

# Blind Source Separation in Dynamic Cell Imaging Using Non-Negative Matrix Factorization Applied to Breast Cancer Biopsies

Diana Mandache<sup>\*,1,2</sup>, Emilie Benoit<sup>1</sup>, Jean-Christophe Olivo-Marin<sup>2</sup>, Vannary Meas-Yedid<sup>2</sup>

<sup>1</sup> LLTech SAS    <sup>2</sup> Biological Image Analysis Unit, Pasteur Institute

\* [diana.mandache@pasteur.fr](mailto:diana.mandache@pasteur.fr)

International Symposium on Biomedical Imaging  
IEEE - ISBI 2021



# Overview

1 Clinical Context

2 Non-invasive Histopathology

3 Signal Decomposition

4 Classification & Features

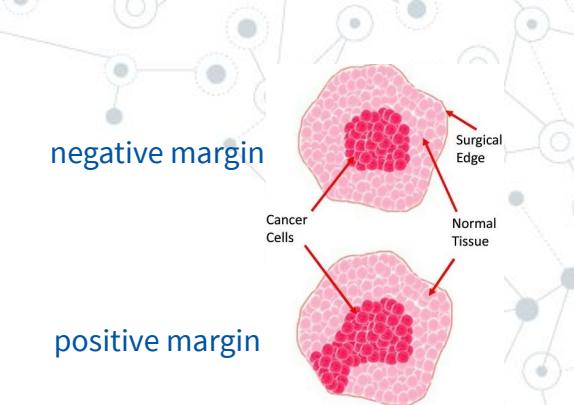
5 Conclusion

# Clinical Context & Motivation

Improve experience of breast cancer patients :

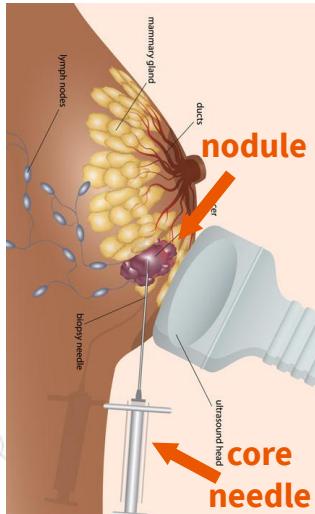
1. Immediate diagnostic after biopsy procedure
2. Surgical margin assessment for breast conserving surgery

