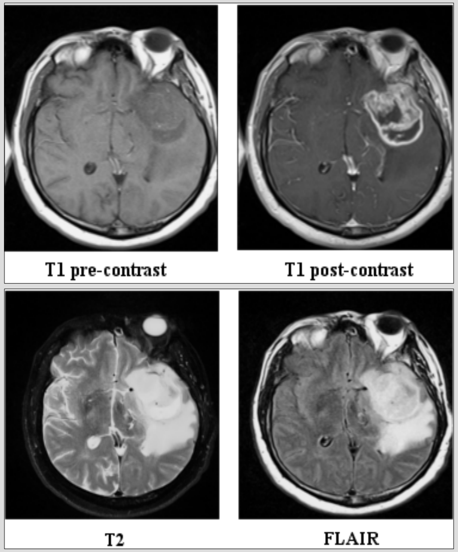


Data (images or features) must be modified before doing multicentric studies

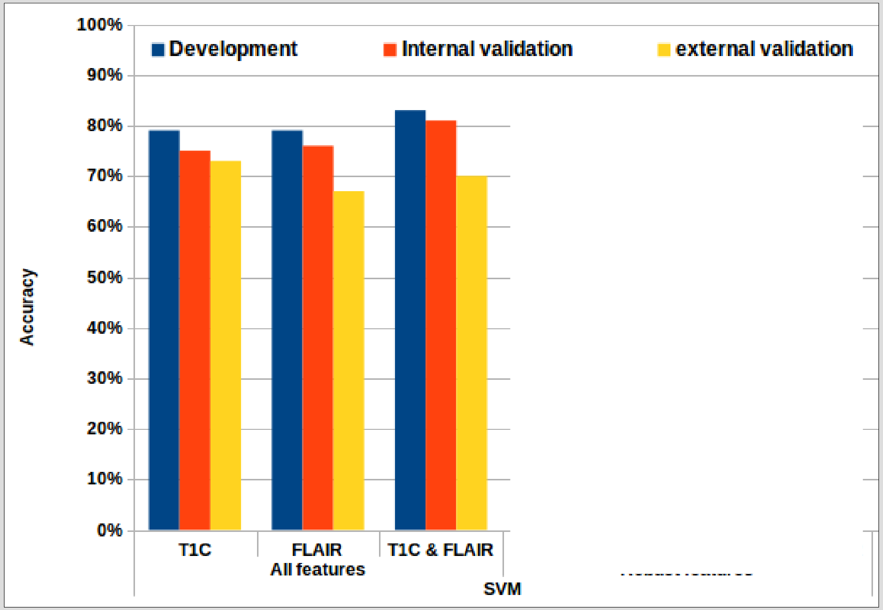
Galavis, et al. Variability of textural features in FDG PET images due to different acquisition modes and reconstruction parameters. Acta Oncol 2010

1st idea : Selection of robust features

→ Problem of choosing the most robust features (human choice)



Glioblastoma Multiforme
Prediction of overall survival

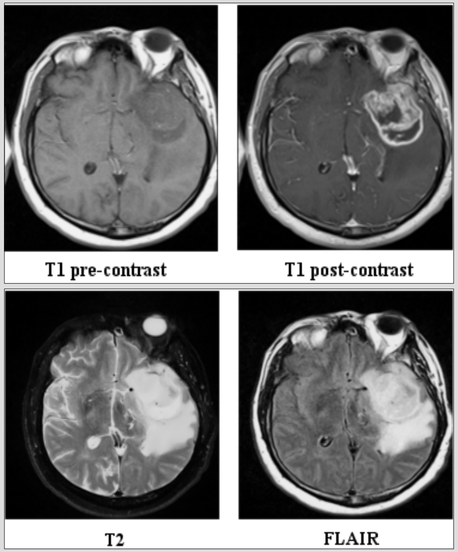


3 – Deep Learning in Radiomics

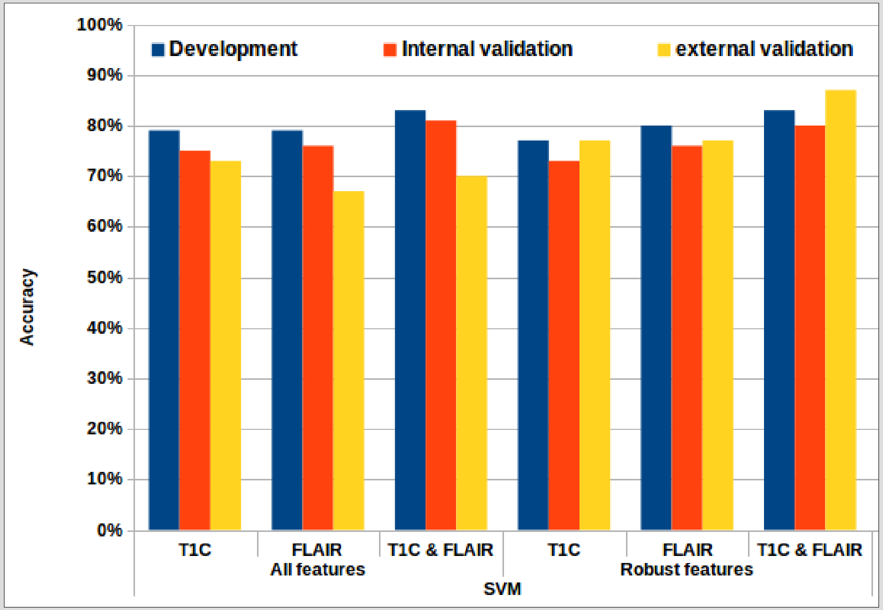
Data Harmonization

1st idea : Selection of robust features

→ Problem of choosing the most robust features (human choice)



Glioblastoma Multiforme
Prediction of overall survival



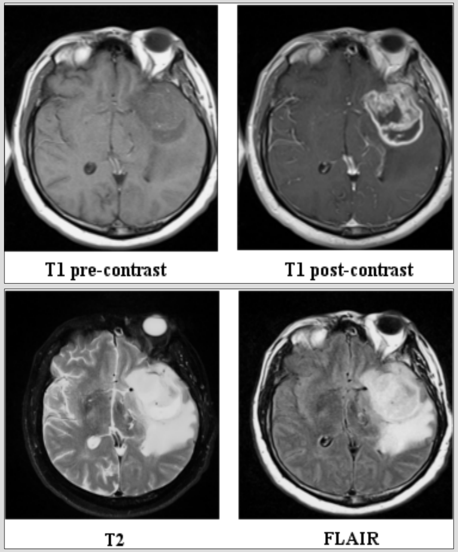
Upadhaya, et al. Machine Learning with Embedded Radiomics Feature Selection Methods: Study on Prognostication of Glioblastoma Multiforme using Multimodal MRI. Submitted 2020

3 – Deep Learning in Radiomics

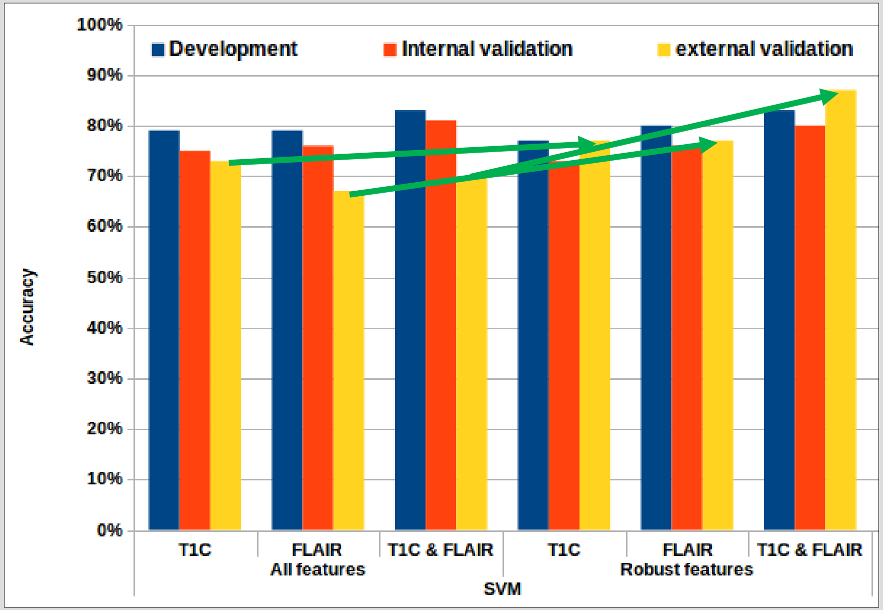
Data Harmonization

1st idea : Selection of robust features

→ Problem of choosing the most robust features (human choice)



Glioblastoma Multiforme
Prediction of overall survival



Upadhaya, et al. Machine Learning with Embedded Radiomics Feature Selection Methods: Study on Prognostication of Glioblastoma Multiforme using Multimodal MRI. Submitted 2020

3 – Deep Learning in Radiomics

Data Harmonization : ComBat

2nd idea : ComBat Harmonization

- Widely used since 2017
- Applied on extracted features for MRI / CT / PET
- Better performances observed for most studies

$$y_{ij} = \alpha + \gamma_i + \delta_i \varepsilon_{ij}$$

Assumption for ComBat

Johnson, et al. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007

3 – Deep Learning in Radiomics

Data Harmonization : ComBat

2nd idea : <u>ComBat Harmonization</u>

- Widely used since 2017
- Applied on extracted features for MRI / CT / PET
- Better performances observed for most studies

$$y_{ij} = \alpha + \gamma_i + \delta_i \varepsilon_{ij}$$

Feature y measured
in scanner i
for volume j

Assumption for ComBat

2nd idea : <u>ComBat Harmonization</u>

- Widely used since 2017
- Applied on extracted features for MRI / CT / PET
- Better performances observed for most studies

Average value of y

$$y_{ij} = \alpha + \gamma_i + \delta_i \varepsilon_{ij}$$

Feature y measured
in scanner i
for volume j

Assumption for ComBat

Johnson, et al. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007

3 – Deep Learning in Radiomics

Data Harmonization : ComBat

2nd idea : ComBat Harmonization

- Widely used since 2017
- Applied on extracted features for MRI / CT / PET
- Better performances observed for most studies

Average value of y

$$y_{ij} = \alpha + \gamma_i + \delta_i \varepsilon_{ij}$$

Feature y measured in scanner i for volume j

Additive effect of scanner i

Assumption for ComBat

Johnson, et al. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007

3 – Deep Learning in Radiomics

Data Harmonization : ComBat

2nd idea : <u>ComBat Harmonization</u>

- Widely used since 2017
- Applied on extracted features for MRI / CT / PET
- Better performances observed for most studies

Average value of y

Multiplicative effect
of scanner i

$$y_{ij} = \alpha + \gamma_i + \delta_i \varepsilon_{ij}$$

Feature y measured
in scanner i
for volume j

Additive effect
of scanner i

Assumption for ComBat

3 – Deep Learning in Radiomics

Data Harmonization : ComBat

2nd idea : ComBat Harmonization

- Widely used since 2017
- Applied on extracted features for MRI / CT / PET
- Better performances observed for most studies

Average value of y

Multiplicative effect of scanner i

$$y_{ij} = \alpha + \gamma_i + \delta_i \varepsilon_{ij}$$

Feature y measured in scanner i for volume j

Additive effect of scanner i

Error term

Assumption for ComBat

Johnson, et al. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007

3 – Deep Learning in Radiomics

Data Harmonization : ComBat

2nd idea : <u>ComBat Harmonization</u>

- Widely used since 2017
- Applied on extracted features for MRI / CT / PET
- Better performances observed for most studies

Average value of y

Multiplicative effect of scanner i

$$y_{ij} = \alpha + \gamma_i + \delta_i \varepsilon_{ij}$$

Feature y measured in scanner i for volume j

Additive effect of scanner i

Error term

Assumption for ComBat

Multiple Measurements

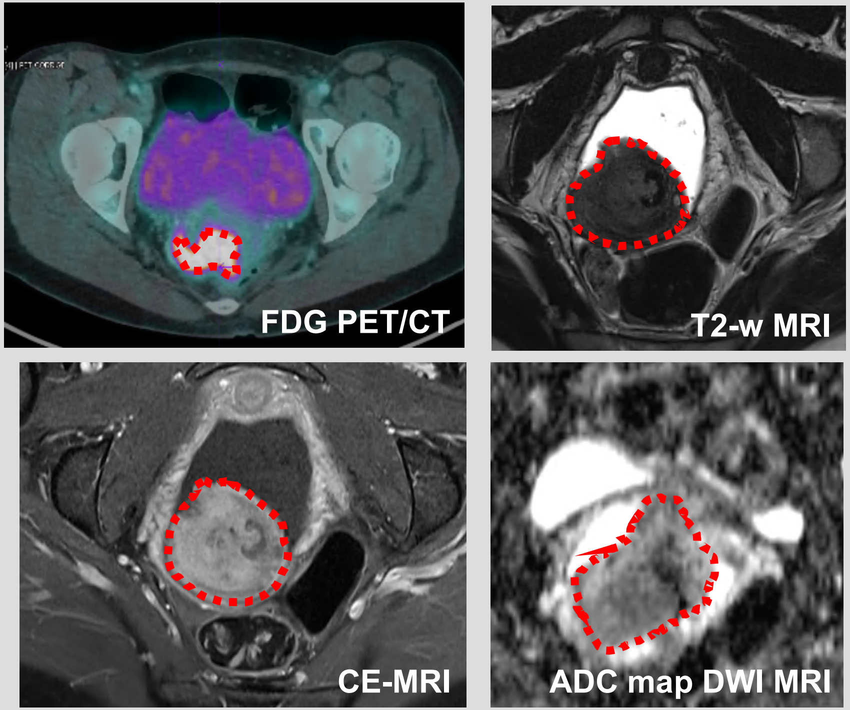
$$y_{ij}^{ComBat} = \frac{y_{ij} - \hat{\alpha} - X_{ij}\hat{\beta} - \gamma_i^*}{\delta_i^*} + \hat{\alpha} + X_{ij}\hat{\beta}$$

Correction of External Influence

Johnson, et al. Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics. 2007

3 – Deep Learning in Radiomics

Data Harmonization : ComBat

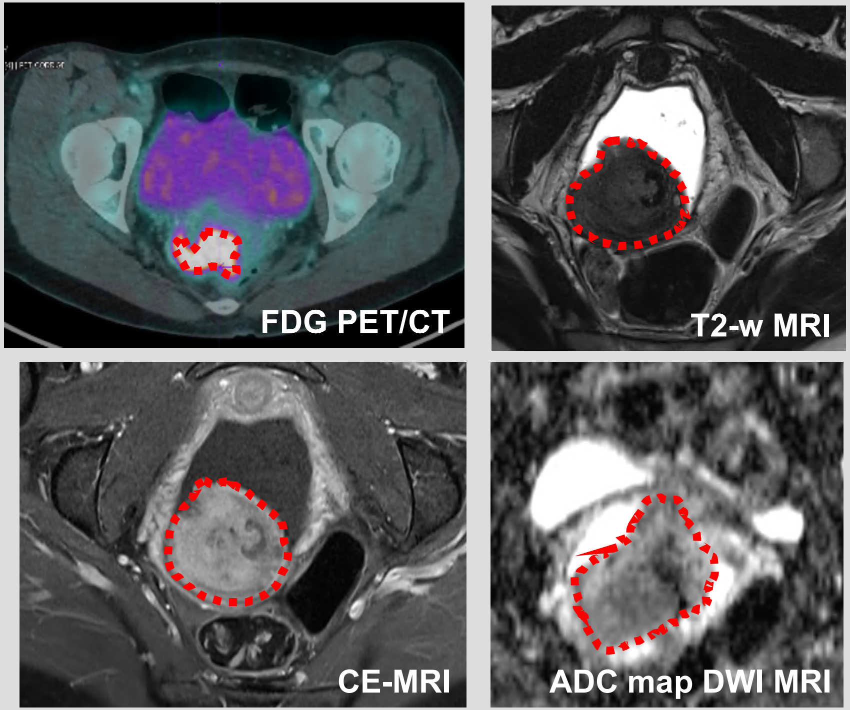


Total : 197 patients from 3 centers
119 (Brest, France)
50 (Nantes, France)
28 (McGill, Canada)

