Cytomine for collaborative and semantic analysis of multi-gigapixel images

Raphaël Marée (PhD), University of Liège, Belgium

cytomine

cytomine

@cytomine

github.com/cytomine

www.cytomine.be

info@cytomine.be

+32 4 366 26 44

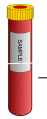
June 9th, 2016. ISP Group @ UCL

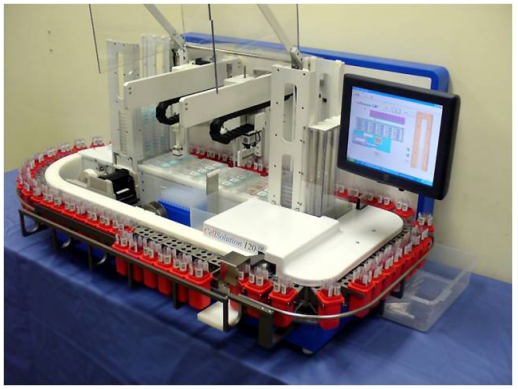
MONTEFIORE INSTITUTE
Department of Electrical Engineering and Computer Science

GIGA

Previous work : cervical cancer screening

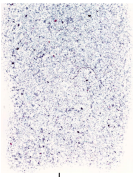
(specific, proprietary classifier developed with CellSolutions : 2008-2014)



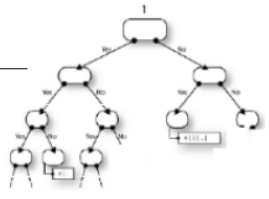


120 slides/hour

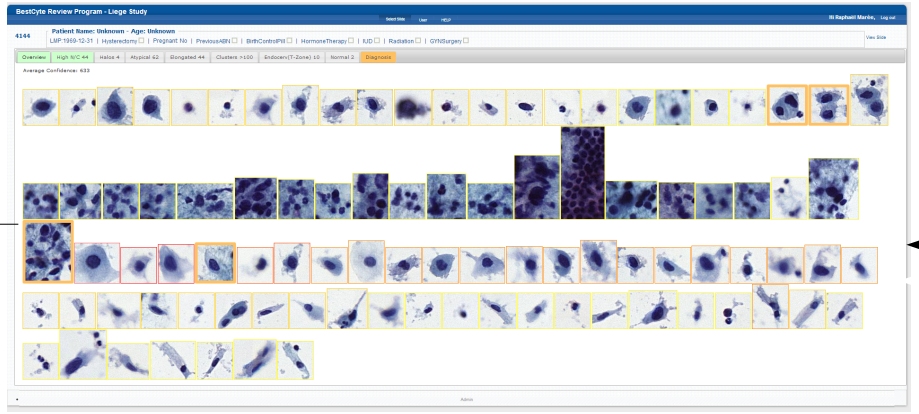
~ 60K cells per slide
~ 65K x 135K pixels



20X



Classify & rank every cells (3min/slide)



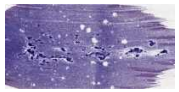
Cytotechnologist + pathologist web review of most suspicious cells according to our image recognition algorithms

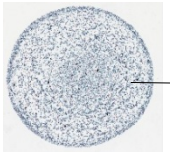
Final diagnosis

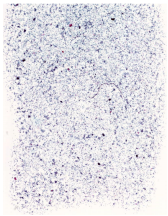
Evaluation of CellSolutions BestPrep Automated Thin-Layer Liquid-Based Cytology Papanicolaou Slide Preparation and BestCyte Cell Sorter Imaging System,
A. Delga, F. Goffin, F. Kridelka, R. Marée, C. Lambert, P. Delvenne,
Acta Cytologica, 2014;58(5):469-77

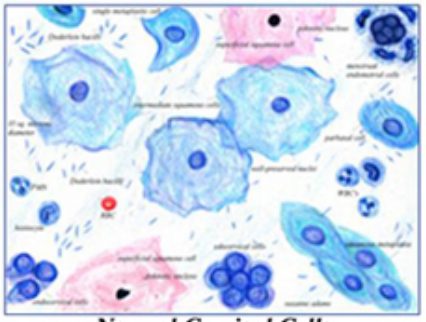
The world of cytology

Many sample types, preparation protocols, cell types, ...

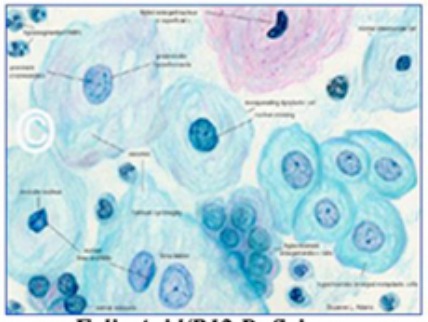




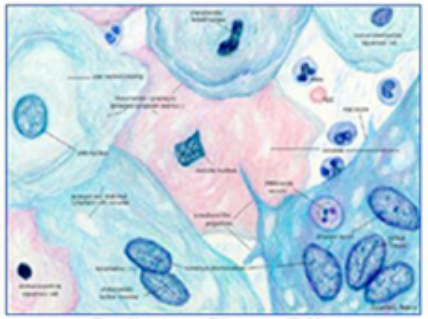




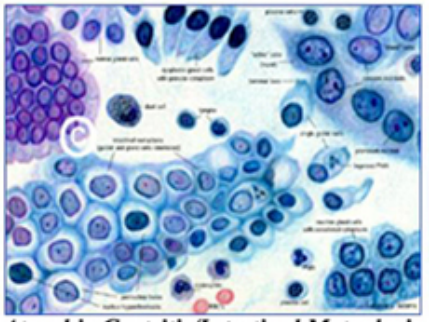
Normal Cervical Cells



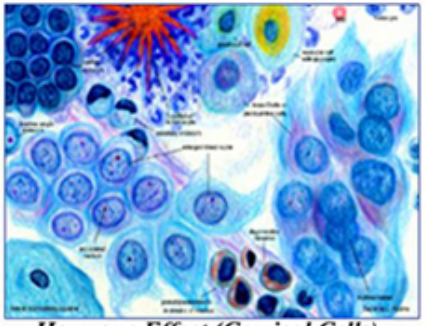
Folic Acid/B12 Deficiency



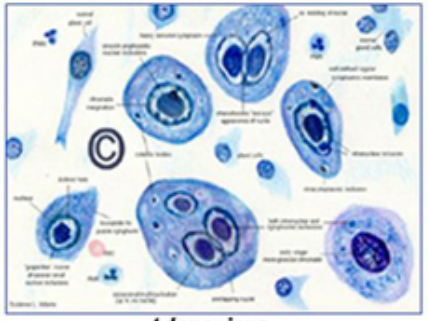
Radiation/Chemo Effect



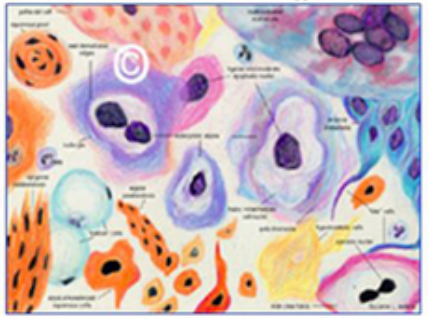
Atrophic Gastritis/Intestinal Metaplasia



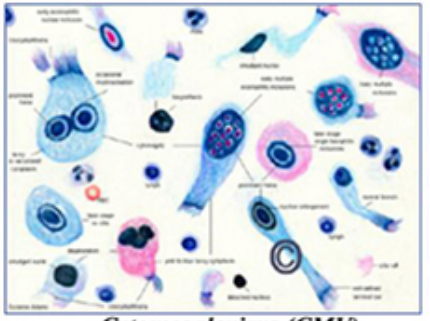
Hormone Effect (Cervical Cells)



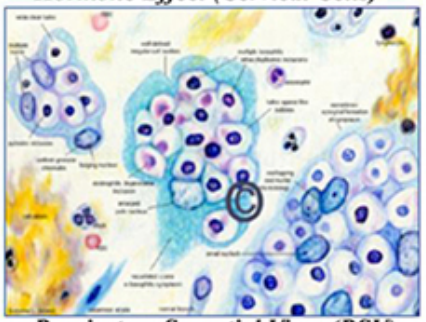
Adenovirus



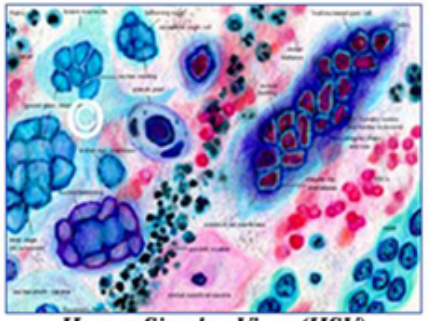
Human Papilloma Virus (HPV)



Cytomegalovirus (CMV)

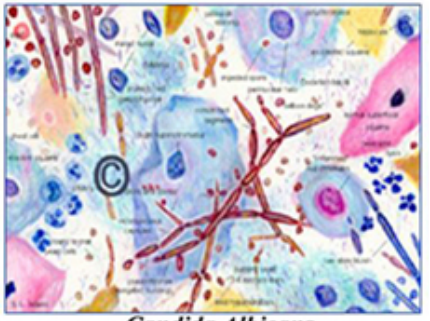


Respiratory Syncytial Virus (RSV)

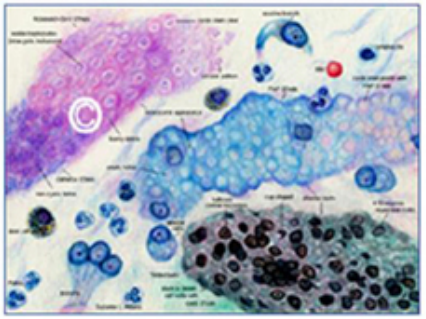


Herpes Simplex Virus (HSV)

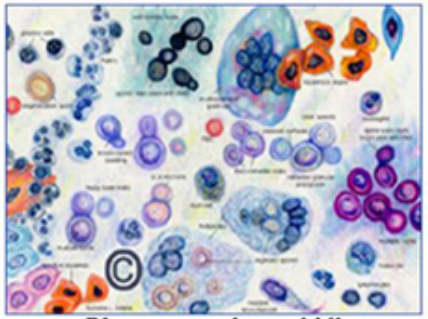
Moluscum Contagiosum



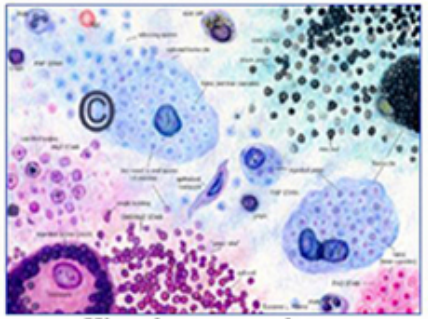
Candida Albicans



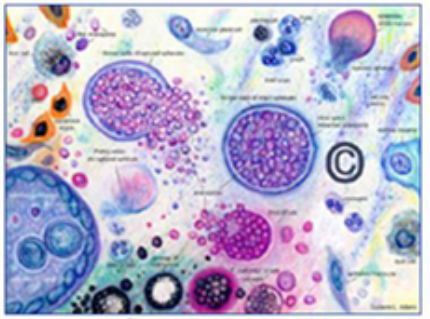
Pneumocystis carinni



Blastomyces dermatitidis



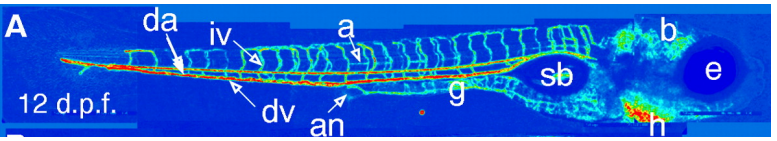
Histoplasma capsulatum

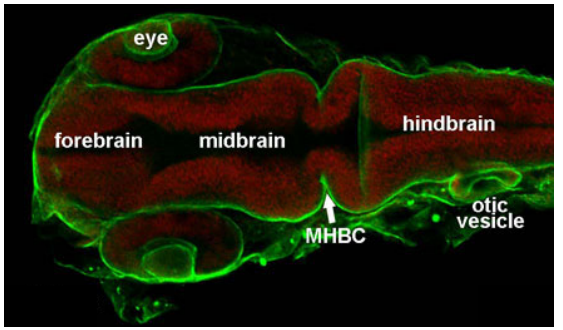


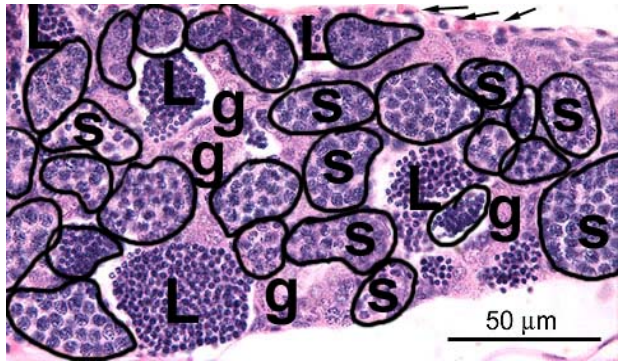
Coccidioides immitis

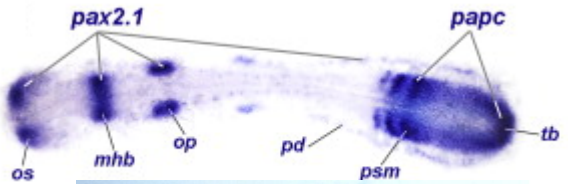
(Source : The Art of Cytology, Suzanne L. Adams)

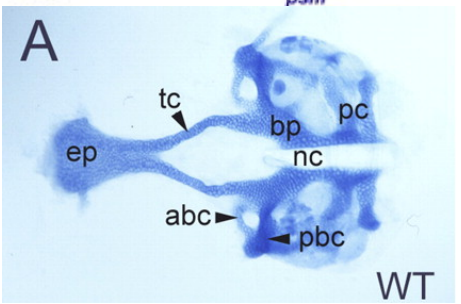
Biological/biomedical research heavily rely on **semantic annotation & quantification** of images

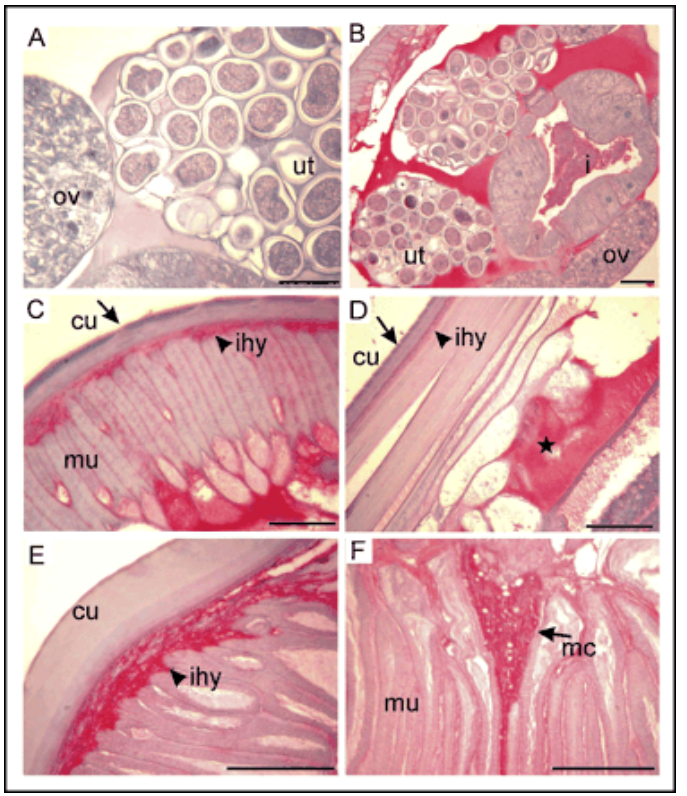


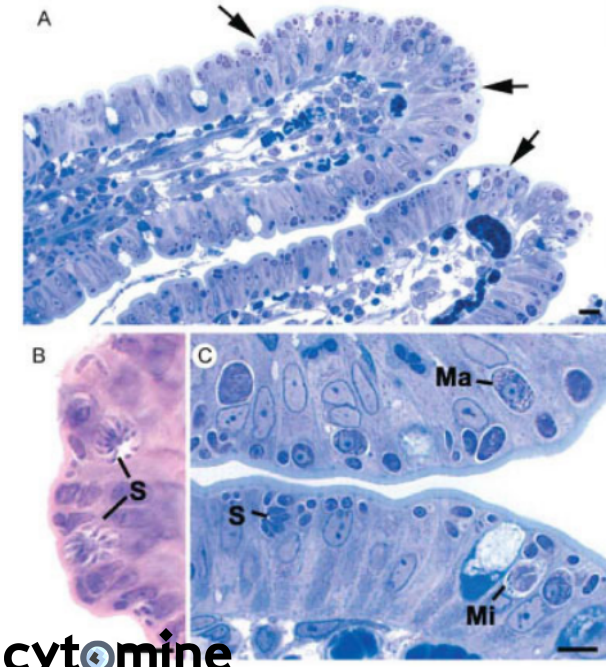


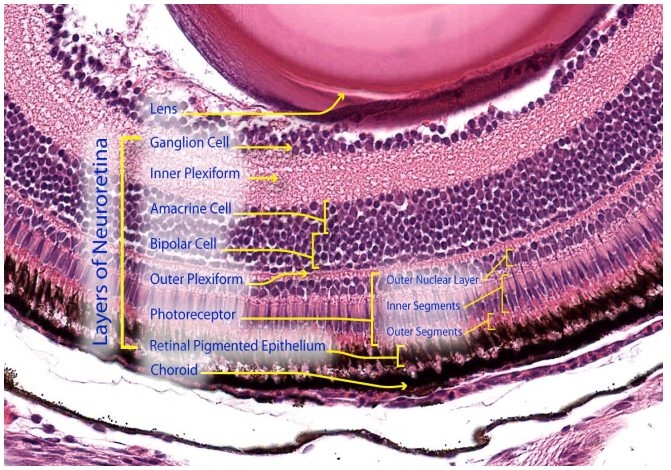


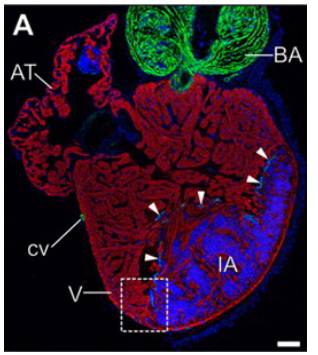


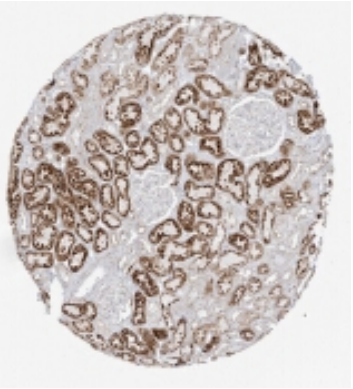












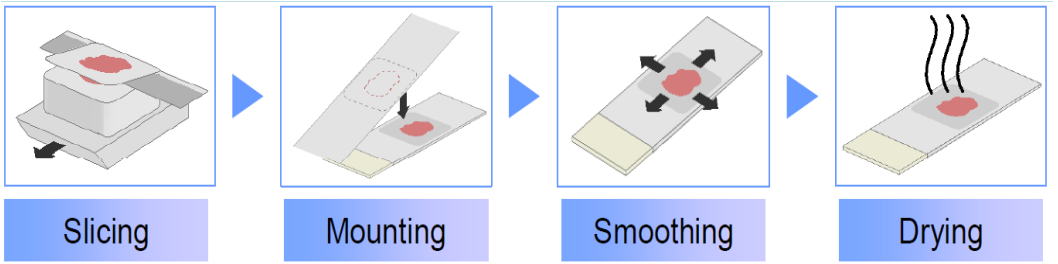
cytomine

(Source : Google Images)