negative margin

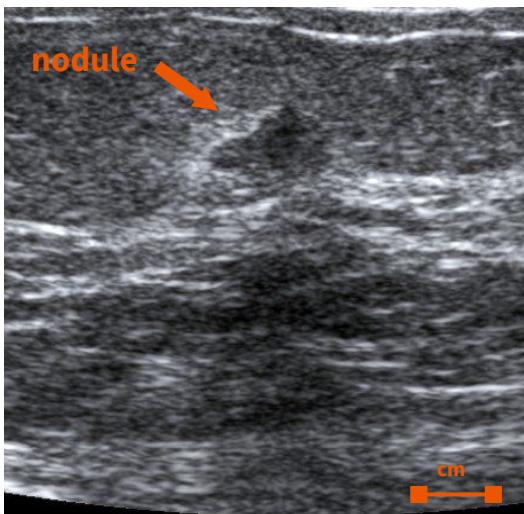


positive margin

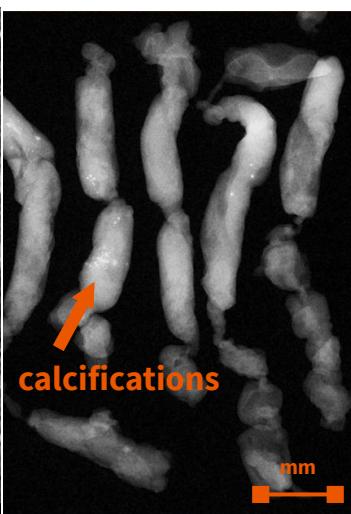
breast biopsy



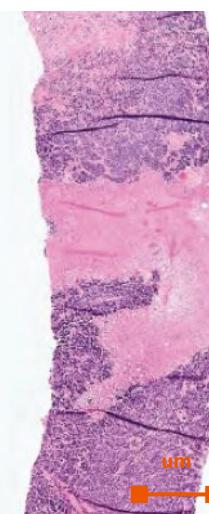
ultrasound guidance



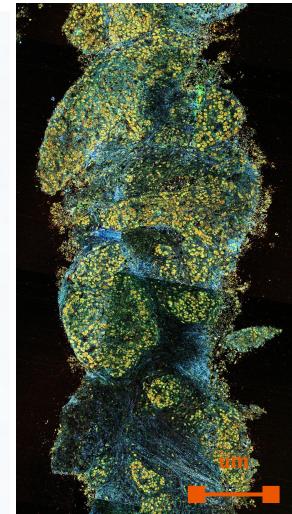
specimen radiography



specimen histology



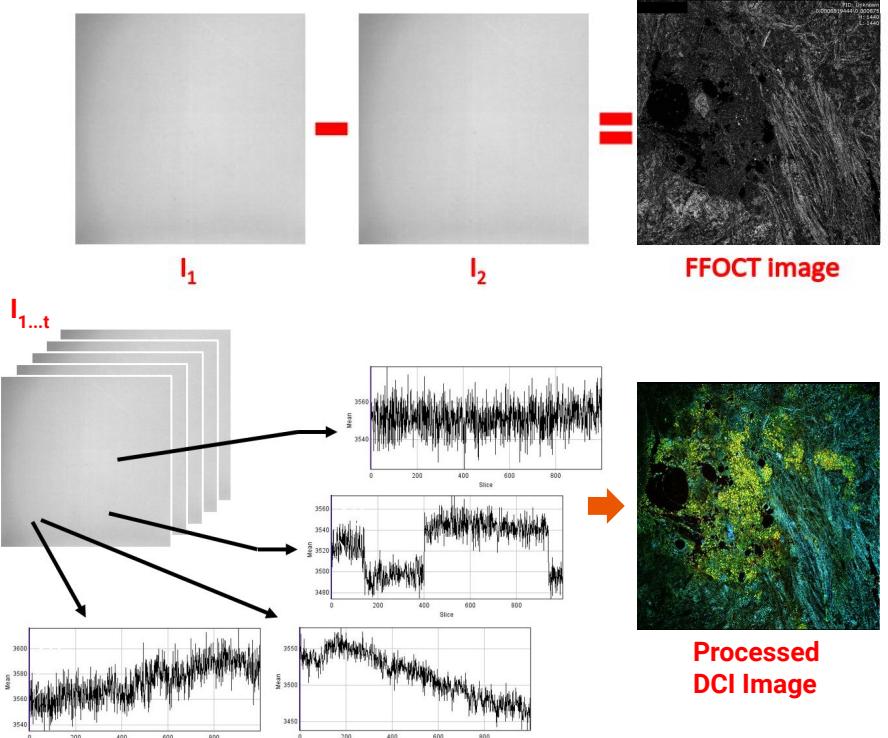
specimen DCI



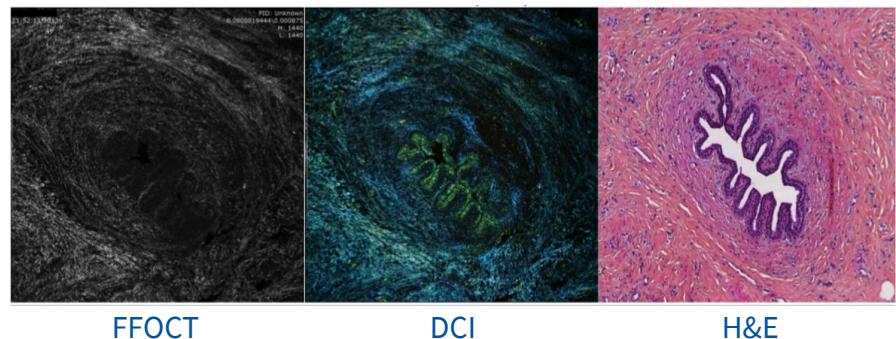
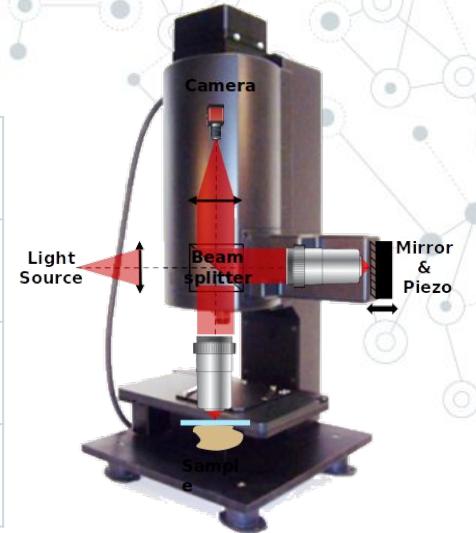
⌚ days

⌚ minutes 3/12

## FFOCT and DCI

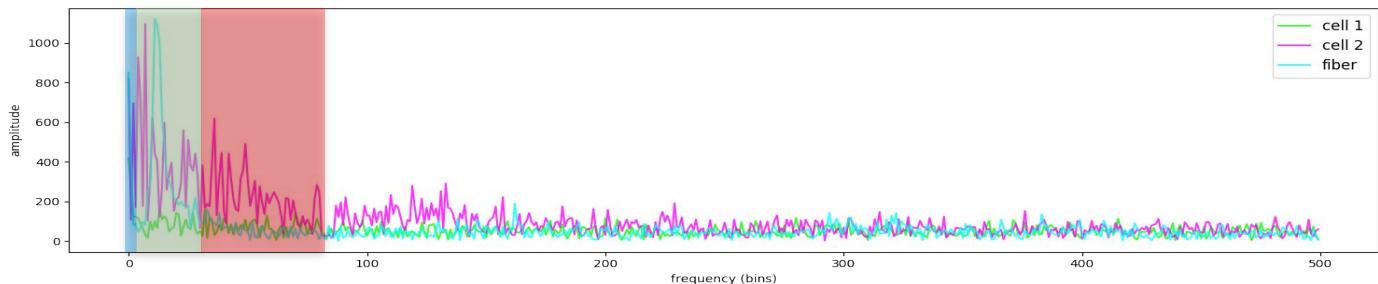
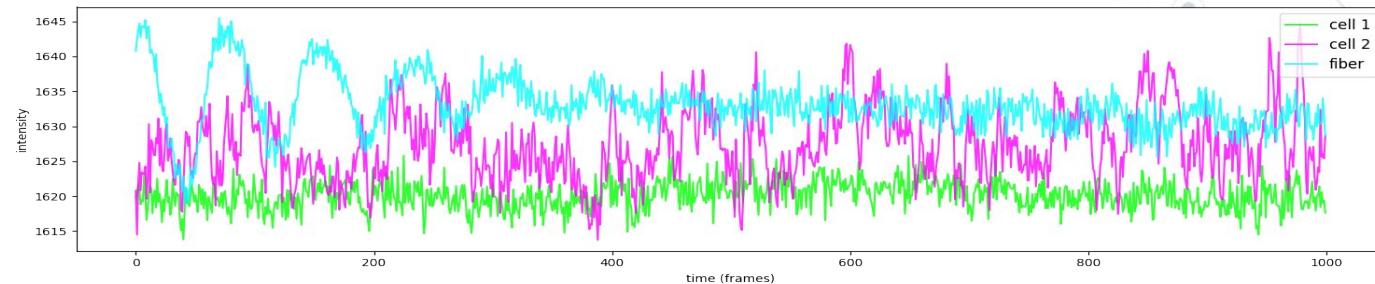
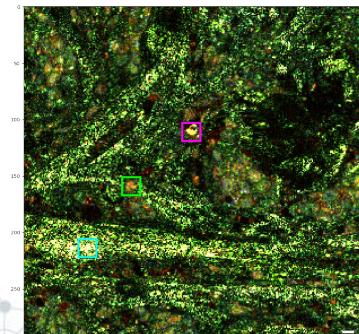
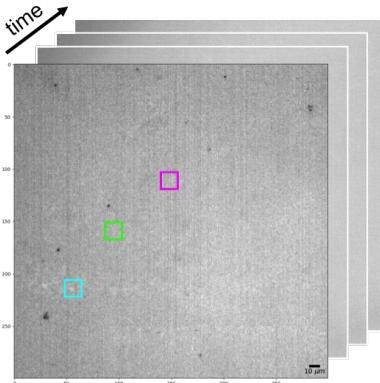


XYZ Resolution	< $1 \mu\text{m}^3$ 10X objective
Acquisition Speed	1 min/cm <sup>2</sup> FFOCT 10 min/cm <sup>2</sup> DCI
Acquisition Frequency	150 Hz DCI
Light Source	$\lambda = 565 \text{ nm}$ $\Delta\lambda = 150 \text{ nm}$



[FFOCT] Dubois, Boccara et al. Applied Optics. 2002  
[DCI] Apelian, Boccara et al. Biomedical Optics Express. 2016

## DCI Signal Encoding



Signal integration by frequency bands : [0.15 Hz] [0.3 - 2.55 Hz] [2.7 - 12 Hz]  
Reveals oscillations from 80ms to 6s

## Blind Source Separation with NMF

Separate signals produced by different phenomena :

- intrinsic sample signal vs. external perturbations
- multiple types of scatterers with different dynamics

Non-negative Matrix Factorization = unsupervised part-based decomposition

Every data point in  $\mathbf{X}$  is a linear combination of a composing **basis H** with a corresponding weighting **contribution W**