Lucia, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. Eur J Nucl Med Mol Imaging 2019

3 – Deep Learning in Radiomics

Data Harmonization : ComBat



Total : 197 patients from 3 centers
119 (Brest, France)
50 (Nantes, France)
28 (McGill, Canada)

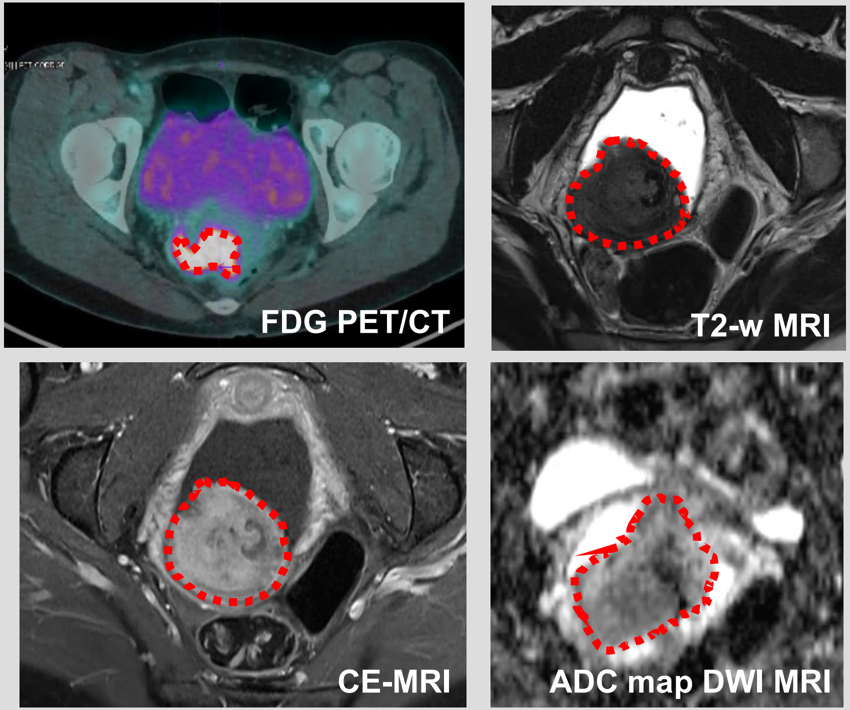
Training / internal validation
119 (Brest, France)
70 / 49 split

External validation
50 (Nantes, France)
28 (McGill, Canada)

Lucia, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. Eur J Nucl Med Mol Imaging 2019

3 – Deep Learning in Radiomics

Data Harmonization : ComBat



Total : 197 patients from 3 centers
119 (Brest, France)
50 (Nantes, France)
28 (McGill, Canada)

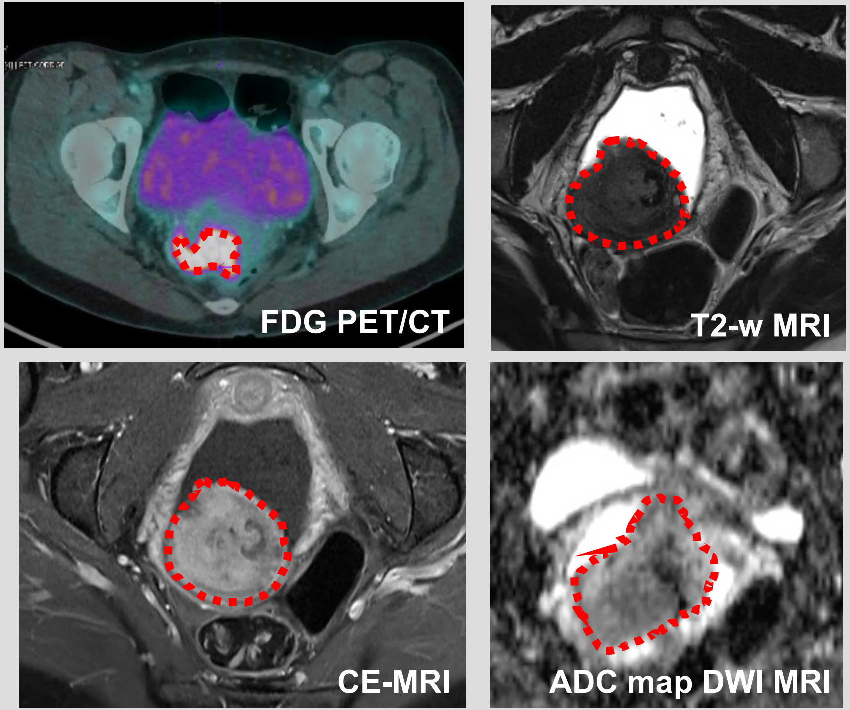
Training / internal validation
119 (Brest, France)
70 / 49 split

External validation
50 (Nantes, France)
28 (McGill, Canada)

- Same PET/CT scanners in Brest and Nantes but different protocols

3 – Deep Learning in Radiomics

Data Harmonization : ComBat



Total : 197 patients from 3 centers
119 (Brest, France)
50 (Nantes, France)
28 (McGill, Canada)

Training / internal validation
119 (Brest, France)
70 / 49 split

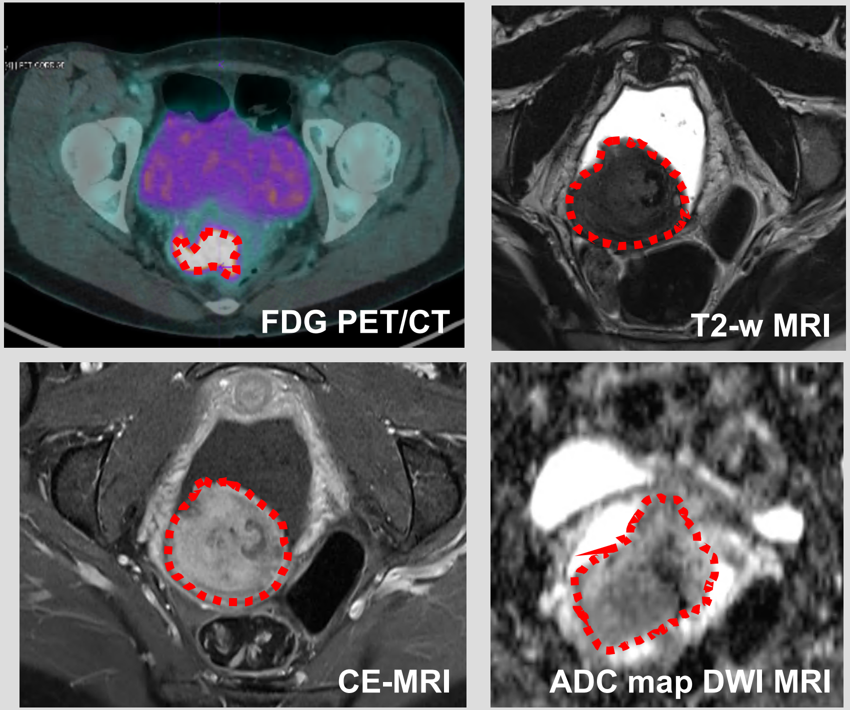
External validation
50 (Nantes, France)
28 (McGill, Canada)

- Same PET/CT scanners in Brest and Nantes but different protocols
- Different PET/CT scanner and protocol in McGill

Lucia, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. Eur J Nucl Med Mol Imaging 2019

3 – Deep Learning in Radiomics

Data Harmonization : ComBat



Total : 197 patients from 3 centers
119 (Brest, France)
50 (Nantes, France)
28 (McGill, Canada)

Training / internal validation
119 (Brest, France)
70 / 49 split

External validation
50 (Nantes, France)
28 (McGill, Canada)

- Same PET/CT scanners in Brest and Nantes but different protocols
- Different PET/CT scanner and protocol in McGill
- Different MRI scanners and protocols in all 3 centers

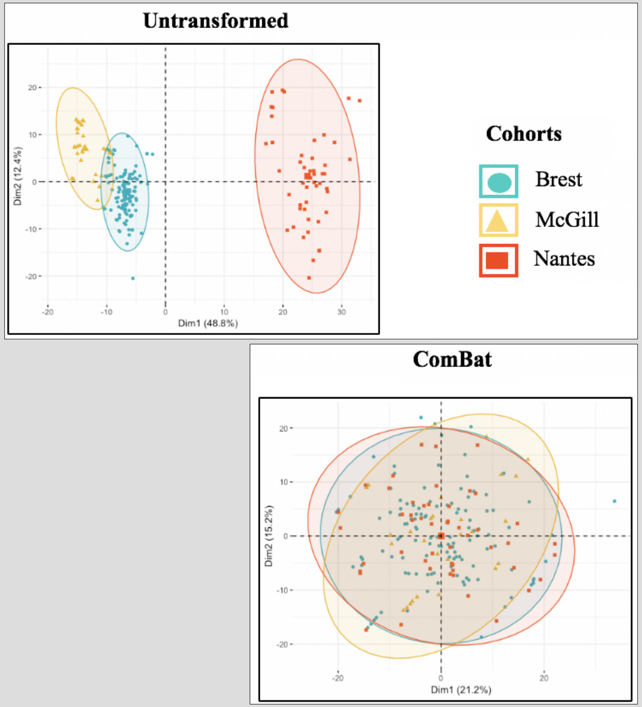
Lucia, et al. External validation of a combined PET and MRI radiomics model for prediction of recurrence in cervical cancer patients treated with chemoradiotherapy. Eur J Nucl Med Mol Imaging 2019

3 – Deep Learning in Radiomics

Data Harmonization : ComBat

(Top) No Harmonization

(Bottom) With ComBat



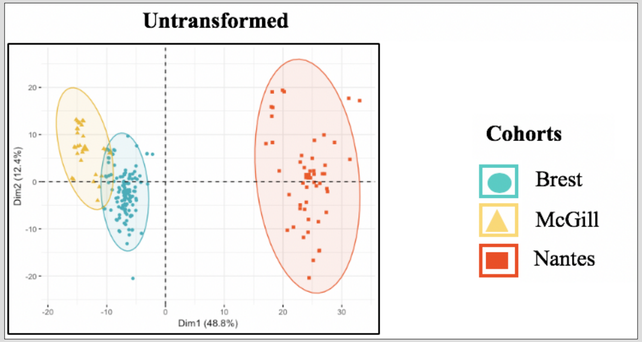
Cervical Cancer Recurrence Prediction on FDG PET + MRI ADC Radiomics

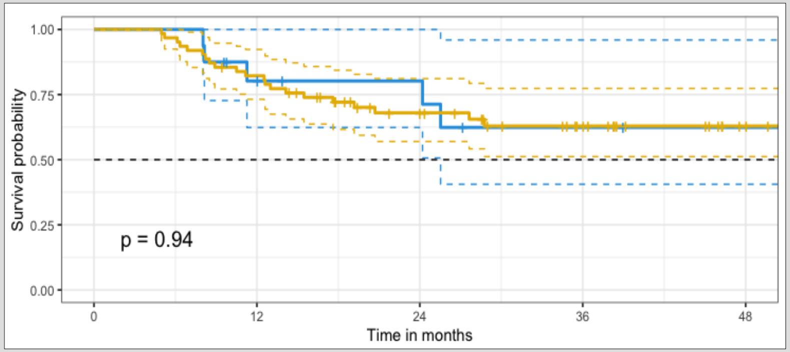
Ronrick, et al. Performance comparison of modified ComBat for harmonization of radiomic features for multicentric studies. Sci Rep 2020

3 – Deep Learning in Radiomics

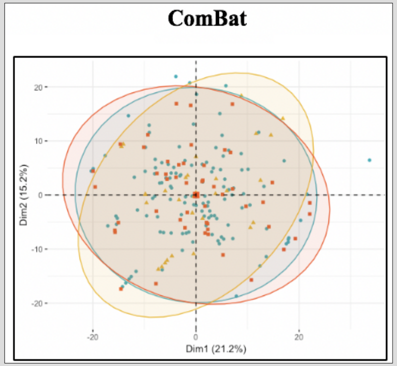
Data Harmonization : ComBat

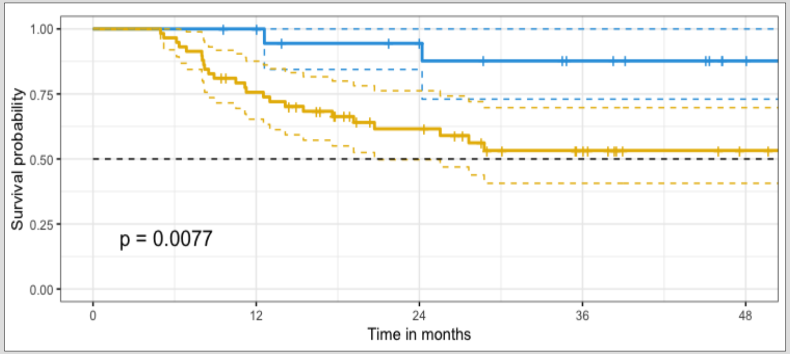
(Top) No Harmonization





(Bottom) With ComBat





Cervical Cancer Recurrence Prediction on FDG PET + MRI ADC Radiomics

Ronrick, et al. Performance comparison of modified ComBat for harmonization of radiomic features for multicentric studies. Sci Rep 2020

One issue remains → based on assumptions

→ Data Harmonization

3rd Idea : <u>Deep Learning</u>

- Before processing and data extraction of the images
- Registration of images to a pre-determined type :
  - Reconstruction Kernel
  - Voxel Size
  - Specific Scanner
  - …

→ Example of Conversion of Reconstruction Kernels for CT

Choe, et al. Deep Learning-based Image Conversion of CT Reconstruction Kernels Improves Radiomics Reproducibility for Pulmonary Nodules or Masses. Radiology 2019

3 – Deep Learning in Radiomics

Data Harmonization : DL

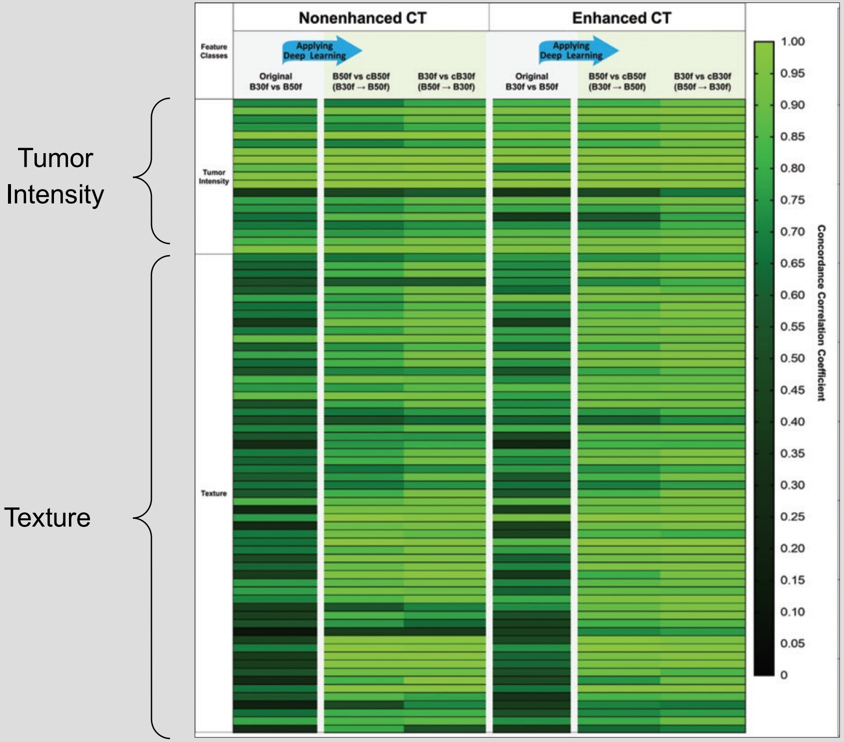
One issue remains → based on assumptions