Automated tissue bioimaging

Sample preparation

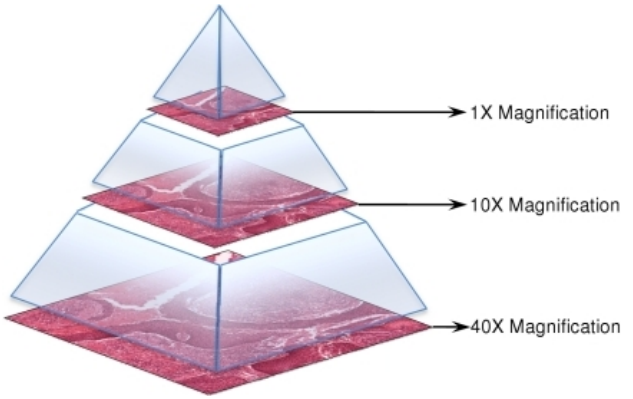
Kurabo AS-200 auto slide preparation system
100 slides / hour



200 GByte/hour → 5TByte/day

Image acquisition

Nanozoomer (Hamamatsu) whole-slide scanner
320 slide loader, **35 s / slide** (area: 15 mm × 15 mm),
0.23 µm/pixel. One image ~ (100K x 100K pixels)



+/- 2 GByte / image

cytomine

Scientists and pathologist's daily work

Slide annotations/quantifications are still often :

- performed manually (through a microscope)

- performed within tissue subregions in small sample groups
  (might not be statistically significant or one might miss specific patterns)

- created by isolated experts and stored locally
  (sometimes lost)

- not saved, or in proprietary format so hardly reusable
  (e.g. Photoshop draws for paper figures)

cytomine

Objective : *to make life easier for both life scientists, pathologists, and computer scientists*

- Continuous development since 2010

- Academic project but **industrial-grade software**
(modern & proven software libraries, user-driven, code quality control,…)

- **General-purpose**

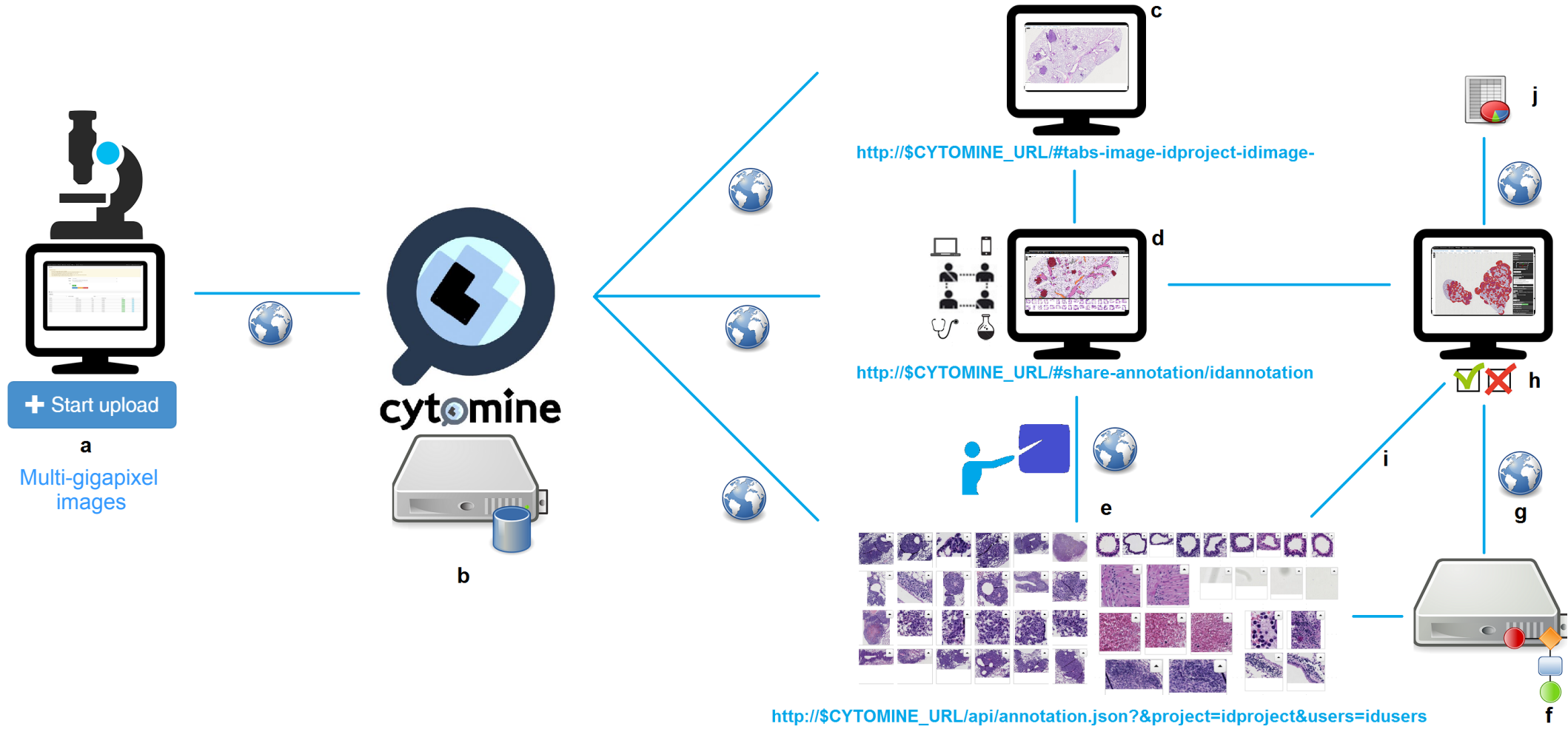
- A rich internet application (**web-based)**
  - Google Maps-like visualization of digital slides
  - **Collaborative, semantic**, annotation of regions of interest
  - **Semi-automated analysis**
    - Generic machine learning algorithms
    - Proofreading

- User-authentification for **secure sharing** of images, annotations, results

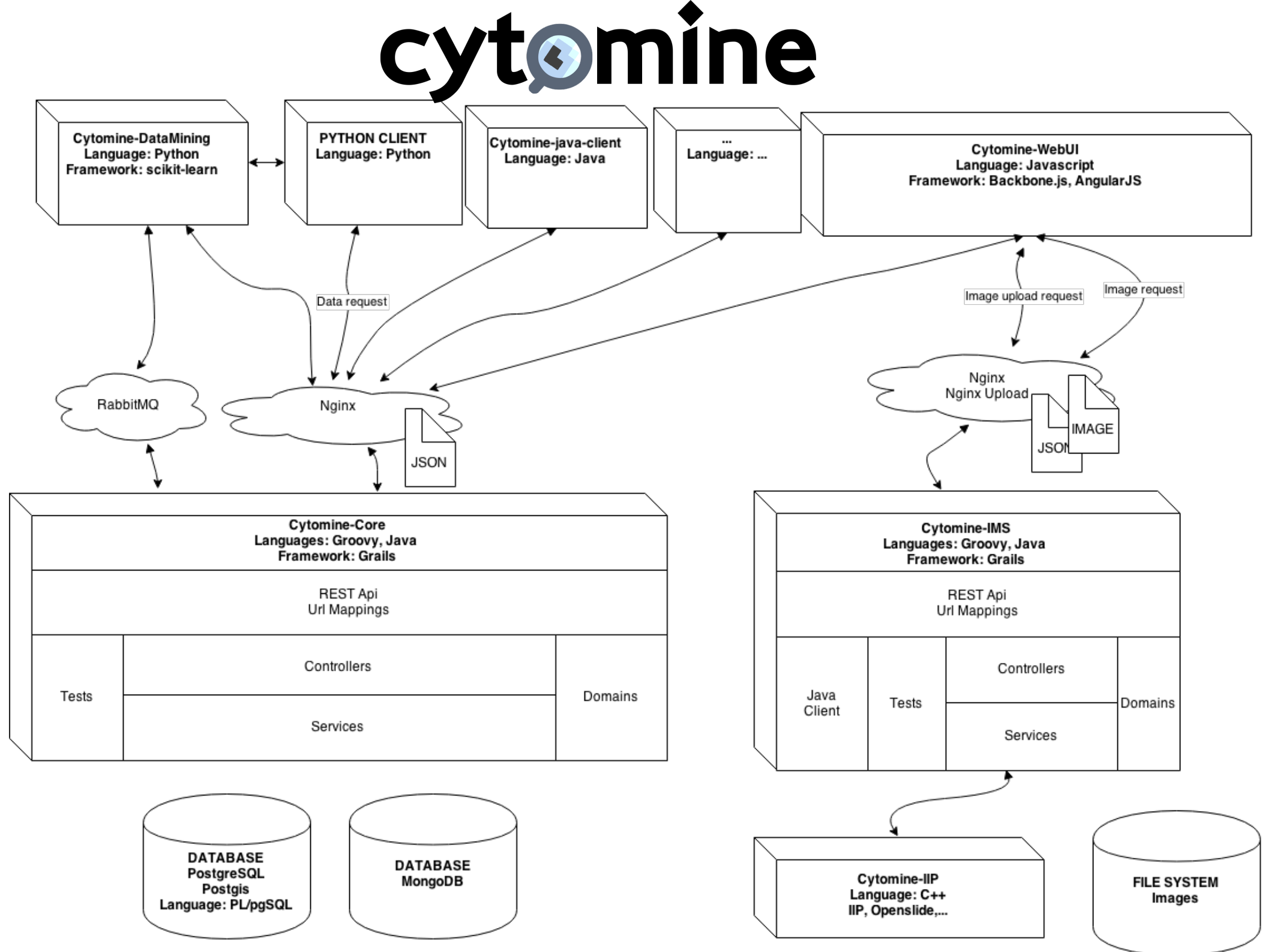
- Documented, **Free** and **Open-source (www.cytomine.be)**

*A computer software with its **source code made available** with a license in which the copyright holder provides the **rights to study, change, and distribute the software to anyone and for any purpose**.*

cytomine


+ Start upload
a
Multi-gigapixel images
cytomine
b
c
http://$CYTOMINE_URL/#tabs-image-idproject-idimage-
d
http://$CYTOMINE_URL/#share-annotation/idannotation
e
http://$CYTOMINE_URL/api/annotation.json?&project=idproject&users=idusers
f
g
h
i
j

cytomine


Cytomine-DataMining
Language: Python
Framework: scikit-learn
PYTHON CLIENT
Language: Python
Cytomine-java-client
Language: Java
...
Language: ...
Cytomine-WebUI
Language: Javascript
Framework: Backbone.js, AngularJS
Data request
Image upload request
Image request
RabbitMQ
Nginx
JSON
Nginx
Nginx Upload
IMAGE
JSON
Cytomine-Core
Languages: Groovy, Java
Framework: Grails
REST Api
Url Mappings
Tests
Controllers
Services
Domains
Cytomine-IMS
Languages: Groovy, Java
Framework: Grails
REST Api
Url Mappings
Java
Client
Tests
Controllers
Services
Domains
DATABASE
PostgreSQL
Postgis
Language: PL/pgSQL
DATABASE
MongoDB
Cytomine-IIP
Language: C++
IIP, Openslide,...
FILE SYSTEM
Images

cytomine

Html 5

Javascript

Jquery

Backbone.js

AngularJS

Openlayers

Cytomine-WebUI

Scikit-learn

OpenCV

Redis

Content-based Image retrieval
Cell Classifier
Landmark Detection
Tissue Segment
Annotation Stats
...

Java

Python

Cytomine-DataMining

Python

Cytomine-clients

Java

REST API HTTP NGINX

RabbitMQ
Message broker

REST API HTTP NGINX

Grails

Groovy

Java

Spring

Tomcat

Postgis

PostgreSQL

Cytomine-Core

MongoDB

Grails

MEMCACHED

Java

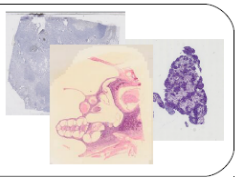
IIP + OpenSlide

C++

Cytomine-IMS

GlusterFS

File system



Docker

Cytomine Install/Deploy

Node.js

MongoDB

Cytomine Monitoring

cytomine

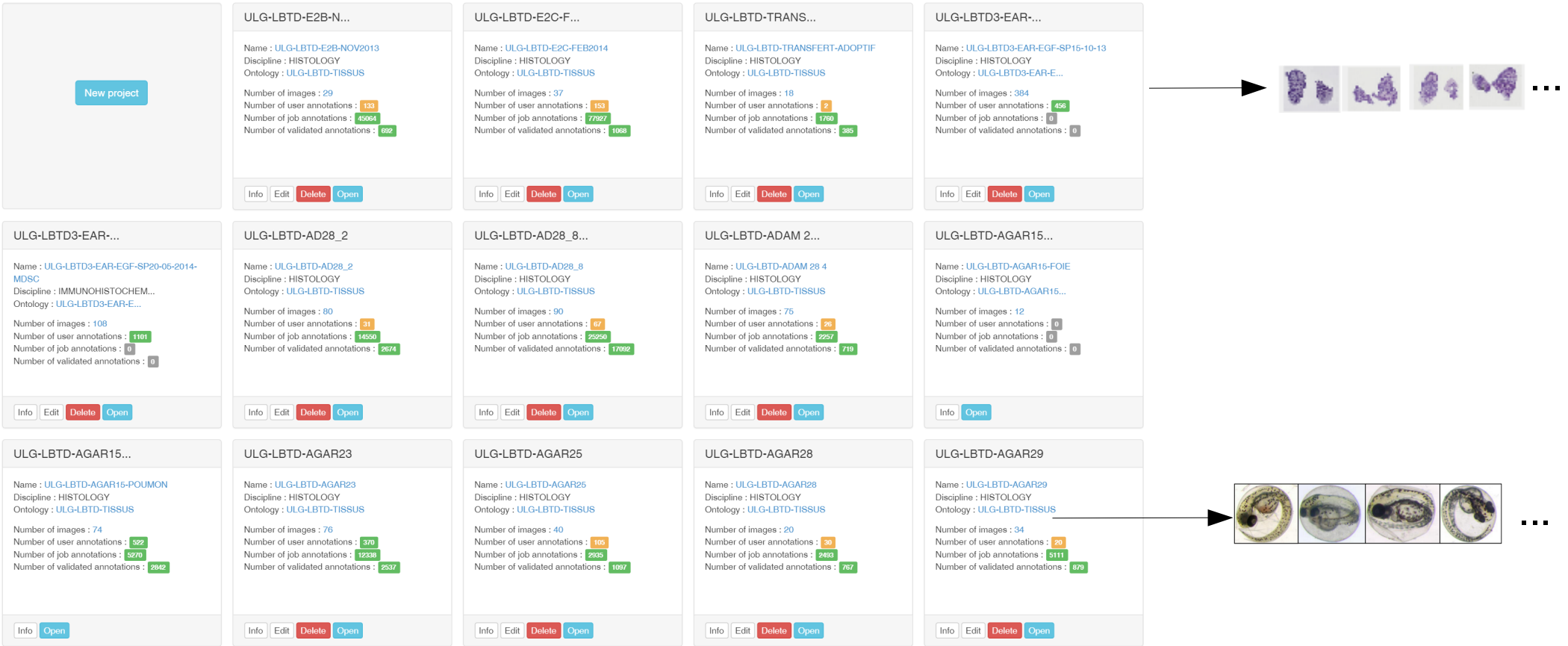
Collaborative analysis of gigapixel images

www.cytomine.be

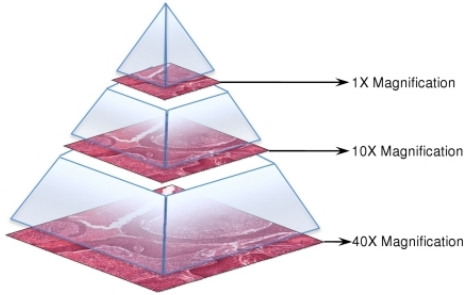
cytomine features :
Organize and centralize on the web

