An original voxel-wise supervised analysis of tumors with multimodal radiomics to highlight predictive biological patterns

Thibault Escobar[1,2], Sébastien Vauclin[2], Fanny Orlhac[1], Christophe Nioche[1], Pascal Pineau[2], Irène Buvat[1]

[1]Laboratoire d'Imagerie Translationnelle en Oncologie (LITO), Université Paris Saclay, U1288 Inserm, Institut Curie, Orsay, France [2]DOSIsoft SA, Cachan, France

Background and Objectives

- Translation of predictive and prognostic image-based learning models to **clinical applications** are **challenging** due in part **to their lack of interpretability.**
- Deep-learning-based Class Activation Maps (CAM) give **information about the regions driving the models.**
- Yet, due to the high-level abstraction of deep features, **deep CAM** can be unstable and **difficult to interpret**, and **low sample size** can lead to **instabilities** and **sub-optimal convergence** for **complex models** such as CNN.

⇒ We propose and validate a method that combines the **interpretability** of handcrafted radiomics with a voxel-wise analysis and facilitates the **biological interpretation of models.**

Materials and Methods

- Publicly available dataset of **51 soft tissue sarcomas** (STS) of the extremities containing **FDG PET** (reconstruction: OSEM, median voxel size: 5.47 × 5.47 × 3.27 mm³, in-plane voxel size range: 3.91-5.47 mm) and **CT images** (voxel size: 9.8 × 9.8 × 3.27 mm³) from the **same PET/CT machine** (GE Discovery ST), and clinical and **follow-up information** [1]
- 19/51 patients have developed **lung metastases** 2 years after the diagnosis.
- **Voxel-based probabilistic supervised machine learning** pipeline to predict the **risk of lung metastasis occurrence at 2 years** (Figure 1: steps 1 to 5)
  ⇒ **Radiomic model M1** (P = modelized probability of metastasis occurrence)
- Backprojection of the **coefficients of M1 at the voxel level** (Figure 1: step 6)
  ⇒ **M1 handcrafted-radiomics CAM**
- Radiomics CAM **interpretation** ⇒ **identification of tumor sub-regions** ⇒ feature engineering to define **simple features** and build a **radiologically interpretable surrogate model**
  ⇒ **Radiological model M2**

Results

- **Radiomic model M1** showed **higher classification performance (ROC AUC)** compared to conventional biomarkers (Figure 2, Table 1).
- **Radiomics CAM** (Figure 3) highlighting the **biologically interpretable sub-regions that drive the model decision**
  ⇒ **Voxel-wise local interpretability**
- By interpreting the radiomics CAM jointly with the PET and CT images for all patients, we found out that the model M1 (Equation 1) has captured patterns that can be related to tumor **necrosis**, and **high and localized FDG uptakes** in the tumor. These results are **consistent with [1] and the soft tissue sarcoma grading system based on the biopsy [7].**
- This made it possible to define a **new feature** reflecting the **volume of the tumor sub-region devoid of metabolic activity (SUV < 40% SUVmax)** or with **reduced Hounsfield units (< 20HU)**. This corresponds to suspected necrosis in PET/CT.
- The simpler **radiological model M2** (Equation 2) **that we built from the interpretation of M1** showed **the 2nd best performance**, following M1 (Figure 2, Table 1).
- M1's handcrafted features-based signature (Equation 1)

PREDICTIVE BIOMARKER	MEAN (± 1 STD) OOB ROC AUC
Radiomic model M1	**0.85 (± 0.09)**
Radiological model M2	**0.83 (± 0.08)**
SUVmax	0.80 (± 0.09)
TLG	0.73 (± 0.11)
Anatomical Tumor Volume	0.69 (± 0.11)
MTV	0.60 (± 0.12)

Table 1: Mean (± 1 std) OOB ROC AUC of models M1, M2, and conventional biomarkers.

− 0.627 (± 0.601) * PET_original_glcm_ClusterShade
− 0.548 (± 0.546) * CT_original_glcm_Correlation
− 0.550 (± 0.782) * PET_original_glcm_InverseVariance
+ 2.76 (± 0.895) * PET_original_fo_Skewness
− 1.58 (± 0.690) * PET_original_gldm_SmallDepLowGrayLevelEmphasis
− 1.31 (± 0.771) * PET_original_gldm_GrayLevelNonUniformity
− 0.667 (± 0.539)

Equation 1. Decision function of the bagging logistic model M1 with standard deviation of the coefficients.