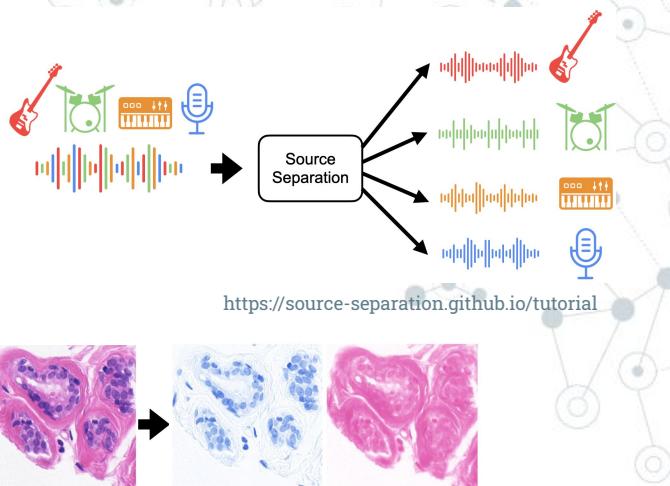
Choice of **rank k** is empirical

$$\min_{W \in \mathbb{R}^{n \times k}, H \in \mathbb{R}^{k \times m}} \|X - WH\|_2 + \alpha_1 \|W\|_2 + \alpha_2 \|H\|_2 \text{ s.t. } W, H \geq 0$$

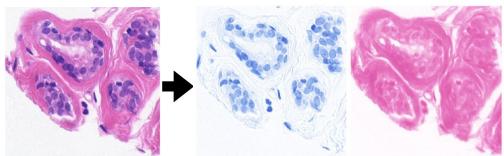
Solved with **Multiplicative Update** rule :

$$W \leftarrow W \cdot \frac{(XH^T)}{(WHH^T)}$$

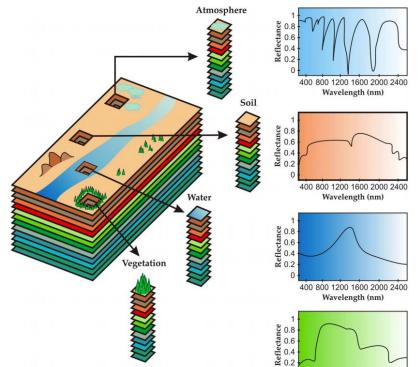
$$H \leftarrow H \cdot \frac{(W^TX)}{(W^TWH)}$$



<https://source-separation.github.io/tutorial>

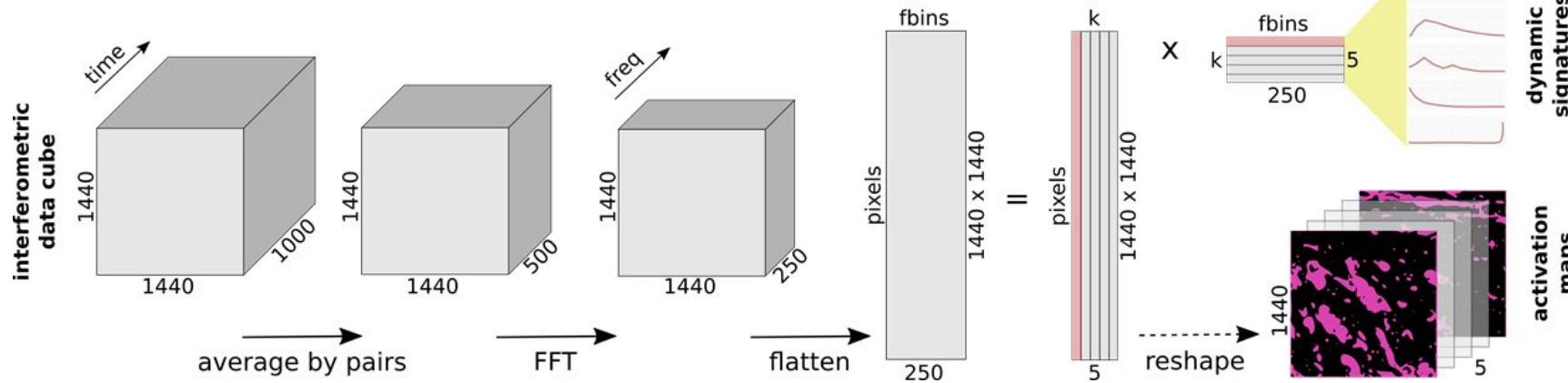


Sha et al. SPIE. 2017



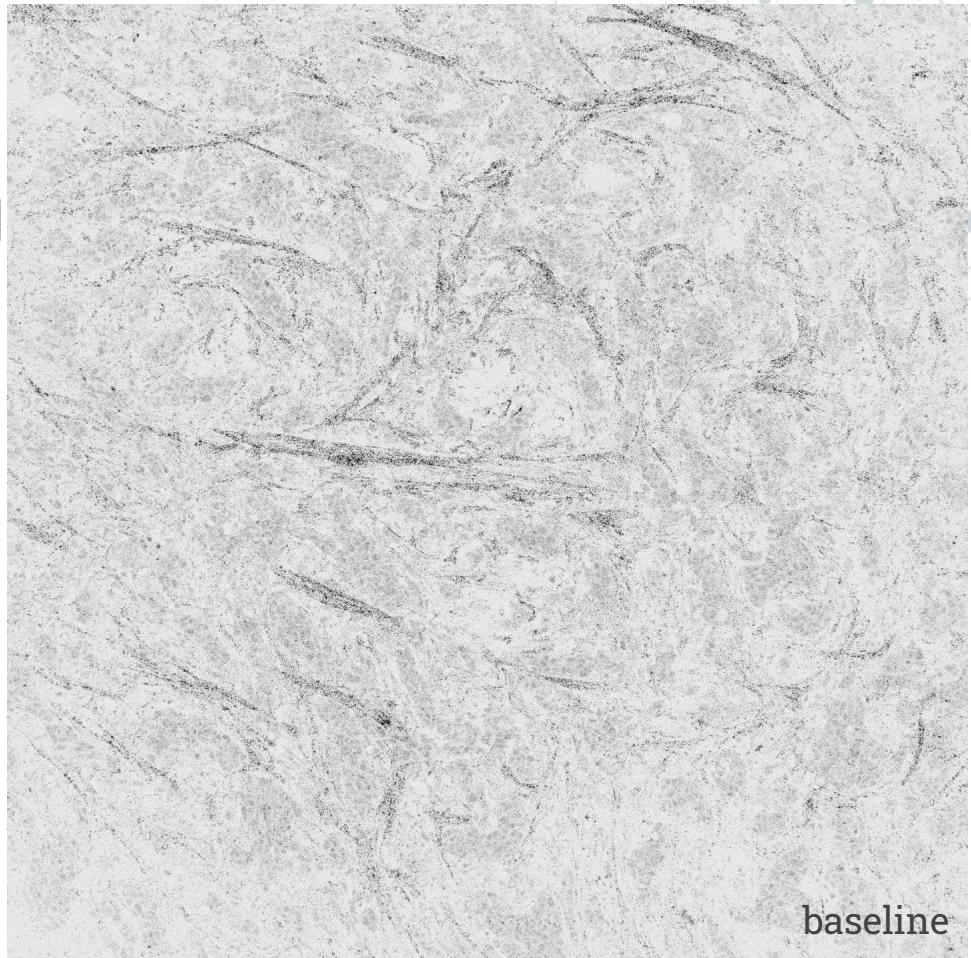
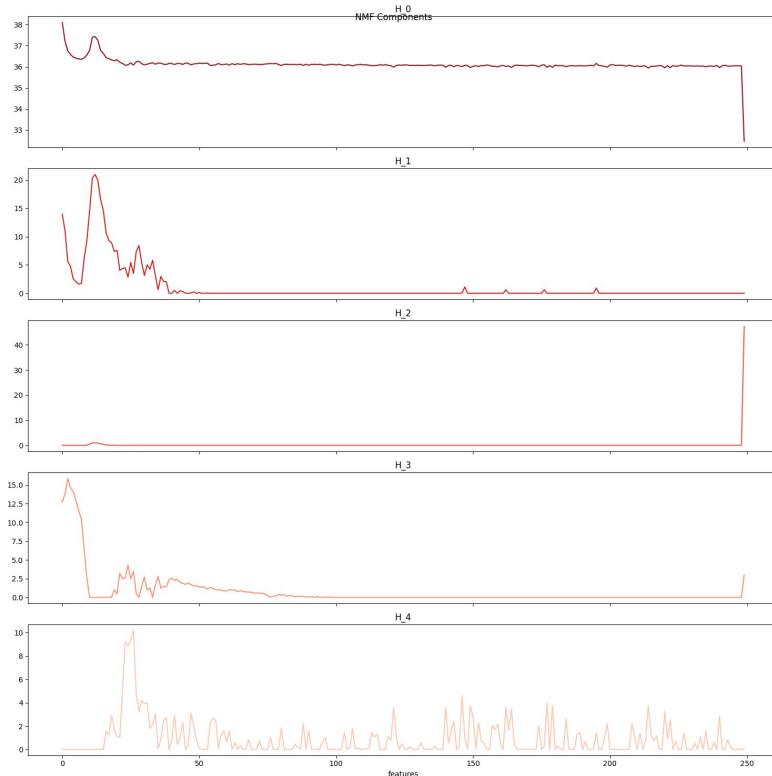
Bioucas et al. JSTARS. 2012