→ Data Harmonization

3rd Idea : Deep Learning

- Before processing and data extraction of the images
- Registration of images to a pre-determined type :
  - Reconstruction Kernel
  - Voxel Size
  - Specific Scanner
  - …

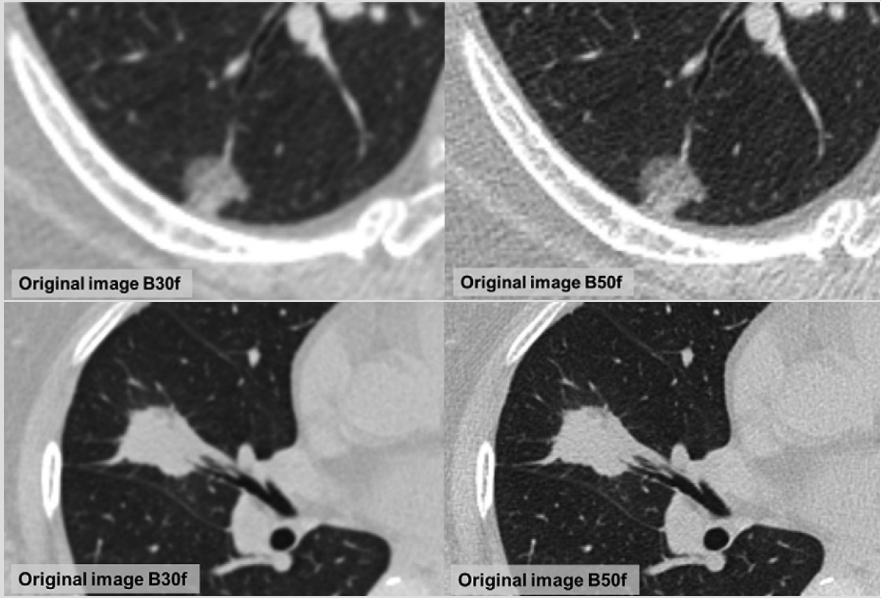
→ Example of Conversion of Reconstruction Kernels for CT



Concordance Correlation Coefficient

3 – Deep Learning in Radiomics

Data Harmonization : DL



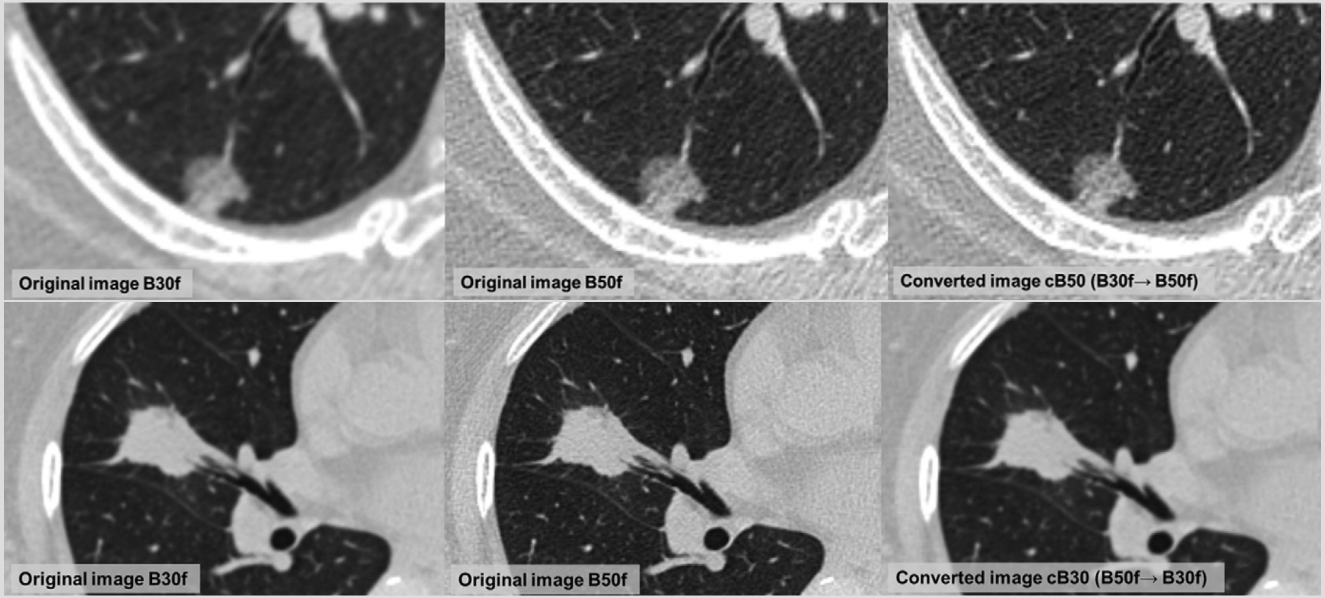
Original Images acquired with different reconstruction kernels

Harmonization (transfer) from one kernel to the other

Choe, et al. Deep Learning-based Image Conversion of CT Reconstruction Kernels Improves Radiomics Reproducibility for Pulmonary Nodules or Masses. Radiology 2019

3 – Deep Learning in Radiomics

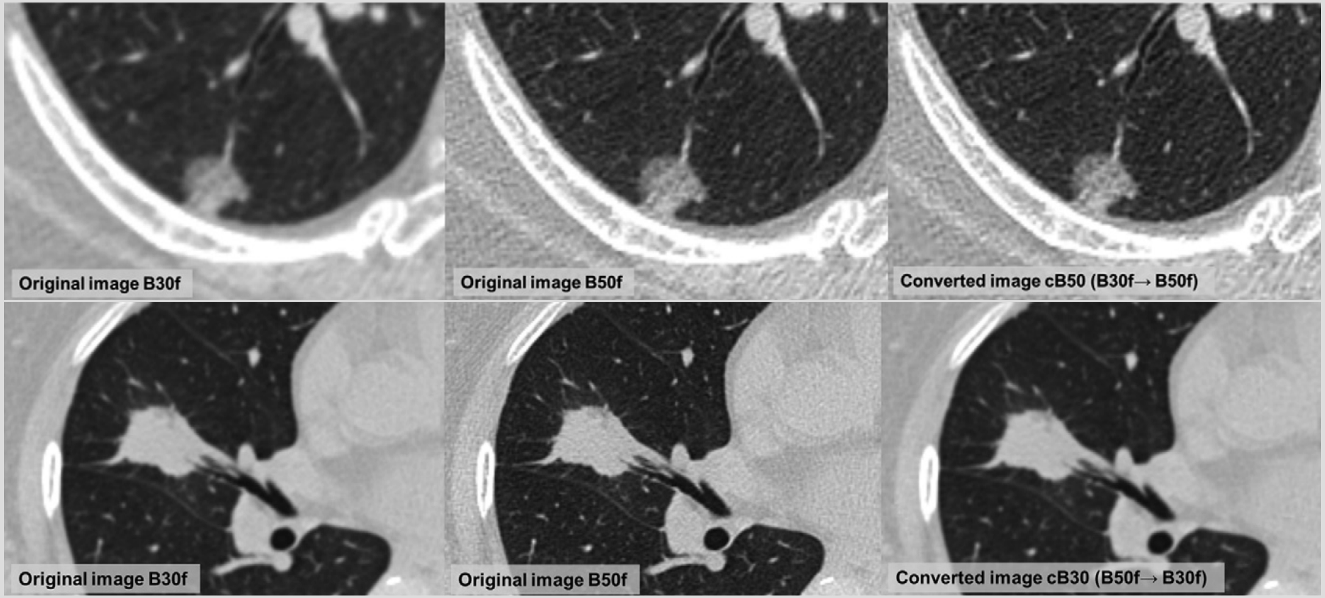
Data Harmonization : DL



Original Images acquired with different reconstruction kernels

Harmonization (transfer) from one kernel to the other

→ **ComBat** reduces the impact of external parameters on the measurements of features



Original Images acquired with different reconstruction kernels

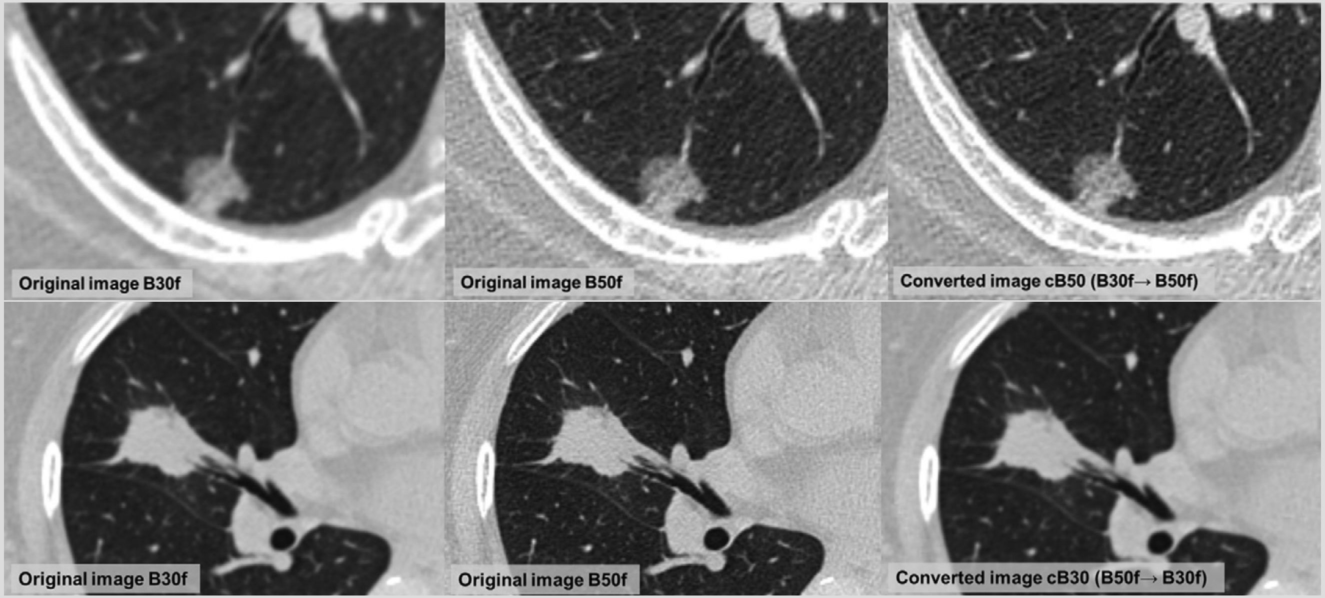
Harmonization (transfer) from one kernel to the other

3 – Deep Learning in Radiomics

Data Harmonization : DL

→ **ComBat** reduces the impact of external parameters on the measurements of features

→ **DL approaches** directly convert images (registration), then features are measured



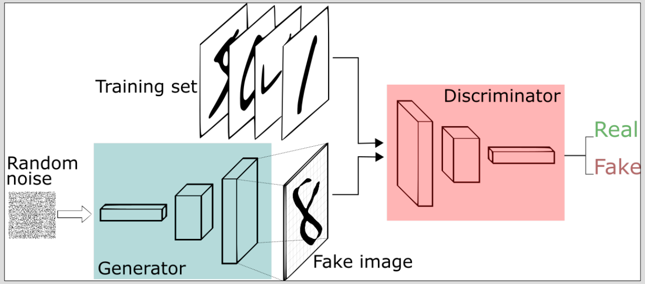
Original Images acquired with different reconstruction kernels

Harmonization (transfer) from one kernel to the other

Choe, et al. Deep Learning-based Image Conversion of CT Reconstruction Kernels Improves Radiomics Reproducibility for Pulmonary Nodules or Masses. Radiology 2019

3 – Deep Learning in Radiomics

Data Harmonization : DL



GAN functioning illustration

DL Harmonization (usually) uses GANs

- Generator tries to create convincing images
- Discriminator tries to distinguish real & fake images

Example

- Brain Tumor Image Segmentation (BRATS) benchmark
- MRI images from 19 different centers

New idea : Diffusion Models

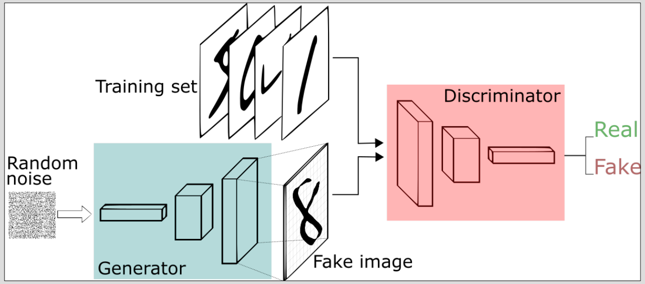
Hognon, Clément, et al. "Standardization of multicentric image datasets with generative adversarial networks." IEEE Nuclear Science Symposium and Medical Imaging Conference 2019. 2019.

3 – Deep Learning in Radiomics

Data Harmonization : DL

DL Harmonization (usually) uses GANs

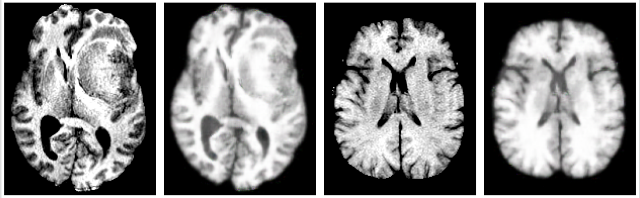
- Generator tries to create convincing images
- Discriminator tries to distinguish real & fake images



GAN functioning illustration

Example

- Brain Tumor Image Segmentation (BRATS) benchmark
- MRI images from 19 different centers



Harmonized images from BRATS dataset

New idea : Diffusion Models

Hognon, Clément, et al. "Standardization of multicentric image datasets with generative adversarial networks." IEEE Nuclear Science Symposium and Medical Imaging Conference 2019. 2019.