+ 1.52 (± 0.803) * SUVmax
+ 0.809 (± 0.577) * log_{10}(hypodense ∪ inactive volume)
− 0.126 (± 0.255)

Equation 2. Decision function of the bagging logistic model M2 with standard deviation of the coefficients.

- **Simple and interpretable surrogate model** (Equation 2) ⇒ **Model interpretability**
- Logistic regression performed using the **coordinate descent** algorithm ⇒ **Algorithm transparency**

⇒ **Fully interpretable** machine learning pipeline and model

- Thanks to the **linear** nature of the ROI-GAP and the decision function, **no distortion of information occurred when backprojecting the model's coefficients at the voxel level.** ⇒ **Probabilistic quantification preserved**

⇒ Our methodology identifies **tumor sub-regions associated to the prediction task** and thus permits to **highlight predictive patterns at the voxel level.** ⇒ **Voxel-level mapping of the model**

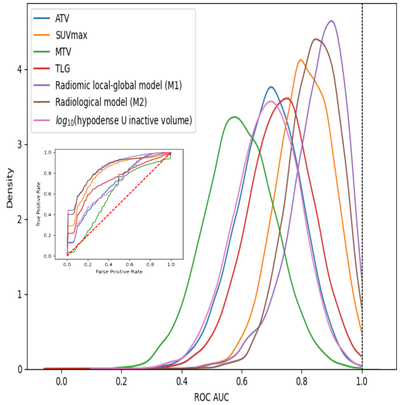

Figure 2: OOB ROC AUC of models and biomarkers for predicting lung metastasis 2 years after the diagnosis. Distribution of the 1000 bootstrap drawings. Due to the small size of the dataset and thus models variance, differences between AUC did not reach p < 0.05 statistical significance (DeLong test p-value: 0.052 to 0.790).

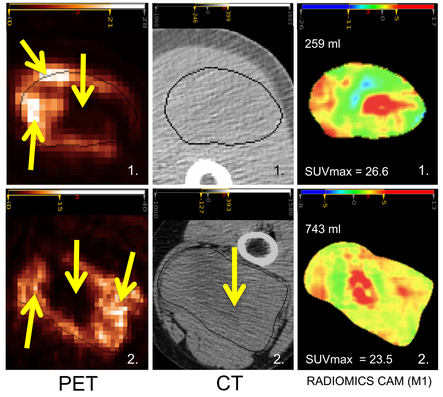

Figure 3: Axial slice examples of voxel-level model mapping with PET, CT, and the radiomics CAM of the model M1 for two patients. Yellow arrows indicate the M1-highlighted biological patterns : localized high FDG uptakes, homogeneous inactive region, tumor hypodensities.

Conclusions

- We describe a method based on **locally-calculated handcrafted radiomic features** to **highlight the sub-regions and biological signal driving the model predictions.**

⇒ In a situation where the number of **data is limited**, we demonstrate how that method makes it possible to spatially and quantitatively **interpret radiomic models and design simple and robust biomarkers amenable to a biological interpretation** for patient stratification.

- This approach is **applicable to any question compatible with image-based classification and prediction.**

- **Detailed pipeline:**

M1 HANDCRAFTED-RADIOMICS CAM

1 FEATURE MAPS COMPUTATION

Voxel-wise computation of 154 **handcrafted** radiomic feature maps using a 9³-voxel kernel with the **IBSI** [2] compliant software **Pyradiomics** [3]

2 "ROI-GLOBAL-AVERAGE-POOLING"

Average aggregation of voxels in each feature map within each tumor region to yield one **global 154-feature vector** per patient (ROI-GAP)

3 MULTICOLLINEARITY REDUCTION-BASED FEATURE SELECTION

ROI-GAP **feature selection** by iterative removal of the ones with the **highest Variance Inflation Factor (VIF)** [4] until the **max VIF** in the feature set is **< 10** using the **R** package **car** [5]