Create and manage multiple **projects** :

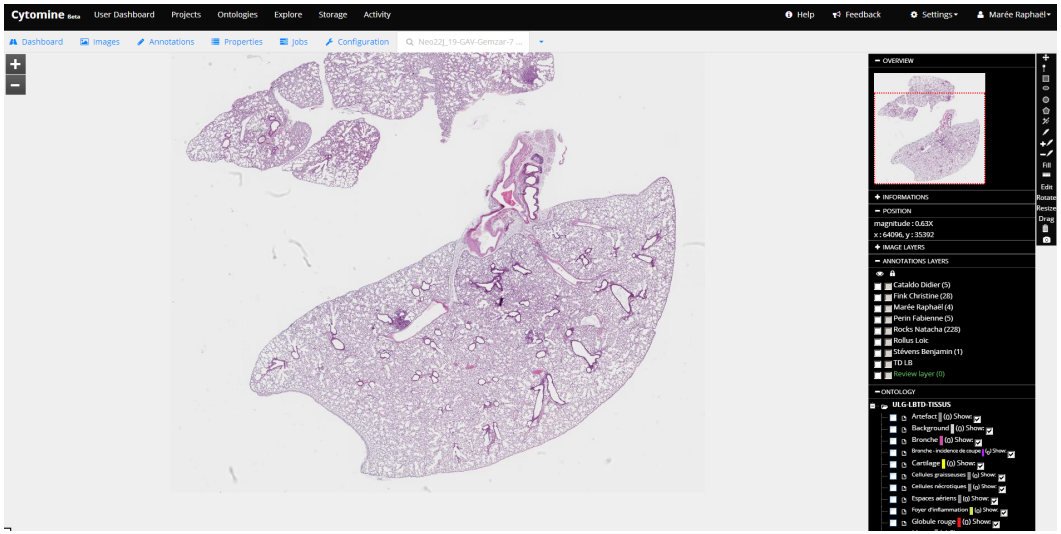
- **Upload** images to centralized server or keep data local (distributed image tile servers)
- Support for **various formats** (TIFF, JP2000, Aperio SVS, Hamamatsu NDPI/VMS, 3DHistech MRXS, Leica SCN, Roche TIF...)
- Users with **authentification** (e.g. LDAP), **access rights,** and **roles**
- **Specific ontologies** with user-defined, vocabulary terms

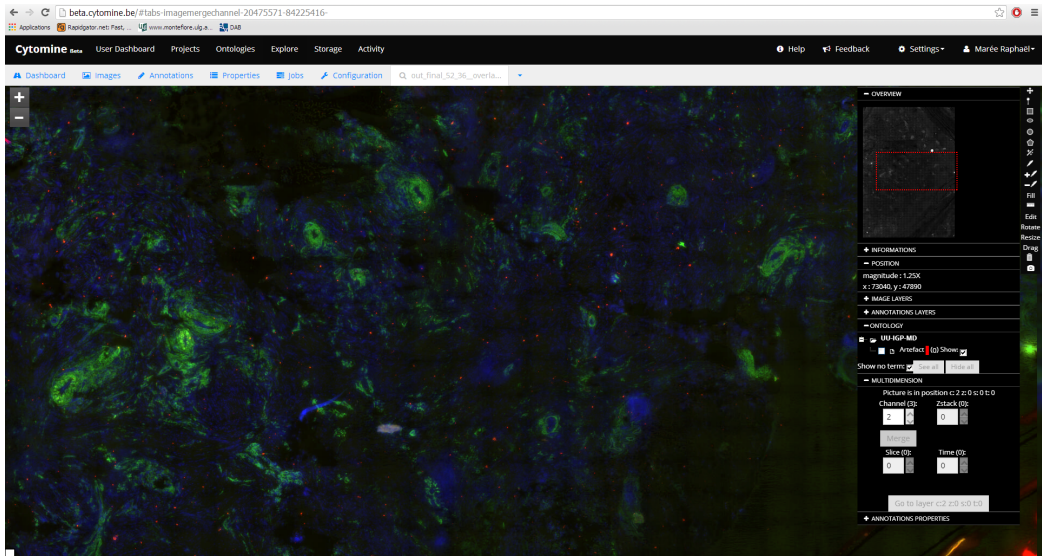


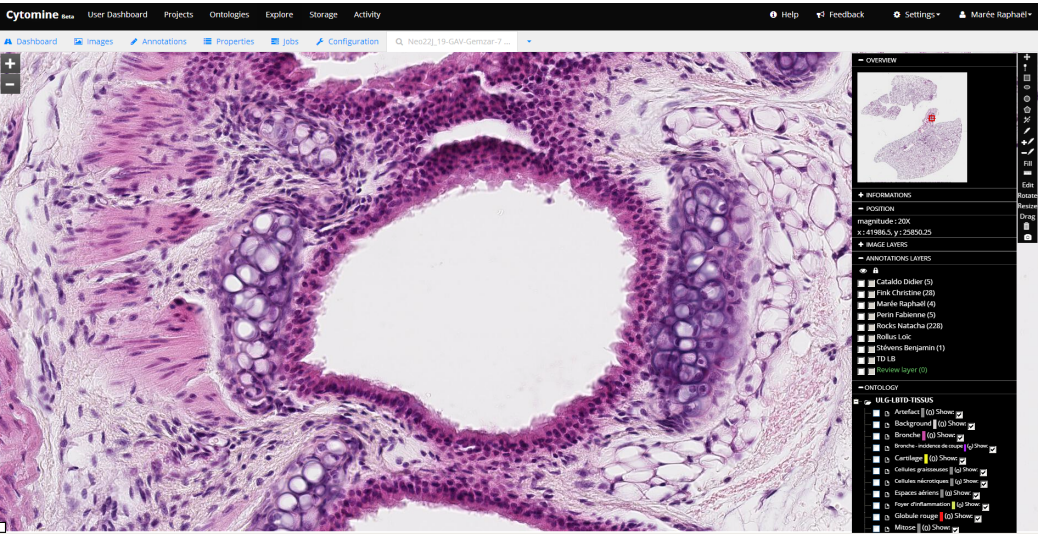
cytomine features : Visualize

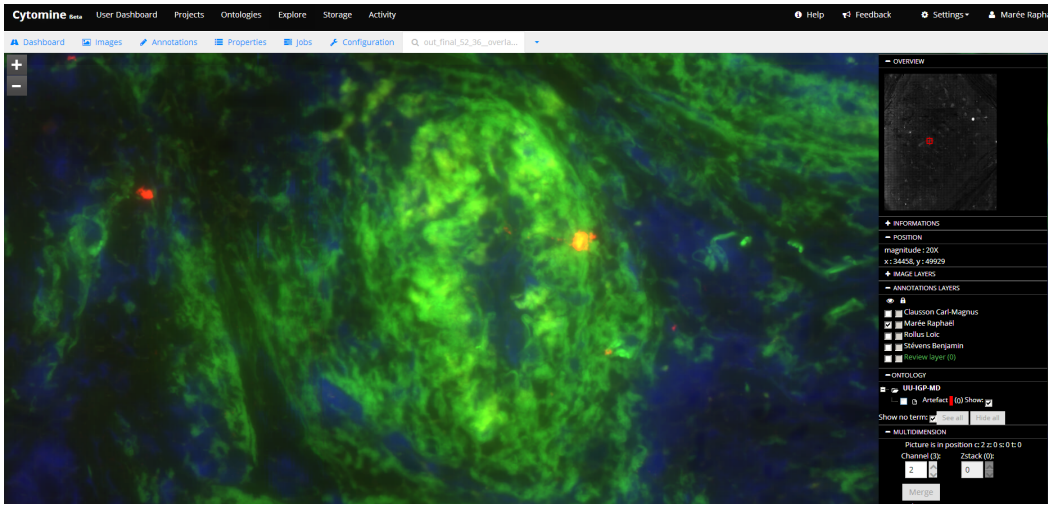


- Explore **large** (>gigabyte) images at multiple resolutions, **remotely**
- GoogleMaps/OpenStreetmap browsing style (zoom in/out, pyramid tile-based)







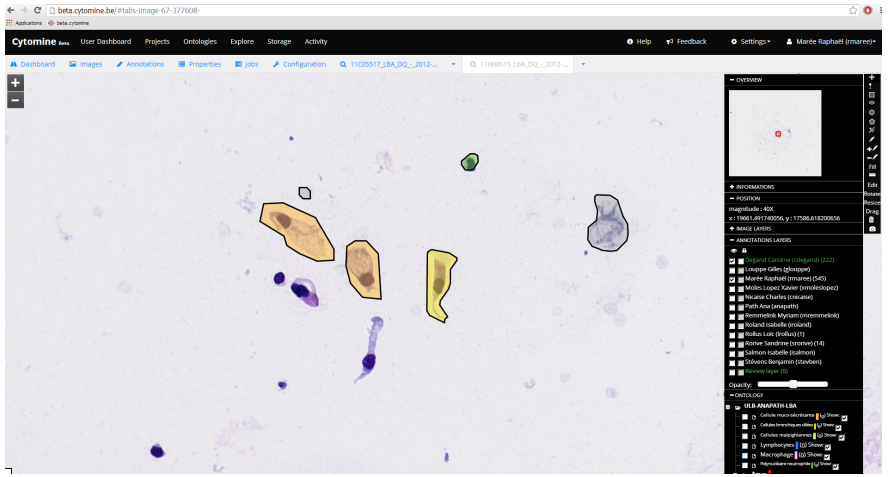


1 tissue slice = 35000 x 30000 pixels (0.23µm/pixel)

4 fluo channels 83000 x 100 000 pixels = 4 x 16GB image

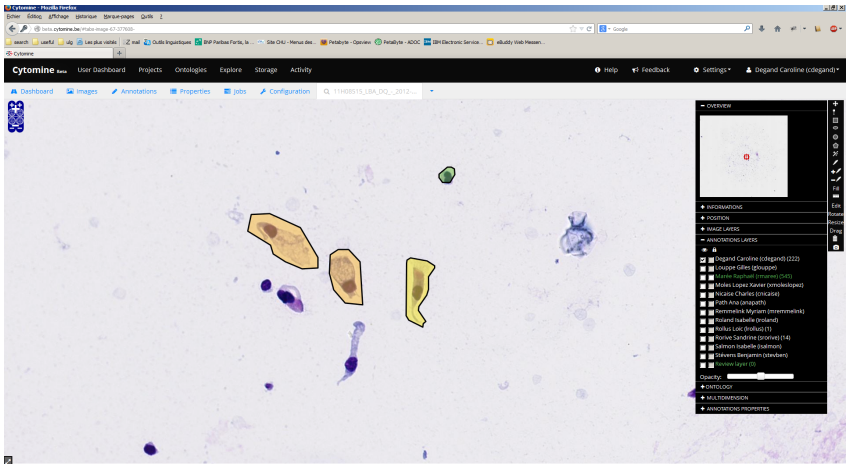
(image from O.Söderberg group, Uppsala)

cytomine features : Live broadcast



User A

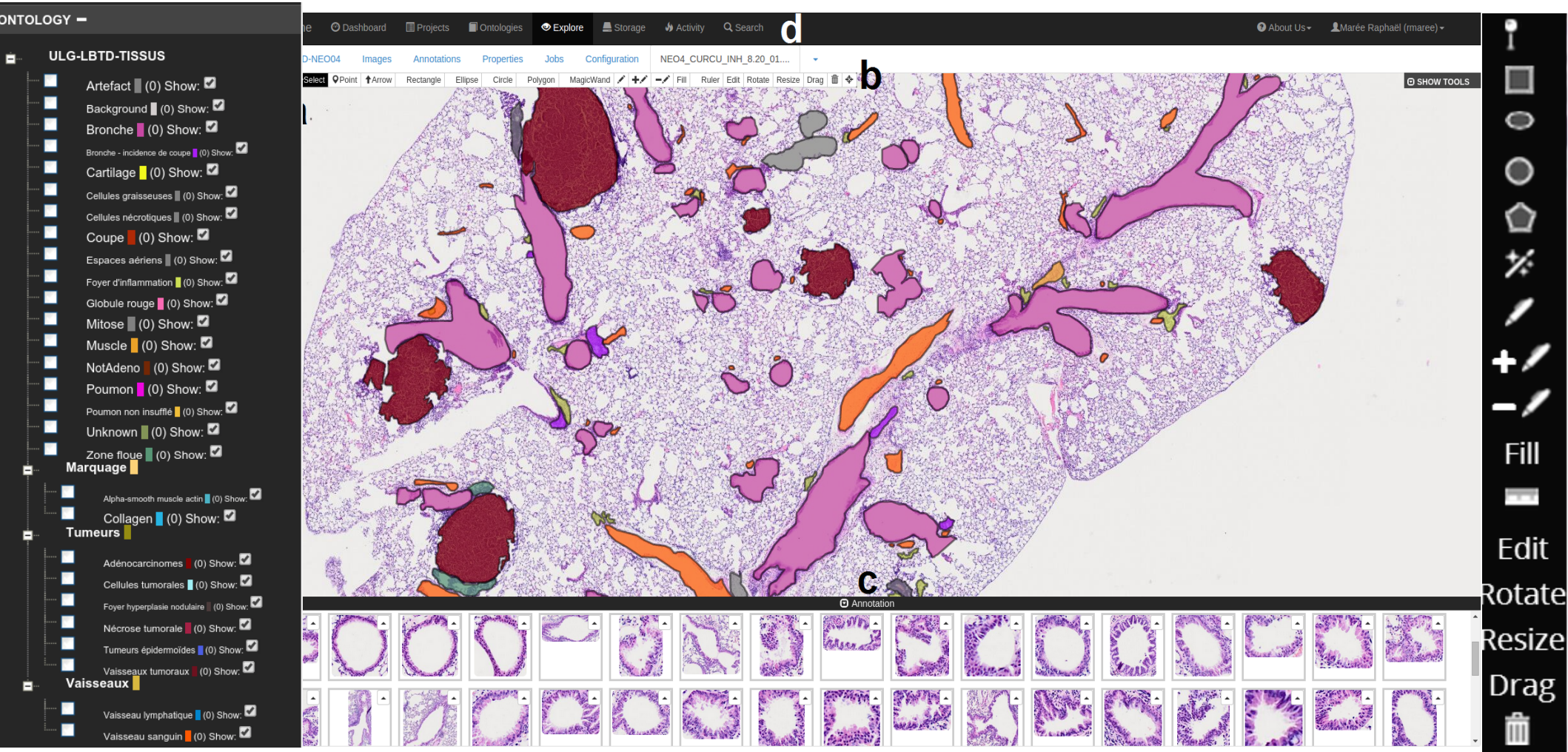
User B



Start following Marée Raphaël (rmaree)

cytomine features : Annotate

- **Annotate** images using various **drawing tools**, with **user-specific layers**
- Describe ROIs with **ontology terms**
- Describe images and ROIs with any **key-value properties** or **text description**



cytomine features : Ontology editor

Ontologies + Add | Refresh

ALCIAN BLUE
ALIZARIN RED
ATEST
AUSTRALIAN-MUSEUM-INSECT-SOUP
BLABLA
BM-01
BUDÉSONIDE 1 BALF
BUDÉSONIDE 1 SCORE INFLAMMATOIRE
CELLSOLUTIONS-PROSTATE
CMU-BE-DEMO
CVM
CellSolutions-PAP
DEFAUTSSHEARIOGRAPHIE
DEMO
DEMO-OXFORD-ZEGAMI
DEMO-UMINN-TPENGO
DEMO-VARIOUS
DEMO_2014
EDEMA
FRAUNHOFER-IIS-DEMO
GHDC-PLANCANCER
HISTOLOGIE DIVERS
HUBBLE
IPASTEUR-DEMO
LABSETDEMO
LS4-NEWANNOTATIONS
MEDEX-CHEST-CCD
MEDUNIGRAZ-BONE-MARROW
MUMC-CARIM-PATHOLOGY
NANJING-DEMO
OPENCONNECTOME-KASTHURI11
PHYSIO 1 BALF
PROJET_DEMO
RHTEST
ROSTOCK-HJTHIESEN-KIDNEY-TMA
RWTH-METRIPOL-CRYSTAL
SEGMENT_FISH
TB
TEST-CZI
TEST-MERGE
TEST-MERGE-TIFF