## NMF Workflow on DCI



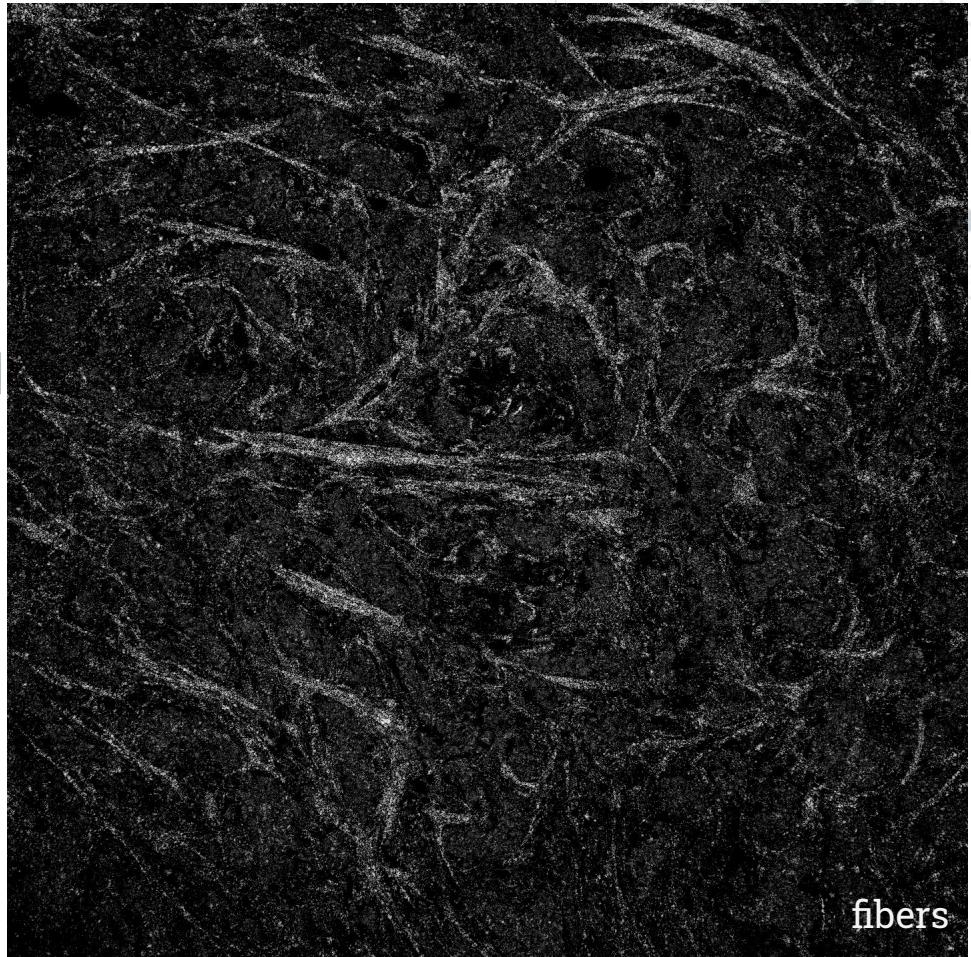
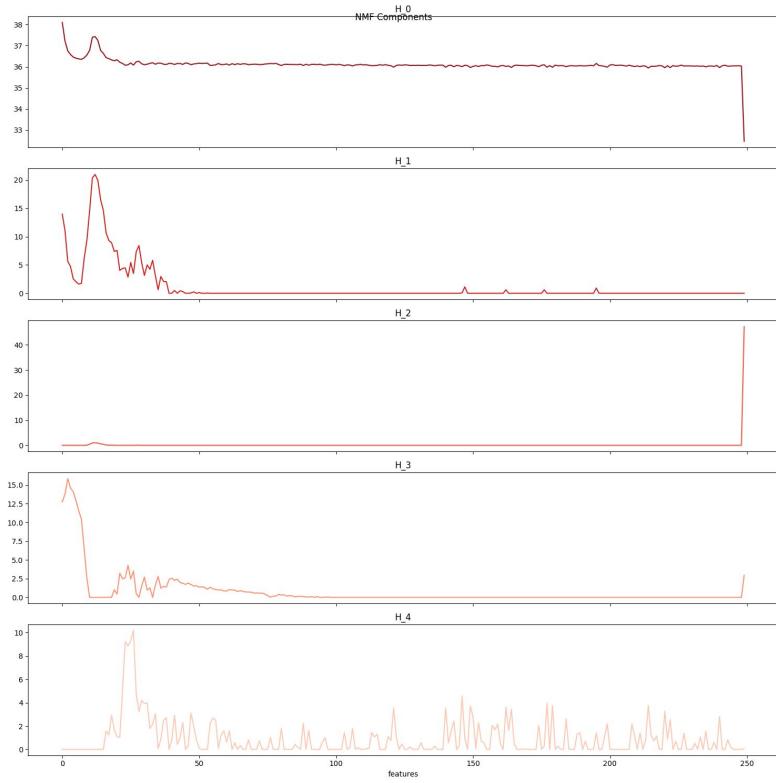
$n = 1440 \times 1440 = 2M$  pixels  
 $m = 250$  frequency bins  
 $k = 5$  components