4 LASSO REGULARIZED LOGISTIC REGRESSION

Grid-search (200 × 5 folds repeated stratified) **cross-validation tuning** of the **regularization term**, and **ROC AUC**-based **forward stepwise selection** using the **Python** package **scikit-learn** [6]

5 BAGGING FINAL MODEL

Bootstrap aggregation of the **decision functions** using **1000 bootstrap drawings** to build the **final radiomic model M1**, and comparison of the Out-Of-Bag (OOB) **ROC AUC** with usual biomarkers **SUVmax, Metabolic Tumor Volume (MTV), Total Lesion Glycolysis (TLG)**, and **Anatomical Tumor Volume (ATV)**

6 BACKPROJECTION OF THE COEFFICIENTS

Backprojection of the **regression coefficients** of M1's bagging decision function at the **voxel level**

Figure 1: Detailed pipeline of the radiomics CAM calculation.

JNM The Journal of Nuclear Medicine 2021; 62: 1404 (supplement 1) ASSOCIATED ABSTRACT: JNM SNMMI DOWNLOAD THE POSTER (PDF)

References

[1] M. Vallières *et al.*, "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities," *Phys. Med. Biol.*, 2015.

[2] A. Zwanenburg *et al.*, "The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping," *Radiology*, 2020.

[3] J. J. M. Van Griethuysen *et al.*, "Computational radiomics system to decode the radiographic phenotype," *Cancer Res.*, 2017.

[4] J. Fox *et al.*, "*Applied Regression Analysis and Generalized Linear Models*," *3rd Edition, Sage*, 2016.

[5] J. Fox and S. Weisberg, "*An R Companion to Applied Regression*," *Third Edition, Sage*, 2018.

[6] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, 2011.

[7] L. Guillou *et al.*, "Comparative study of the National Cancer Institute and French Federation of Cancer Centers Sarcoma Group grading systems in a population of 410 adult patients with soft tissue sarcoma," *J. Clin. Oncol.*, 1997.

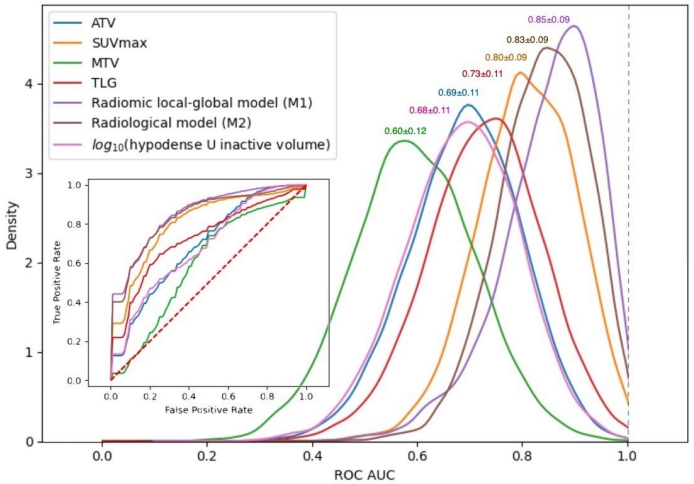

Figure 1. Distributions of the 1000 bootstrap ROC AUC of OOB models' predictions and imaging biomarkers for predicting lung metastasis 2 years after diagnosis.