ULG-LBTD-TISSUS

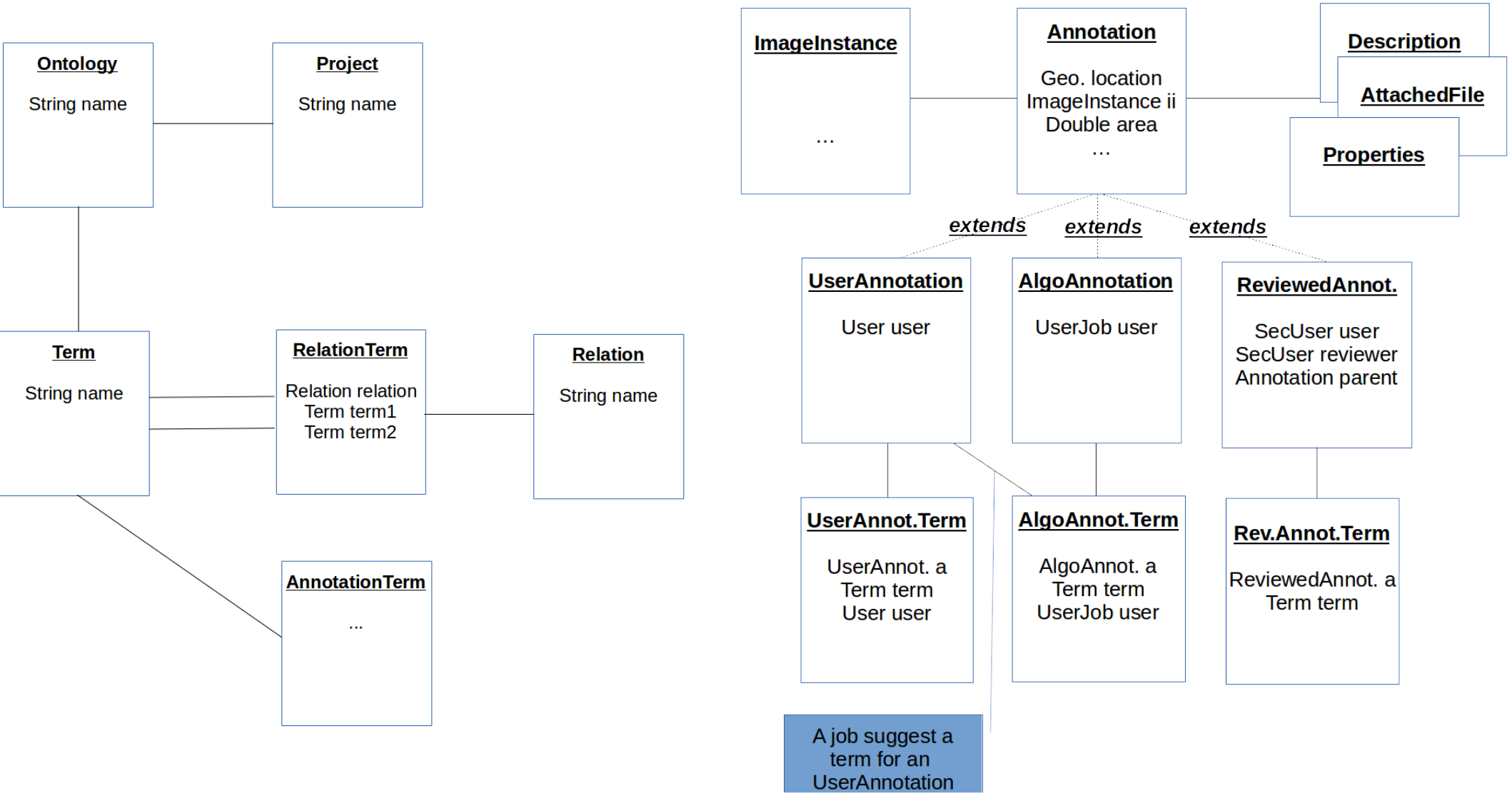
Ontology | Term

- Add
- Edit
- Delete

Ontology : ULG
Creator : Stéve
Shared with :
Projects : ULG D-NEO13, ULG-TEST-PHL, ULG-LBTD-NEO22, ULG-LBTD-HDM1-HE, ULG-LBTD-CS7, BORDET, ULG-LBTD-AGAR15-POUMON, ULG-LBTD-W18-HE, ULG-LBTD-W18-CR, ULG-LBTD-CR1-HISTO, ULG-LBTD-O3, ULG-LBTD-PGP, ULG-LBTD-G1-HE POUM, PGPTEST, ULG-LBTD2-BALBC-4T1, ULG-LBTD-LIDOCAINE, ULG-LBTD-AS6, ULG-LBTD-AGAR23, ULG-LBTD-AGAR25, TESTAGA, ULG-LBTD-AGAR28, ULG-LBTD-SH1-IV-POUMON, ULG-LBTD-SH1-SC-POUMON-TUMEUR, ULG-LBTD-E2A-FEB2013, ULG-LBTD-TRANSFERT-ADOPTIF, ULG-DEMO-HE-LUNG, TESTSTATS, ULG-LBTD-E2B-NOV2013, ULG-LBTD-E2C-FEB2014, ULG-LBTD-AGAR29, ULG-LBTD-AD28_2, ULG-LBTD-E2E-JUN2014, ULG-LBTD-ADAM 28 4, ULG-LBTD-E2F-AOUT 2014, ULG-LBTD-AD28_8 , ULG-LBTD-AGIC3, ULG-LBTD-AGIC1, ULG-LBTD-AGIC2, ULG-LBTD-AGIC6, ULG-LBTD-IT MC IV 4T1 G1, ULG-LBTD-AGIC5, ULG-LBTD-AGIC7, ULG-LBTD-AGIC8, ULG-LBTD-AGCT1, ULG-TEST-RABBIT-SEGMENT, ULG-LBTD-AGDC1bis, ULG-LBTD-AGDC2, ULG-LBTD-AGDC5, ULG-LBTD-AGDC6, ULG-LBTD-MT1, ULG-LBTD-MT3, ULG-LBTD-MT2, ULG-LBTD-AGIC9, ULG-LBTD-AGDC10, ULG-LBTD-MT5, ULG-LBTD-MT5', ULG-LBTD-AGDC9, ULG-LBTD-MT4, ULG-LBTD-IT IL16 10NG GR3

- **ULG-LBTD-TISSUS**
  - Artefact
  - Background
  - Bronche
  - Bronche - incidence de coupe
  - Cartilage
  - Cellules graisseuses
  - Cellules nécrotiques
  - Coupe
  - Espaces aériens
  - Foyer d'inflammation
  - Globule rouge
  - Mitose
  - Muscle
  - NotAdeno
  - Poumon
  - Poumon non insufflé
  - Unknown
  - Zone floue
  - **Marquage**
    - Alpha-smooth muscle actin
    - Collagen
  - **Tumeurs**

cytomine

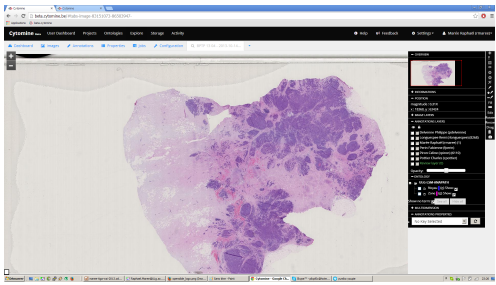

Ontology
String name
Project
String name
Term
String name
RelationTerm
Relation relation
Term term1
Term term2
Relation
String name
AnnotationTerm
...
ImageInstance
...
Annotation
Geo. location
ImageInstance ii
Double area
...
Description
AttachedFile
Properties
extends
extends
extends
UserAnnotation
User user
AlgoAnnotation
UserJob user
ReviewedAnnot.
SecUser user
SecUser reviewer
Annotation parent
UserAnnot.Term
UserAnnot. a
Term term
User user
AlgoAnnot.Term
AlgoAnnot. a
Term term
UserJob user
Rev.Annot.Term
ReviewedAnnot. a
Term term
A job suggest a term for an UserAnnotation

cytomine features : Share

- **Share images** through simple URLs

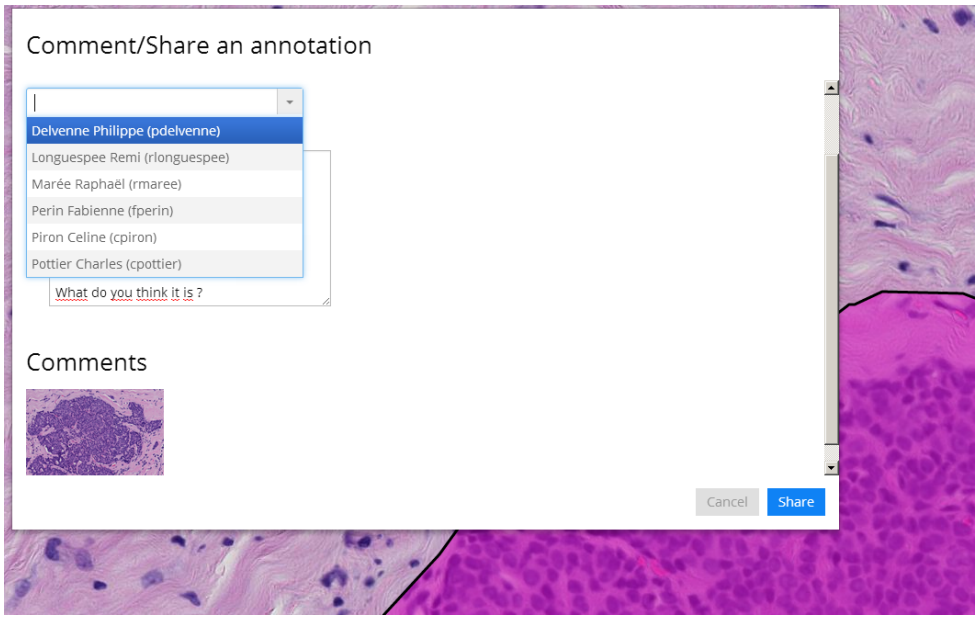
http://beta.cytomine.be/#tabs-image-83151073-86503947- →

→



- **Share** and **comment annotations** through simple URLs & e-mail mechanisms

http://beta.cytomine.be/#share-annotation/77044



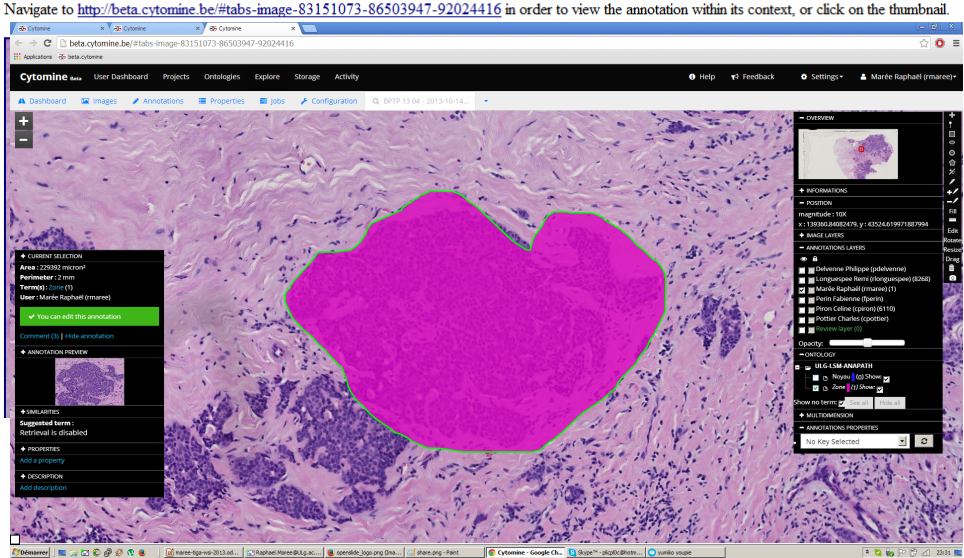
From: cytomine.ulg@gmail.com **To:** raphael.maree@ulg.ac.be
Subject: Cytomine : Marée Raphaël (rmaree) shared an annotation with you **Date:** Fri, 29 Nov 2013 23:12:05 +0100 (CET)

Dear Marée Raphaël (rmaree),

Marée Raphaël (rmaree) shared an annotation with you and said :

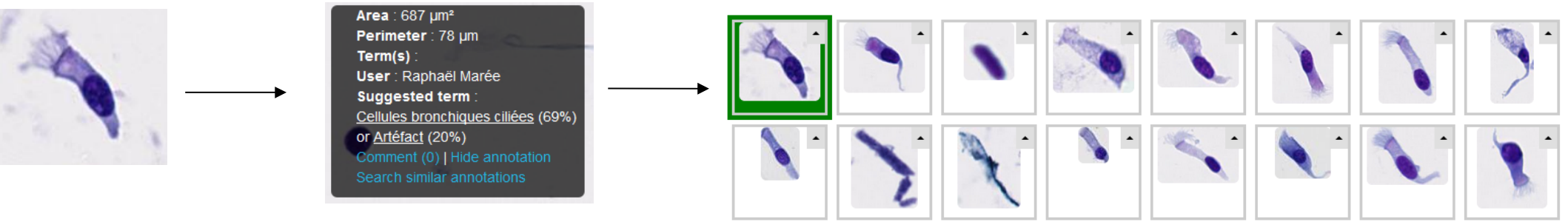
Dear colleague, Do you think it is grade C ?

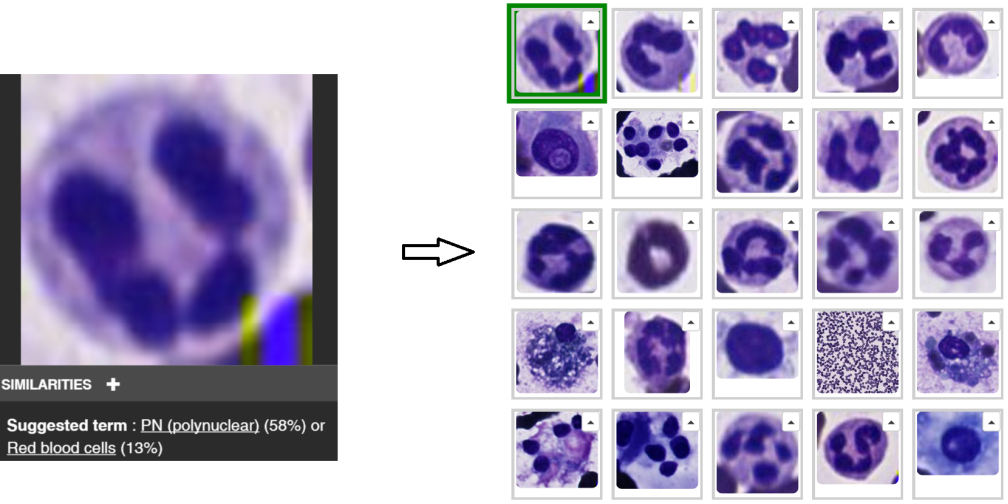
Navigate to http://beta.cytomine.be/#share-annotation/92024416 in order to reply.
Navigate to http://beta.cytomine.be/#tabs-image-83151073-86503947-92024416 in order to view the annotation within its context, or click on the thumbnail.



cytomine features : Search

- **Text search** (projects, images, annotations)
- **Visual search** (CBIR) of regions of interest and **ontology term suggestion**