## NMF Results

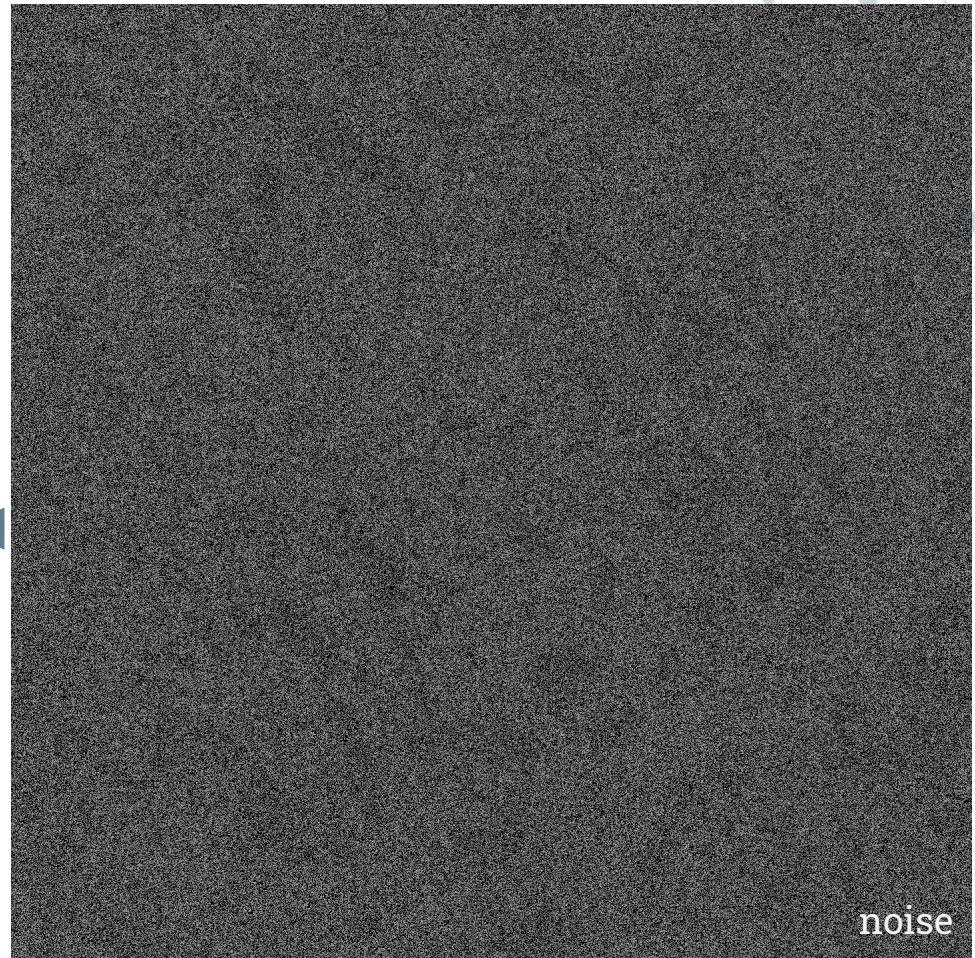
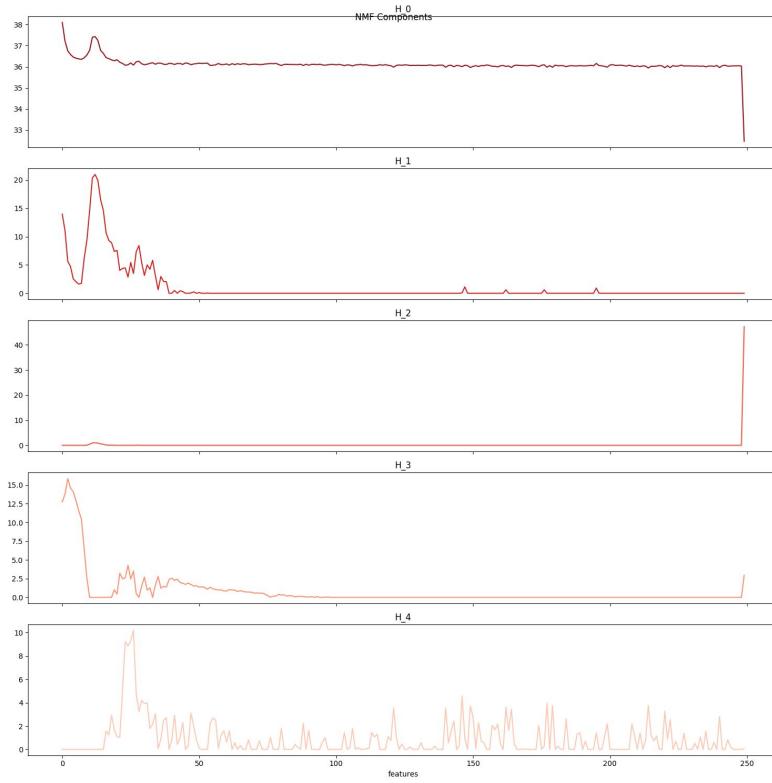