3 – Deep Learning in Radiomics

Data Harmonization : DL

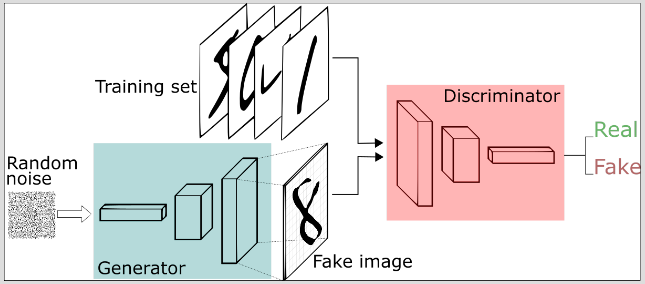
DL Harmonization (usually) uses GANs

- Generator tries to create convincing images
- Discriminator tries to distinguish real & fake images

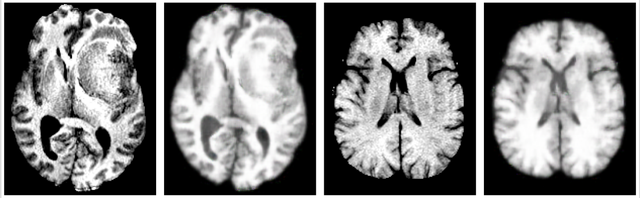
<u>Example</u>

- Brain Tumor Image Segmentation (BRATS) benchmark
- MRI images from 19 different centers

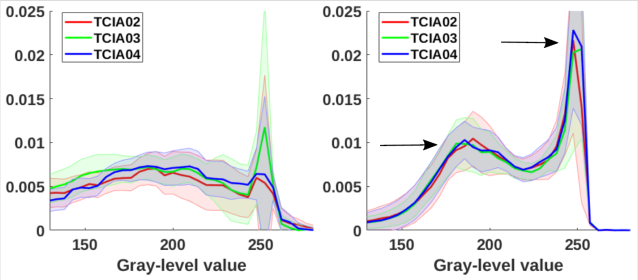
<u>New idea : Diffusion Models</u>



GAN functioning illustration



Harmonized images from BRATS dataset



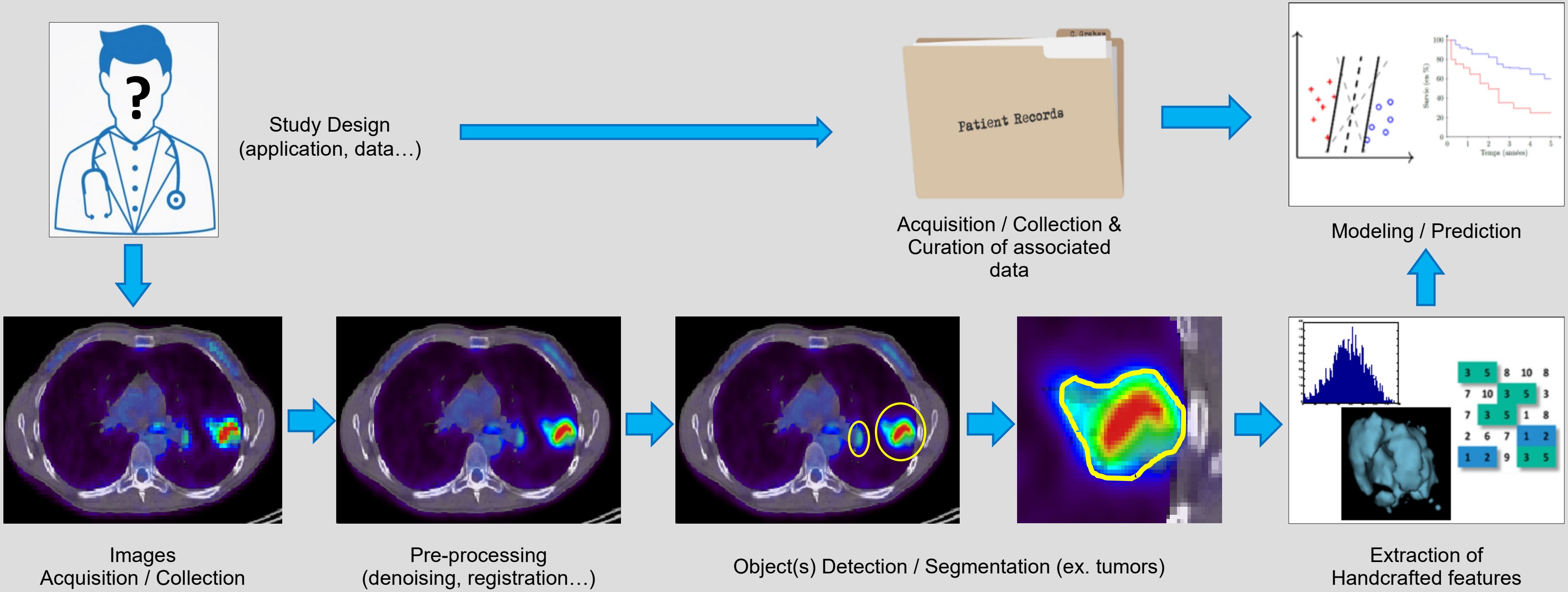
Kernel Density estimations before (left) and after (right) harmonization

Hognon, Clément, et al. "Standardization of multicentric image datasets with generative adversarial networks." IEEE Nuclear Science Symposium and Medical Imaging Conference 2019. 2019.

3 – Deep Learning in Radiomics

Tumor Detection & Segmentation

Automation Objective → framework can work fast on large data quantities

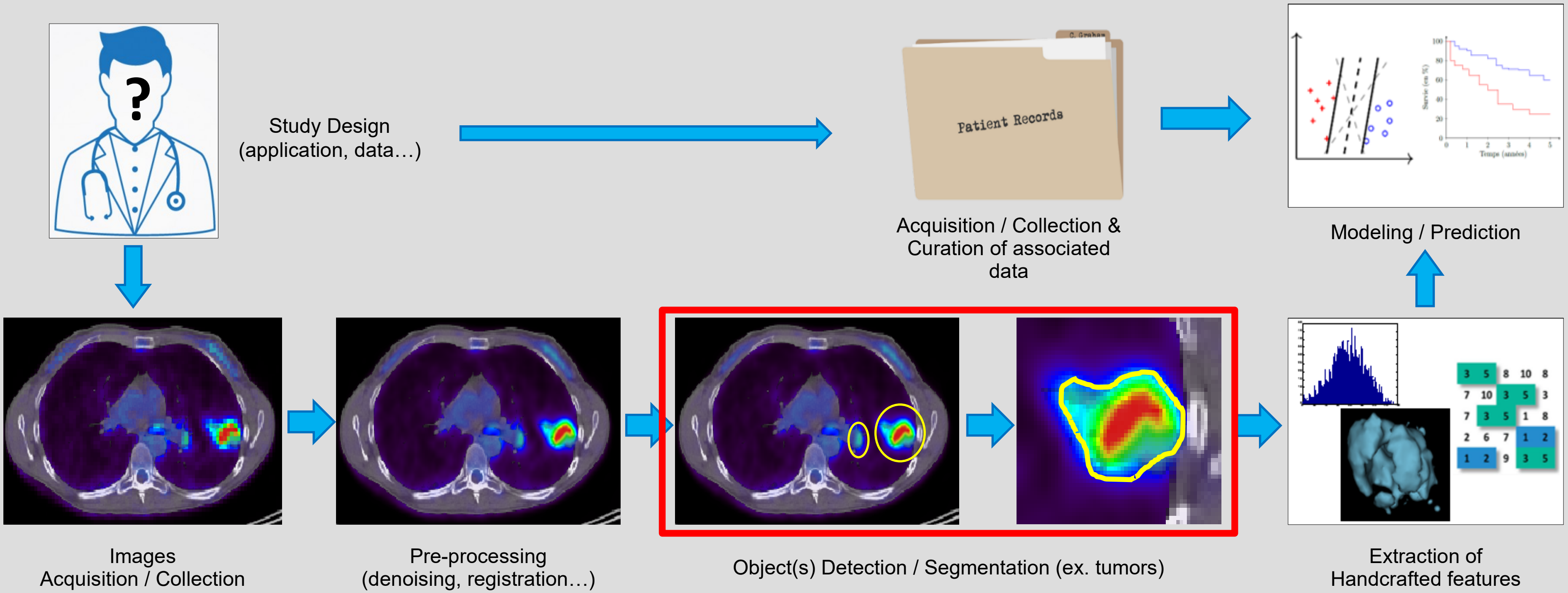


Hatt, et al. Data are also images. J Nucl Med 2019

3 – Deep Learning in Radiomics

Tumor Detection & Segmentation

Automation Objective → framework can work fast on large data quantities



3 – Deep Learning in Radiomics

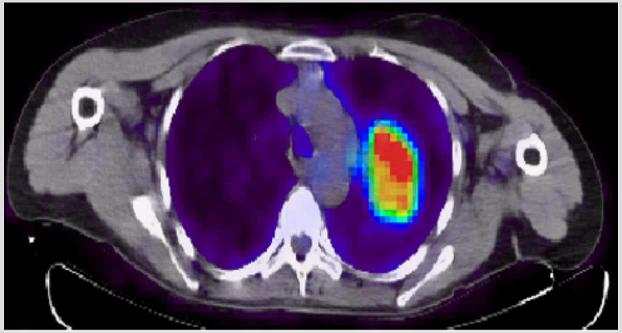
Tumor Detection & Segmentation

Tumor shape features are based on
Manual / Semi-Automatic Segmentation

→ What about inter / intra-study Reproducibility ?
→ How to process large data amount ?

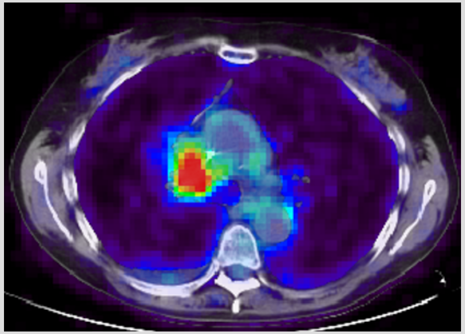
Fully Automatic frameworks based on DL
should give harmonious values across data

Tumor Volume &
Sphericity Computation

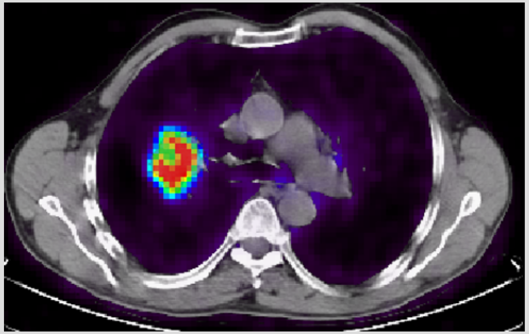


V = 34 cm^3
S = 0.66

V = 36 cm^3
S = 0.81

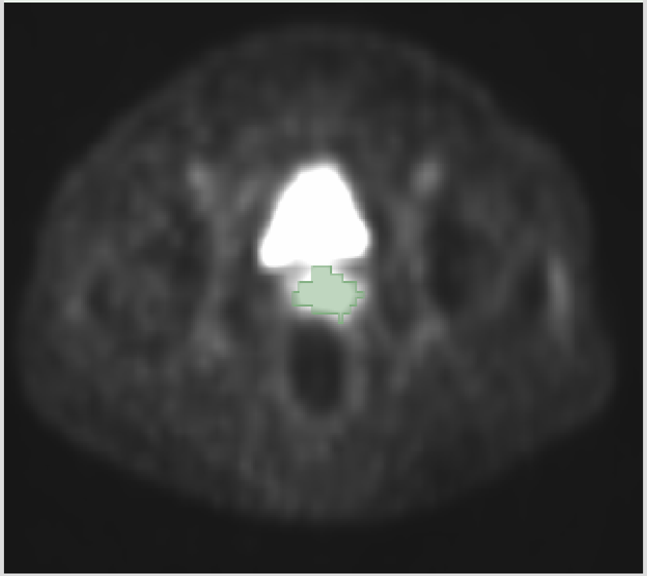


V = 28 cm^3
S = 0.57



3 – Deep Learning in Radiomics

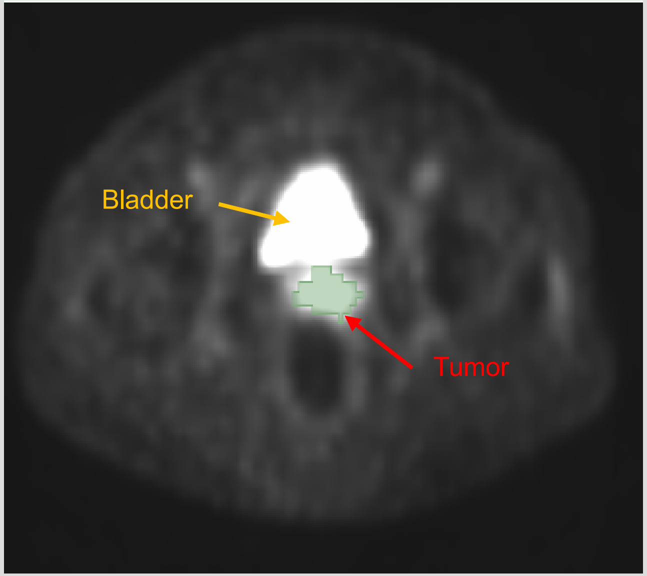
Tumor Detection & Segmentation



Semi-Automatic Tumor Segmentation

3 – Deep Learning in Radiomics

Tumor Detection & Segmentation

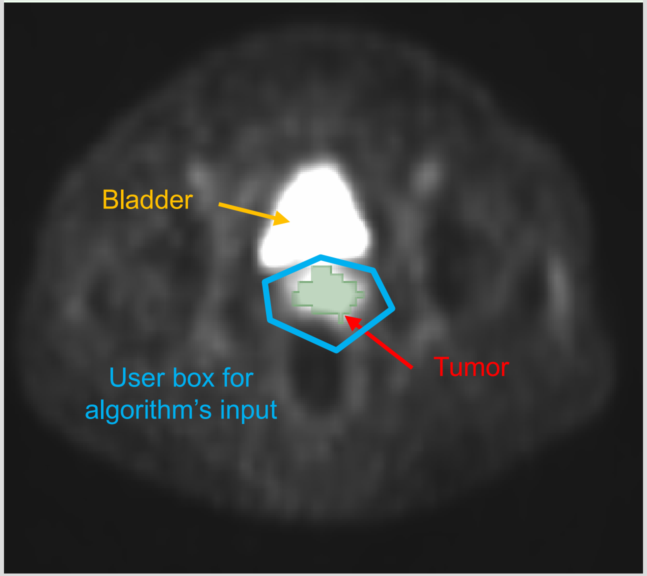


Semi-Automatic Tumor Segmentation

Iantsen, et al. *Convolutional neural networks for PET functional volume fully automatic segmentation: development and validation in a multi-center setting. EJNMMI 2021*

