Supervised Machine Learning (SML), a Branch of Artificial Intelligence, in Microscopic Image Segmentation and Measurement

Journées SOFFOET 24/09/2021

Christo Christov
Histology Department
INSERM U1256, NGERE
Laboratory of Fetal and Placental Pathology

Nancy

christo.christov@univ-lorraine.fr

Overview

- machine learning (ML) in the context of artifical intelligence
- microscopic image segmentation in the context of supervised machine learning (SML)
- examples

Journées SOFFOET 24/09/2021

Machine Learning (ML) as a Branch of Artifical Intelligence

Journées SOFFOET 24/09/2021

Artifical intelligence (AI)

- symbolic AI
  - learning via formulae and rules
  - produces « logical » conclusions
  - mimics repetitive rigid tasks
  - programming (coders)

- subsymbolic AI
  - learning via experience
  - produces associative results
  - mimics the human brain
  - machine learning

- in-between methods

Journées SOFFOET 24/09/2021

Table 1
Symbolic vs Sub-symbolic methods characteristics

Symbolic	Sub-symbolic
Symbols	Numbers
Logical	Associative
Serial	Parallel
Reasoning	Learning
von Neumann machines	Dynamic Systems
Localised	Distributed
Rigid and static	Flexible and adaptive
Concept composition and expansion	Concept creation, and generalization
Model abstraction	Fitting to data
Human intervention	Learning from data
Small data	Big data
Literal/precise input	Noisy/incomplete input

Ilkou & Koutraki. Proceedings of the CIKM 2020 Workshops

Supervised Machine Learning (SML)

Symbolic AI → A paradigm with high explainability but low accuracy performance

- **A cat:**
  - **has 4 paws**
  - **has 2 ears**
  - **has long moustaches**
  - **purs**
  - **meows**
  - **claws**
  - **cuddles up**
  - **drinks milk**
  - **may piss in shoes**
  - **………**
  - **………**
  - **ad infinitum???**

Subsymbolic AI → A paradigm with low explainability but high accuracy performance

- **A cat:**
  - **show some images**

Journées SOFFOET 24/09/2021

BUSINESS

Why coders love the AI that could replace them

https://www.bbc.com/ 07.09.2021

Journées SOFT-OET 24/09/2021

Check for updates

OPINION ARTICLE

Developing open-source software for bioimage analysis: opportunities and challenges [version 1; peer review: 2 approved]

Florian Levet[1,2], Anne E. Carpenter[3], Kevin W. Eliceiri[4], Anna Kreshuk[5], Peter Bankhead[6], Robert Haase[7]

Journées SOFFOET 24/09/2021

Table 1. Size, impact and timeline of a selection of open-source tools.

	Type	Total lines of code	Commits in 2020	Citations in 2020	Development time since project started (in months)	Timeline (start-end)
JACoP[15]	Plugin	2,400	3	358	8	[2005-ongoing]
SR-Tesseler/ PoCA	Software	100,800	1	56	75	[2012-ongoing]
clij/clij2/ assistant	Library	100,000	2,500	12	20	[2018-ongoing]
QuPath	Software	110,000	570	655	60	[2016-ongoing]
ilastik	Software	155,000	910	442	200	[2011-ongoing]
CellProfiler	Software	280,770	492	1,740	216	[2003-ongoing]
Bio-Formats	Library	1,502,214	573	245	180	[2006-ongoing]
ImageJ/FIJI	Software	2,024,516	2,934	44,400	432	[1997-ongoing]
OMERO	Software	2,171,241	3,667	361	420	[2003-ongoing]
IDR	Repository	16,517,904	1,756	76	180	[2016-ongoing]

Free (or not so free) software packages to perform pixel segmentation by SML

EPFL
BIOMEDICAL IMAGING GROUP (BIG)

Ecole Polytechnique Fédérale de Lausanne Pr M. Unser

CellProfiler™
cell image analysis software

Broad Institute Harvard MIT, Carpenter Lab

ORBIT IMAGE ANALYSIS

Idorsia Pharmaceuticals

QuPath
Quantitative Pathology & Bioimage Analysis

Pathology, Genetics and Molecular Medicine, Edinburgh, P. Bankhead

Robo realm
vision for machines

RoboRealm Inc., Denver, Colorado

WEKA
The University of Waikato

Waikato Environment for Knowledge Analysis, New Zealand

GemIdent

Stanford, Pr S. Holmes

histoCAT
Histology Topography Cytometry Analysis Toolbox

Institute of Molecular Life Sciences, University of Zurich, Pr B. Bodenmiller

CytoMAP
By The Gerner Lab

University of Washington, Seattle

Icy
AN OPEN COMMUNITY PLATFORM FOR BIOIMAGE INFORMATICS

Institut Pasteur, Unité d'Analyse d'Images Quantitative, Pr JC Olivo-Marin

Visualization and Intelligent Systems Laboratory
VISLab

University of California

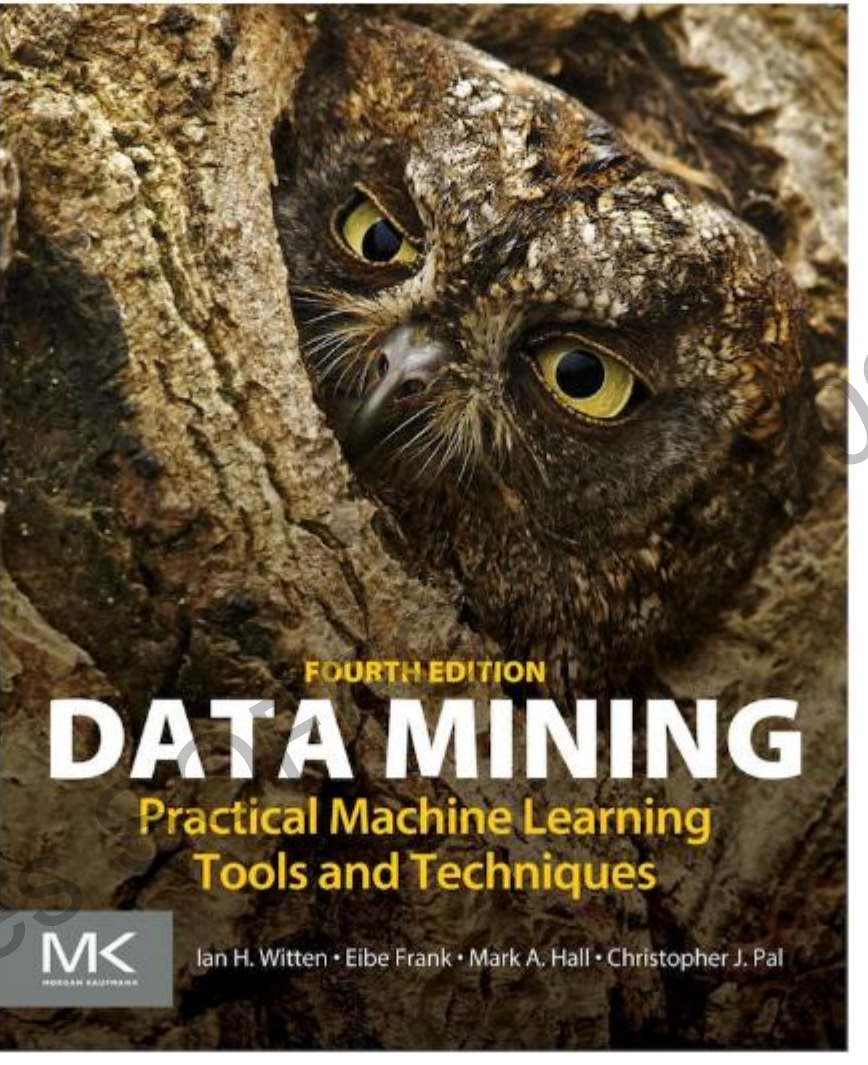
MathWorks®
Image Processing Toolbox

MathWorks [R]

ZEN

Carl Zeiss GMBH

Journées CFFOET 24/09/2021



Waikato Environment for Knowledge Analysis, New Zealand

Journées ... 09/2021

Machine learning models cheat sheet

Supervised learning

Data scientists provide input, output and feedback to build model (as the definition)

EXAMPLE ALGORITHMS:

Linear regressions
- sales forecasting
- risk assessment

Support vector machines
- image classification
- financial performance comparison

Decision tree
- predictive analytics
- pricing

Unsupervised learning

Use deep learning to arrive at conclusions and patterns through unlabeled training data.

EXAMPLE ALGORITHMS:

Apriori
- sales functions
- word associations
- searcher

K-means clustering
- performance monitoring
- searcher intent

Semi-supervised learning

Builds a model through a mix of labeled and unlabeled data, a set of categories, suggestions and exampled labels.

EXAMPLE ALGORITHMS:

Generative adversarial networks
- audio and video manipulation
- data creation

Self-trained Naïve Bayes classifier
- natural language processing

Reinforcement learning

Self-interpreting but based on a system of rewards and punishments learned through trial and error, seeking maximum reward.

EXAMPLE ALGORITHMS:

Q-learning
- policy creation
- consumption reduction

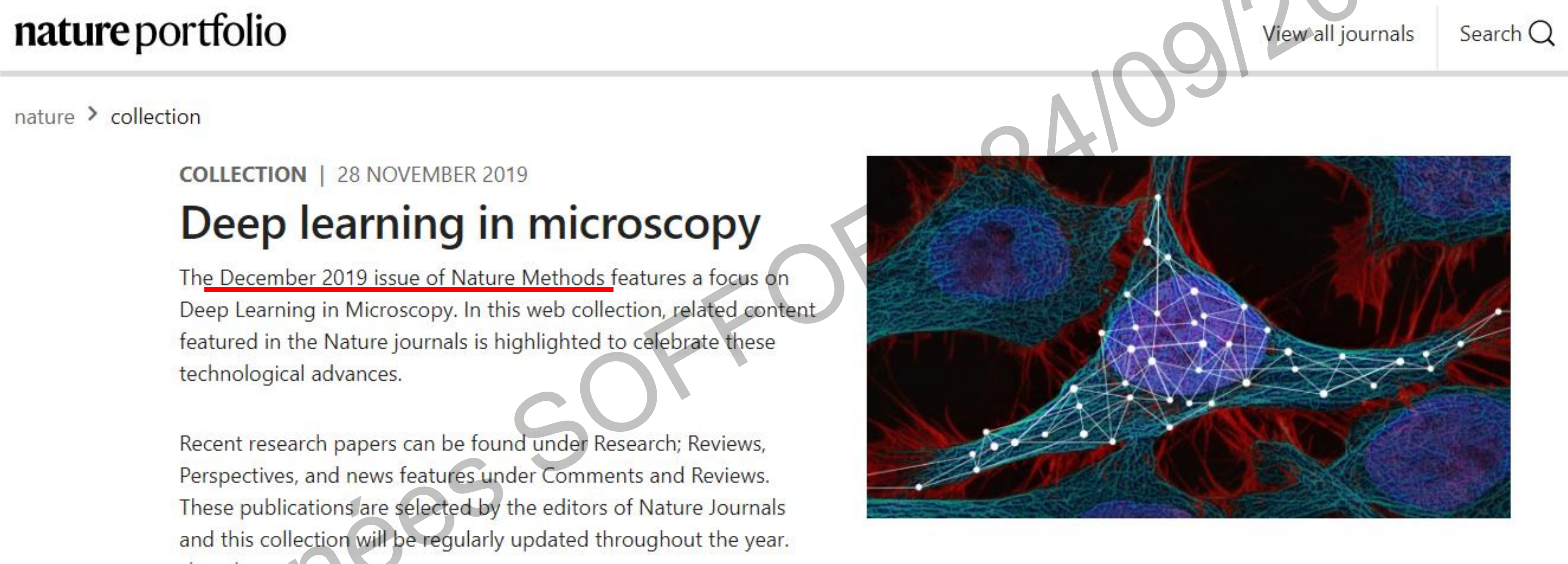
Model-based value estimation
- linear tasks
- estimating parameters

©2020 TECHTARGET. ALL RIGHTS RESERVED TechTarget

Journées SOFFOET 24/09/2021

Deep Learning in Microscopy

Journées SOFOP 24/09/2021



The supervised machine learning (SML) paradigm in image segmentation

1. User trains a classifier on training data set (few images)

2. Trained classifier!!!

3. The trained classifier is applied to measure a real data set (100s or 1000s of images if needed)

4. Final segmentation!!!

Journees SOFFOET 24/09/2021

Vocabulary

- segmentation: identification of objects to be measured
- classifier: a set of rules / properties identifying measurable object
- training data set: few images teaching the machine to identify objects
- real data set: all images to be used for measurement

The supervised machine learning (SML) paradigm in image segmentation combines the best of 2 worlds

1. User trains a classifier on training data set (few images)

→ 2. Trained classifier!!!

→ 3. The trained classifier is applied to measure a real data set (100s or 1000s of images if needed)

→ 4. Final segmentation!!!

The best of 2 worlds:

- users: introduce their natural intelligence (NI) expertise
- SML algorithms: introduce their AI expertise
- SML teaches users about their images

No magic wand: garbage in garbage out!!!