baseline

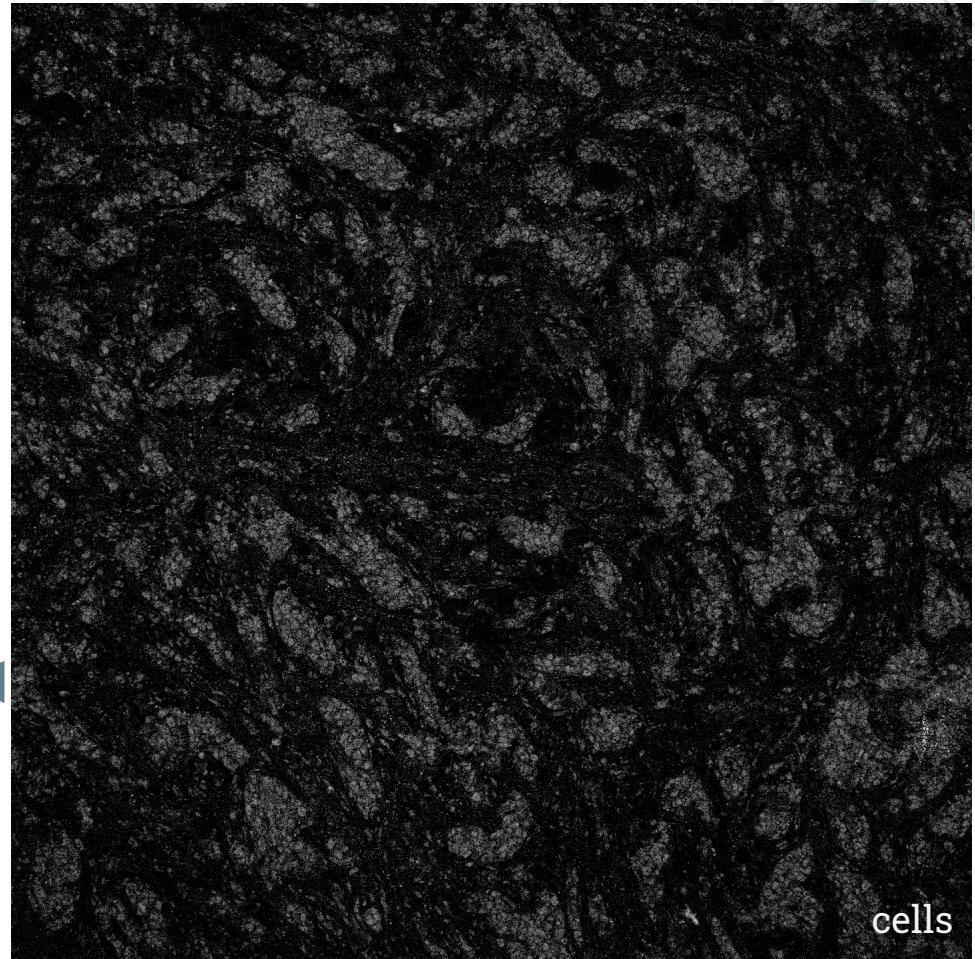
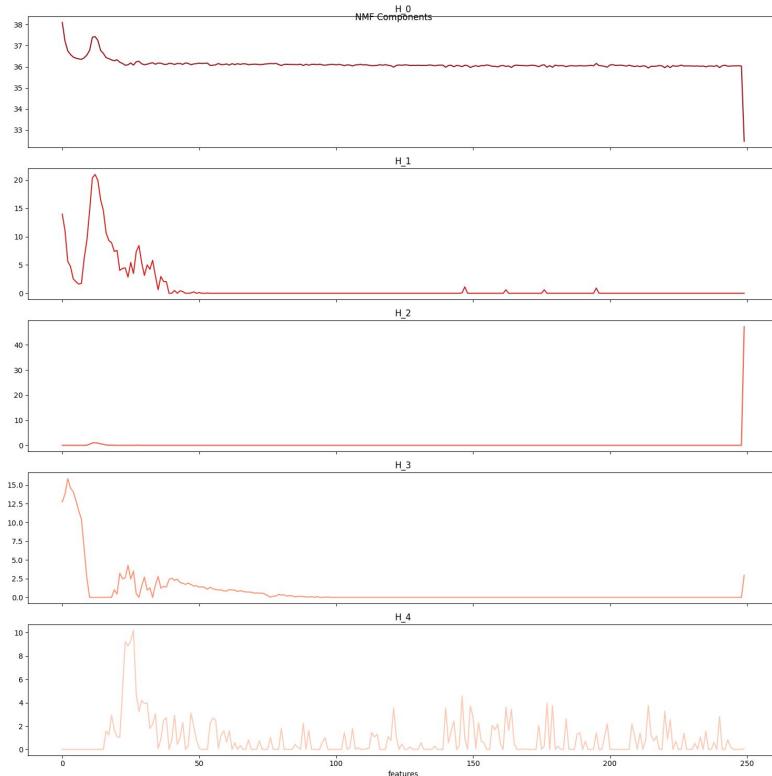
## NMF Results