3 – Deep Learning in Radiomics

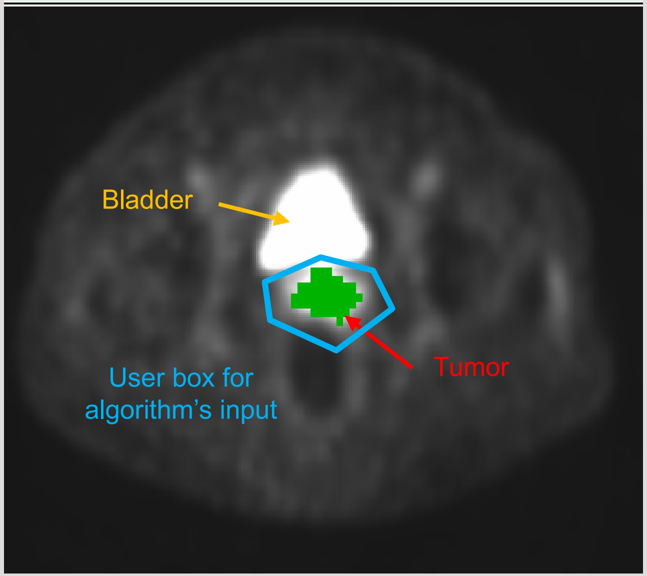
Tumor Detection & Segmentation



Semi-Automatic Tumor Segmentation

3 – Deep Learning in Radiomics

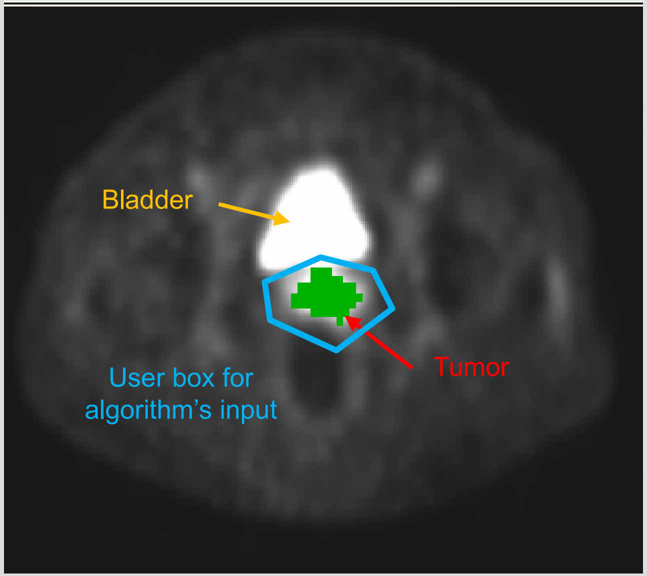
Tumor Detection & Segmentation



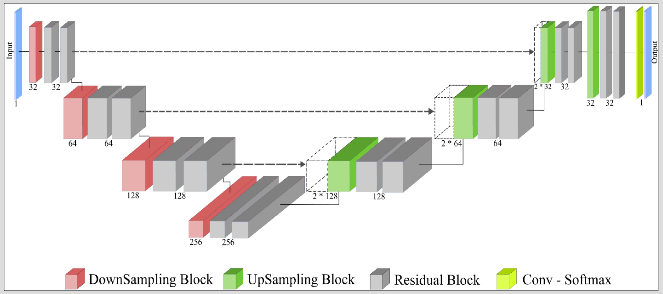
Semi-Automatic Tumor Segmentation

3 – Deep Learning in Radiomics

Tumor Detection & Segmentation



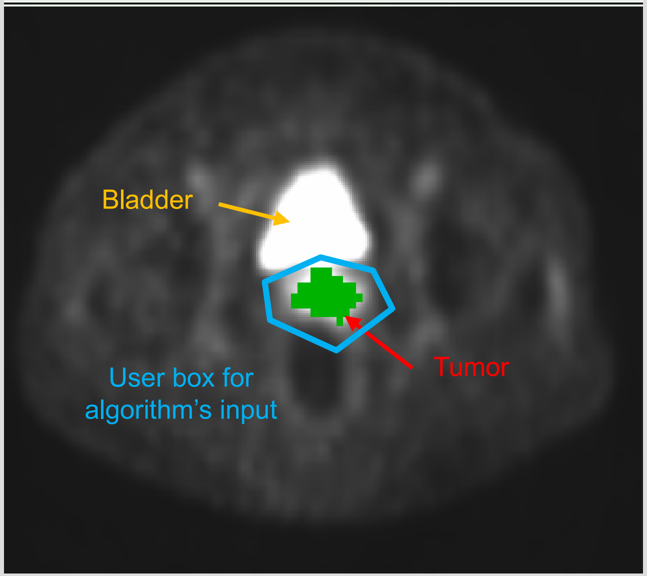
Semi-Automatic Tumor Segmentation



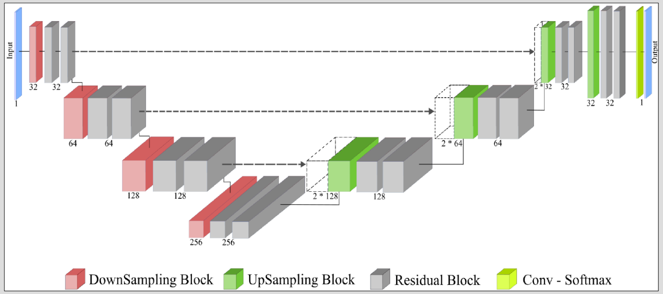
U-NET

3 – Deep Learning in Radiomics

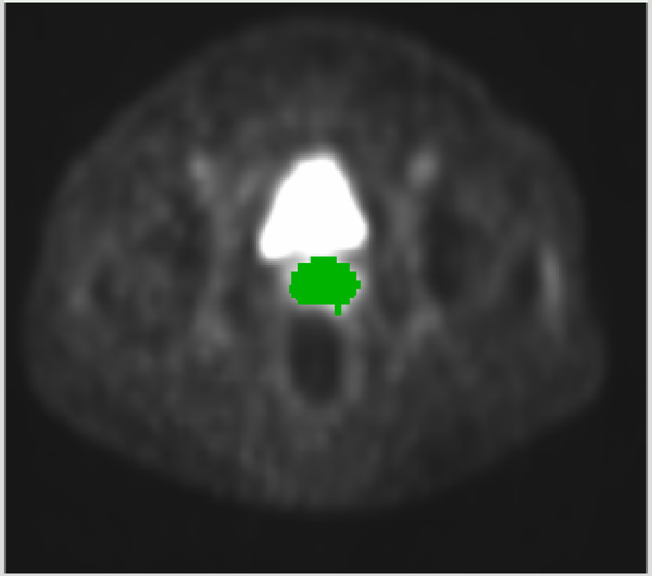
Tumor Detection & Segmentation



Semi-Automatic Tumor Segmentation



U-NET

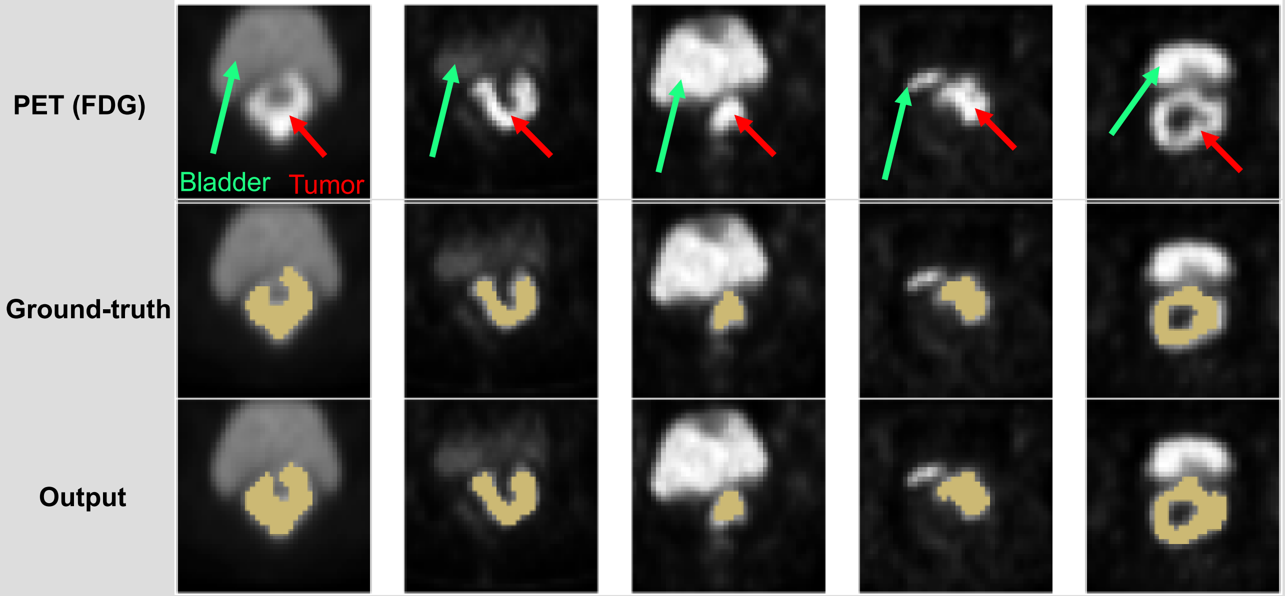


Deep Learning Automated Tumor Segmentation

Iantsen, et al. Convolutional neural networks for PET functional volume fully automatic segmentation: development and validation in a multi-center setting. EJNMMI 2021

3 – Deep Learning in Radiomics

Tumor Detection & Segmentation

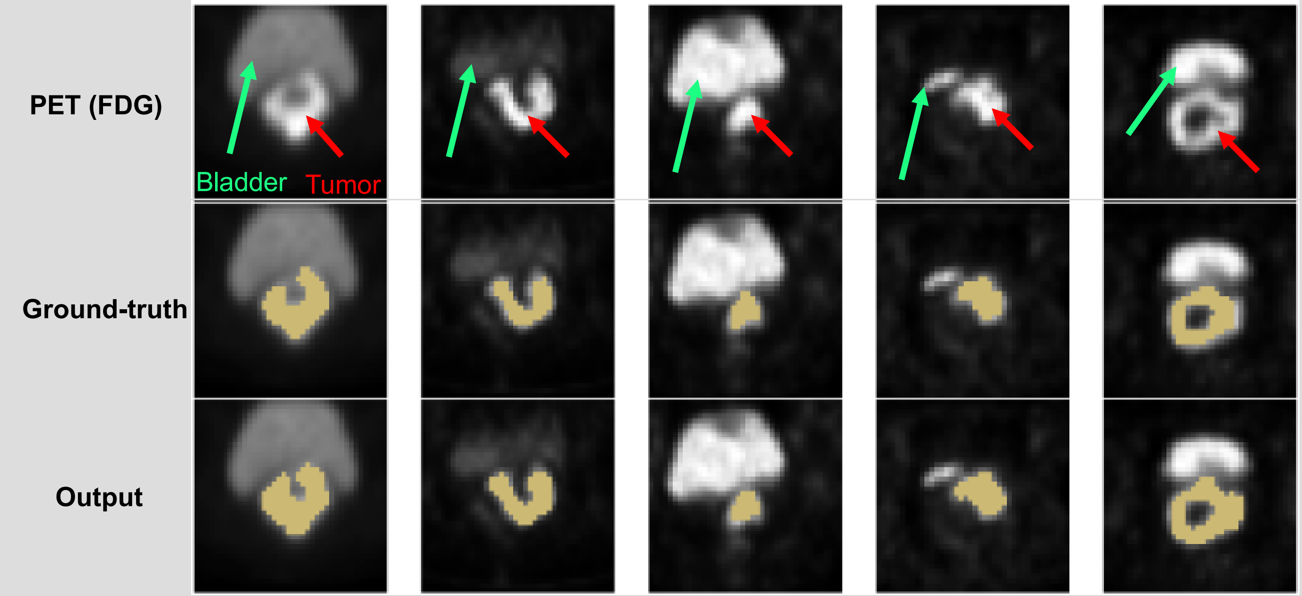


Semi-automatic vs Fully automated Bladder Tumor Segmentation

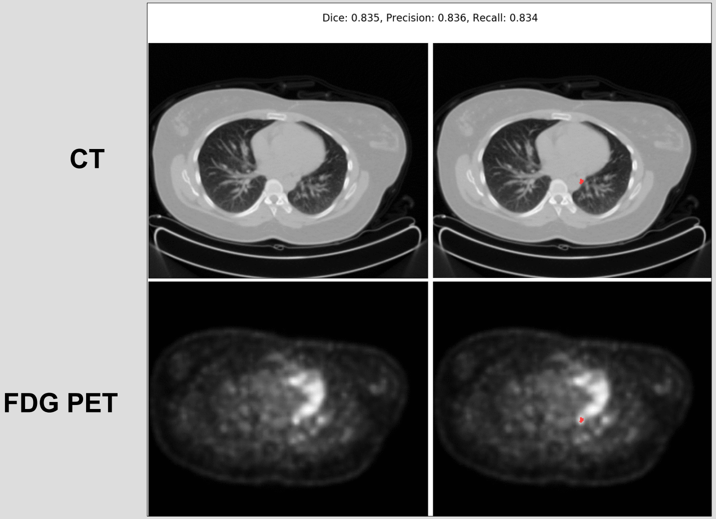
Iantsen, et al. Convolutional neural networks for PET functional volume fully automatic segmentation: development and validation in a multi-center setting. EJNMMI 2021
Iantsen, et al. Fully automated detection and segmentation of hypermetabolic lesions in pretherapeutic [18F]FDG PET / CT images of lymphoma and sarcoidosis patients. EANM 2021 annual meeting

3 – Deep Learning in Radiomics

Tumor Detection & Segmentation



Semi-automatic vs Fully automated Bladder Tumor Segmentation

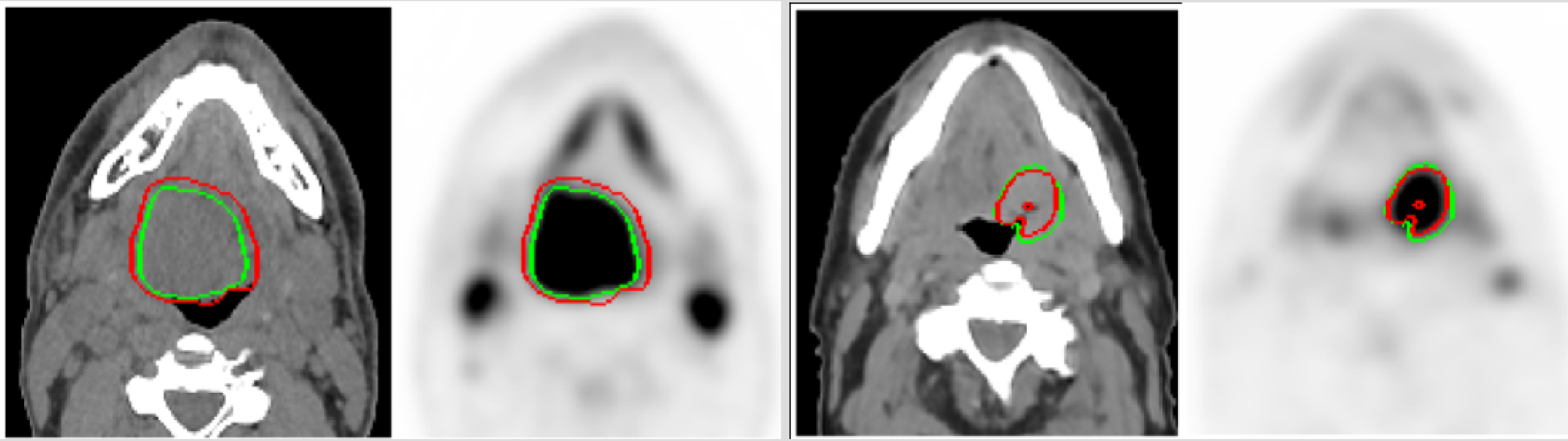


Ground Truth *Segmentation*

Iantsen, et al. Convolutional neural networks for PET functional volume fully automatic segmentation: development and validation in a multi-center setting. EJNMMI 2021
Iantsen, et al. Fully automated detection and segmentation of hypermetabolic lesions in pretherapeutic [18F]FDG PET / CT images of lymphoma and sarcoidosis patients. EANM 2021 annual meeting

3 – Deep Learning in Radiomics

Tumor Detection & Segmentation

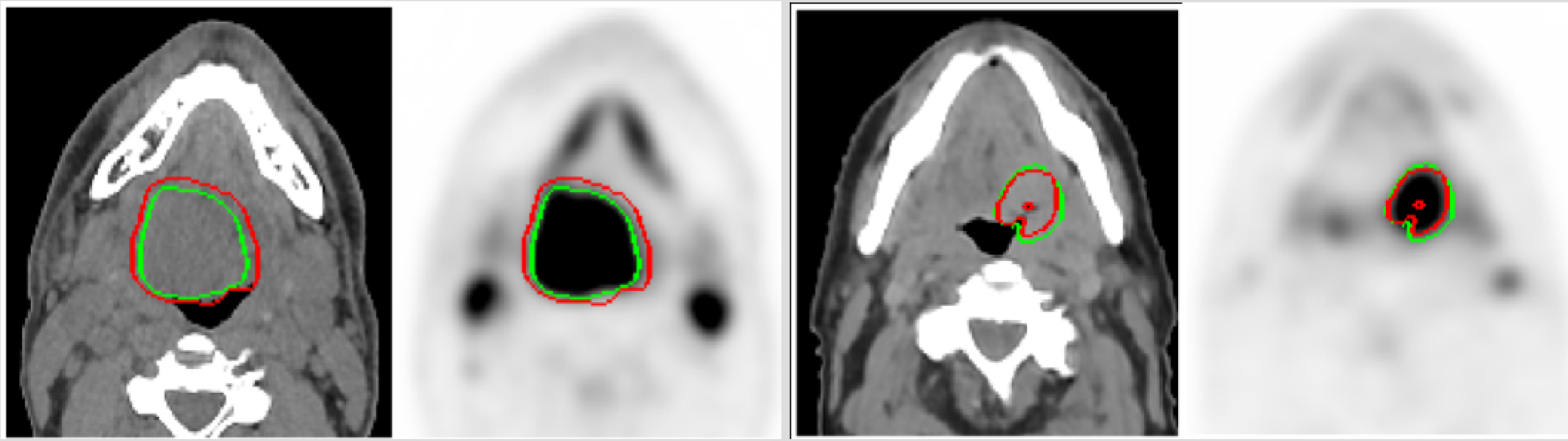


Fully automated Throat Tumor Segmentation on TDM (dark) & PET (clear)