Journees SOFFOET 24/09/2021

Microscopic image segmentation in the context of ML
(segmentation: identification of objects of interest):

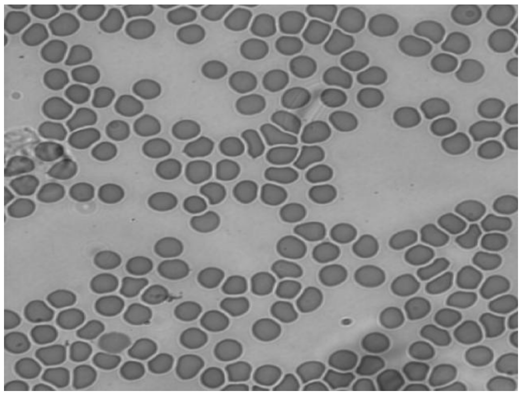
- **semantic (pixels according to their properties)**

- **instance (objects according to their properties) (Deep Learning)**

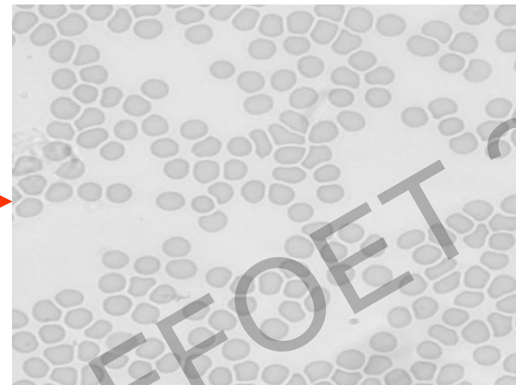
Journées SOFFOET 24/09/2021

Image segmentation: obtain a binary mask (1→4)

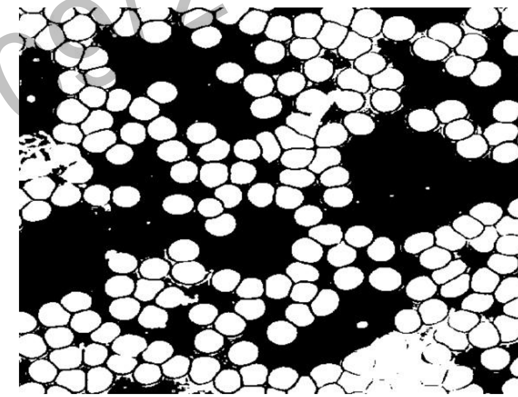
1. Image



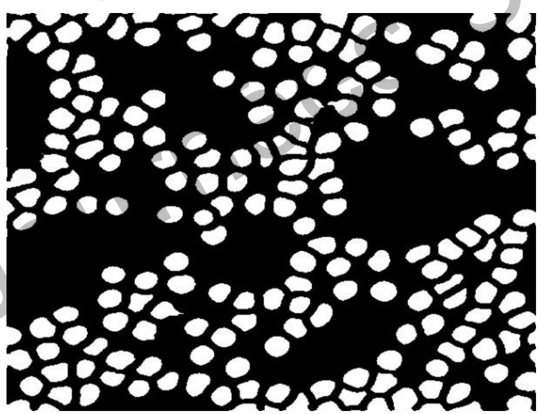
2. Pretreatment (shading correction)



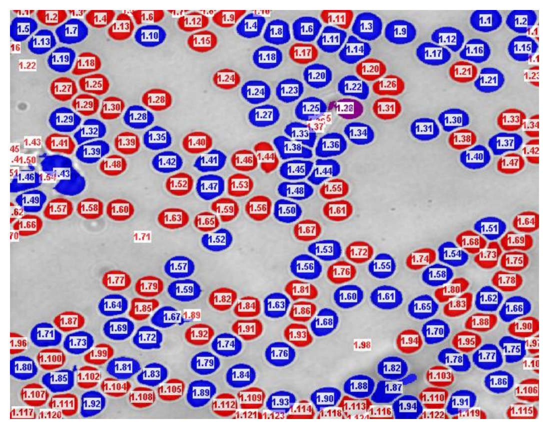
3. Object identification (segmentation)



4. Postreatment



5. Measurement

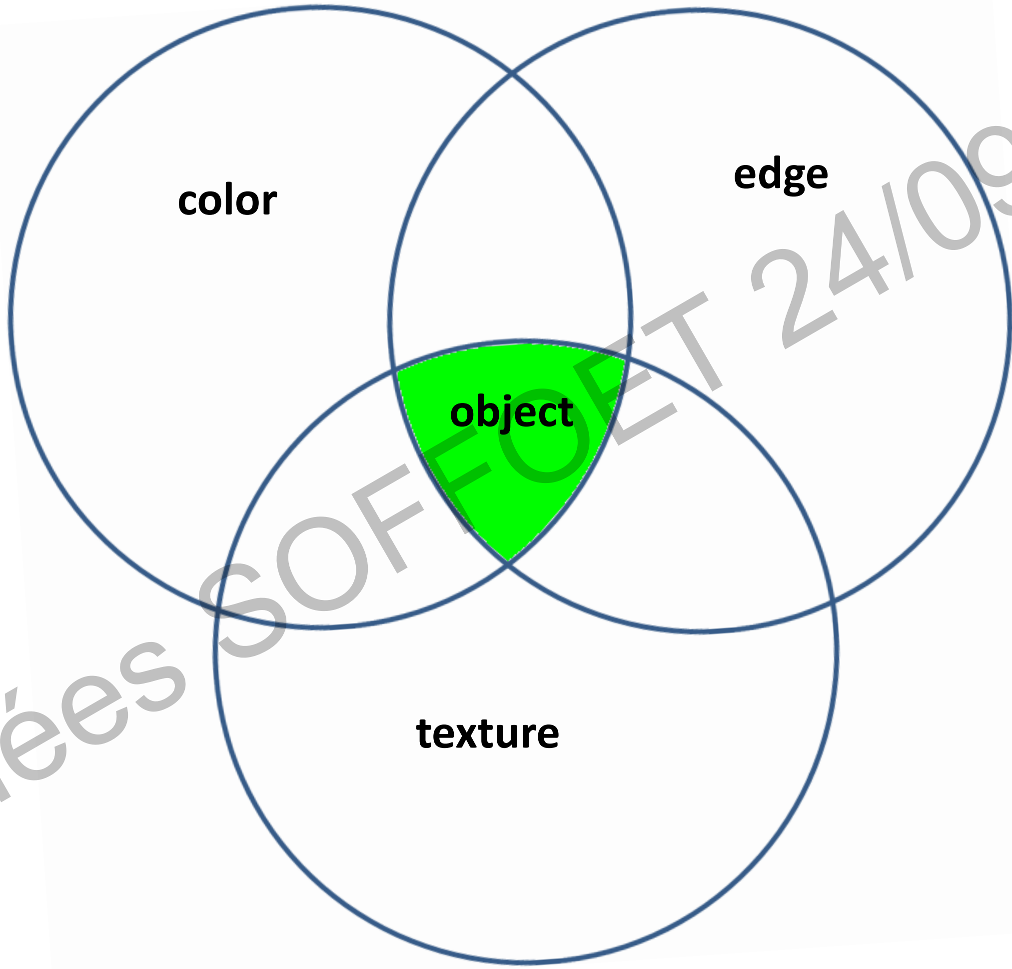


$N_{cells} = 198$

$S_{cell} = 60$ µm2

SOFFOET 24/09/2021

An integrated approach based on color, edge and texture, etc. detection: artificial (robotic) vision



Journées SOFOET 24/09/2021

An integrated approach based on color, edge and texture detection: spectral image analysis, wavelets

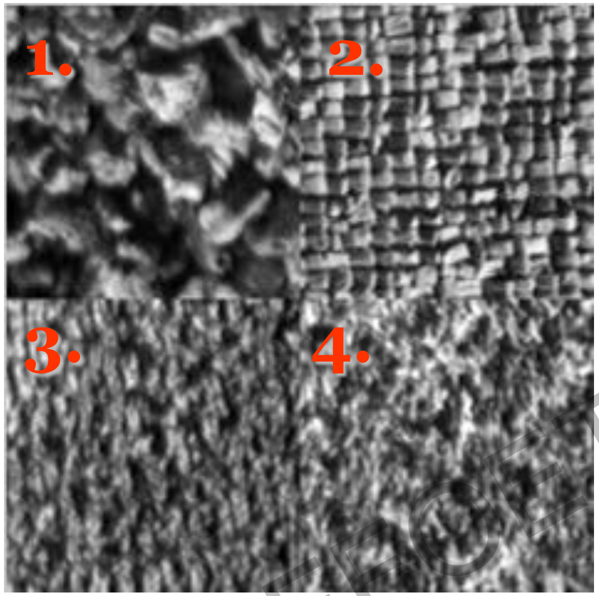
- **set of metrics designed to quantify the perceived structure of an image**

Journees SOFFOET 24/09/2021

A clear definition of texture does not exist, it can be understood as a group of image properties that relate to our intuitive notions of rugosity and smoothness

(Haralick, 1979)

Why and when use texture analysis?



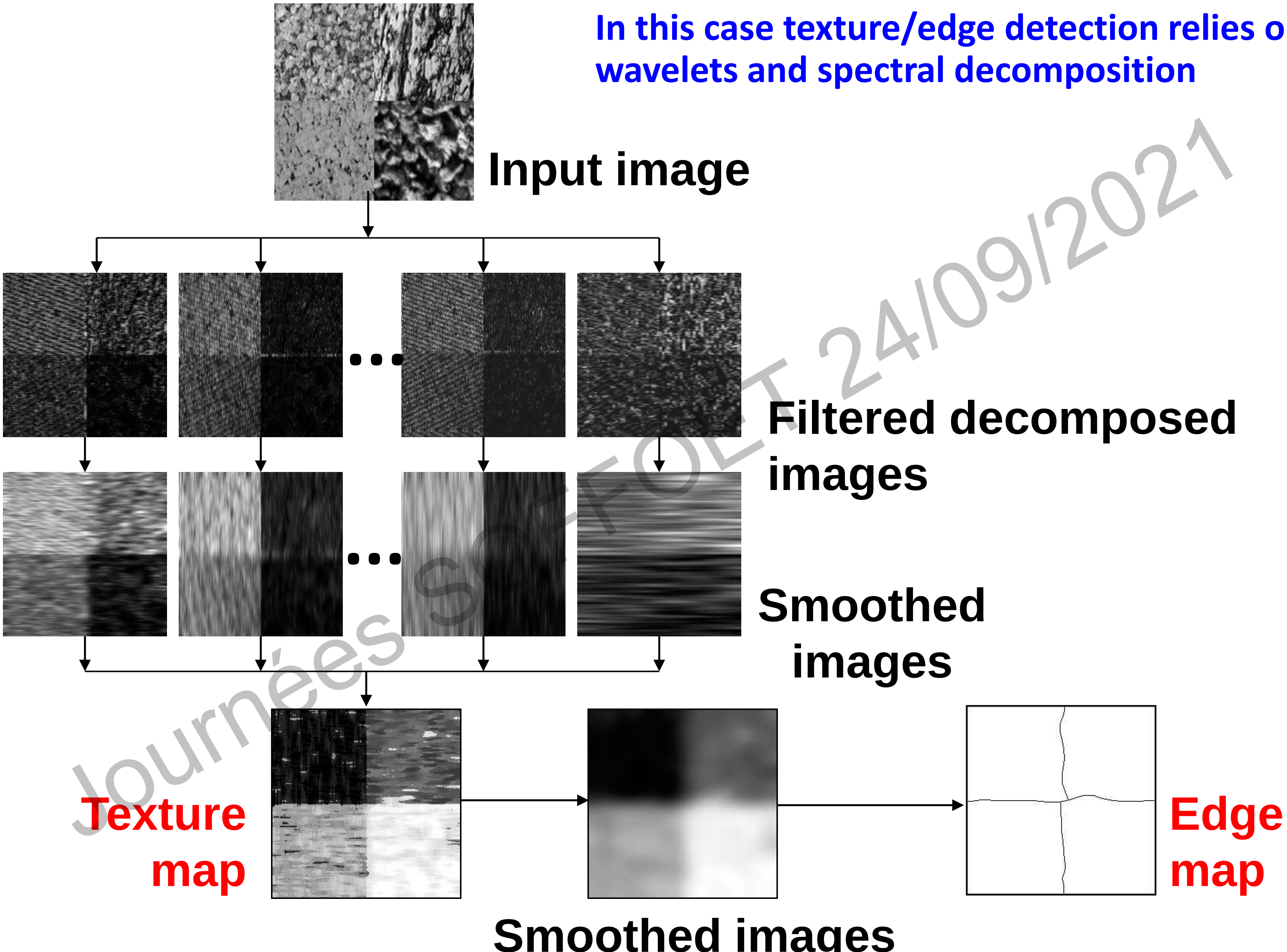
The human eye is able to discriminate between 1-64 textures at a time;

computer programs discriminate 1-4095 textures at a time

SML teaches users about their images

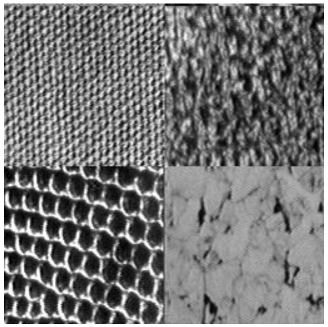
Journées SOFFCOT 24/09/2021

In this case texture/edge detection relies on wavelets and spectral decomposition

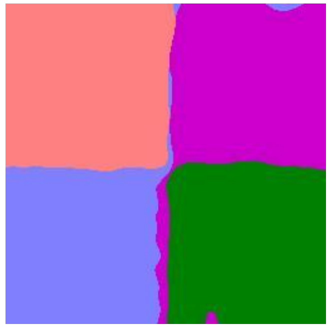


Journées SOFOLT 24/09/2021

Integrating texture and edge information for segmentation: SML may use hundreds of filters impenetrable to the operator