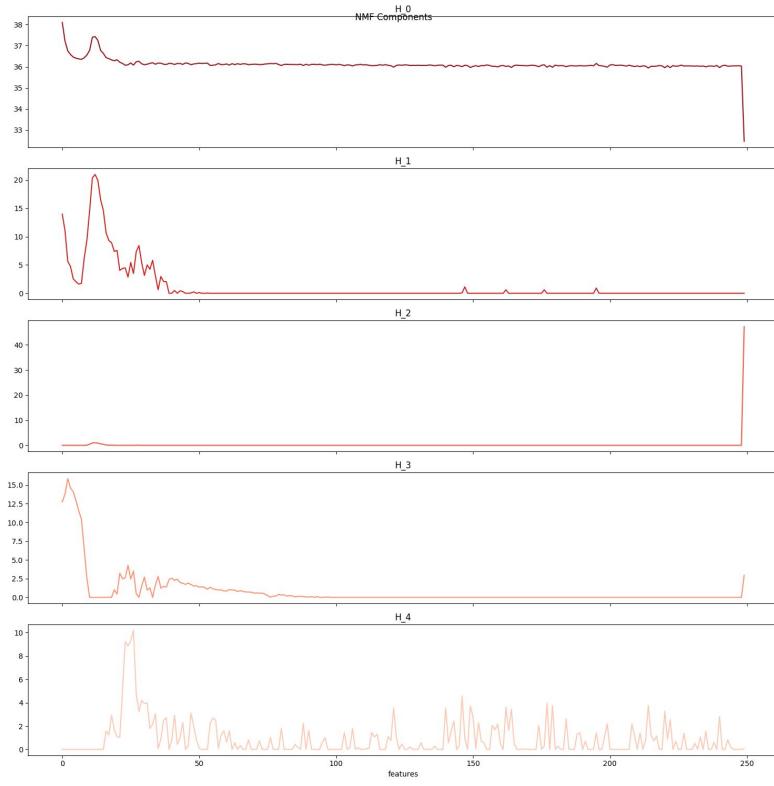
## NMF Results



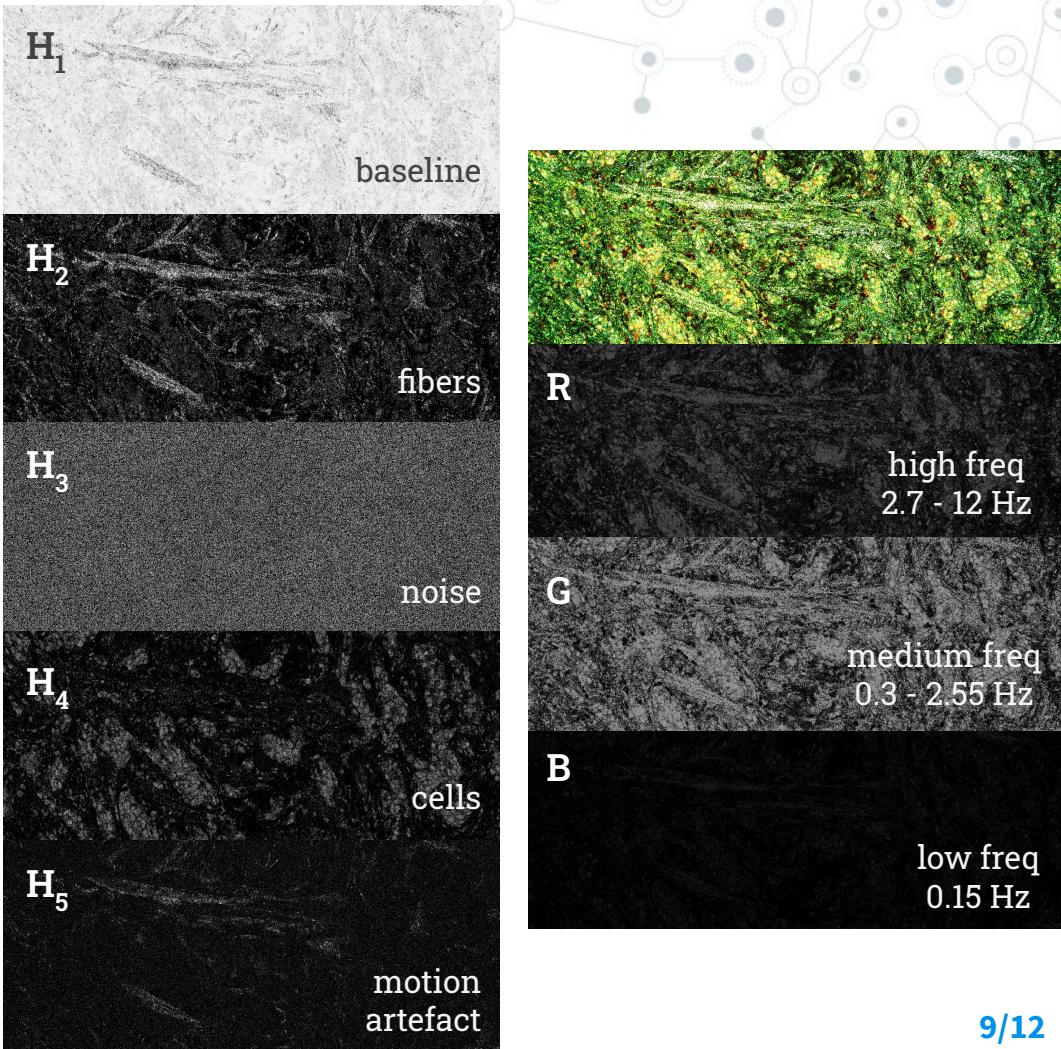
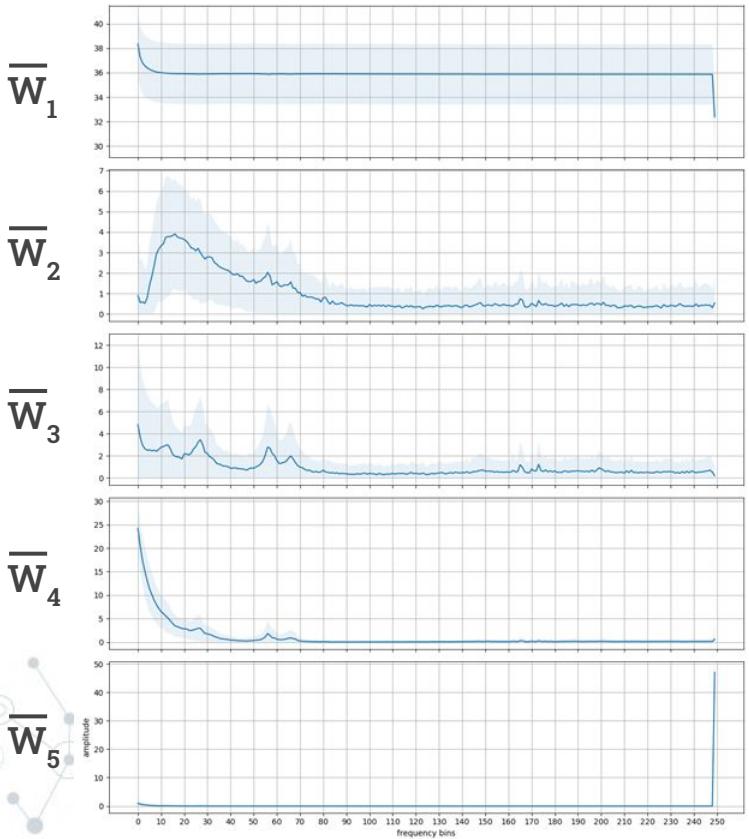
## NMF Results



## NMF Results



## NMF Results



# Classification

Data :

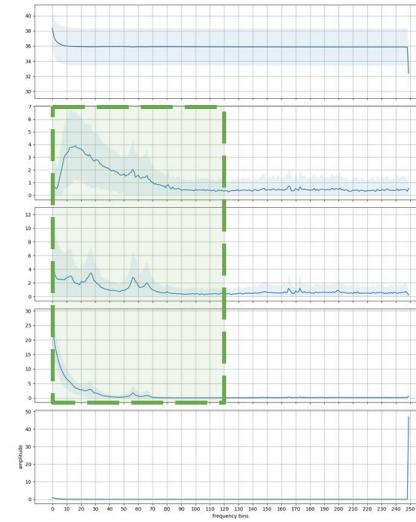
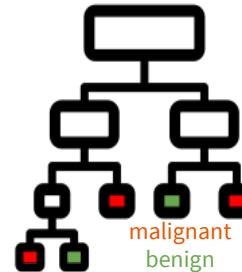
- **382** DCI interferometric cubes (FOVs)
- from **47** samples of **33** patients

Unified feature vector = concatenated  $\frac{3}{5}$  NMF basis components :

- removed baseline and noise
- ordered by energy
- kept lower half of the spectrum  $N_{fbins} = 120$

Training:

- binary classification : **malignant** vs **benign**
- decision tree - based ensemble classifiers
- 4-fold cross-validation
- SMOTE data balancing