Segmentation	Reference

3 – Deep Learning in Radiomics

Tumor Detection & Segmentation



Fully automated Throat Tumor Segmentation on TDM (dark) & PET (clear)

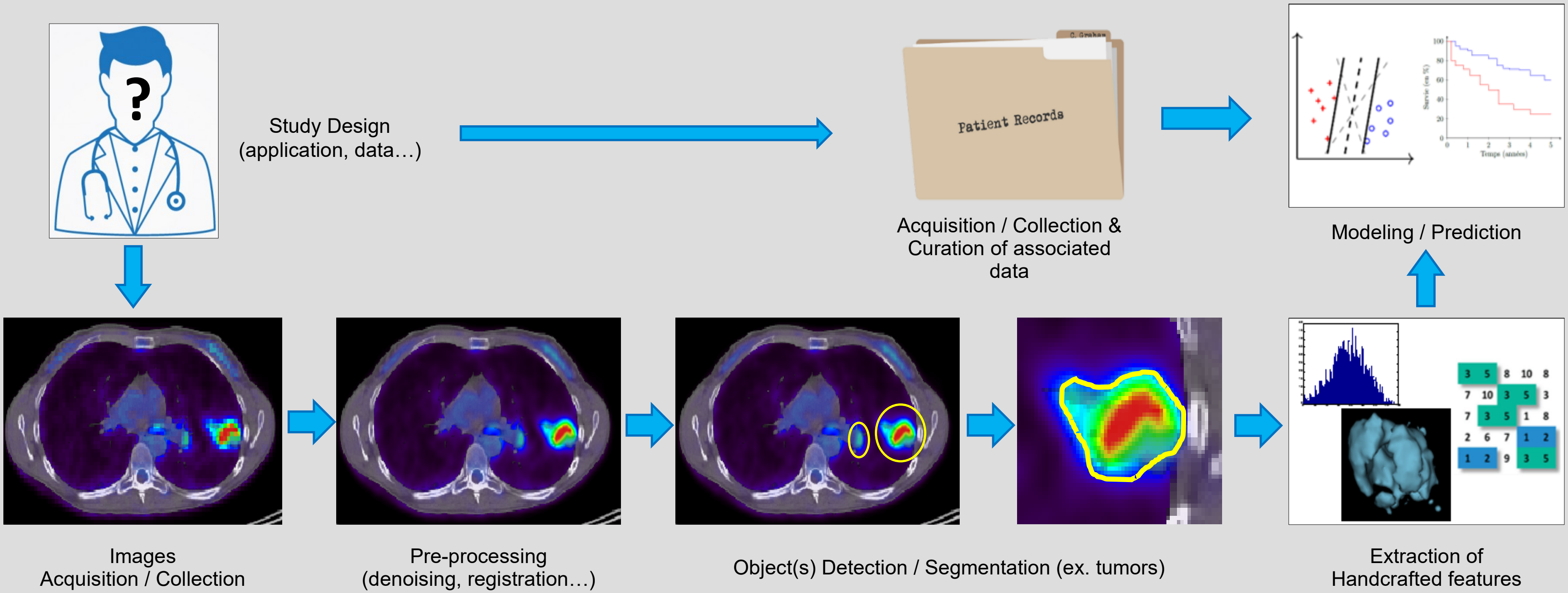
Segmentation	Reference

Segmentation is now widely experimented → high performances

3 – Deep Learning in Radiomics

Features Standardization

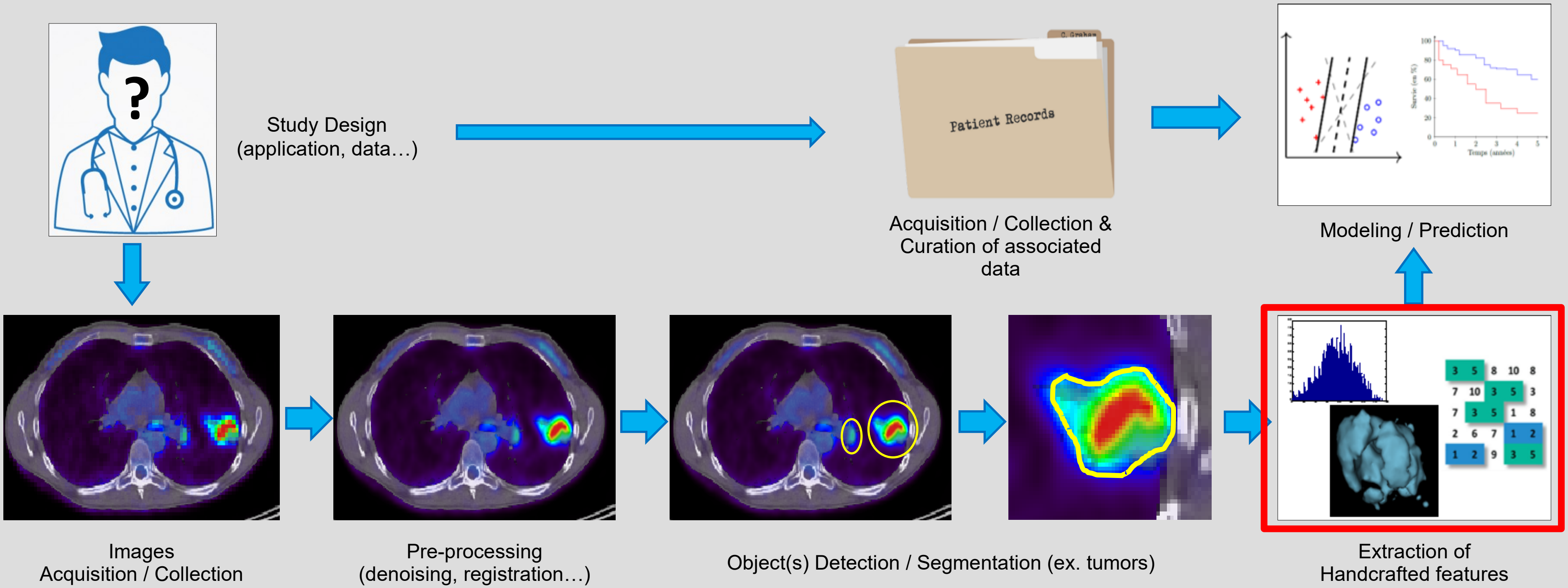
Standardization Objective → Radiomics features must be common to all processes



3 – Deep Learning in Radiomics

Features Standardization

Standardization Objective → Radiomics features must be common to all processes



Hatt, et al. Data are also images. J Nucl Med 2019

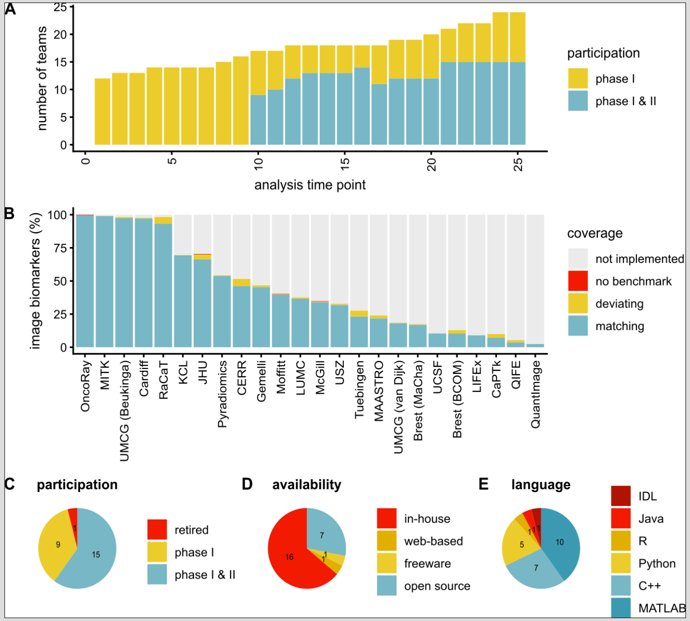
3 – Deep Learning in Radiomics

Features Standardization

<u>IBSI: Image Biomarker Standardization Initiative</u>

Most important study on Feature Standard

BUT no Deep Learning

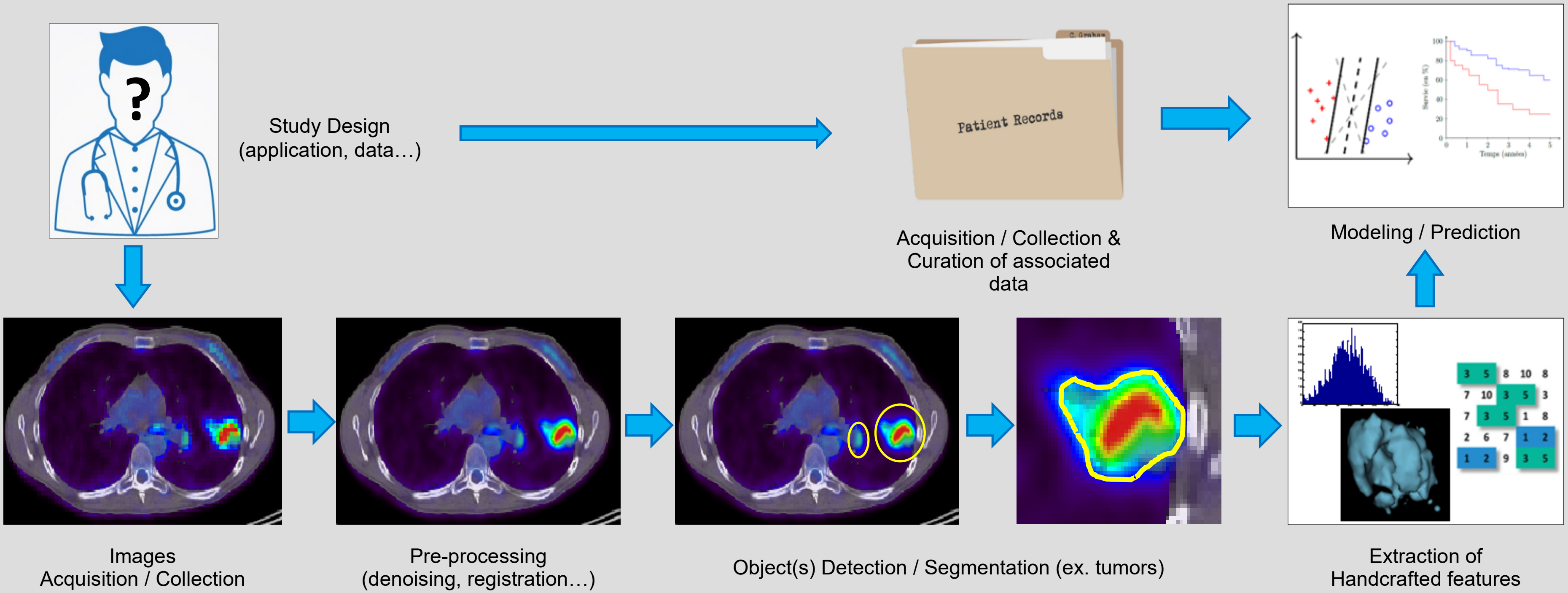


21 participants from 8 countries

3 – Deep Learning in Radiomics

Modeling & Prediction

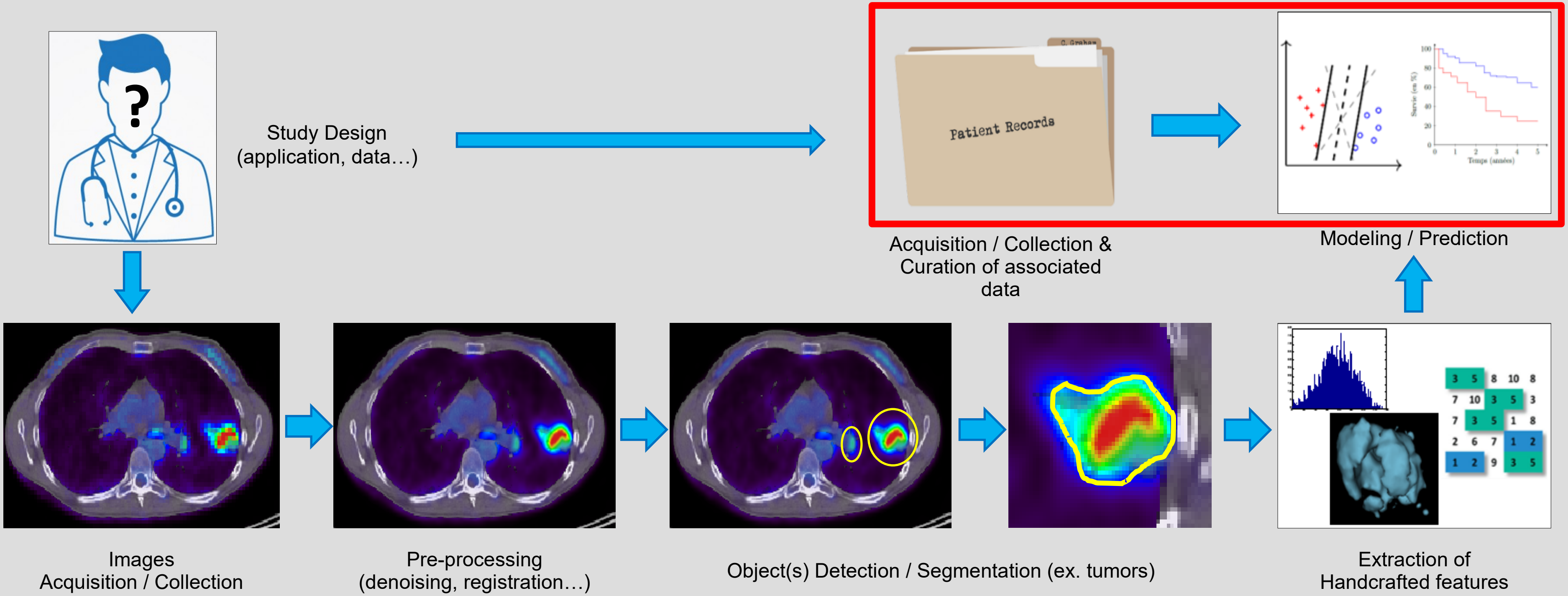
Use **external data & Radiomics features** for automatic Modeling & Prediction



3 – Deep Learning in Radiomics

Modeling & Prediction

Use **external data & Radiomics features** for automatic Modeling & Prediction



Hatt, et al. Data are also images. J Nucl Med 2019

3 – Deep Learning in Radiomics

Modeling & Prediction

Machine Learning algorithms are **widely used** (inputs are standard Radiomics features)

→ Idea of finding relations between features

Training Data (50%)

- v1
- v2
- v3
- v4
- ….

⟷ Finding Correlation between variables

Example

Non-responsive patients have larger tumor volume (v2) than responsive patients, and have ganglions (v3)

3 – Deep Learning in Radiomics

Modeling & Prediction

Machine Learning algorithms are **widely used** (inputs are standard Radiomics features)

→ Idea of finding relations between features

Training Data (50%)

- v1
- v2
- v3
- v4
- ….

↔ Finding Correlation between variables

Validation Data (30%)

f(v2, v3) ↔ Optimizing Model Parameters

Example

Non-responsive patients have larger tumor volume (v2) than responsive patients, and have ganglions (v3) → Max accuracy if Volume > 30 cm^3 AND more than 0 ganglions

3 – Deep Learning in Radiomics

Modeling & Prediction

Machine Learning algorithms are **widely used** (inputs are standard Radiomics features)

→ Idea of finding relations between features

Training Data (50%)

- v1
- v2
- v3
- v4
- ….