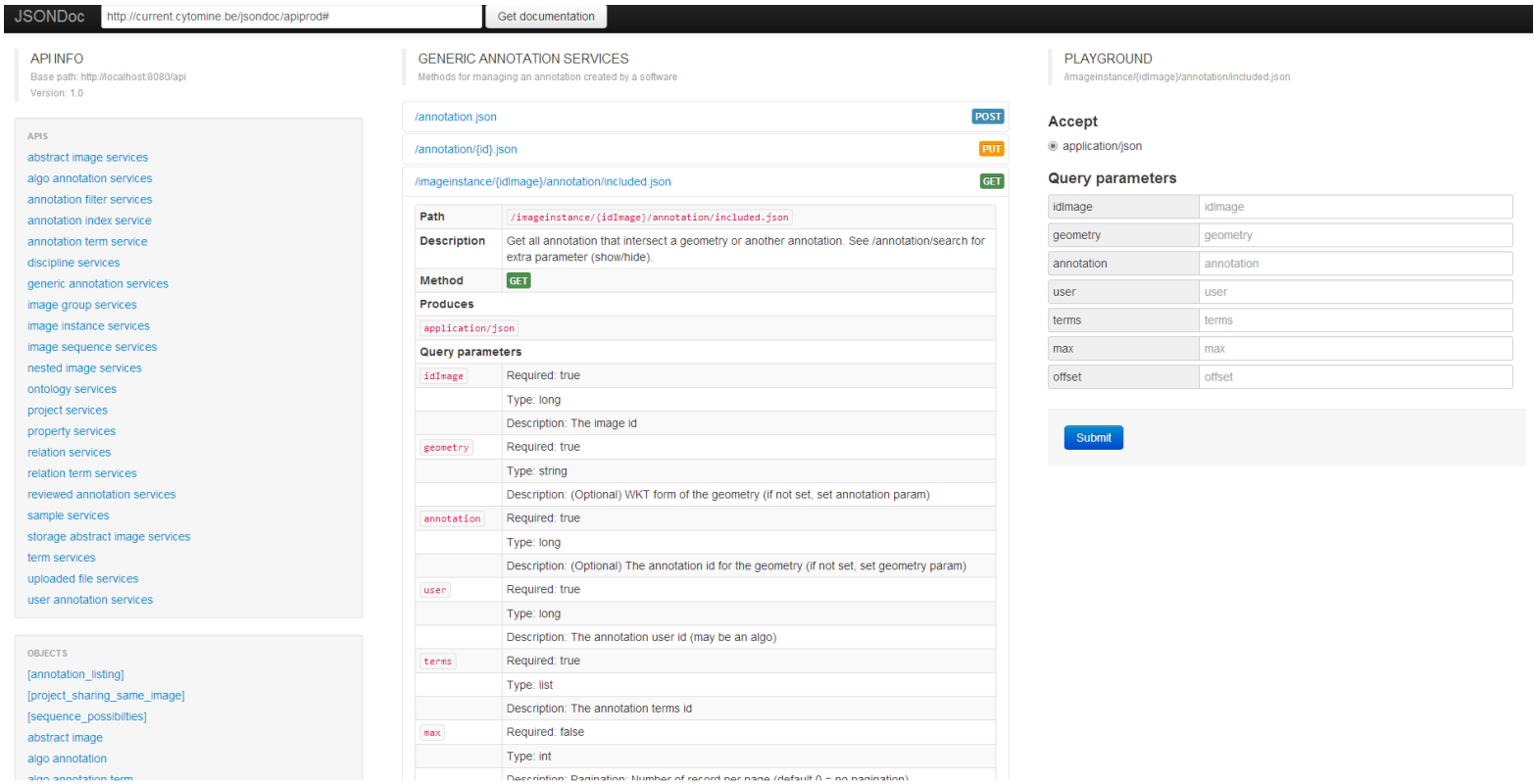


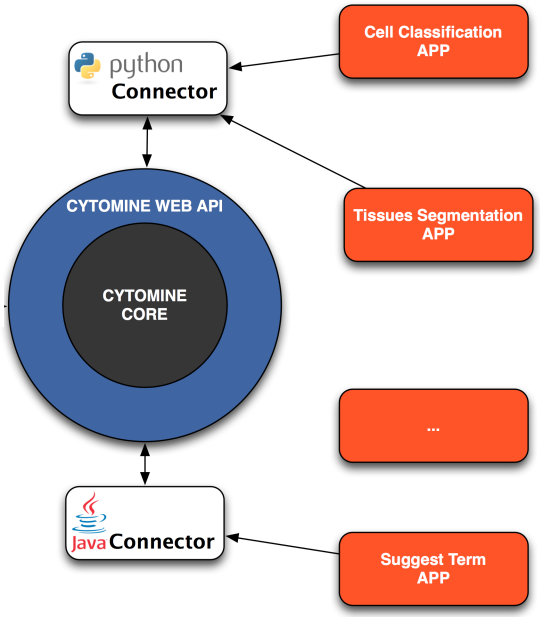
Marée et al., Incremental indexing and distributed image search using shared randomized vocabularies, Proc. MIR 2010

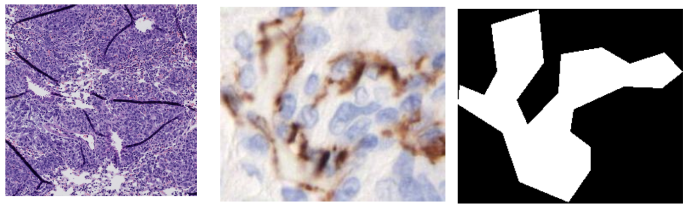
cytomine : Fully open

- **Documentation** wiki (doc.cytomine.be)
- **User guide**
- **Open REST API**
- **Open source** code

- Extensible
- Interoperable





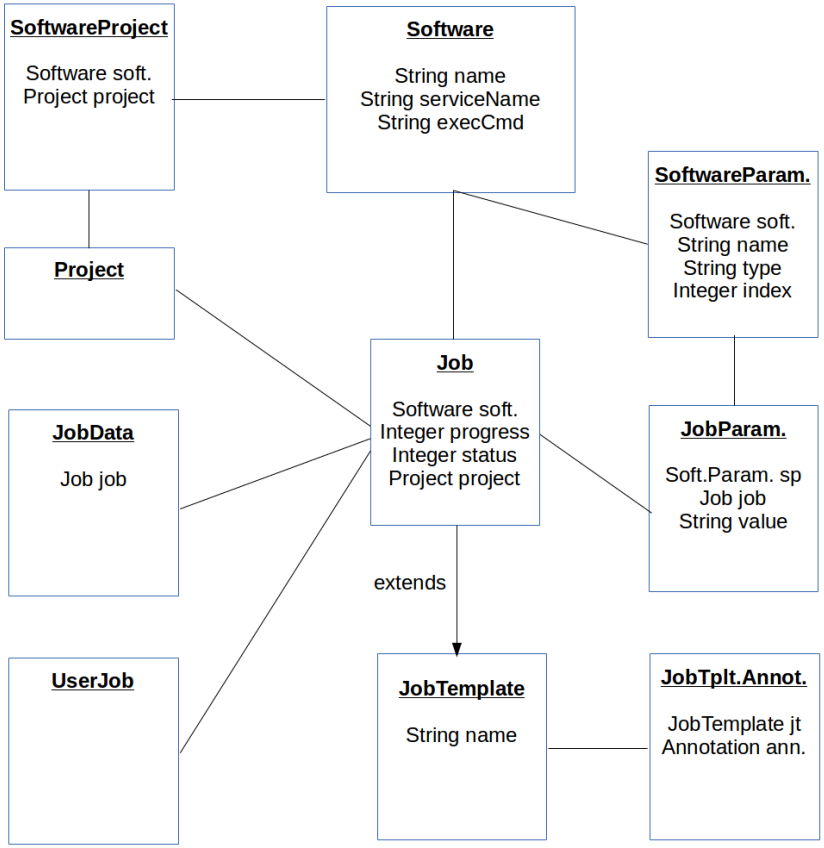


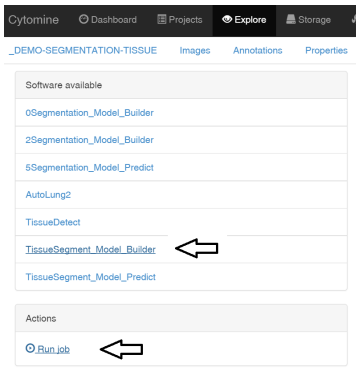
{"class":"be.cytomine.ontology.UserAnnotation","id":95189319,"created":"1390233213025","location":"POLYGON ((83409.75 96490.375, 83412.75 96492.375, 83412.75 96481.375, 83411.75 96476.375, 83408.75 96473.375 … 96489.375, 83409.75 96490.375))", "image":95062074,"project":94152784,"container":94152784, "user":94115159,"area":8953.0,"perimeterUnit":"mm","areaUnit":"micron²","perimeter":1.0, "centroid":{"x":83426.59663300458,"y":96382.22057123308},"term": [95062790],"cropURL":"http://beta.cytomine.be/api/userannotation/95189319/crop.jpg","reviewed":true}

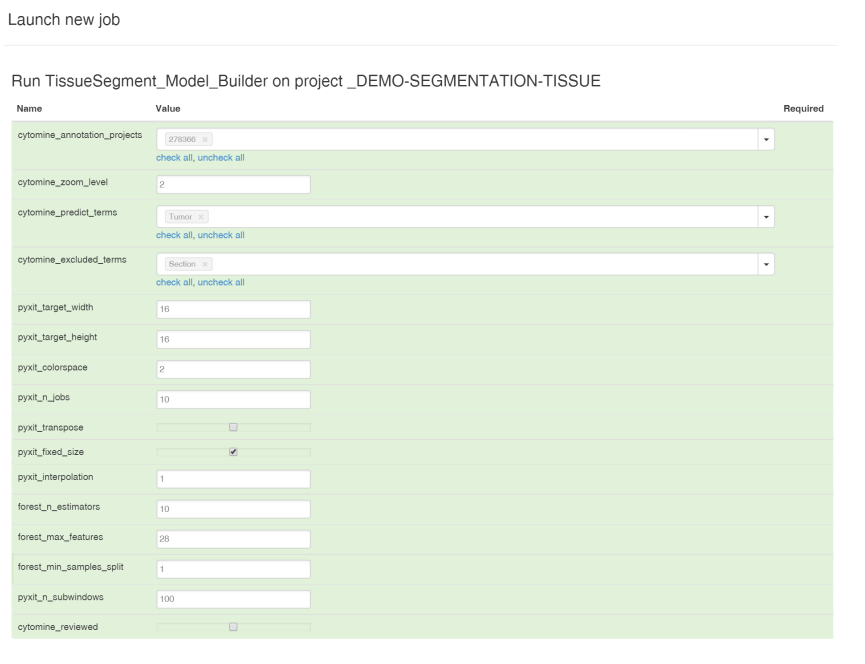
http://$HOST/api/imageinstance/7889452/window-10000-10000-1500-1500.jpg
http://$HOST/api/userannotation/95189319/crop.png?zoom=0
http://$HOST/api/userannotation/95189319/crop.png?mask=true
http://$HOST/api/userannotation/95189319.json

Write and plug your own algorithms (Java/Python/…) or design your own user interfaces

cytomine features :
Third-party software integration





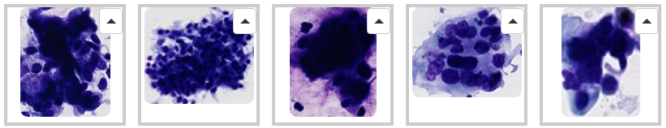


cytomine features : Algorithm evaluation (e.g. class conf matrix)

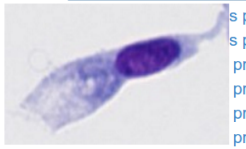
- **Average (per class)** : 76.80

View confusion matrix | View predicted galleries

X	Amas.	Amas. non-.	Amas. susp.	Arté.	Autr.	Bact.	Cell. muco.	Cell. bron. cili.	Cell. malp.	Cell. tumo.	Cham.	Lymp.	Macr.	Mucu.	Non. smal. cell. carc.	Poly. neut.	Smal. cell. carc.	total
Amas.	0																	
Amas. non-.		44						2	7						5	5		70%
Amas. susp.			5												19	3		19%
Arté.		2	1	597		1	5	21				3	12		5	7		91%
Autr.					0													
Bact.		2		7		29	4	3										64%
Cell. muco.		1		9		1	126	27	9				2		1			72%
Cell. bron. cili.		1		6			6	265	1				3		9	7		89%
Cell. malp.		2		5			6	1	320				2		3			94%
Cell. tumo.										0								
Cham.											0							
Lymp.								2				96	3		5	29		71%
Macr.		1		4			2	2	7			1	385		4	7		93%
Mucu.														0				
Non. smal. cell. carc.			1	3				5	5			1	5		270	13		89%
Poly. neut.		1		2				3	2			12	4		10	413		92%
Smal. cell. carc.																	0	



- **98%** : Annotation 2182594 is predicted Cellules malpighiennes instead of Amas non-suspects
- **97%** : Annotation 2511070 is predicted Cellules malpighiennes instead of Amas non-suspects
- **96%** : Annotation 1814853 is predicted Non small cell carcinoma instead of Artéfact
- **96%** : Annotation 1815026 is predicted Non small cell carcinoma instead of Amas suspects
- **96%** : Annotation 59515 is predicted Macrophage instead of Artéfact
- **96%** : Annotation 673307 is predicted Non small cell carcinoma instead of Polynucléaire neutrophile
- **96%** : Annotation 2017832 is predicted Non small cell carcinoma instead of Amas non-suspects
- **96%** : Annotation 2177840 is predicted Cellules malpighiennes instead of Amas non-suspects
- **96%** : Annotation 796313 is predicted Macrophage instead of Artéfact
- **95%** : Annotation 682189 is predicted Cellules bronchiques ciliées instead of Cellule muco-sécrétante

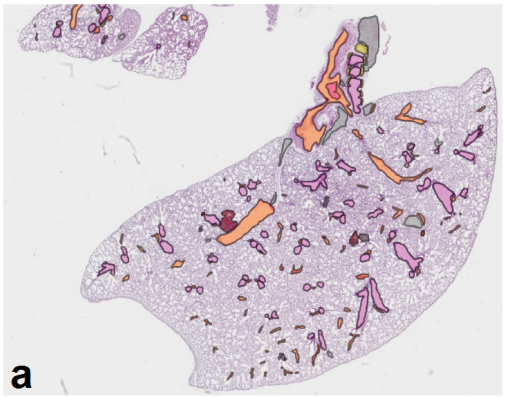


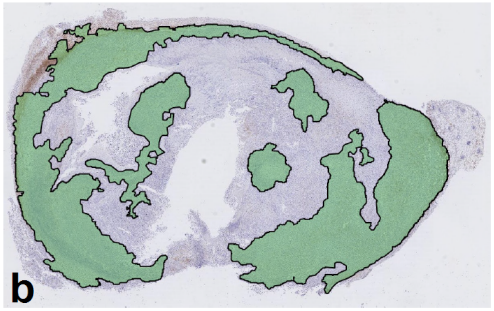
s predicted Non small cell carcinoma instead of Amas suspects
s predicted Polynucléaire neutrophile instead of Non small cell carcinoma
predicted Artéfact instead of Bactérie
predicted Cellules bronchiques ciliées instead of Artéfact
predicted Cellules bronchiques ciliées instead of Cellule muco-sécrétante
predicted Non small cell carcinoma instead of Lymphocytes

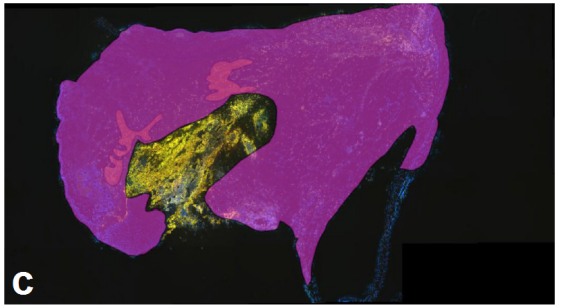
+ proofreading
(see applications)

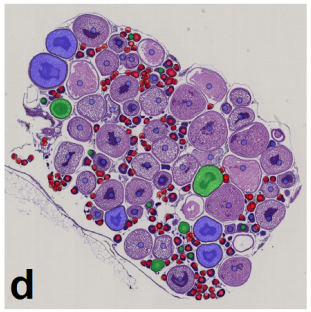
cytomine research applications :

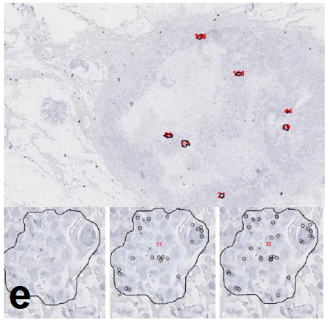
As of March 2016 : > 200 registered users, > 300 projects,
> 21 000 images, > 700 000 manual annotations
> 6 000 000 algorithm annotations

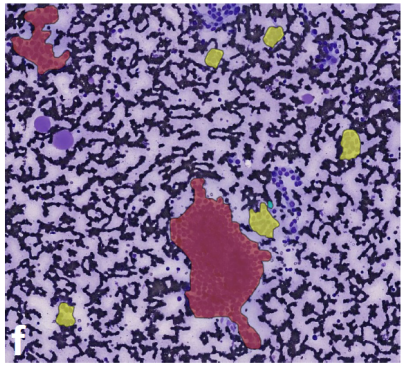


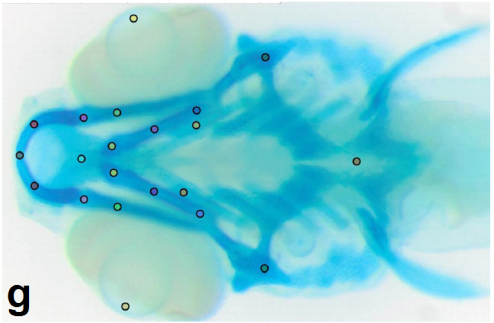


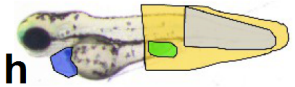


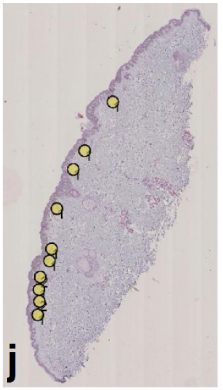


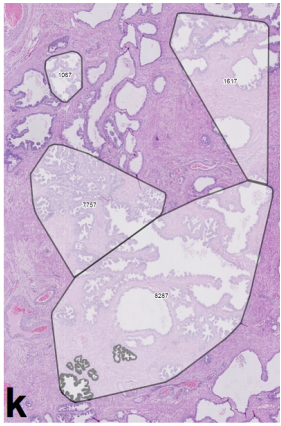








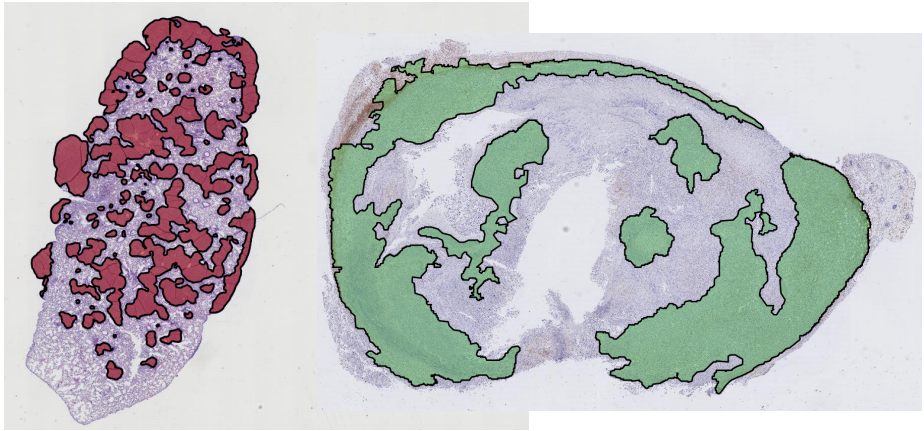




***a**: H&E mice lung cancer research (D.Cataldo's lab, GIGA-Research) @ ISBI 2014.*
***b**: IHC mice lung cancer research (P. Martinive's lab, GIGA-Research) @ OncoTarget 2015.*
***c**: Immunofluorescent mouse ear sponge assays in tumor angiogenesis. @ Journal of Pathology 2015.*
***d**: H&E Chondrostoma nasus sexual maturation research (Gennotte's lab, CEFRA).*
***e**: in situ hybridization assays in human breast cancer research (C.Josse's lab, GIGA-Research) @ BMC Cancer 2015*
***f**: Human thyroid cytology (I.Salmon's lab, ULB Anatomical Pathology Department) @ ECDP 2016.*
***g**: Danio rerio embryo development (M. Muller's lab, GIGA-Research) @ PLoS ONE 2015.*
***h**: Danio rerio toxicology research (M.Muller's lab, GIGA-Research) @ PLoS ONE 2015.*
***i**: IHC renal ischemia/reperfusion research (F.Jouret's lab, GIGA-Research) @ Am J Transl Res. 2015.*
***j**: IHC in melanoma microenvironment research (P.Quatresooz's lab, GIGA-Research).*
***k**: H&E in human breast cancer research (R. Longuespée, GIGA-Research) @ Methods. 2016*