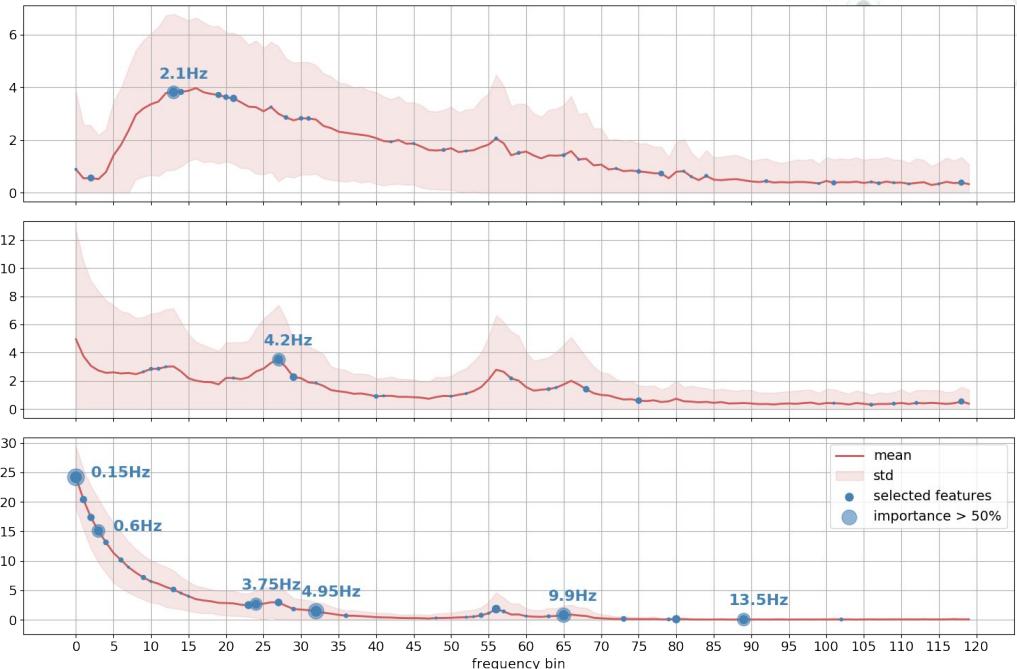


building unified feature vector

Classifier	Train Acc.	Test Acc.	Test Sn.	Test Sp.
AdaBoost	$90.95 \pm 1.39$	<b><math>70.91 \pm 6.38</math></b>	$77.59 \pm 7.21$	<b><math>64.22 \pm 10.37</math></b>
XGBoost	$91.38 \pm 5.66$	$70.69 \pm 5.21$	$82.33 \pm 5.5$	$59.05 \pm 18.48$
RandomForest	$98.13 \pm 0.32$	$65.73 \pm 7.93$	<b><math>83.62 \pm 4.95</math></b>	$47.84 \pm 15.73$
ExtraTrees	$96.77 \pm 1.81$	$65.52 \pm 4.00$	$78.45 \pm 5.52$	$52.59 \pm 10.73$
GradientBoosting	$98.42 \pm 0.72$	$65.09 \pm 4.66$	$76.72 \pm 4.64$	$53.45 \pm 10.20$
DecisionTree	$99.93 \pm 0.12$	$57.54 \pm 3.41$	$64.66 \pm 3.11$	$50.43 \pm 4.46$

## Classification

- Feature importance = contribution to final diagnosis, as selected by training boosting trees algorithm
- Reveal salient oscillation frequencies than could be later correlated with biological or external phenomena



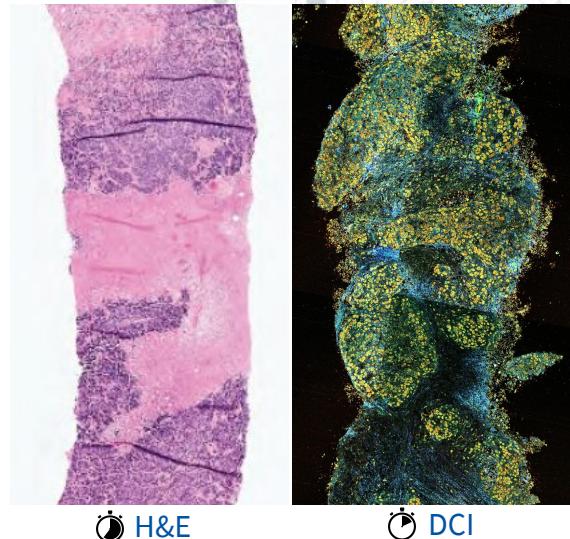
## Conclusion

- **Dynamic FFOCT (aka DCI)** : *non-invasive imaging technique with near-histology resolution and sensitivity*
  - Towards “bedside diagnosis”
- **NMF** : simple yet powerful *source separation algorithm*
  - Towards *quantifiable characterization of cell metabolism*
- Method validated with *70% accuracy* in discriminating malignant from benign breast tissue only based on dynamic signatures

**Perspective** improve accuracy with DL and by including spatial information

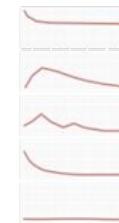
ePoster

Leveraging Global Diagnosis for Tumor Localization in Dynamic Cell Imaging of Breast Cancer Tissue towards Fast Biopsying

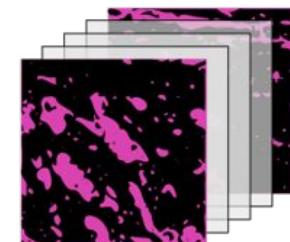


H&E

DCI

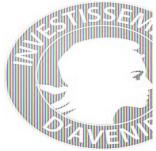


dynamic  
signatures



spatial  
activations

# THANK YOU !



Cifre

anRT  
ASSOCIATION NATIONALE  
RECHERCHE TECHNOLOGIE



Jean-Christophe  
Olivo-Marin  
HEAD OF STRUCTURE



Marie-Anne Lin  
ADMINISTRATIVE STAFF



Vannary Meas-  
Yedid Hardy  
RESEARCH ENGINEER



Elisabeth  
Labrûère  
PERMANENT RESEARCHER



Nancy Guillen  
VISITING SCIENTIST



POST-DOC



Maria MANICH  
RESEARCH ENGINEER



Bertrand  
Le Conte De Poly  
CEO - FOUNDER



Prof. Claude Bocvara  
CSO - FOUNDER



Zoya Volynskaya  
DIRECTOR BUSINESS  
DEVELOPMENT



Diana Mandache  
PHD STUDENT



Marion Louveaux  
RESEARCH ENGINEER



Amandine  
Tournay  
UNDERGRADUATE  
STUDENT



Alba  
PETRACCINI  
GRADUATE STUDENT



Stéphane  
Dallongeville  
RESEARCH ENGINEER



Thibault Lagache  
PERMANENT RESEARCHER



Suvadip  
Mukherjee  
POST-DOC



François Grammont  
DESIGN AND PRODUCTION  
MANAGER



Émilie Benoît  
PRODUCT DEVELOPMENT  
MANAGER



Alban Mounier  
SOFTWARE DEVELOPMENT  
MANAGER



Aleix Boquet-  
Pujadas  
POST-DOC



Daniel Felipe  
González  
Obando  
RESEARCH ENGINEER



Pascal Bochet



Robin  
Chalumeau  
PHD STUDENT



Alexandre  
Bouyssoux  
PHD STUDENT



Samuel Kubler  
PHD STUDENT



Yekta Kesenci  
PHD STUDENT



Erwan Dereure  
PHD STUDENT



Antoine Habis  
PHD STUDENT



Institut Pasteur

Biological  
Image  
Analysis  
Unit

LLTech  
LIGHT FOR LIFE TECHNOLOGIES

diana.mandache@pasteur.fr