↔ Finding Correlation between variables

Validation Data (30%)

f(v2, v3) ↔ Optimizing Model Parameters

Test Data (20%)

v2 > 30 cm³ AND v3 > 0 ↔ Evaluate Performances

Example

Non-responsive patients have larger tumor volume (v2) than responsive patients, and have ganglions (v3) → Max accuracy if Volume > 30 cm³ AND more than 0 ganglions → Accuracy → 74 %

Visvikis, et al. Application of artificial intelligence in nuclear medicine and molecular imaging: a review of current status and future perspectives for clinical translation. Eur J Nucl Med Mol Imaging 2022

3 – Deep Learning in Radiomics

Modeling & Prediction

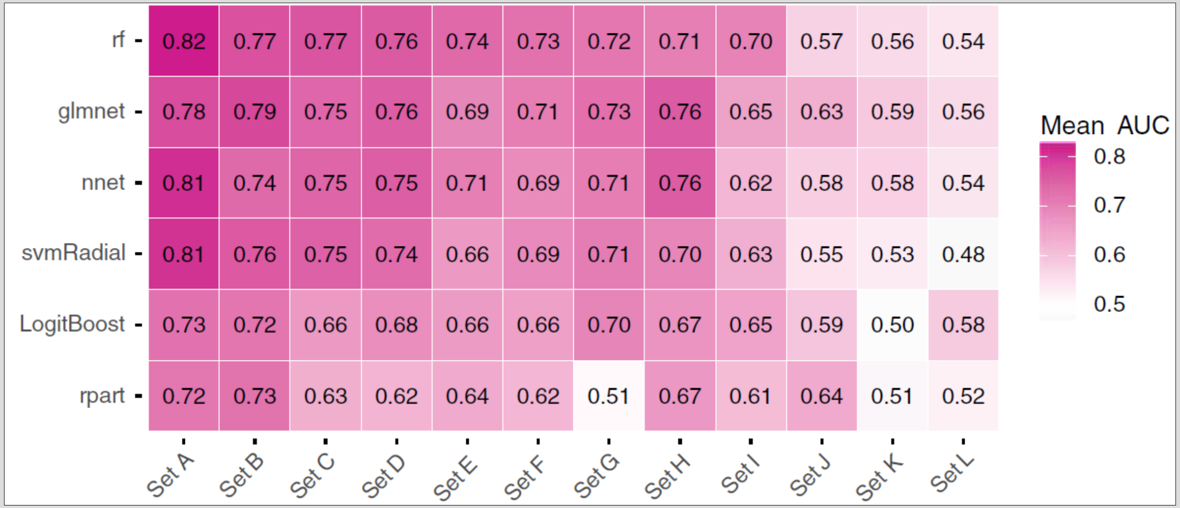
Example : <u>(chemo)radiotherapy</u>

- 12 datasets (3496 patients total)
- Non-Small-Cell Lung Cancer (NSCLC), Head & Neck Cancer, Meningioma

Conclusions :

- Random forest best in 6/12 datasets
- Elastic net logistic regression best in 4/12 sets

→ <u>No single best classifier across all datasets</u>



AUC for multilpe Machine Learning algorithms, on 12 separate datasets

How to select a machine learning algorithm?
→ Fusion / Ensemble ?

- Stage II and III NSCLC
- PET + CT Radiomics
- Classification as poor prognosis

Algorithms :

- Random Forest
- Support Vector Machine
- Logistic regression / LASSO

VS

- Fusion by majority voting
- Fusion by average of probas

→ What about a full end-to-end Deep Learning process ?

Table 2. Performance of the various ML methods and their consensus.

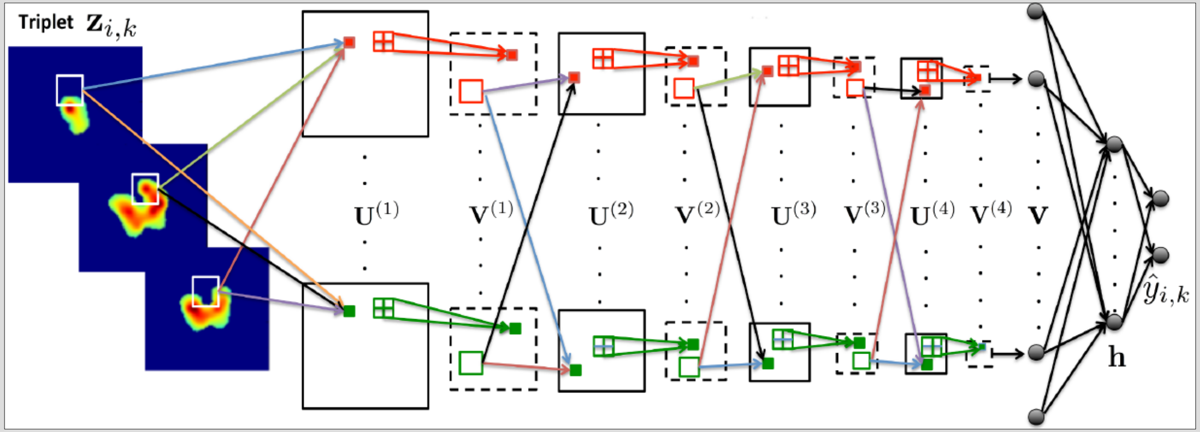
Classification	Accuracy (Median OS) (%)			Balanced Accuracy (OS < 6 Months) (%)		
	Training	Testing	# of Patients Correctly Classified	Training	Testing	# of Patients Correctly Classified
Stage 2 vs. 3	61 [95% CI 53–69]	58 [95% CI 50–66]	30	59 [95% CI 51–57]	53 [95% CI 45–61]	27
RF	89 [95% CI 84–94]	67 [95% CI 59–75]	34	100 [95% CI 100–100]	80 [95% CI 73–87]	41
SVM	100 [95% CI 100–100]	64 [95% CI 56–72]	33	92 [95% CI 87–97]	75 [95% CI 68–82]	38
LR	72 [95% CI 65–79]	63 [95% CI 55–71]	32	84 [95% CI 78–90]	78 [95% CI 71–85]	40
Fusion (majority voting)	100 [95% CI 100–100]	71 [95% CI 63–79]	36	100 [95% CI 100–100]	84 [95% CI 78–90]	43
Fusion (average of probabilities)	100 [95% CI 100–100]	78 [95% CI 71–85]	40	100 [95% CI 100–100]	89 [95% CI 84–94]	45

Performance of stage, SUV measurements and volume in testing: 55-61%

Sepehri, et al. Comparison and fusion of machine learning algorithms for prospective validation of PET/CT radiomic features prognostic value in stage II-III Non-Small Cell Lung Cancer. Diagnostics 2021

1st proposed full DL-based (2015)

- Directly from Image to Prediction of non-responders for Chemotherapy
- 96 patients for training, 11 for testing
- 2 Networks tested : Triplets (3S-CNN) or single slice (1S-CNN)
- Data augmentation → 5316 triplets for both responders and non-responders



CNN architecture for fusion of 3 adjacent 18F-FDG PET intra slices into a vector

Ypsilantis, et al. Predicting Response to Neoadjuvant Chemotherapy with PET Imaging Using Convolutional Neural Networks. PLoS One 2015

3 – Deep Learning in Radiomics

Towards a Full DL Process

→ It is possible to get **higher performances** without using standard Radiomics

Table 2. Classification results: each figure is the average of three independent experiments using different training and test datasets.

Method	Sensitivity	Specificity	Accuracy
3S-CNN	80.7±11.5	81.6±9.2	73.4±5.3
1S-CNN	77.9±12.9	58.3±4.2	66.4±5.9
GB	70.5±6.0	63.8±6.1	66.7±5.2
GB with PCA	68.1±7.9	46.8±16.2	66.8±6.0
RF	61.0±8.6	36.4±18.4	57.3±7.8
RF with PCA	65.8±7.5	52.0±28.9	65.7±5.6
SVM	66.9±8.5	38.4±19.2	55.9±8.1
SVM with PCA	67.4±10.3	50.9±5.0	60.5±8.0
Logistic Reg.	60.4±6.2	38.3±7.3	51.4±3.0
Logistic Reg. with PCA	58.9±4.9	38.9±12.5	48.4±8.0
SUV max with threshold	33.0±33.0	35.2±10.2	41.0±4.5
SUVmax median threshold	81.5±1.5	53.0±13.0	67.7±4.2

Proposed DL architectures against various ML processes

Ypsilantis, et al. Predicting Response to Neoadjuvant Chemotherapy with PET Imaging Using Convolutional Neural Networks. PLoS One 2015

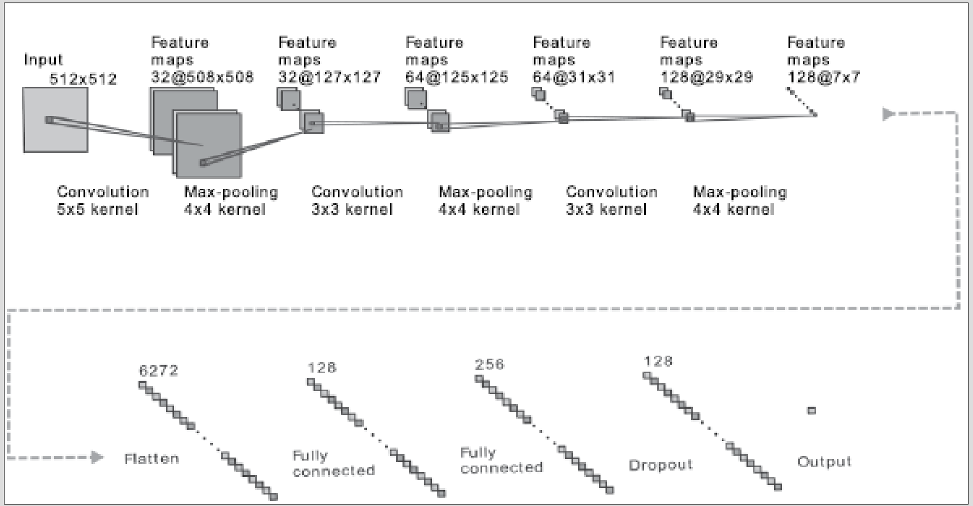
3 – Deep Learning in Radiomics

Towards a Full DL Process

Other example

- Head & Neck Cancer treatment outcome prediction
- Public dataset from TCIA
- Based on pre-treatment CT images
- CNN using only one image as input (512x512)

→ Better metrics scores than *Vallières et al.* Standard Radiomics



	Specificity		Sensitivity		Balanced Accuracy	
	Present study	Vallières *et al.*[25]	Present study	Vallières *et al.*[25]	Present study	Vallières *et al.*[25]
DM	0.89	0.77	0.86	0.79	88%	77%
LRF	0.67	0.61	0.65	0.39	66%	58%
OS	0.67	0.67	0.68	0.55	68%	62%

3 – Deep Learning in Radiomics

Towards a Full DL Process

Another Possibility

→ Combine Standard & DL-based Radiomics processes

- Breast Cancer Diagnosis
- Tested on 3 imaging modalities :
  - Dynamic Contrast Enhanced-MRI
  - Full-Field Digital Mammography
  - Ultrasound

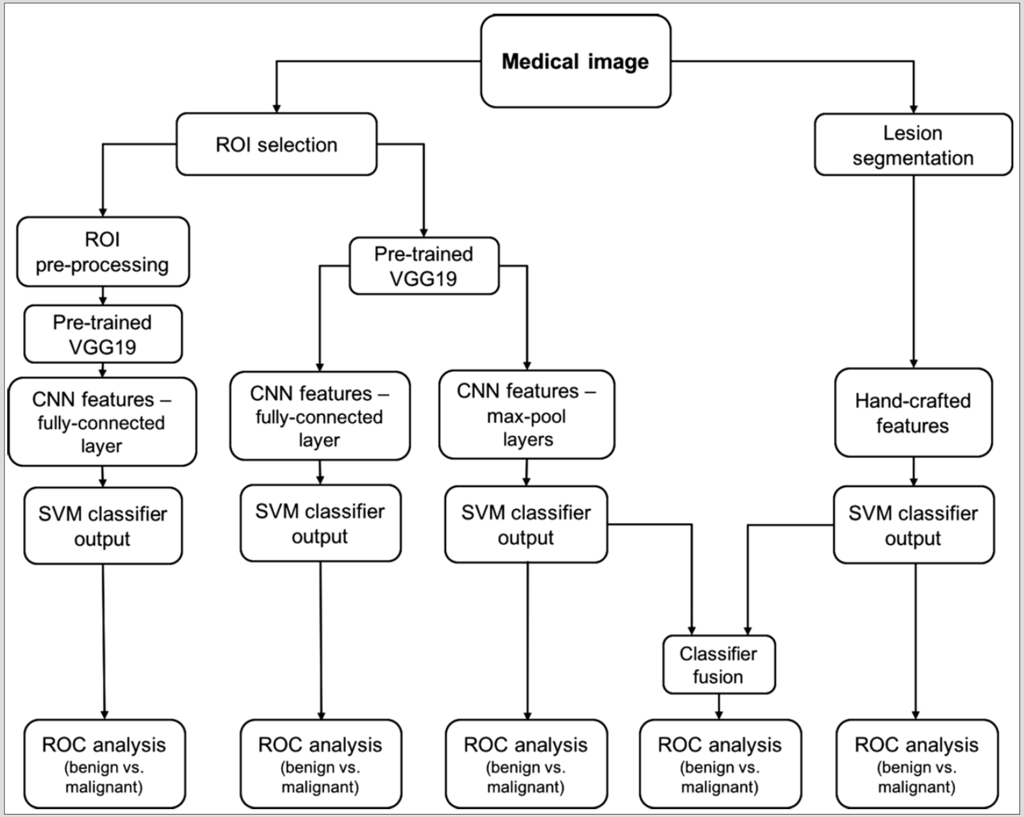
3 – Deep Learning in Radiomics

Towards a Full DL Process

Another Possibility

→ Combine Standard & DL-based Radiomics processes

- Breast Cancer Diagnosis
- Tested on 3 imaging modalities :
  - Dynamic Contrast Enhanced-MRI
  - Full-Field Digital Mammography
  - Ultrasound



Antropova, et al. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Med Phys 2017