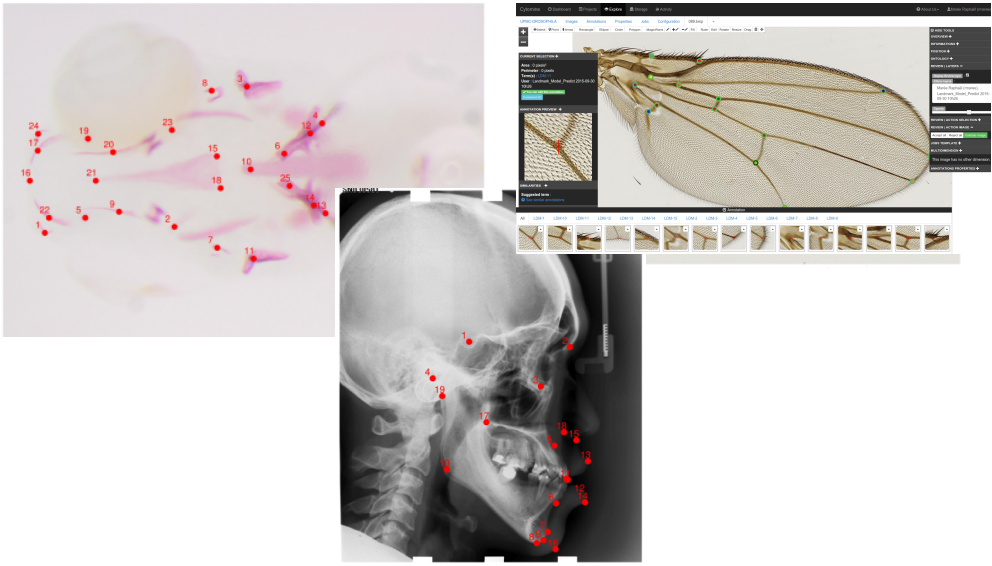
GIGA

cytomine features : Semi-automated analysis using machine learning and proofreading



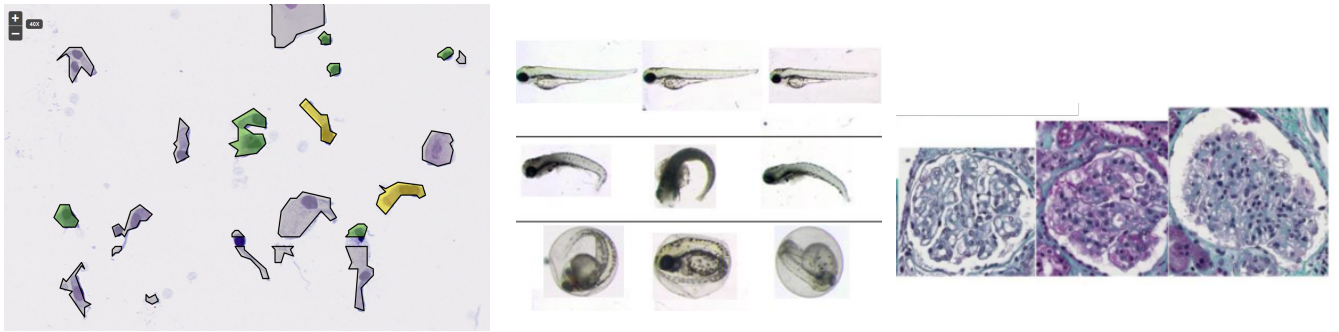
Tissue segmentation

Dumont et al.,VISAPP 2009
Marée et al., ISBI 2014
Leroi et al., Oncotarget 2015



Landmark detection

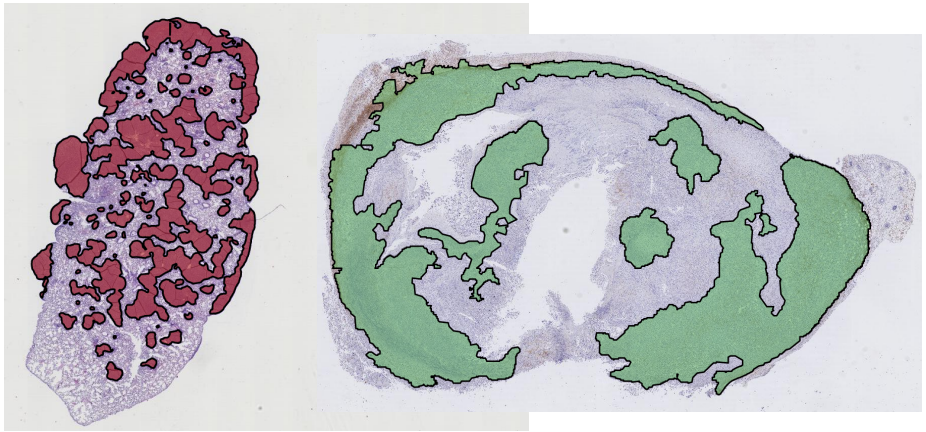
Vandaele et al., Submitted
Huang et al., IEEE TMI 2015



Object classification

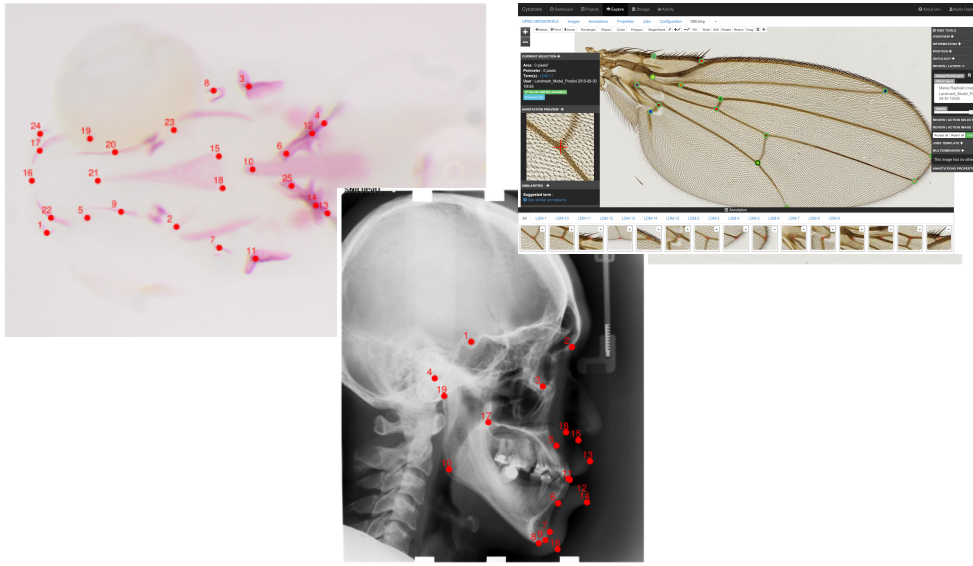
Marée et al. Pattern Recognition Letters 2016 ; ISBI 2016
Delga et al., 2014 ; Jeanray et al., PLoS ONE 2015 ;

cytomine features : Semi-automated analysis using machine learning and proofreading



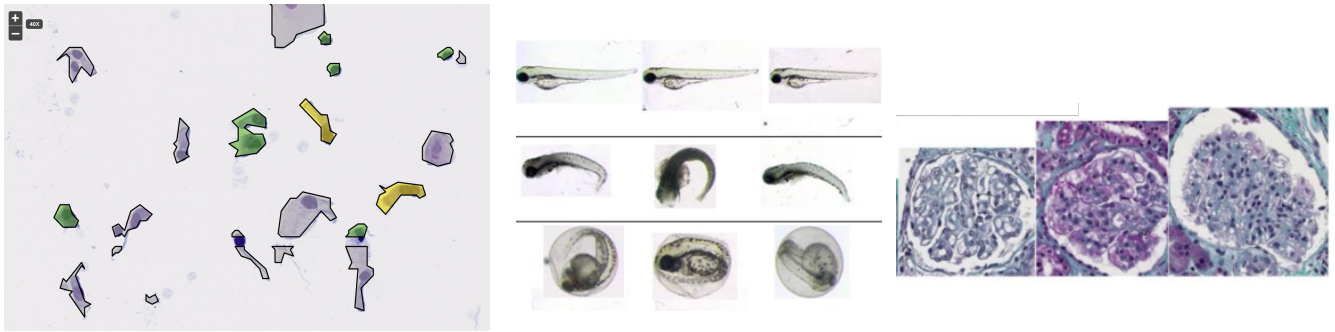
Tissue segmentation

Dumont et al.,VISAPP 2009
Marée et al., ISBI 2014
Leroi et al., Oncotarget 2015



Landmark detection

Vandaele et al., Submitted
Huang et al., IEEE TMI 2015

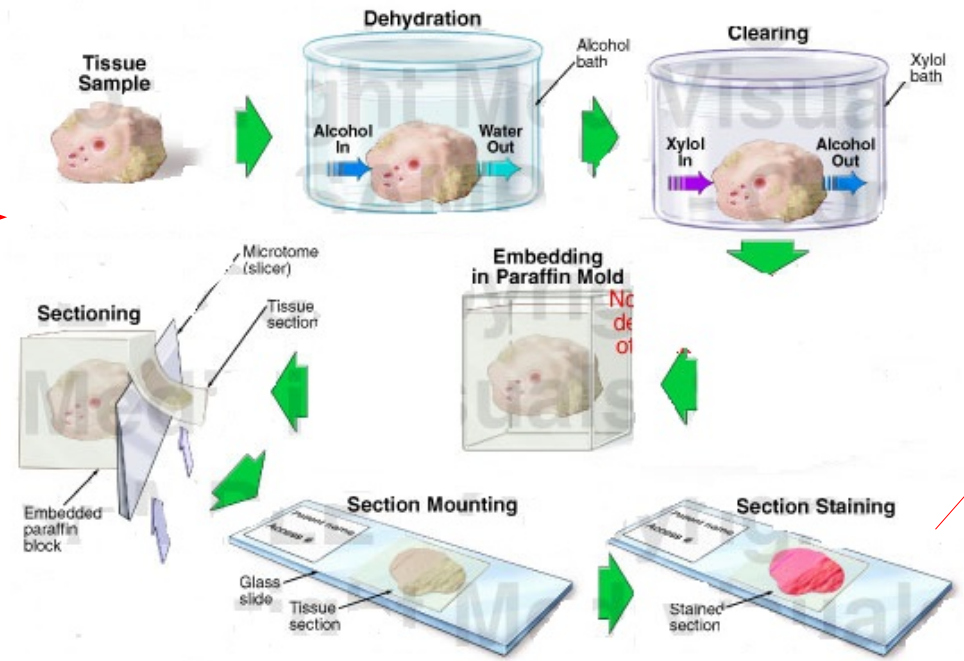


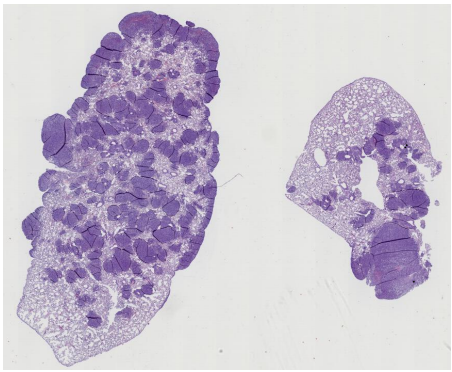
Object classification

Marée et al. Pattern Recognition Letters 2016 ; ISBI 2016
Delga et al., 2014 ; Jeanray et al., PLoS ONE 2015 ;

Tissue recognition : Hybrid human-computer workflow

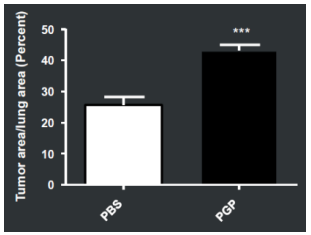
Treated with X

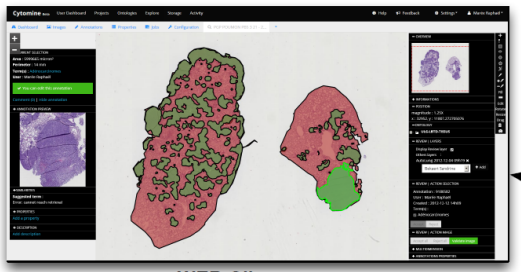




40K x 30K pixels

WEB





WEB Client

PIXEL CLASSIFICATION
python
Java

HTTP(s)
HTTP(s)
HTTP(s)

WEB

distributed processing

CYTOMINE CACHE
CYTOMINE API
CYTOMINE WEB
GRAILS
HIBERNATE

WEB Server

SOCKET

CYTOMINE DATABASE
PostgreSQL
PostGIS

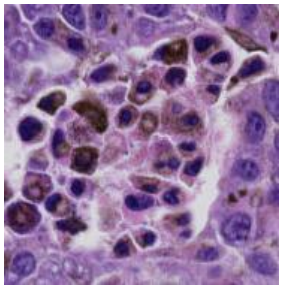
Spatial DB Server

A hybrid human-computer approach for large-scale image-based measurements using web services and machine learning,
Marée et al.
Proc. IEEE International Symposium on Biomedical Imaging, 2014

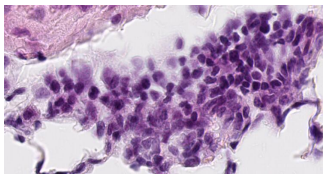
cytomine

Hybrid human-computer workflow

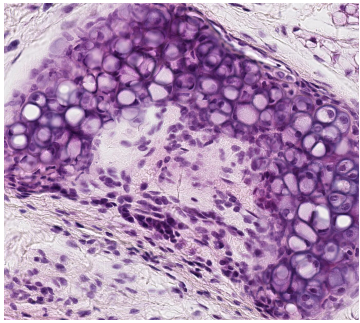
Manual region contouring and labelling to provide training examples