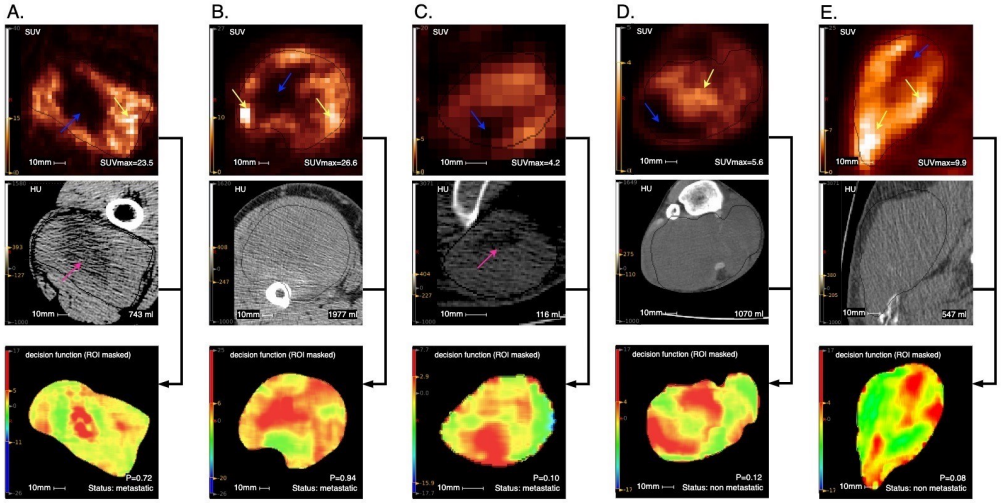

Figure 2. Axial slices examples of voxel-level model mapping.
First row: FDG PET, second row: CT, third row: radiomics CAM.
Arrows indicate the biological patterns highlighted by the radiomics CAM:
Yellow: high FDG uptakes, blue: homogeneous inactive regions, pink: tumoral hypodensities.
P: predicted probability of developing lung metastasis within two years after diagnosis.
A: patient 17 (True positive: Metastatic, P=0.72),
B: patient 9 (True positive: Metastatic, P=0.94),
C: patient 18 (False negative: Metastatic, P=0.10),
D: patient 49 (True negative: Non metastatic, P=0.12),
E: patient 7 (True negative: Non metastatic, P=0.08).

- 0.627 (± 0.601) * PET_original_glcm_ClusterShade
- 0.548 (± 0.546) * CT_original_glcm_Correlation
- 0.550 (± 0.782) * PET_original_glcm_InverseVariance
+ 2.76 (± 0.895) * PET_original_fo_Skewness
- 1.58 (± 0.690) * PET_original_gldm_SmallDepLowGrayLevelEmphasis
- 1.31 (± 0.771) * PET_original_gldm_GrayLevelNonUniformity
- 0.667 (± 0.539)

Equation 1. Decision function of the bagging logistic model M1 with standard deviation of the coefficients.

+ 1.52 (± 0.803) * SUVmax
+ 0.809 (± 0.577) * log_{10}(hypodense ∪ inactive volume)
- 0.126 (± 0.255)

Equation 2. Decision function of the bagging logistic model M2 with standard deviation of the coefficients.

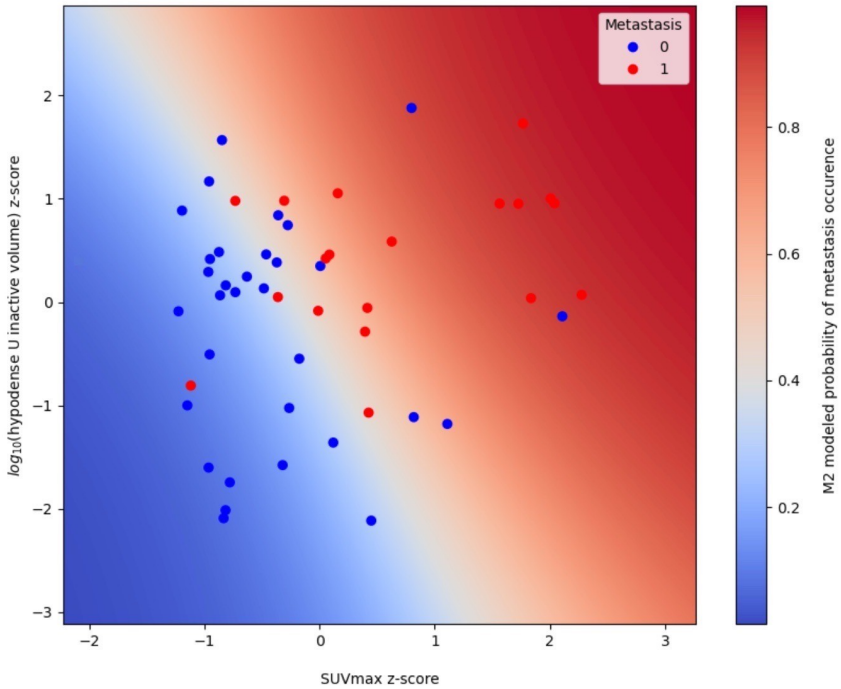

Figure 3. Bagging radiological model M2's logistic regression decision surface and overlayed 2-features z-score standardized scatter plot.