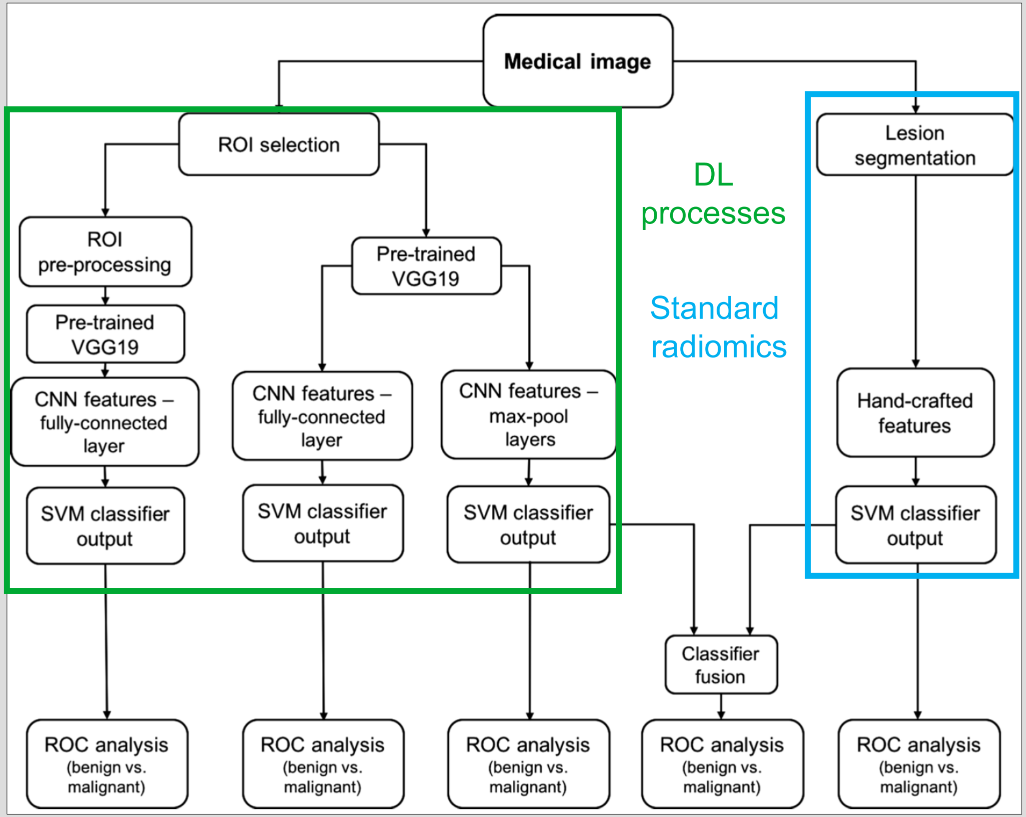
3 – Deep Learning in Radiomics

Towards a Full DL Process

Another Possibility

→ Combine Standard & DL-based Radiomics processes

- Breast Cancer Diagnosis
- Tested on 3 imaging modalities :
  - Dynamic Contrast Enhanced-MRI
  - Full-Field Digital Mammography
  - Ultrasound

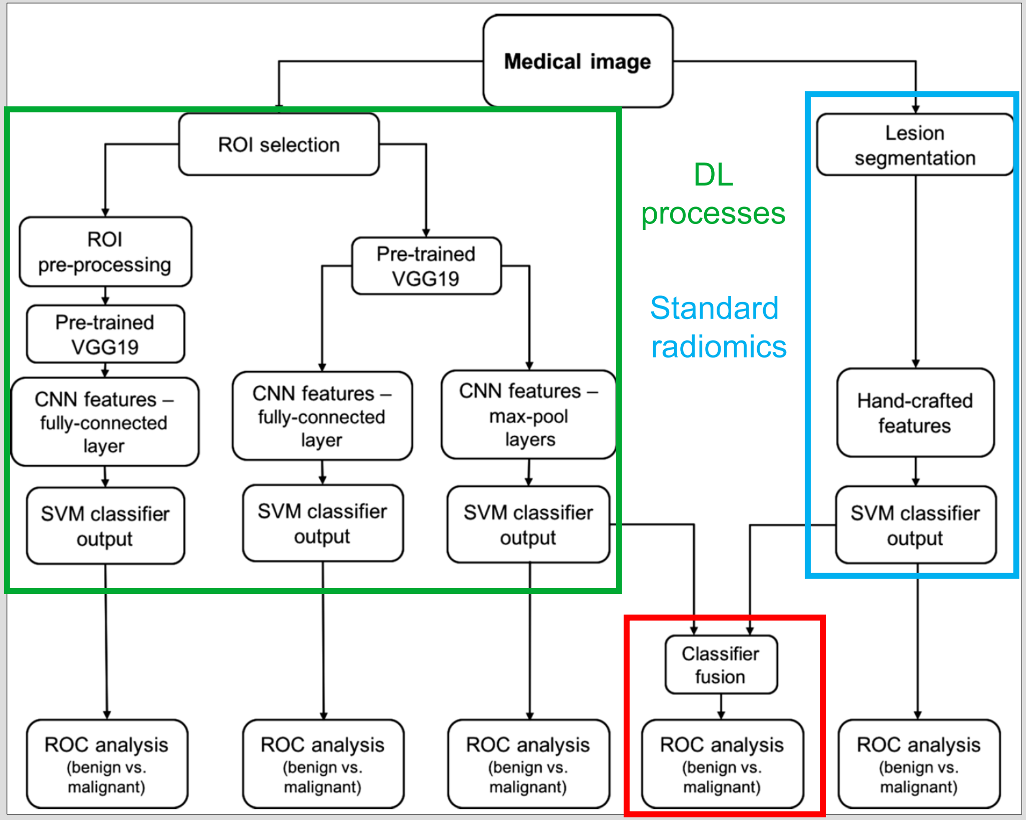


Antropova, et al. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Med Phys 2017

Another Possibility

→ Combine Standard & DL-based Radiomics processes

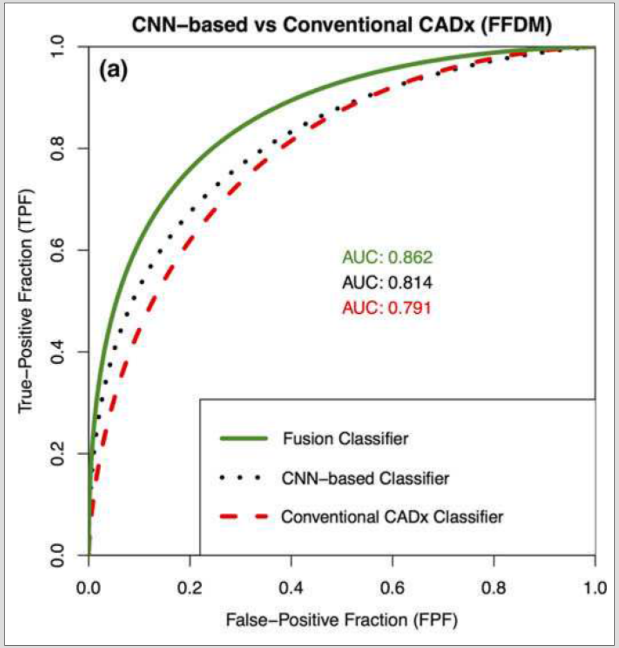
- Breast Cancer Diagnosis
- Tested on 3 imaging modalities :
  - Dynamic Contrast Enhanced-MRI
  - Full-Field Digital Mammography
  - Ultrasound



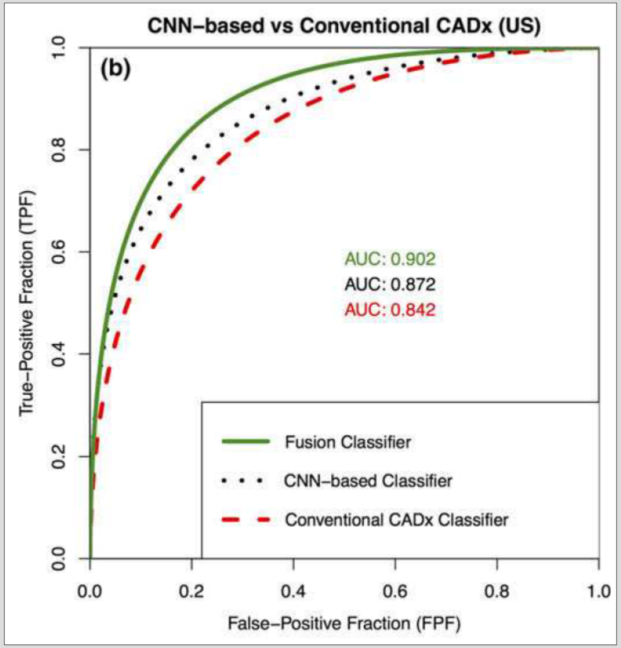
Antropova, et al. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Med Phys 2017

3 – Deep Learning in Radiomics

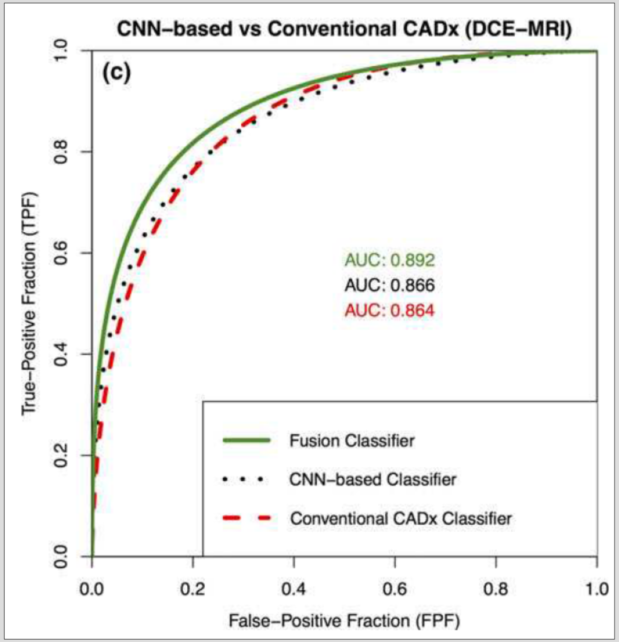
Towards a Full DL Process



Full field digital mamography (FFDM)
N=245



Ultrasound (US)
N=1125



Dynamic contrast enhanced-MRI (DCE-MRI)
N=690

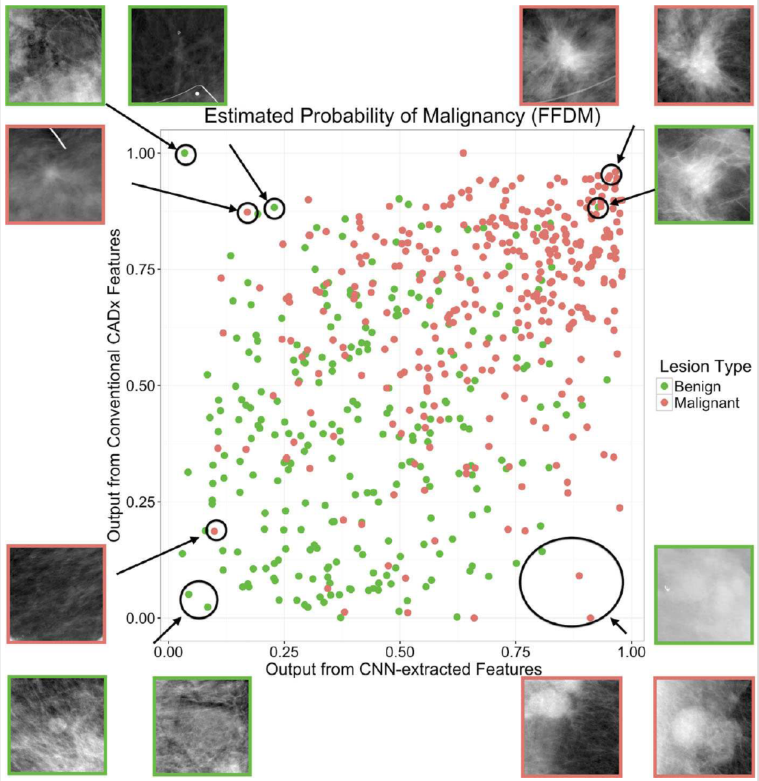
Antropova, et al. A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets. Med Phys 2017

3 – Deep Learning in Radiomics

Towards a Full DL Process

Plotting CNN output against Standard Radiomics output

→ **Moderate agreement** leading to <u>better performance of fusion</u>



Diagonal classifier agreement plot between the CNN-based classifier and the conventional CADx classifier for FFDM

3 – Deep Learning in Radiomics

Clinical Value : Importance of Challenges

Challenges allow for fair evaluation

<u>2 tasks :</u>

- Automatic Segmentation of the tumor
- Survival Prediction without recurrence

<u>Available data :</u>

- PET / CT images
- 2021 → 6 centers, 425 patients
- 2022 → 9 centers, 800 patients

Objectively computed performances (no tinkering the results)

Allows for comparison of different processes (standard Radiomics / DL-based)

Andrearczyk, et al. Overview of the HECKTOR Challenge at MICCAI 2021: Automatic Head and Neck Tumor Segmentation and Outcome Prediction in PET/CT Images, arXiv 2201.04138 2022

3 – Deep Learning in Radiomics

Clinical Value : Importance of Challenges

Deep Learning only

Large set of Engineered Radiomics features + ML

Team	C-index Task 2	Pre-processing						Segment.			Image features				Modeling and training approach										
		Iso-resampling	CT clipping	Min-max norm.	Standardization	PET/CT fusion	Further cropping	Relies on Task 1	Additional segm.	No segmentation	Deep features	Large radiomics set	Volume, shape	IBSI compliant	Ensembling	Deep model	Algo. RF, SVM ...	Feature selection	PET as input	CT as input	PET/CT fusion	Use clinical var.	Imputed missing	Cross-val.	Augmentation
BioMedIA [54]	0.7196			✓		✓	✓			✓	✓					✓					✓	✓			
Fuller MDA [49]	0.6938		✓	✓	✓					✓	✓				✓	✓			✓	✓		✓			✓
Qurit Tecvico [51]	0.6828			✓				✓				✓		✓	✓		✓	✓			✓	✓		✓	
BMIT_USYD [43]	0.6710		✓	✓	✓		✓	✓			✓				✓	✓			✓	✓		✓			✓
DMLang [32]	0.6681		✓	✓				✓			✓					✓			✓	✓		✓			✓
TECVICO_C. [17]	0.6608					✓	✓	✓				✓		✓			✓	✓			✓	✓		✓	
BAMF Health [46]	0.6602		✓	✓	✓			✓	✓			✓		✓			✓	✓	✓	✓		✓	✓	✓	
in-h-ai [56]	0.6592											✓		✓			✓	✓	✓	✓		✓		✓	
Neurophet [33]	0.6495							✓					✓				✓					✓		✓	
UMCG [40]	0.6445		✓	✓	✓	✓				✓	✓				✓	✓			✓	✓	✓	✓		✓	✓
Aarhus Oslo [27]	0.6391																					✓		✓	
RedNeucon [6]	0.6280							✓	✓			✓					✓	✓				✓		✓	
Emmanuelle B. [9]	0.6223	✓	✓		✓			✓			✓	✓		✓		✓	✓	✓	✓	✓		✓	✓	✓	
BCIOQurit	0.6116 [65]	✓	✓		✓			✓			✓	✓		✓			✓	✓	✓	✓		✓		✓	✓
Vokyj [45]	0.5937	✓	✓		✓			✓					✓				✓					✓			✓
Xuefeng [21]	0.5510	✓	✓		✓			✓					✓				✓								
DeepX [67]	0.5290	✓	✓		✓			✓				✓		✓			✓		✓	✓		✓			

Table 4: Synthetic comparison of outcome prediction methods. More details are available in Section 4.
task 3 also participated in task 2.

Performances similar (or lower) to clinical variables-only model (around 0.64 – 0.65)

Partial Results for Outcome Prediction Methods

3 – Deep Learning in Radiomics

Clinical Value : Importance of Challenges

Brain Tumor Segmentation (BraTS) challenge

<u>Also 2 tasks :</u>

- Brain Tumor Segmentation
- Prediction of the MGMT promoter methylation status

<u>Available data :</u>

- Multi-Parametric MRI images
- 2020 → 660 cases
- 2021 → 2,000 cases

Andrearczyk, et al. Overview of the HECKTOR Challenge at MICCAI 2021: Automatic Head and Neck Tumor Segmentation and Outcome Prediction in PET/CT Images, arXiv 2201.04138 2022

Conclusions

- **DL is increasingly used** in Radiomics
  - Automation of simple tasks (Detection & Segmentation)
  - Harmonization of images
  - Directly for end-to-end processes
  - Frequently using CNN, but not only (GANs, Diffusion Models, …)

- **DL could replace standard Radiomics**, but some limitations persist
  - Maturity & Evaluation of results → more time
  - Available Data, needed in large quantities → few-shot learning ?
  - Multi-centers studies are required but not easy to coordinate → better communication

One last limitation to assess → **Interpretability / Explainability of Deep Networks in Radiomics** → more this afternoon

Thank You for your attention

Questions ?

Remarks ?