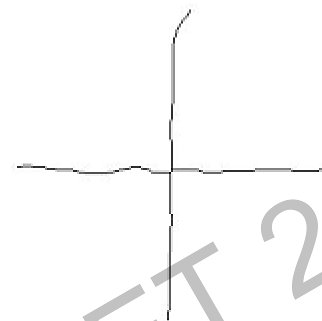


Original image

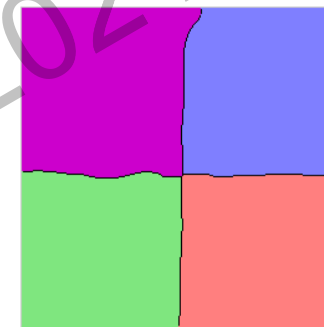


Texture map

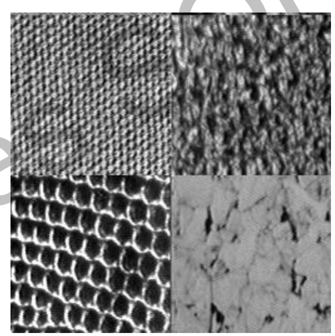
+



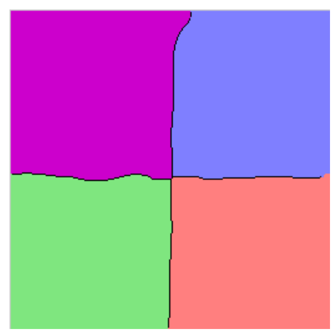
Edge map



Final map



Original image



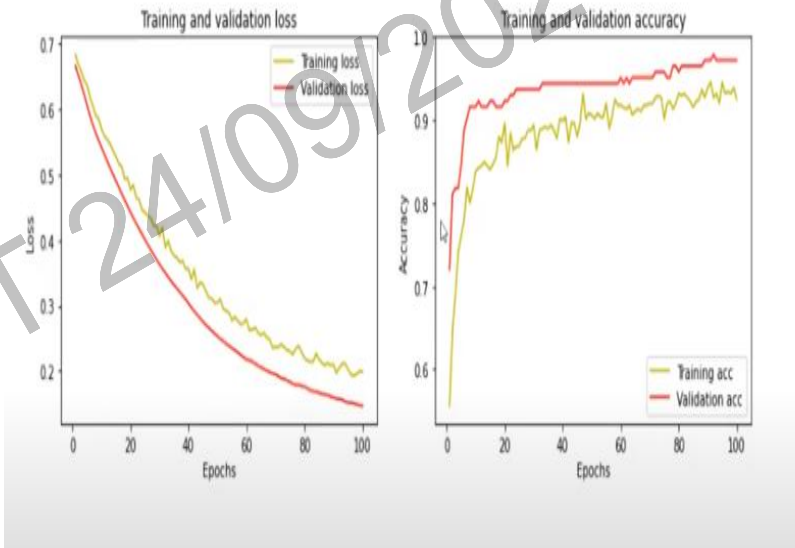
Final map

Supervised machine learning → A paradigm with low explainability but high accuracy performance

Journée OFFOET 24/09/2021

Integrating texture and edge information for segmentation: SML may use hundreds of filters impenetrable to the operator

symmetry_mean	569.0	1.811619e-01	...	1.957000e-01	3.040000e-01
fractal_dimension_mean	569.0	6.279761e-02	...	6.612000e-02	9.744000e-02
radius_se	569.0	4.051721e-01	...	4.789000e-01	2.873000e+00
texture_se	569.0	1.216853e+00	...	1.474000e+00	4.885000e+00
perimeter_se	569.0	2.866059e+00	...	3.357000e+00	2.198000e+01
area_se	569.0	4.033708e+01	...	4.519000e+01	5.422000e+02
smoothness_se	569.0	7.040979e-03	...	8.146000e-03	3.113000e-02
compactness_se	569.0	2.547814e-02	...	3.245000e-02	1.354000e-01
concavity_se	569.0	3.189372e-02	...	4.205000e-02	3.960000e-01
concave.points_se	569.0	1.179614e-02	...	1.471000e-02	5.279000e-02
symmetry_se	569.0	2.054230e-02	...	2.348000e-02	7.895000e-02
fractal_dimension_se	569.0	3.794904e-03	...	4.558000e-03	2.984000e-02
radius_worst	569.0	1.626919e+01	...	1.879000e+01	3.604000e+01
texture_worst	569.0	2.567722e+01	...	2.972000e+01	4.954000e+01
perimeter_worst	569.0	1.072612e+02	...	1.254000e+02	2.512000e+02
area_worst	569.0	8.805831e+02	...	1.084000e+03	4.254000e+03
smoothness_worst	569.0	1.323686e-01	...	1.460000e-01	2.226000e-01
compactness_worst	569.0	2.542650e-01	...	3.391000e-01	1.058000e+00
concavity_worst	569.0	2.721885e-01	...	3.829000e-01	1.252000e+00
concave.points_worst	569.0	1.146062e-01	...	1.614000e-01	2.910000e-01
symmetry_worst	569.0	2.900756e-01	...	3.179000e-01	6.638000e-01
fractal_dimension_worst	569.0	8.394582e-02	...	9.208000e-02	2.075000e-01



ZEN

Carl Zeiss GMBH

Les Voies du Seigneur sont impénétrables :)

Journées SOFFOET 24/09/2021

Examples:

- original image + a pseudocolor map segmented and used for measurement
- measured parameters
- statistical comparisons
- coloration / IHC

Journées SOFFOET 24/09/2021

Placenta: evaluation of trophoblast hyperplasia:

- **measured parameters:**
  - **% trophoblast area of villous area**
  - **ratio of trophoblast outer to inner perimeter**

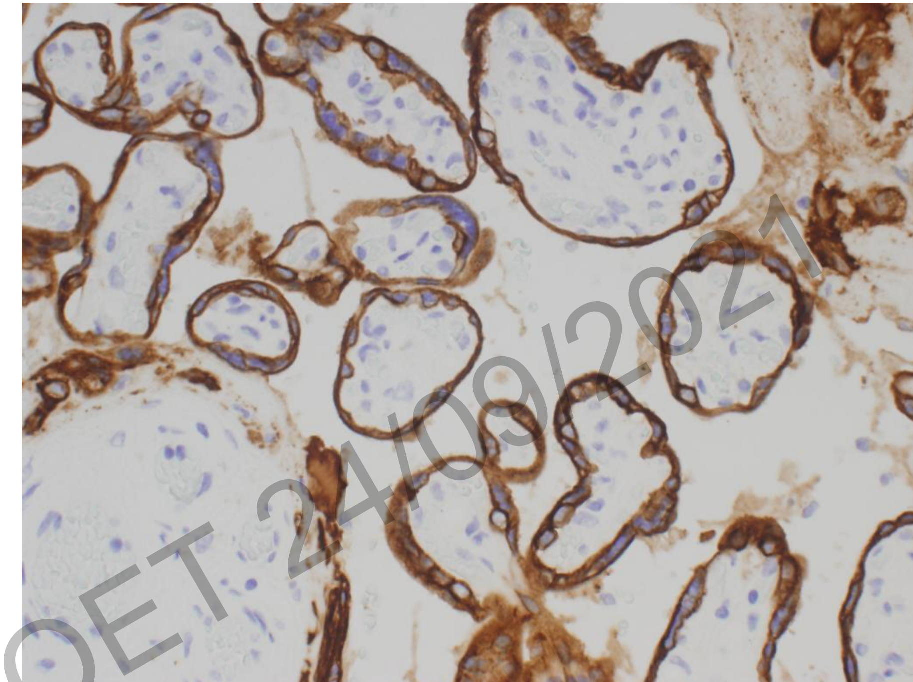
- **CK7 immunostaining**

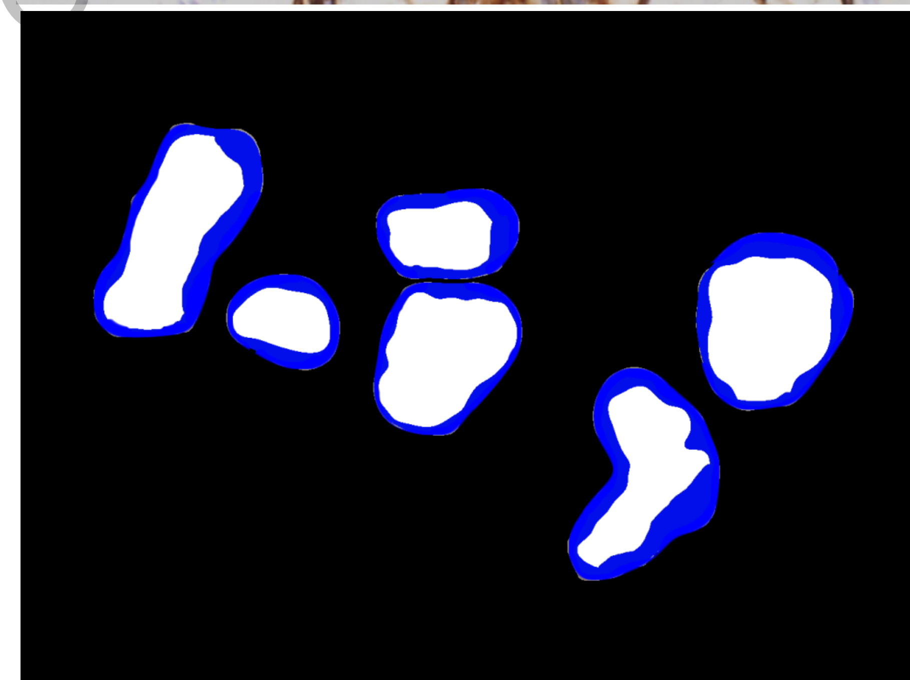
Images and IHC: courtesy to Dr Jelena Martinovic

Journées SOFFOET 24/09/2021

#1

Control example images

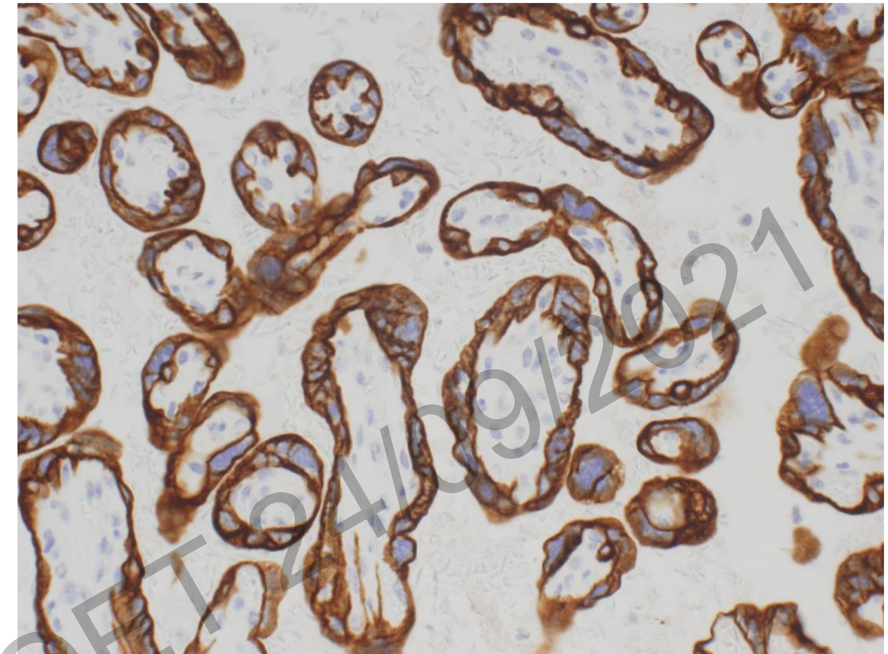




Journées SOFFOET 24/09/2021

#1

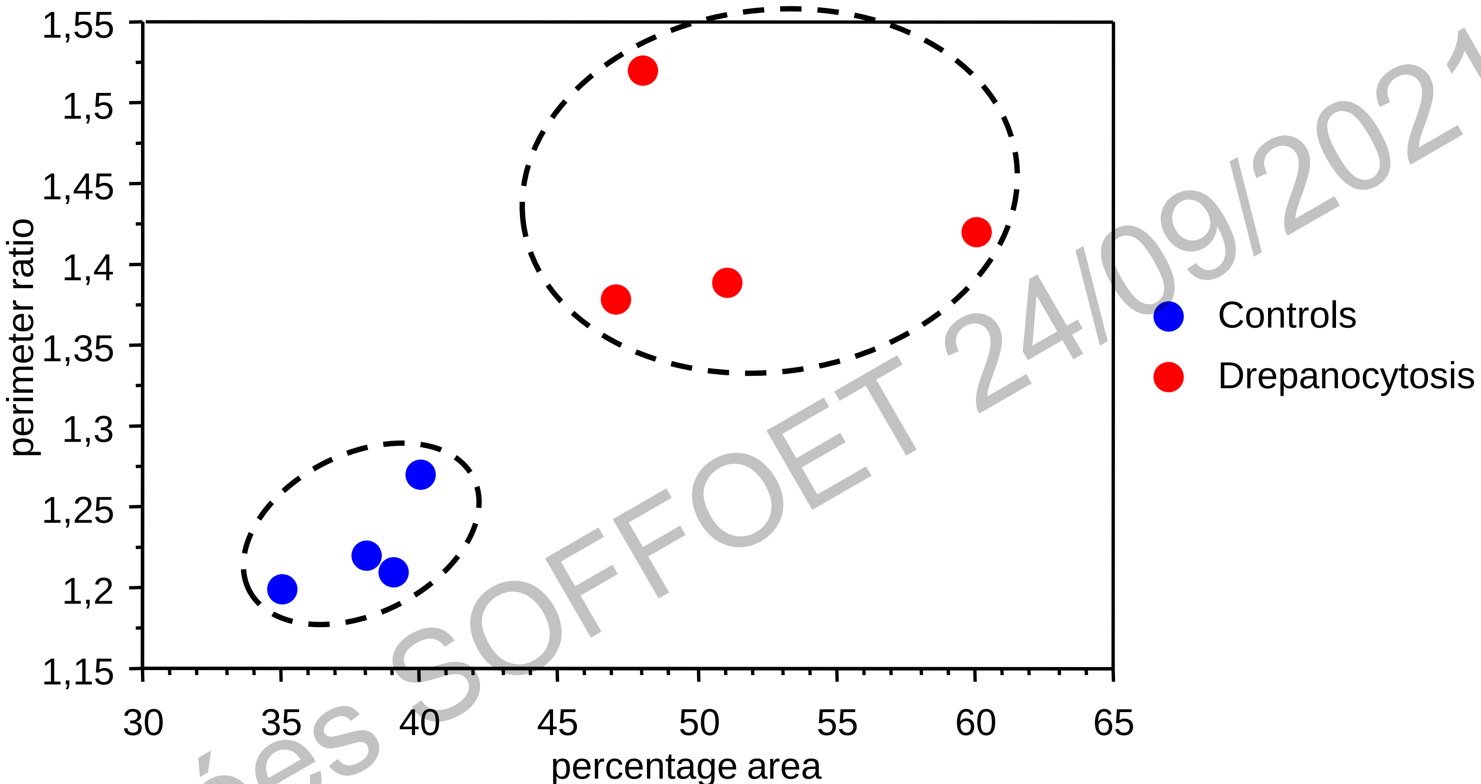
Drepanocytosis example images



Journées SOFFOET 24/09/2021

#1

Results cases (4 + 4)