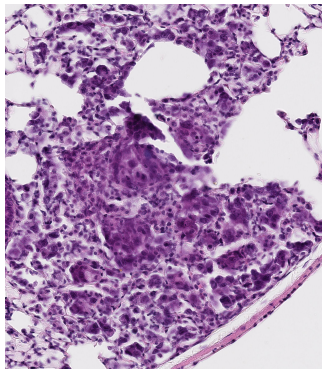
necrosis



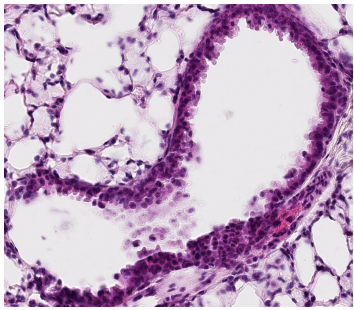
inflammatory cells



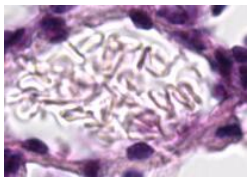
cartilage



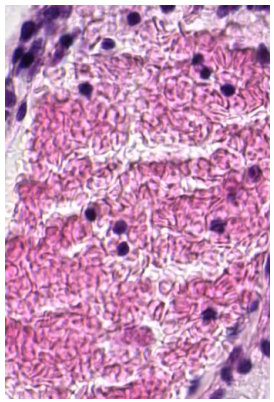
adenocarcinoma



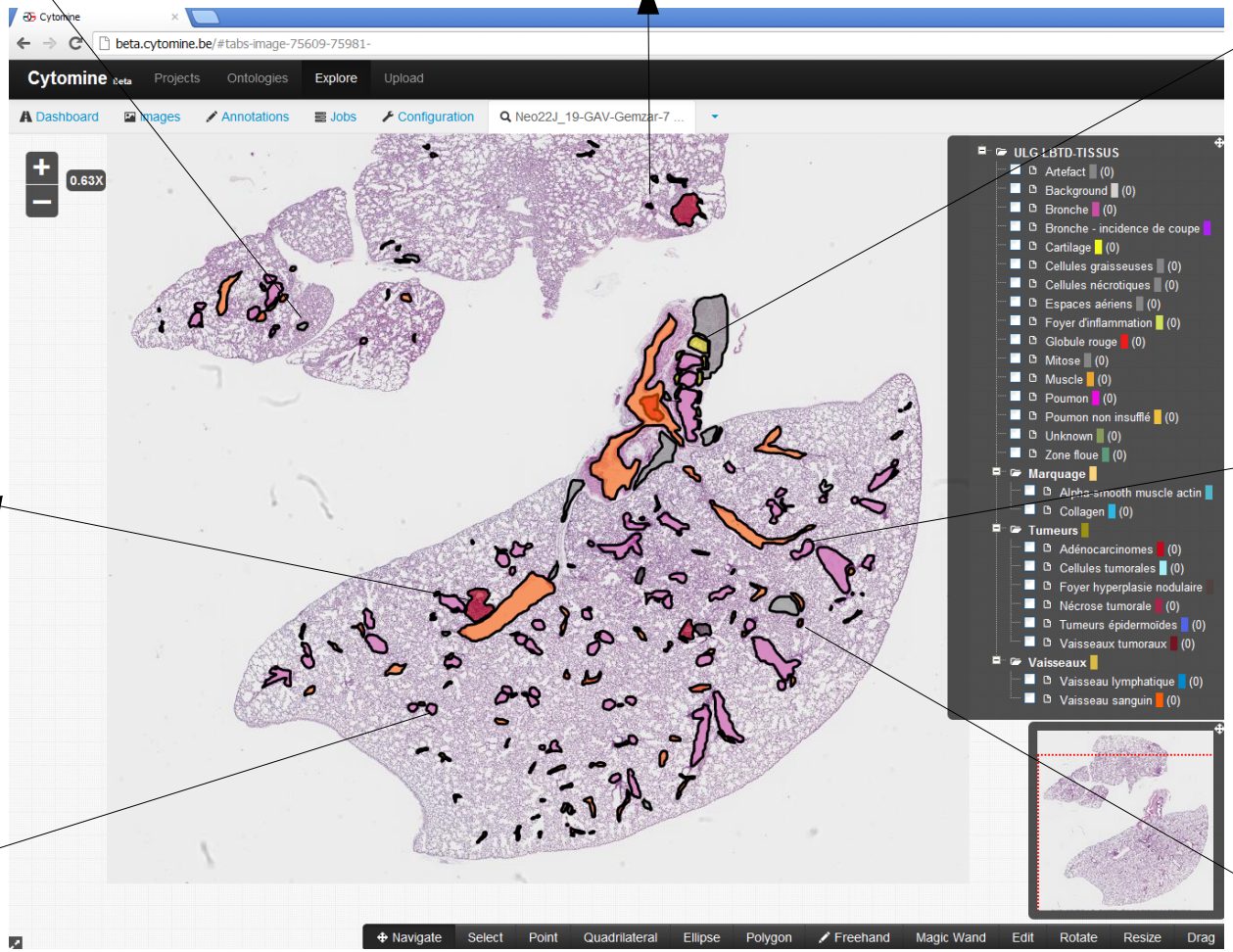
bronchus



blood vessel



red-blood

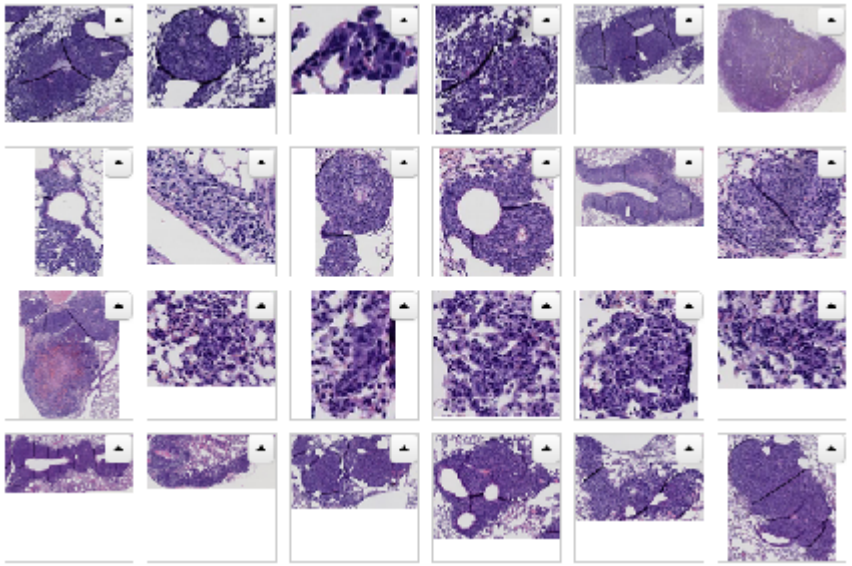


One image ~ 40 K x 30 K pixels

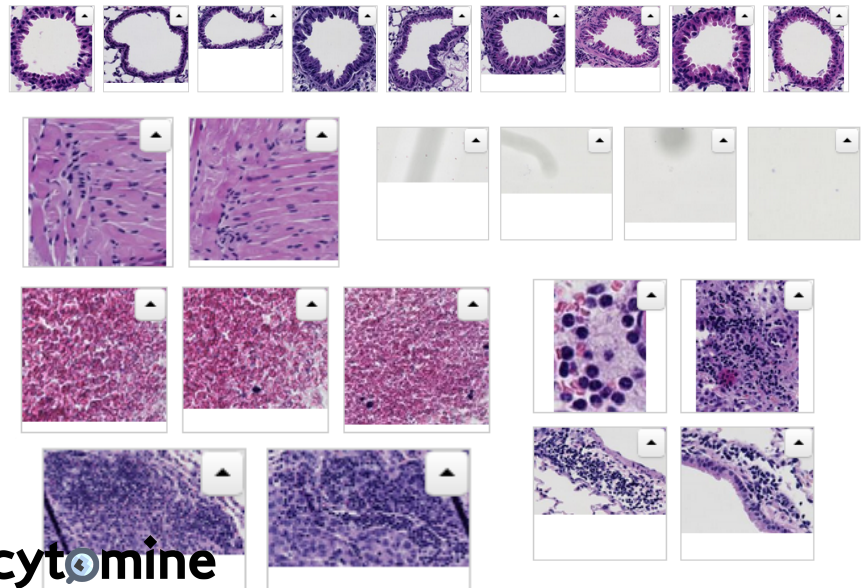
cytomine

Hybrid human-computer workflow

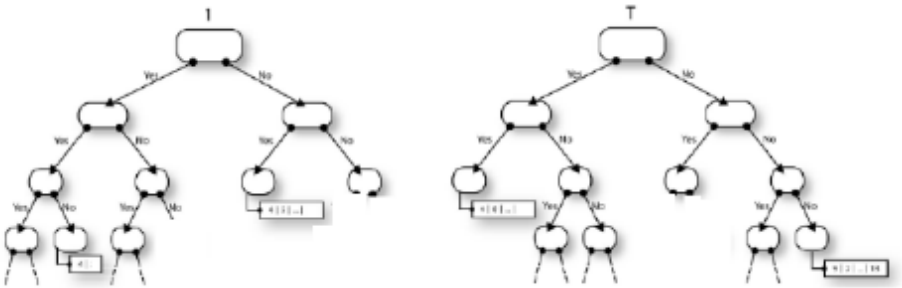
Automatic training of image recognition model based on training examples



VS



A machine-learnt classifier that recognizes tumor/nontumor pixels using local patches



Key ingredients :

Standardized preparation/acquisition protocols

Random Subwindows & Extremely randomized trees with multiple outputs (Dumont et al., 2009)

Rapid tuning of parameters :
number/size of subwindows, color space, number of trees, ...

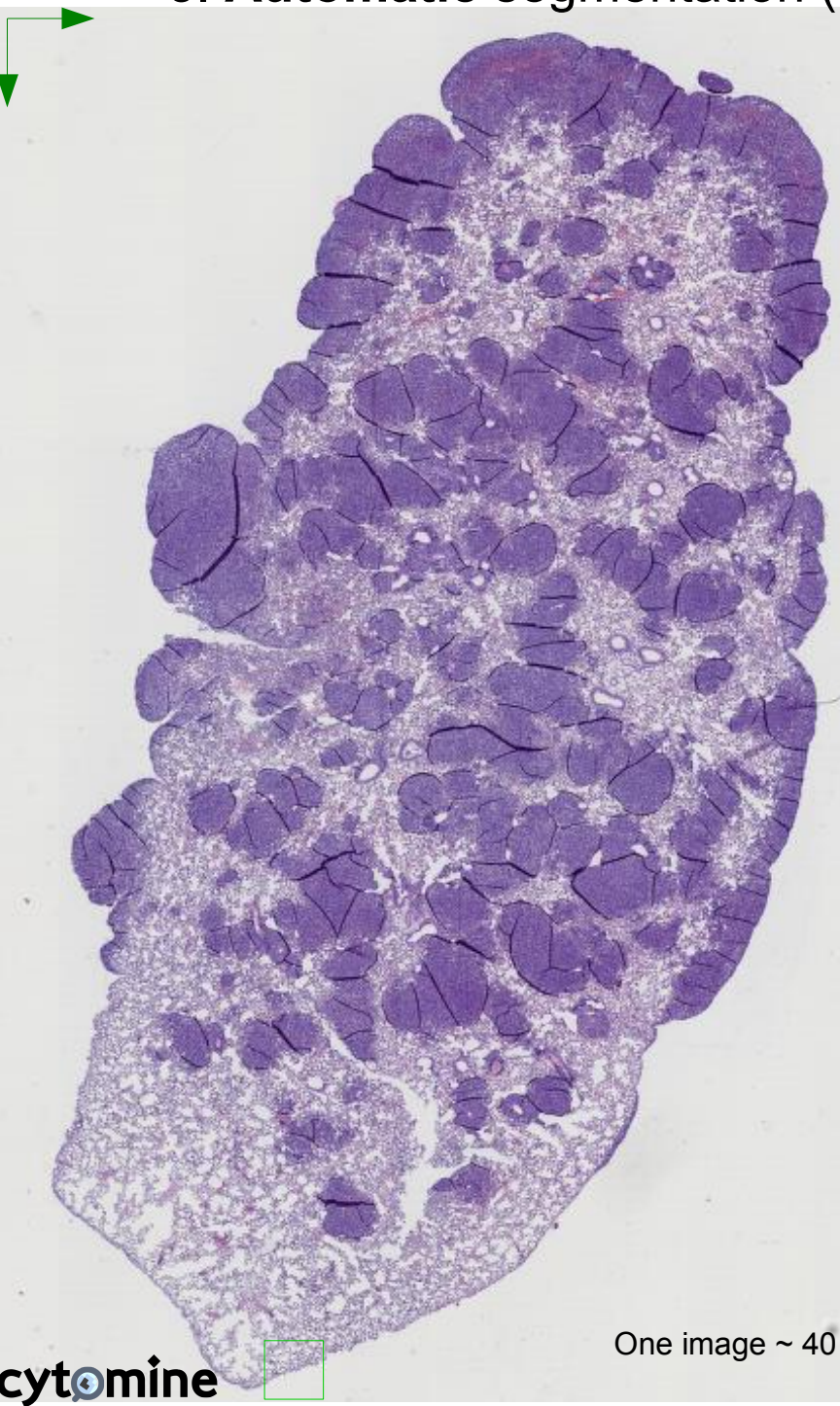
Multi-threaded implementation

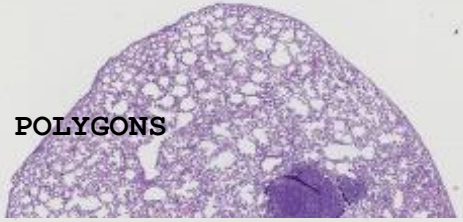
cytomine

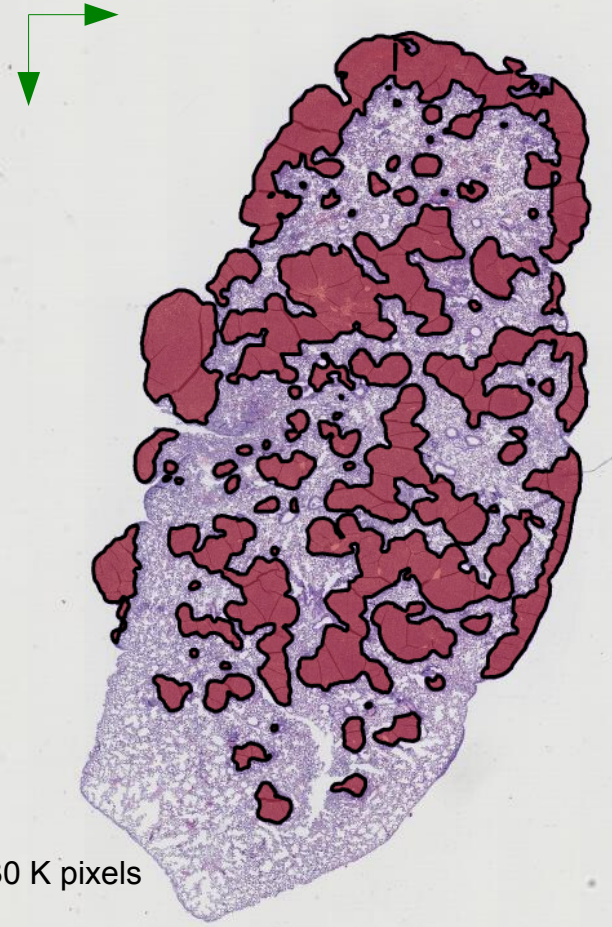
Hybrid human-computer workflow

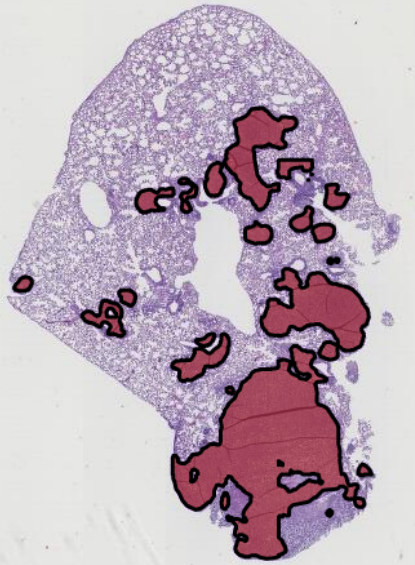
3. **Automatic** segmentation (pixel classification) of tumors in new images

```
Fetch http://is.cytomine.be/.../TileGroup97/2-1-7.jpg

...{processing : Random Subwindows and Extra-Trees}...

POST http://beta.cytomine.be/api/annotation.json
DATA image: 7889452  location: "POLYGON((29672 26176,
34980 24902,29580 …))"

...

Union of POLYGONS
```