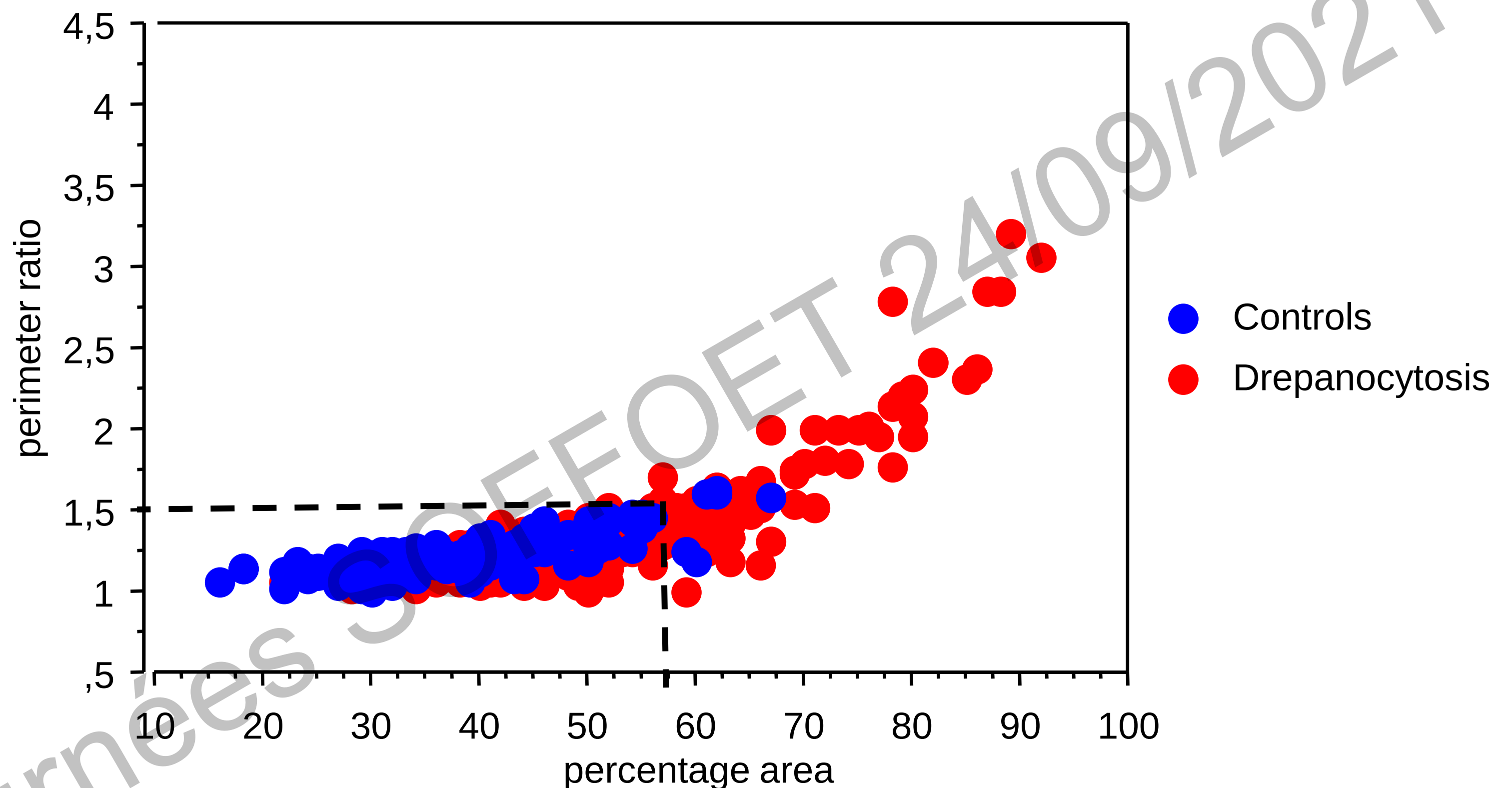


Journées SOFFOET 24/09/2021

Points are mean values per case

#1

Results villi (160 + 232)



Points are individual villus values

Journées SOFFOET 24/09/2021

Myocardium: myocardial fibrosis in a KO mouse model of fibrotic cardiomyopathy:

- measured parameter: ratio fibrosis / cardiac muscle (%)

- coloration Sirius Red

Journées SOFFOET 24/09/2021

Myocardial fibrosis in a KO mouse model Sirius Red

#2

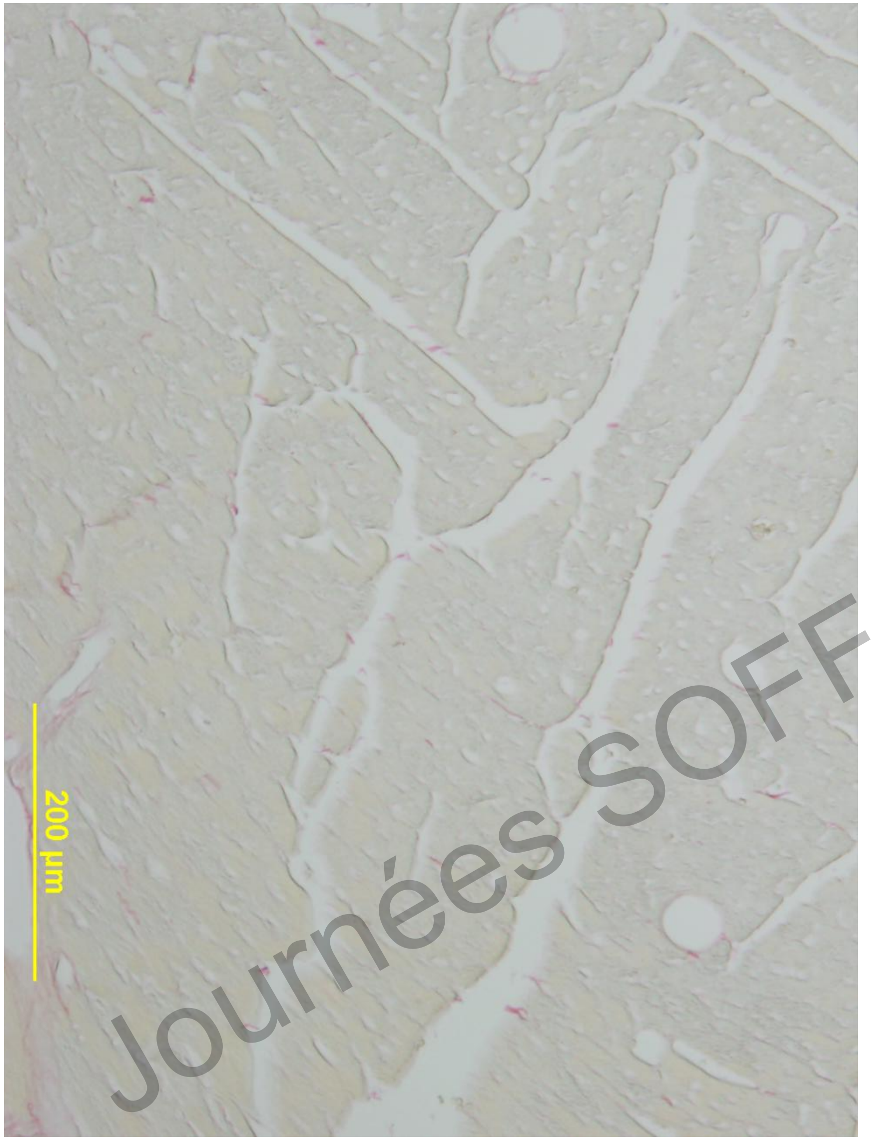
Journées SOFFOET 24/09/2021

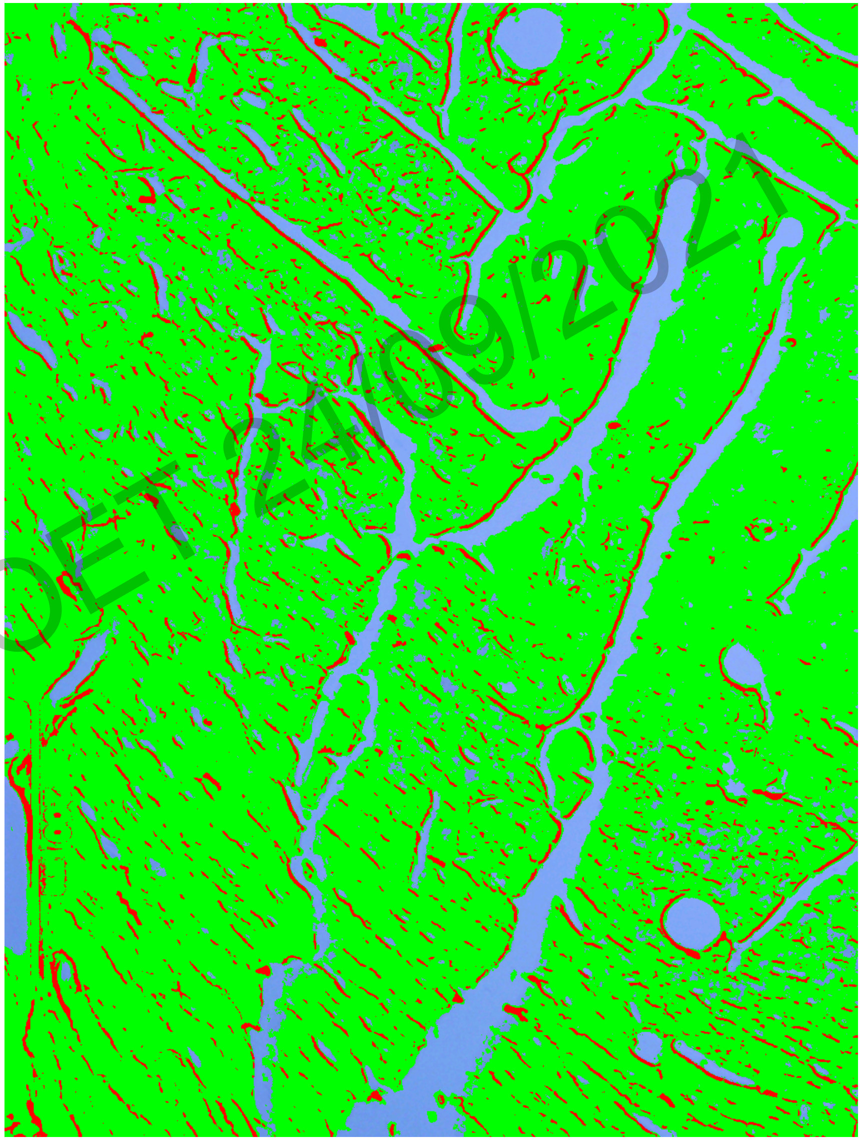
SML teaches users about their images

Green/Red*100 (fibrosis) = 1.0 %

Myocardial fibrosis in a KO mouse model: Sirius Red

#2





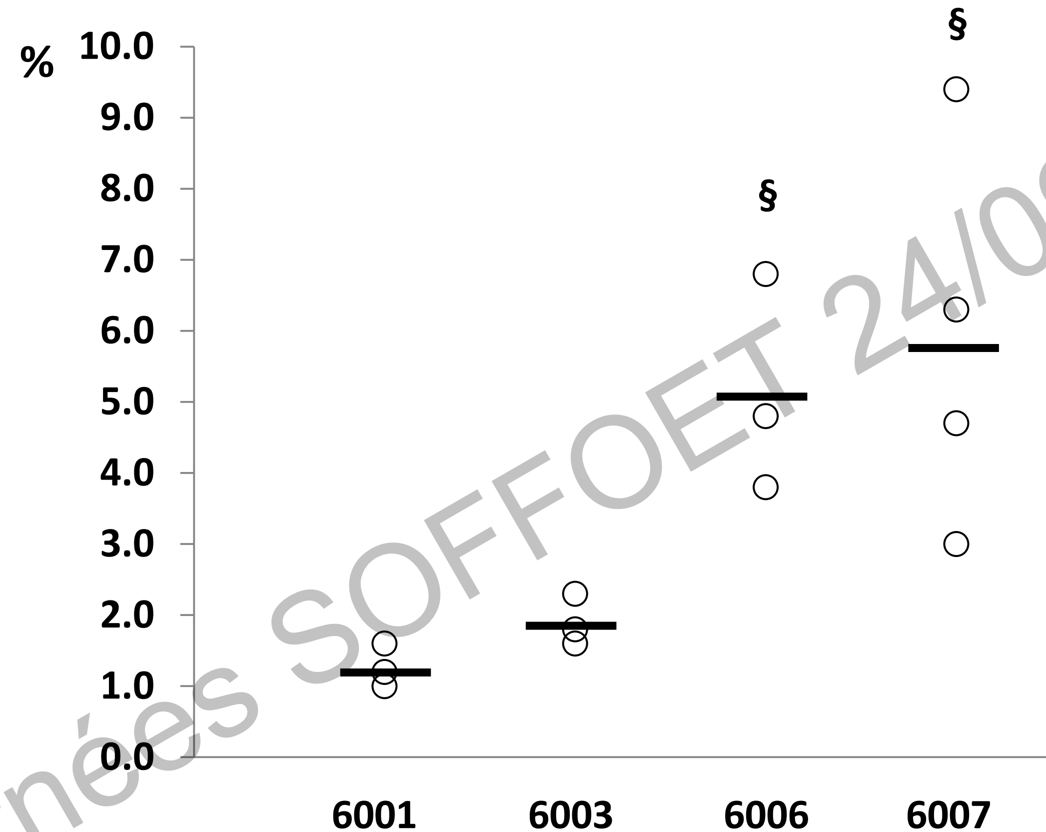
Journées SOFFOET 24/09/2021

SML teaches users about their images

Red/Green*100 (fibrosis) = 6.3 %

#2

Summary fibrosis: ○ are animals and — are medians



§ $p < 0.001$ compared to 6001 & 6003

Journées SOFFOET 24/09/2021

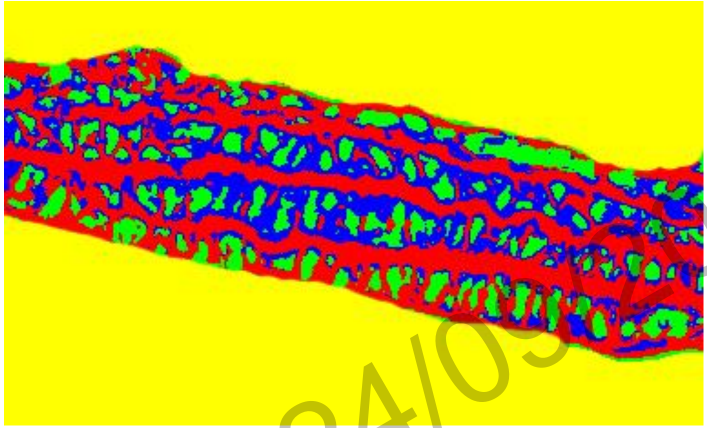
Aorta: desorganization of elastic fibers in a KO mouse model aortic aneurysms:

- measured parameter: % areas of the 3 elastic fiber compartments in an elastic arteria media

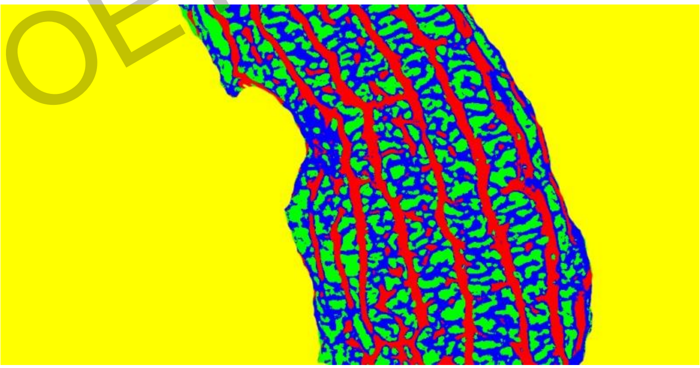
- coloration: Elastica – van Gieson

Journées SOFFOET 24/09/2021

CONTROL CASE A14 M2



EXPERIMENTAL CASE C36 M1



Lamellae plus struts

Interlamellar elastic fibers

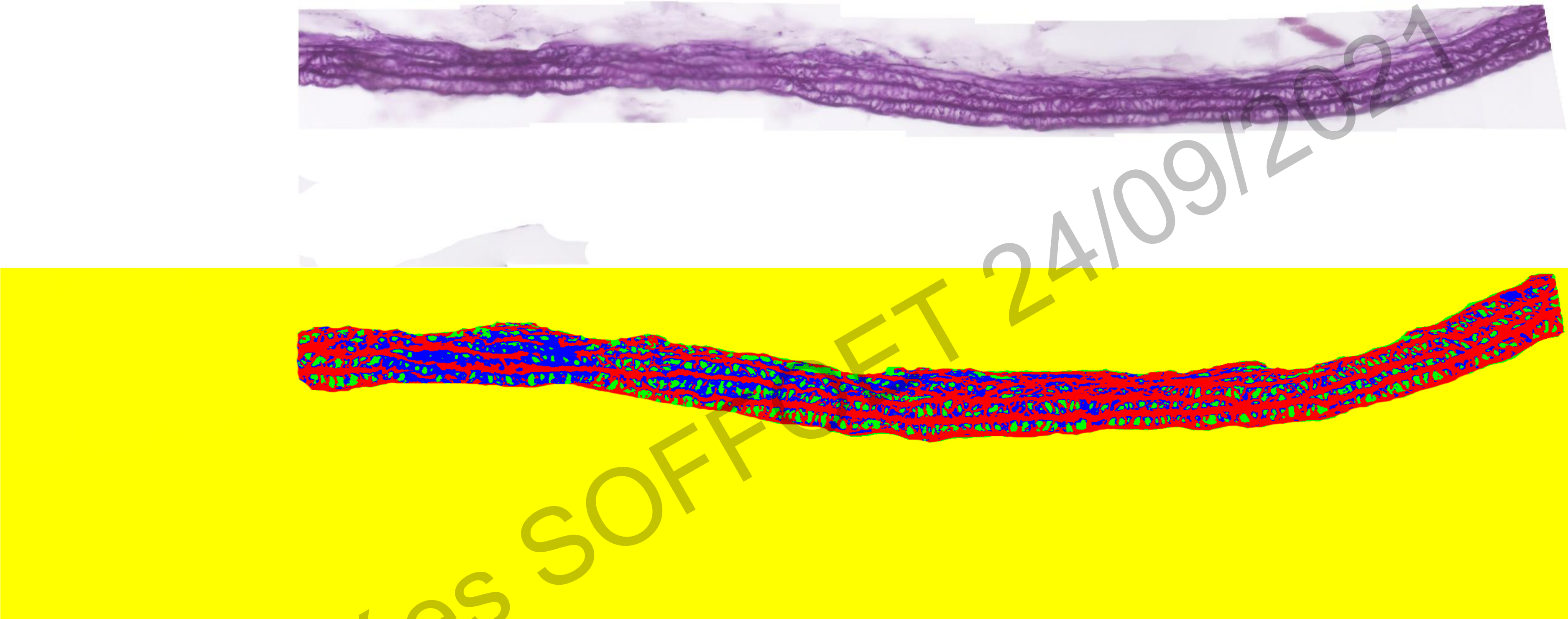
Smooth muscle cells

SML teaches users about their images

Journées SOFOET 24/03/2021

The 3 compartments of elastic fibers in the aorta

#3



Lamellae plus struts

Interlamellar elastic fibers

Smooth muscle cells

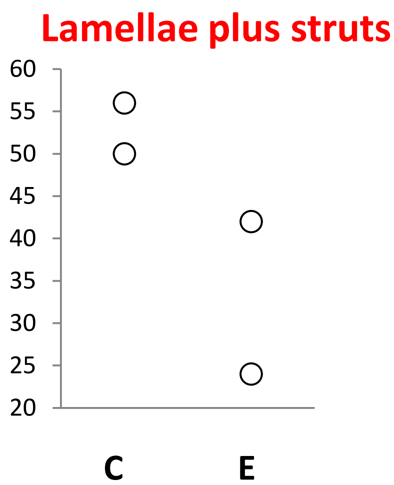
Image represents the color mask of a mosaic of 12 adjacent x40 images (a reconstruction of the entire tissue section) used to assess area ratios

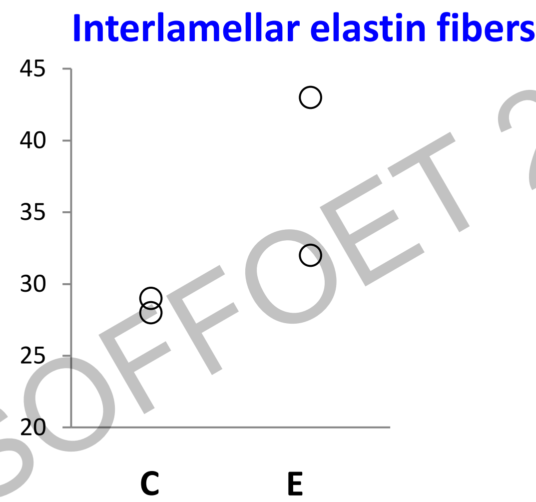
Journées SOFFCET 24/09/2021

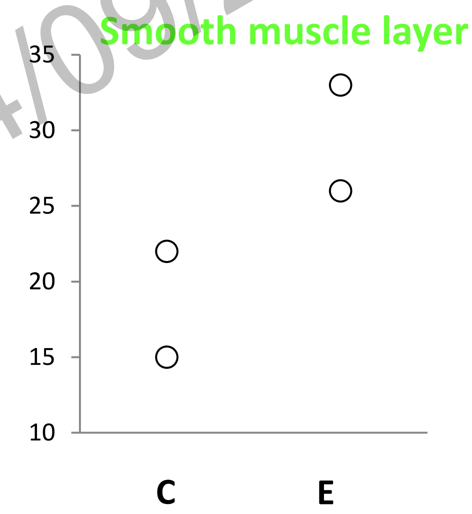
#3

RESULTS: AREAS % OF ALL MEDIA AREA

Two control (A14M2, A18F1) and two experimental (C36M1, C29F1) cases of all 5 cases had sufficient coloration quality to be examined



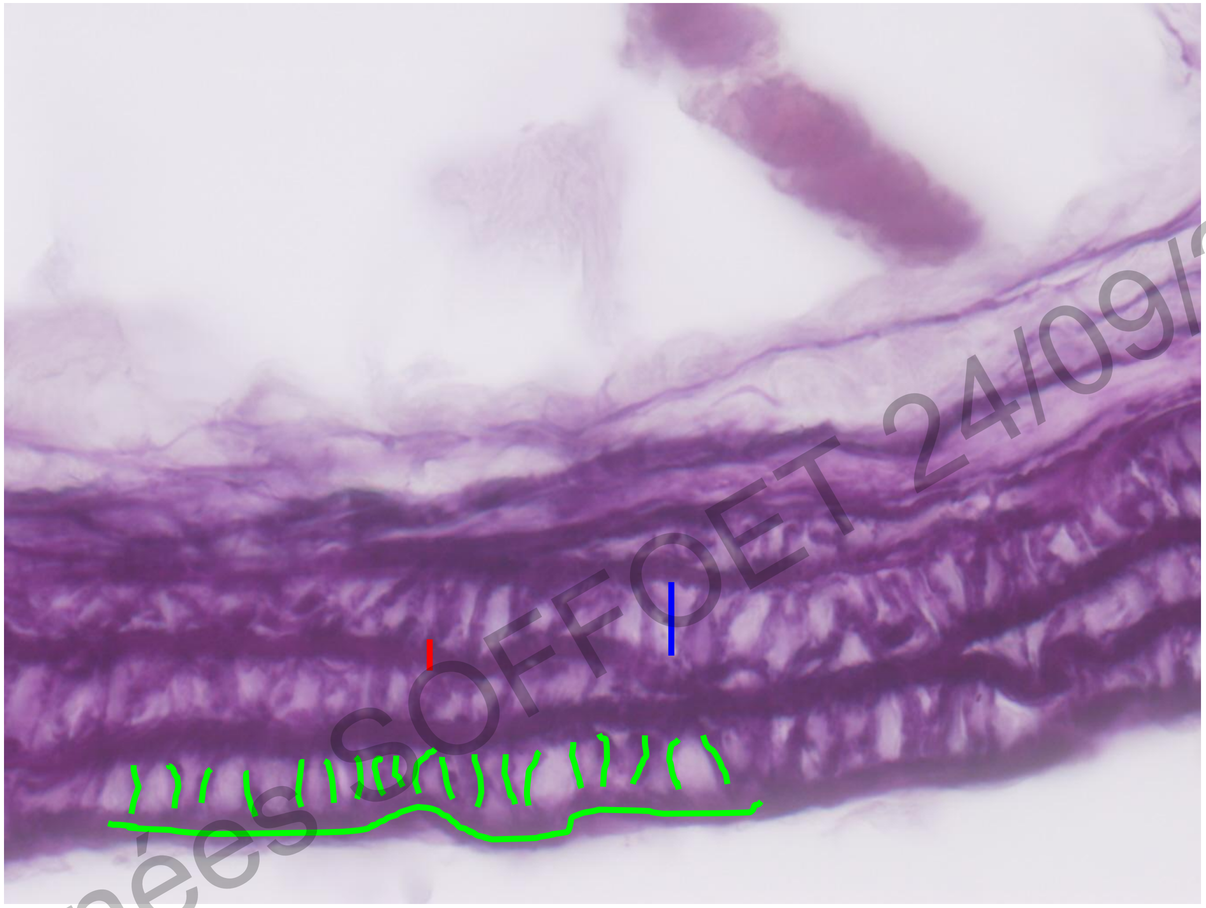




It seems that the experimental group shows a relative augmentation of interlamellar elastin at the expense of the thinning of lamellae and struts

Journées SOFFOET 24/09/2021

MEASURES: THICKNESSES AND NUMBERS



Journées SOFFOET 24/09/2021

Thickness lamellae

Thickness smooth muscle cell layer

Number of elastin struts per 100 µm of lamellar length

Liver: steatosis and fibrosis in a mouse model of NASH:

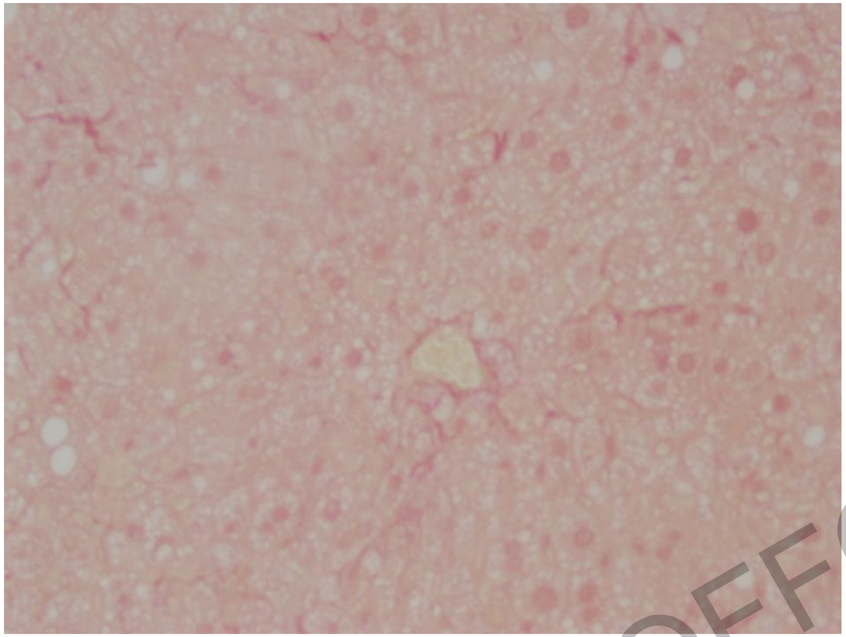
measured parameters

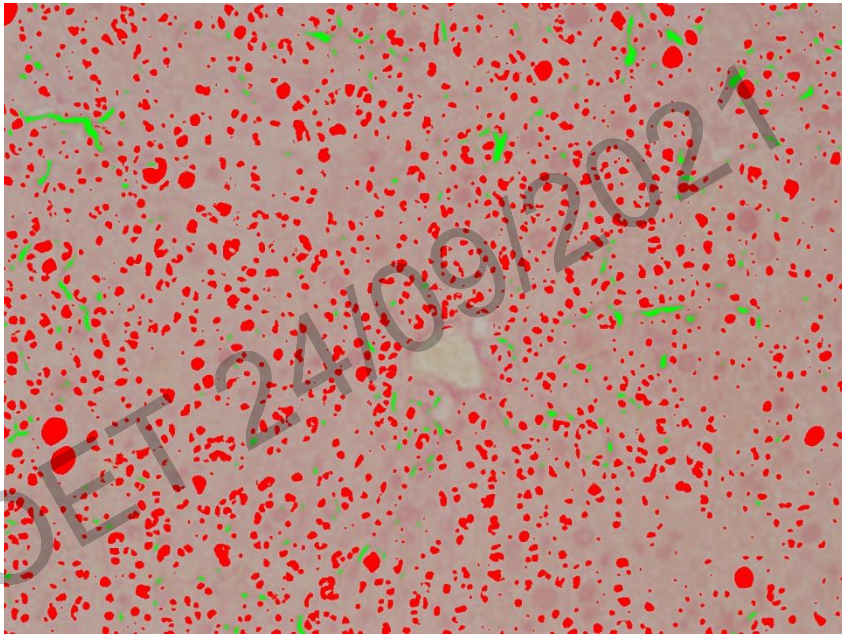
- **fibrosis (%)**
- **steatosis (%)**
- **macrovesicular steatosis of all steatosis (%)**
- **size macrovacuoles**

Sirius Red

Journées SOFFOET 24/09/2021

Example KOMCDHF (43/14)





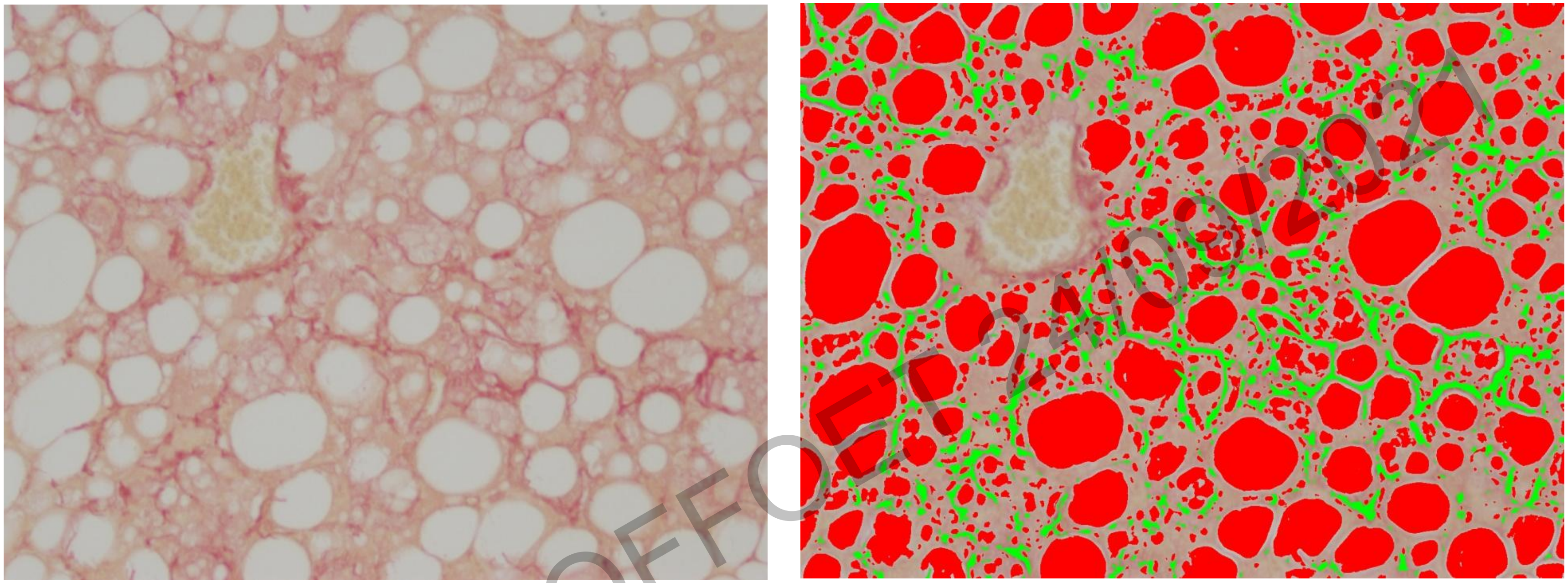
Journées SOFFOET 24/09/2021

Steatosis MF: 12%
Macrovesicular steatosis: 3%
Vacuoles: 1698
Mean size vacuoles: 9 µm^2
Mean size macrovacuoles: 104 µm^2

Fibrosis MF: 0.9%

Example WTMCDHF (35/15)

#4

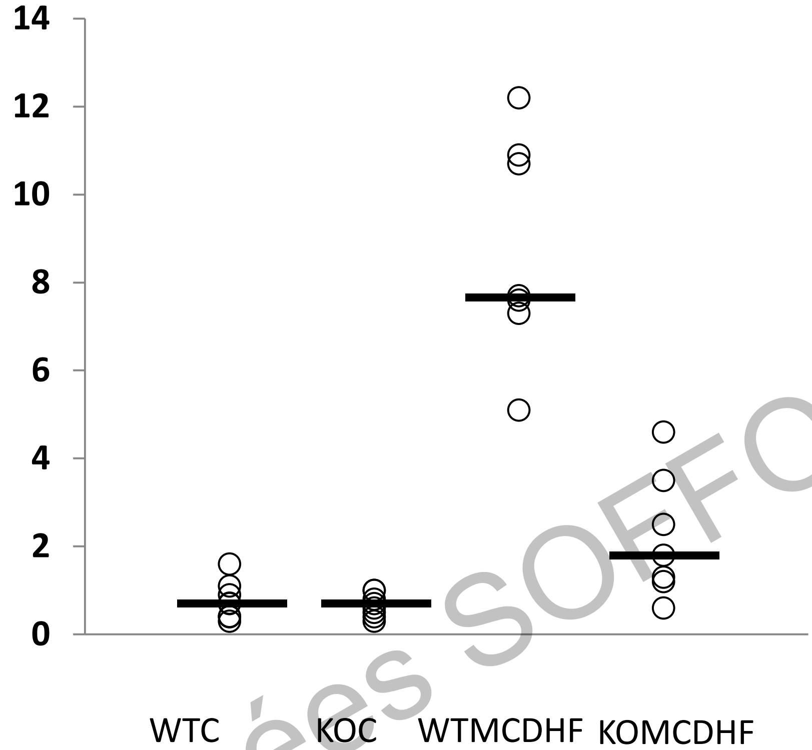


Journées SOFFOET 24/09/2021

Steatosis MF: 45%
Macrovesicular steatosis: 83%
Vacuoles: 1025
Mean size vacuoles: 53 µm^2
Mean size macrovacuoles: 368 µm^2

Fibrosis MF: 7.0%

Fibrosis: % of microscopic field area



WTMCDHF > WTC ($p < 8e-7$)
WTMCDHF > KOC ($p < 8e-7$)
WTMCDHF > KOMCDHF ($p < 3e-3$)

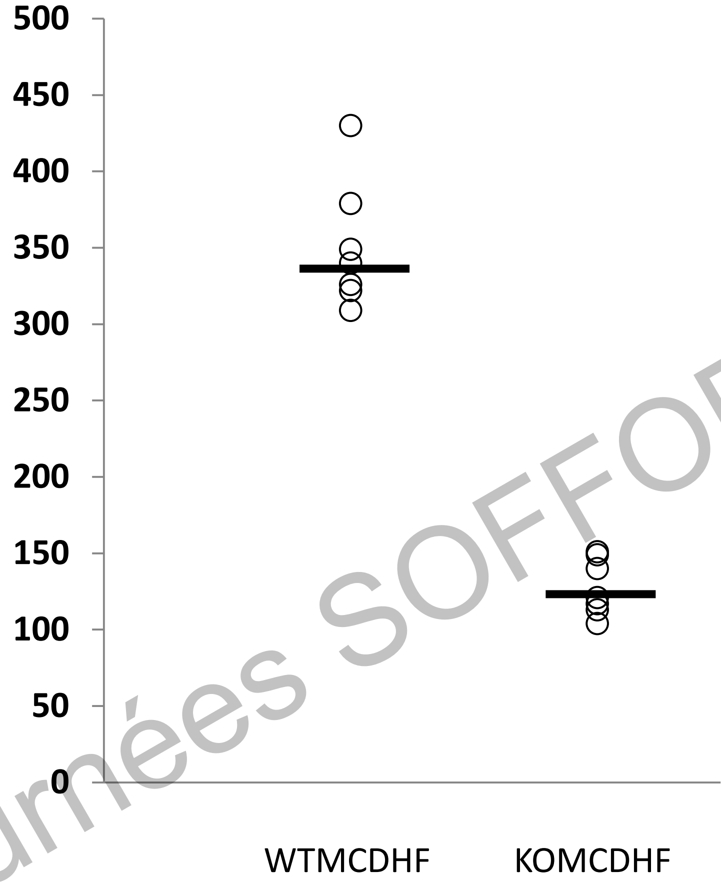
KOMCDHF > WTC ($p < 3e-3$)
KOMCDHF > KOC ($p < 1.8e-3$)

Rounds are mean values per animal and lines are medians; Conover post-test p-values, further adjusted by the Benjamini-Hochberg FDR method, after a significant Kruskal-Wallis test ($p < 0.0001$)

Journées SOFFOET 24/09/2021

Mean size steatotic macrovacuoles (µm^2)

#4



p < 6e-4

Mann & Whitney U-test

Journées SOFFOET 24/09/2021

Invading breast cancer: a competition example image

measured parameters

- **numbers of cells expressing 3 different antigenes**

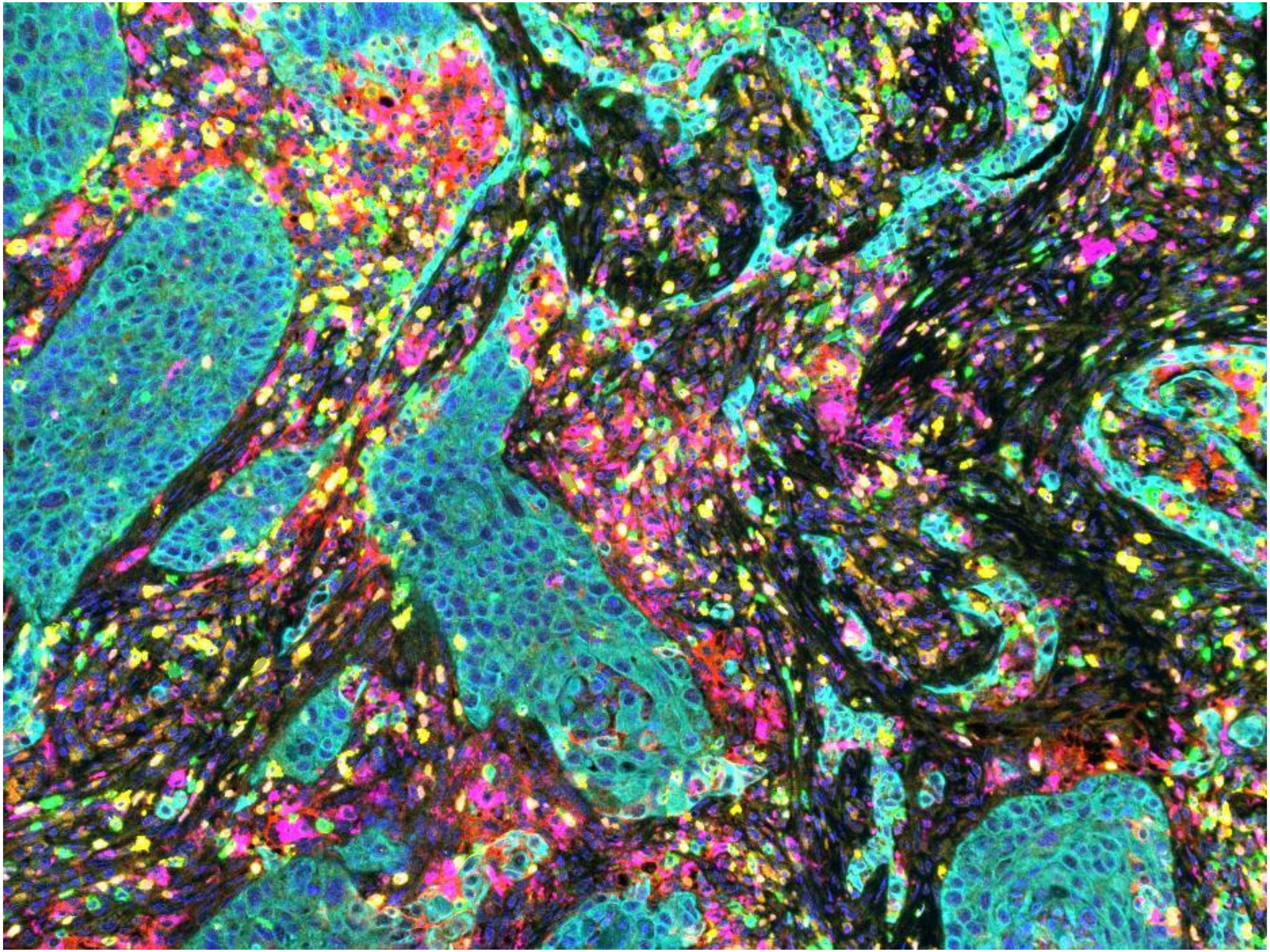
Multiplexed IHC: 6-10 labels, ISH, others

https://www.humancellatlas.org

Journées SOFFOET 24/09/2021

Multiplexed immunohistochemistry

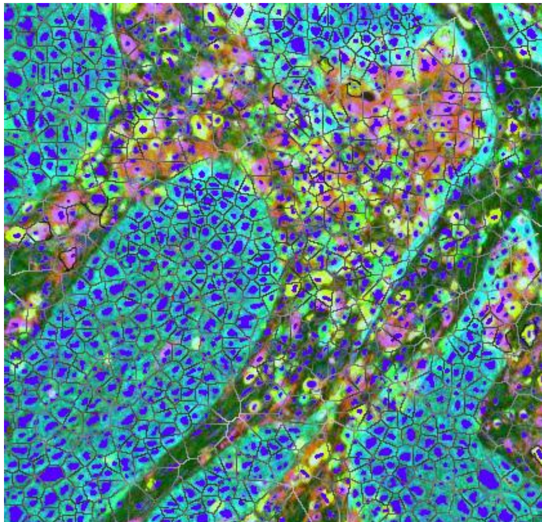
#5



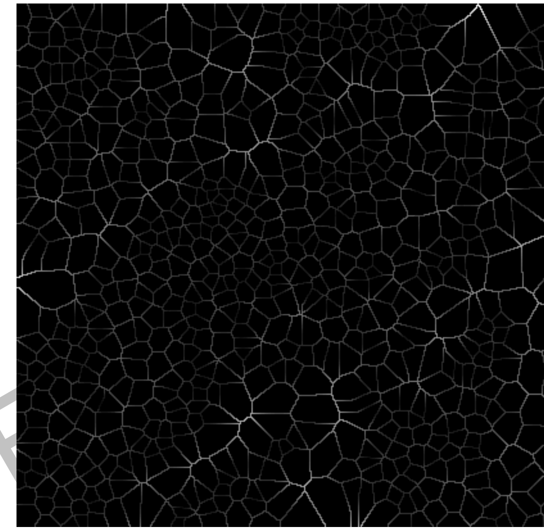
Multiplexed immunohistochemistry

#5

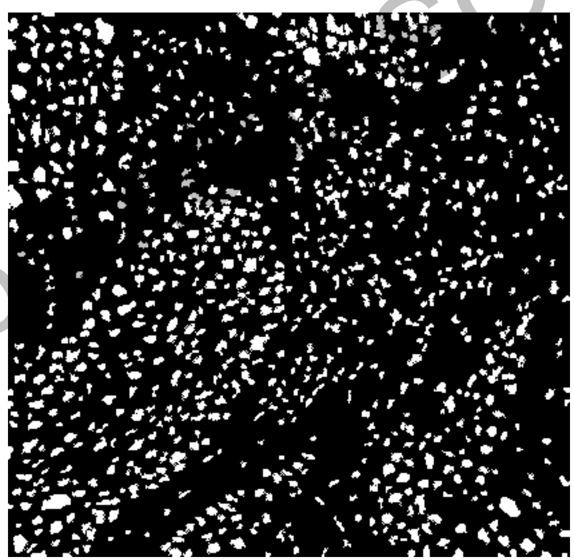
Final result



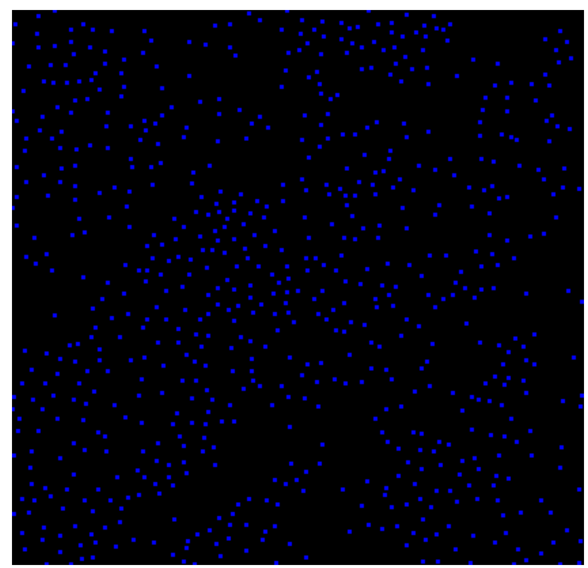
Modified Voronoi diagram of cell borders



Nuclei segmentation

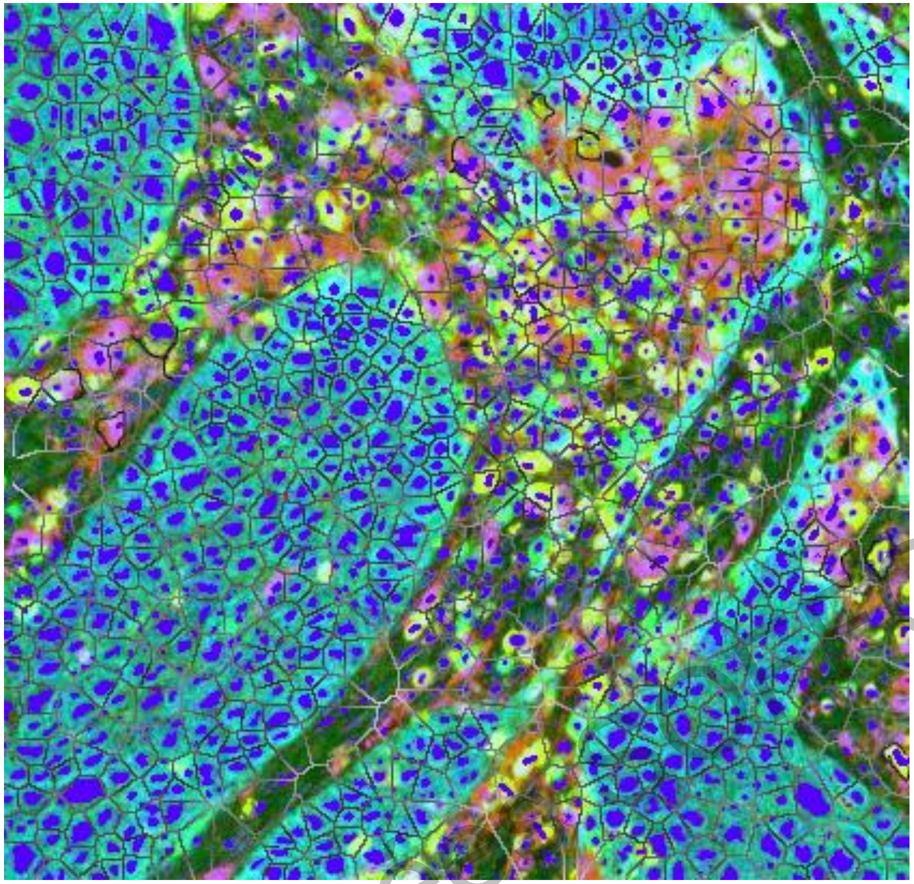


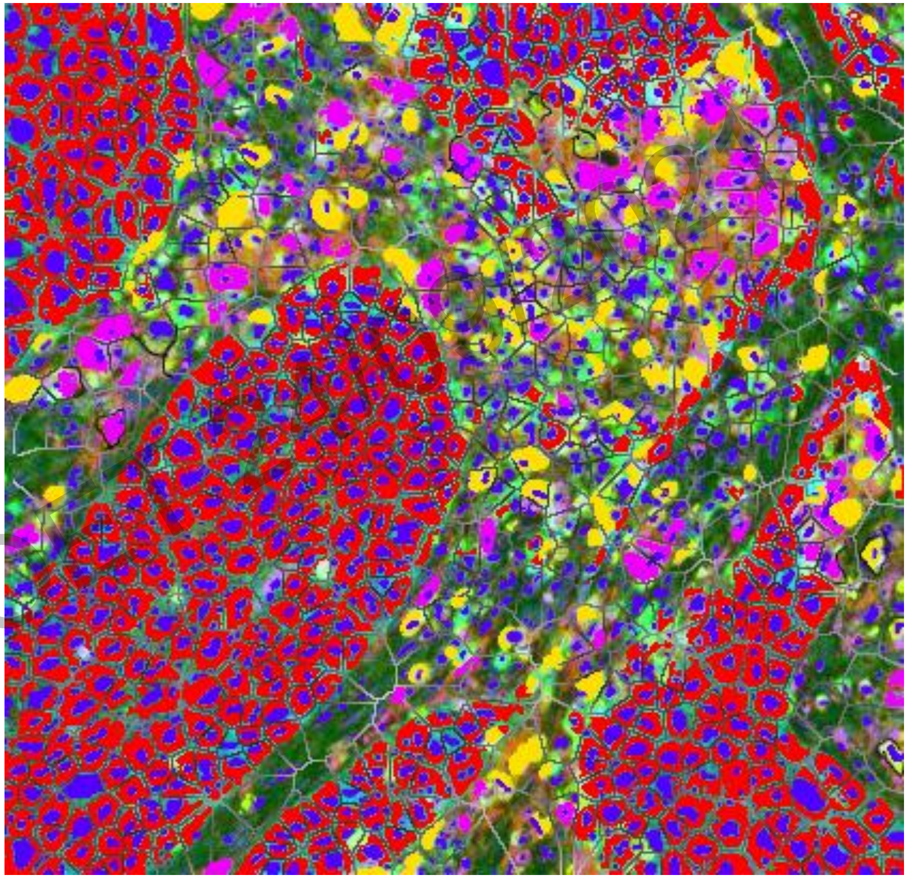
Nuclei reduced to 1 px centroids: seeds



Multiplexed immunohistochemitry

#5





Light blue: 511 (76%)

Yellow: 98 (15%)

Purple: 62 (9%)

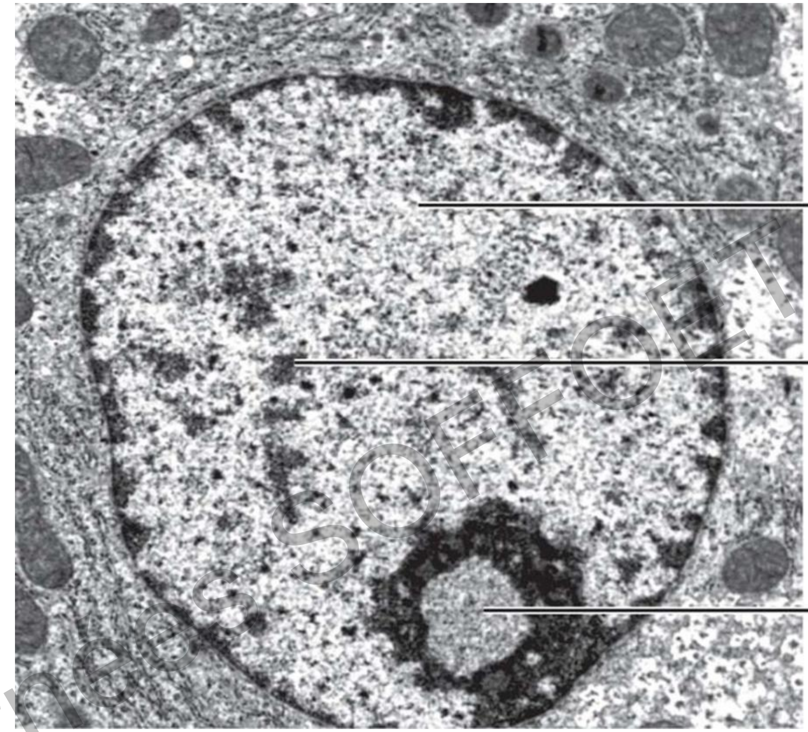
Journal Pre-proof

Stem cell nucleus:

- measured parameter: eu-/heterochromatin/interchromatin space

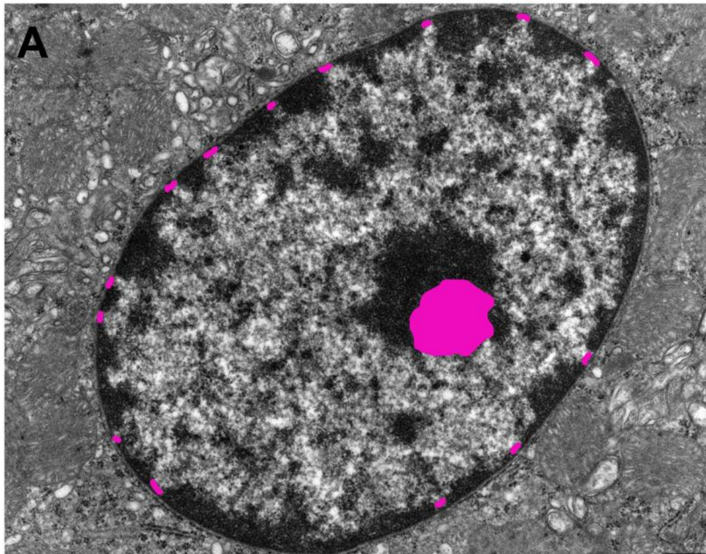
Electron microscopy

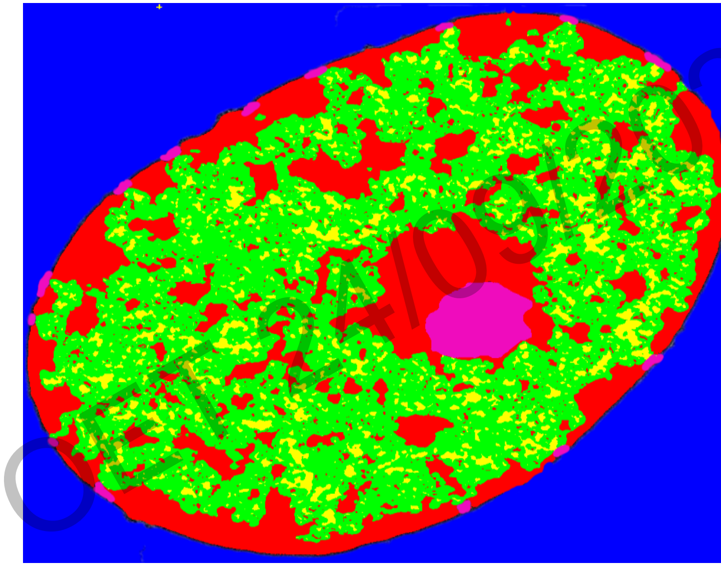
Journées SOFFOET 24/09/2021

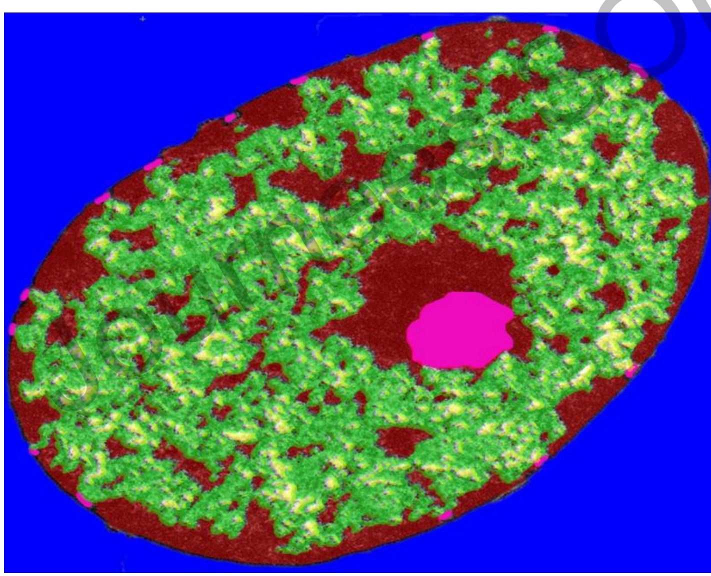


Interchromatin space? **SML teaches users about their images!**

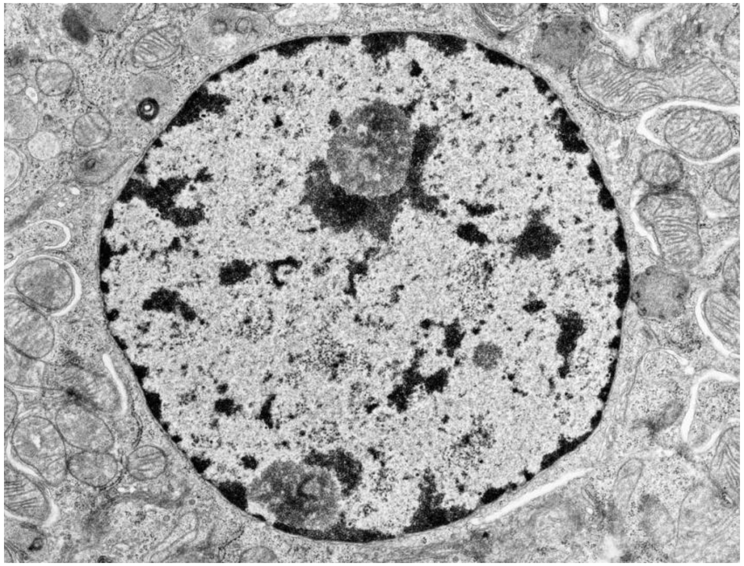
Journal Pre-proof 24/09/2021

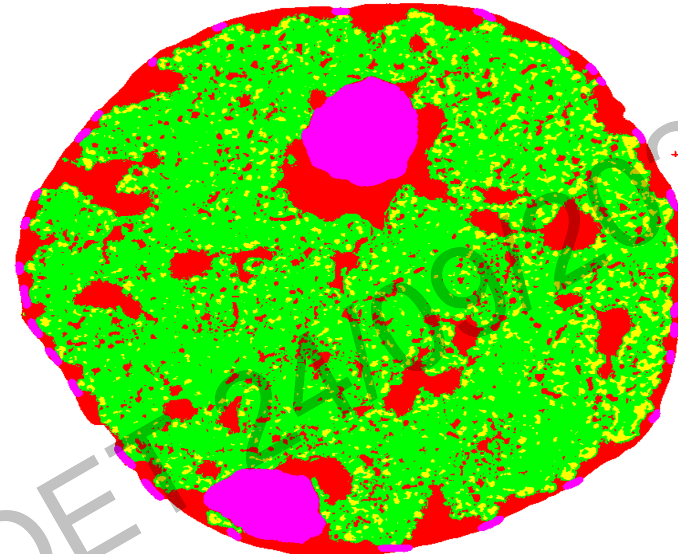


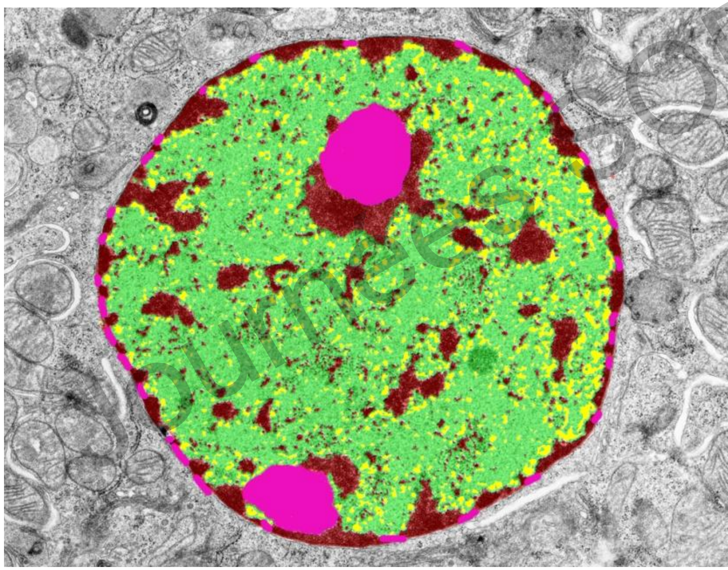




Heterochromatin	34%
Euchromatin	56%
Interchromatin space	10%
Nucleolus	







Heterochromatin	24%
Euchromatin	67%
Interchromatin space	9%
Nucleolus	

Journées OFFOET 24 mai 2021

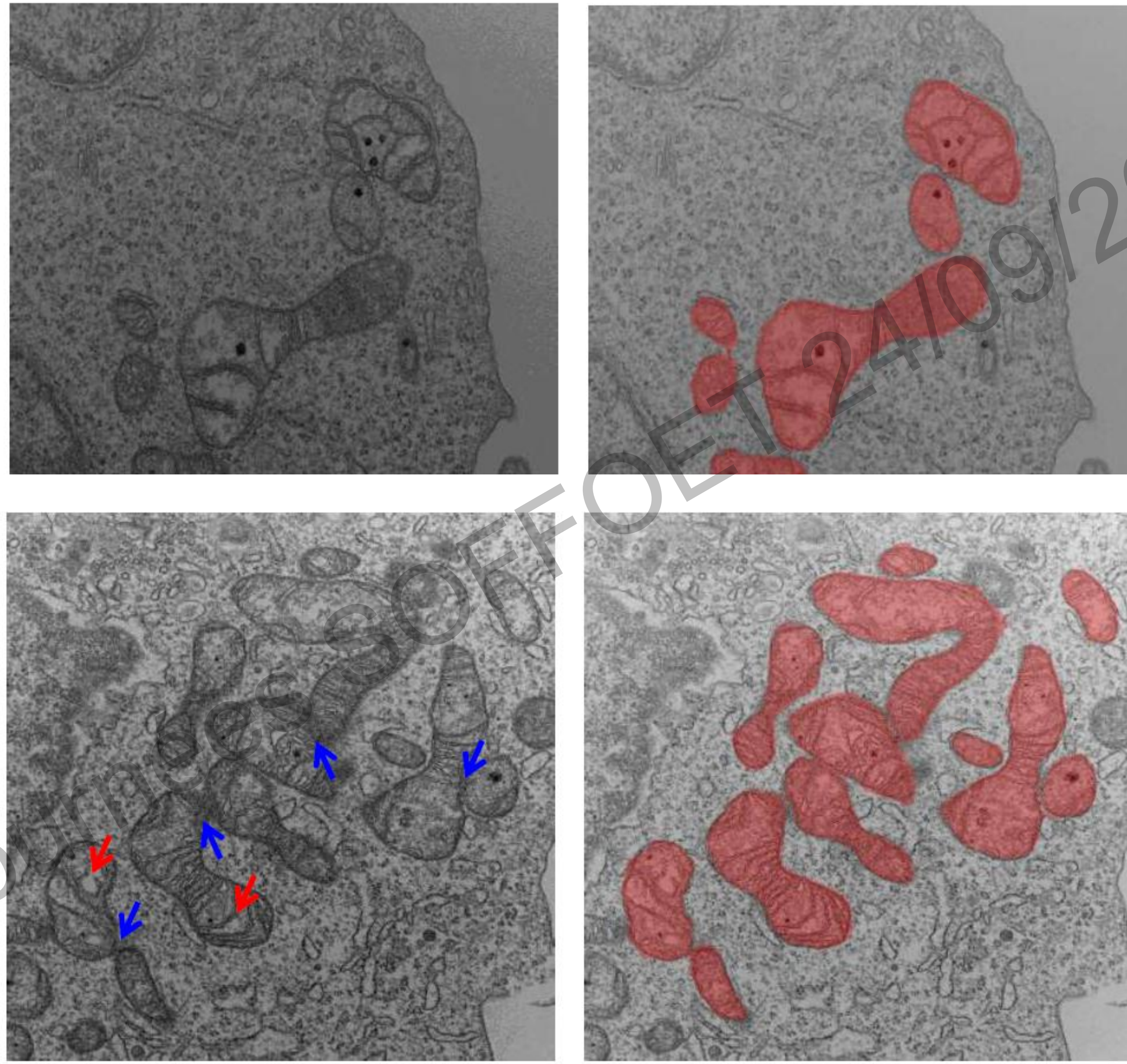
Microscopic image segmentation in the context of ML
(segmentation: identification of objects of interest):

- **semantic (pixels according to their properties)**

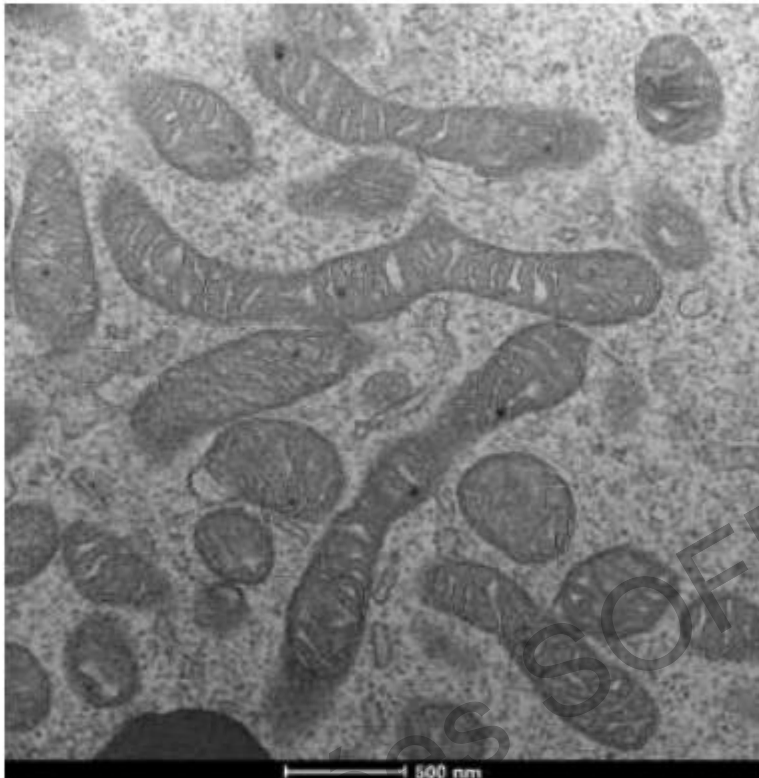
- **instance (objects according to their properties) (Deep Learning)**

Journées SOFFOET 24/09/2021

Measuring mitochondria: electron microscopy



SOFFOET 24/09/2021



Journées SOFFOT 24/09/2021

Journées SOFFOET

MITOCHONDRIA : SHAPE (also Bulthuis 2018)

PARAMETER	NOTE	REFERENCES
Form factor (FF) Aspect ratio (AR) Roundness (R)	FF = mitochondrial complexity and branching AR = mitochondrial elongation R = mitochondrial fragmentation	Bulthuis2018, Cadete2016, Chen2017 Choi2013, Dagda2008, Durand2019, Haileselassie2019, Hirabayashi2017, Jiang2018, Khraiwesh2013,Leduc-Gaudet 2015,2020, Morsci2016, O'Rourke2018, Pereira 2020, Picard2013,2015,Tang2020, Takamura2012, Twaroski2016, Zhang2017
Mean area / perimeter ratio	High values indicate high mitochondrial interconnectivity	Dagda2008
Mitochondrial fragmentation index (MFI)	High values indicate high mitochondrial fragmentation	Archer2012, Durand2019

FF = perimeter^2/ 4π*area = 1/circularity
AR = major/minor axis
R = 4*area/π*major axis^2
MFI = No. mitochondria x1 000/No. mitochondria pixels

Journées SOFFOET 24/09/2021