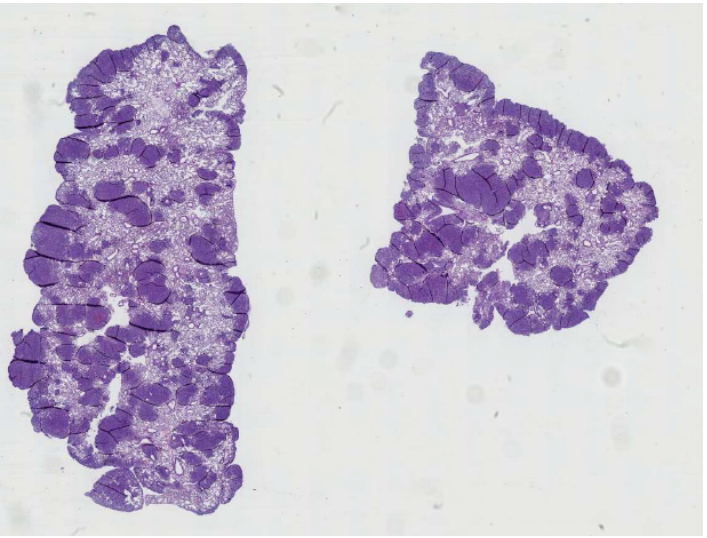


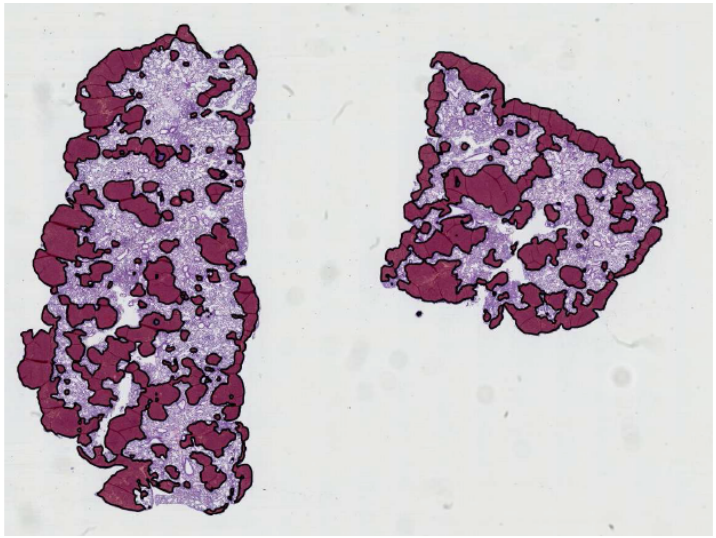
One image ~ 40 K x 30 K pixels

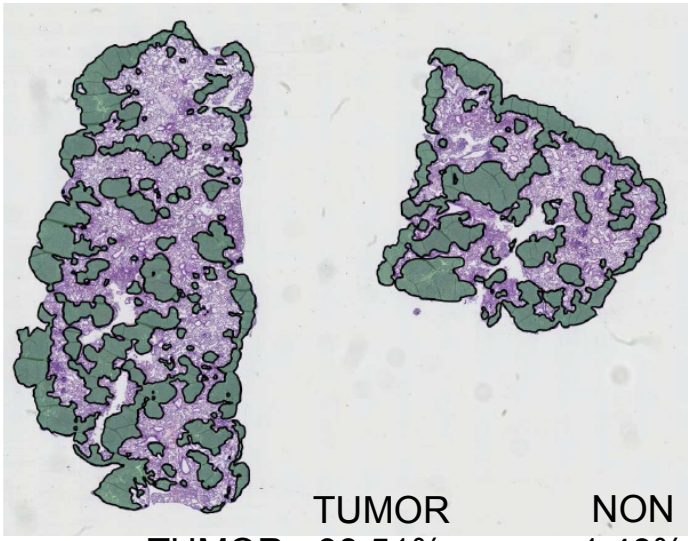
cytomine

Hybrid human-computer workflow

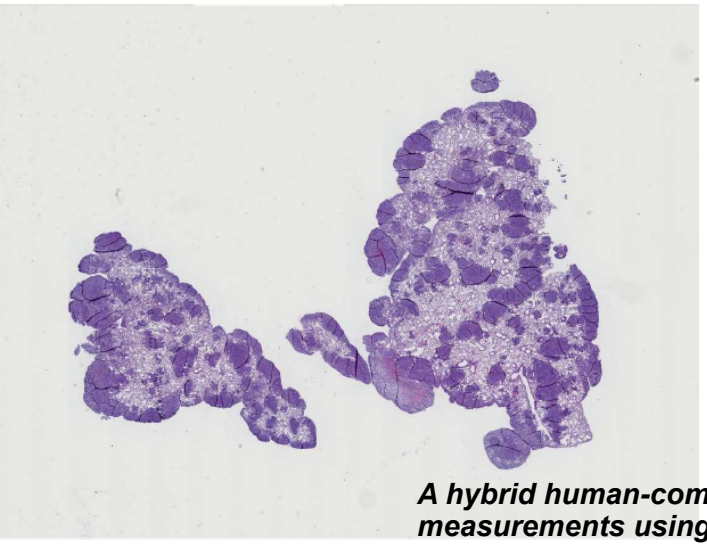
Recognition performances

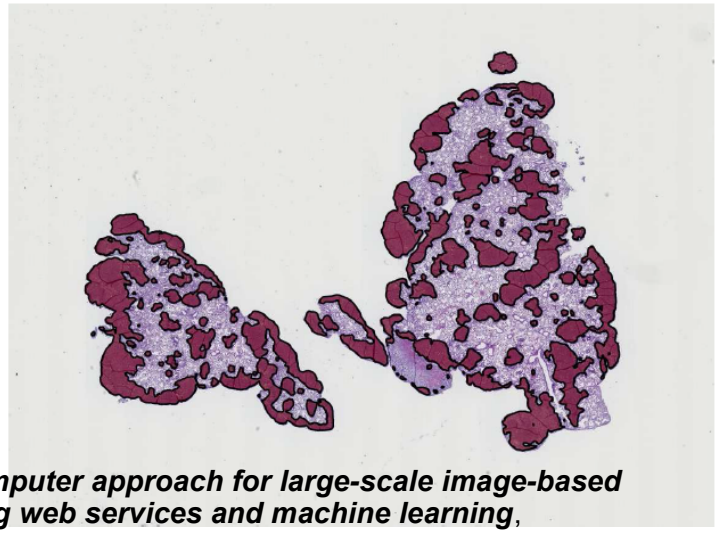


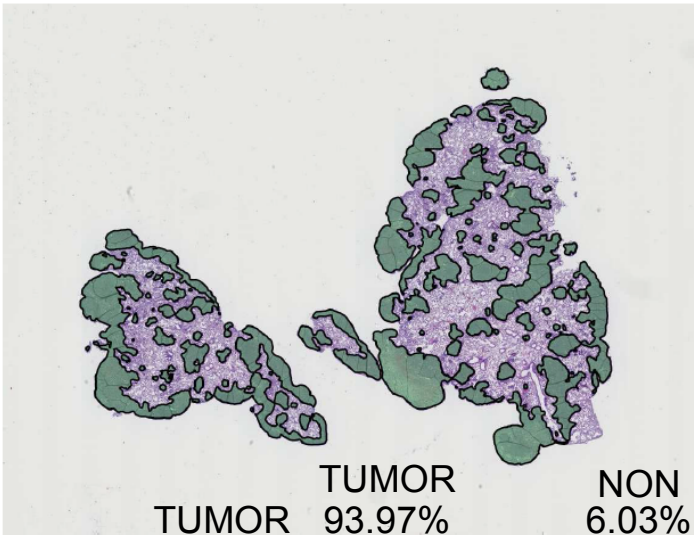




	TUMOR	NON
TUMOR	98.51%	1.49%
NON	2.78%	97.22%





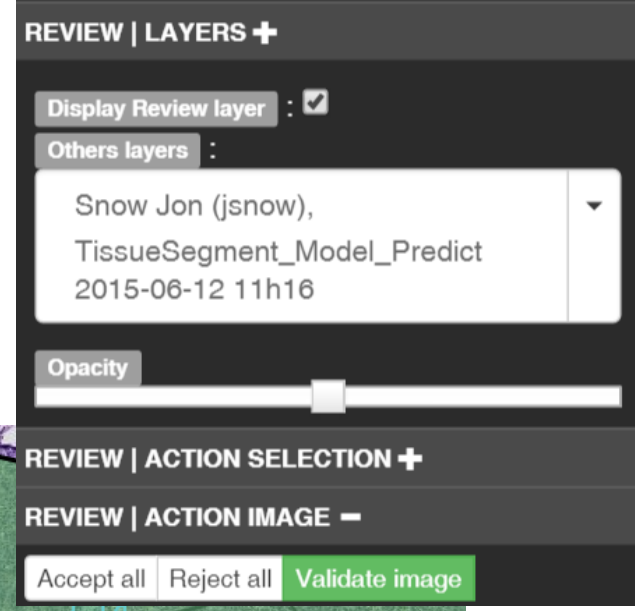


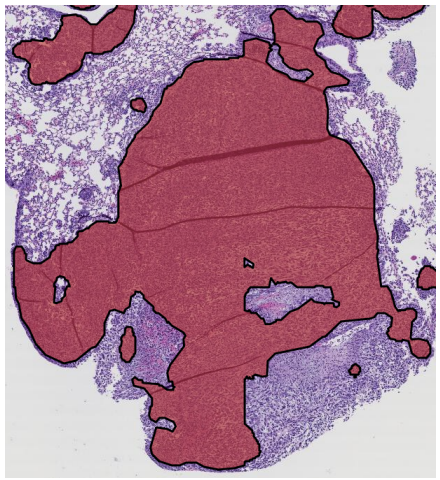
	TUMOR	NON
TUMOR	93.97%	6.03%
NON	0.01%	99.99%

cytomine

A hybrid human-computer approach for large-scale image-based measurements using web services and machine learning,
Marée et al.
Proc. IEEE International Symposium on Biomedical Imaging, 2014

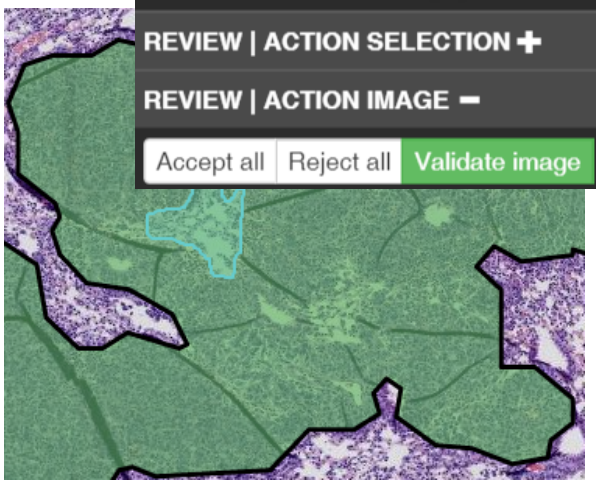
cytomine features : Algorithm proofreading



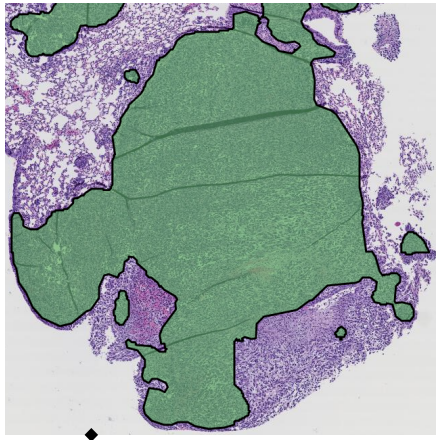


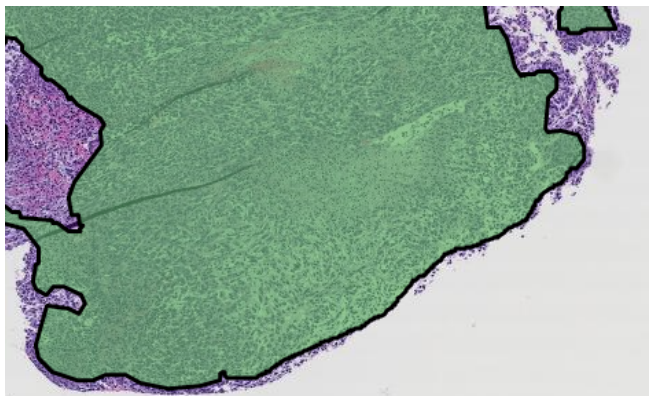
Fill

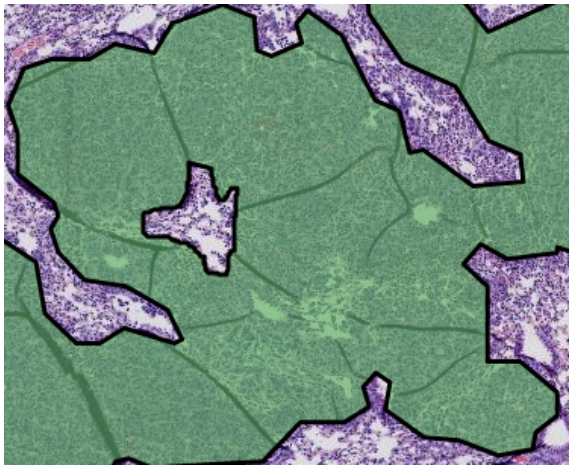
+ Extend



− Subtract

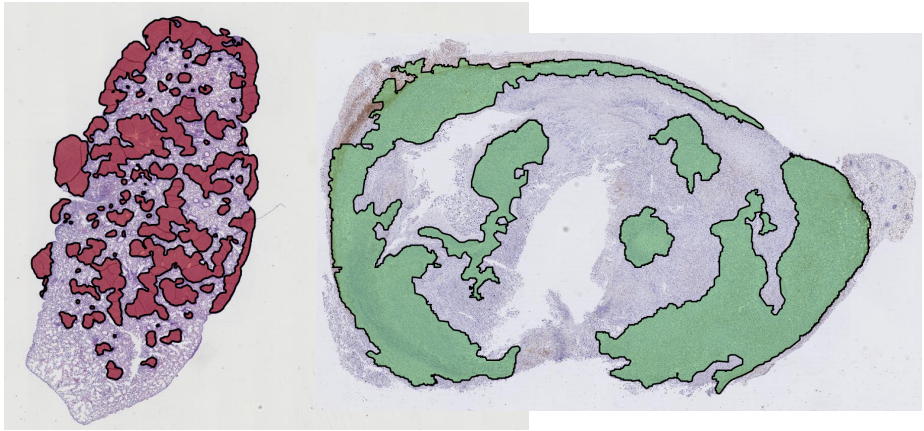






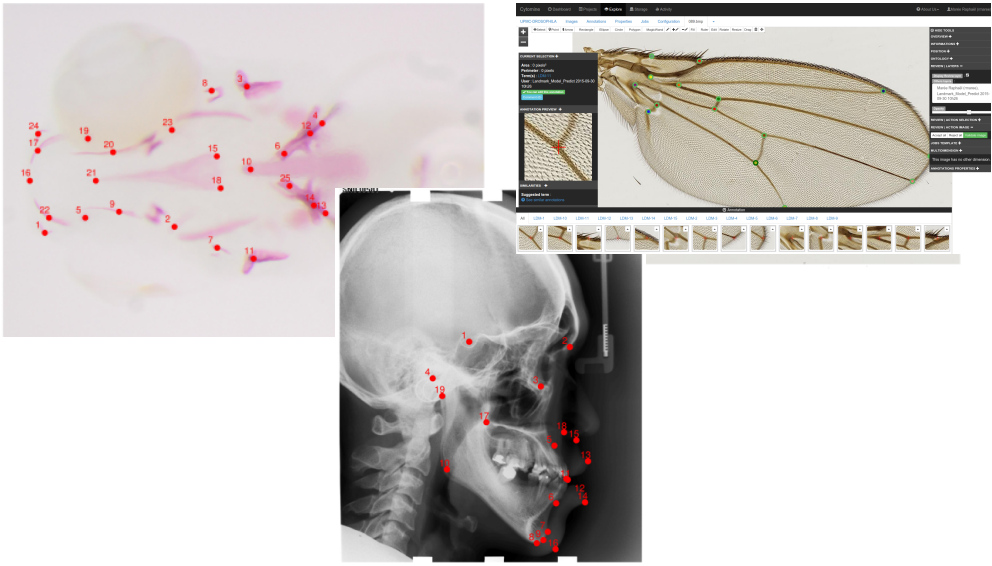
cytomine

cytomine features : Semi-automated analysis using machine learning and proofreading



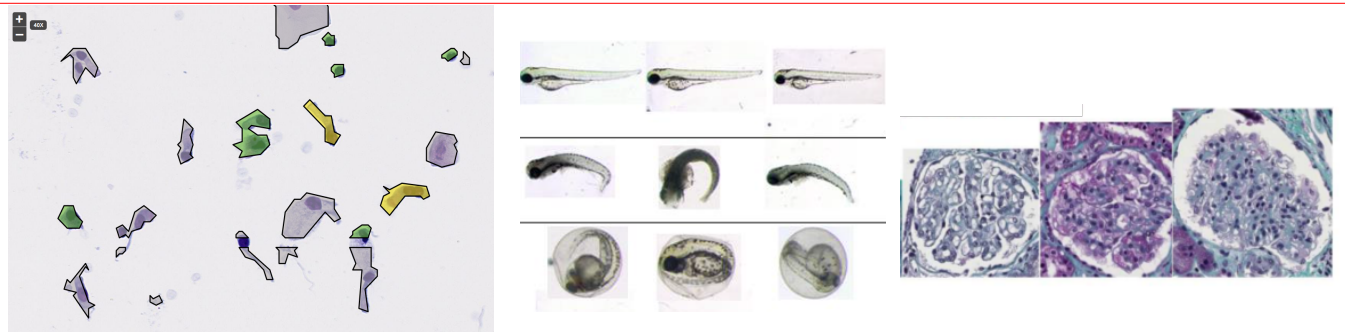
Tissue segmentation

Dumont et al.,VISAPP 2009
Marée et al., ISBI 2014
Leroi et al., Oncotarget 2015



Landmark detection

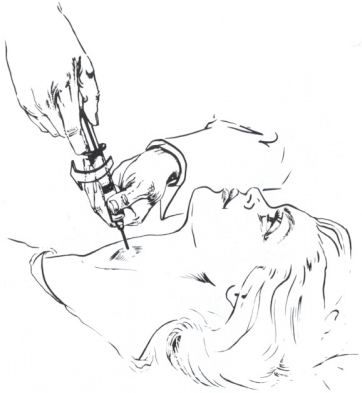
Vandaele et al., Submitted
Huang et al., IEEE TMI 2015

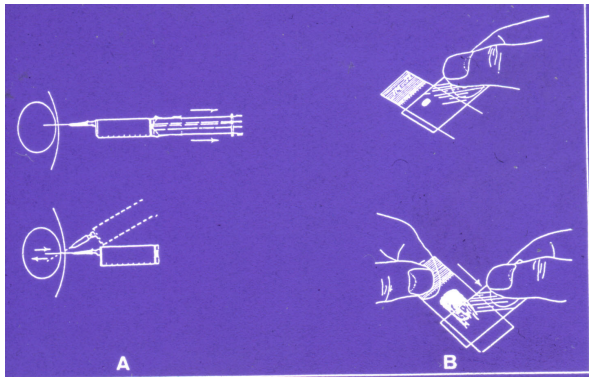


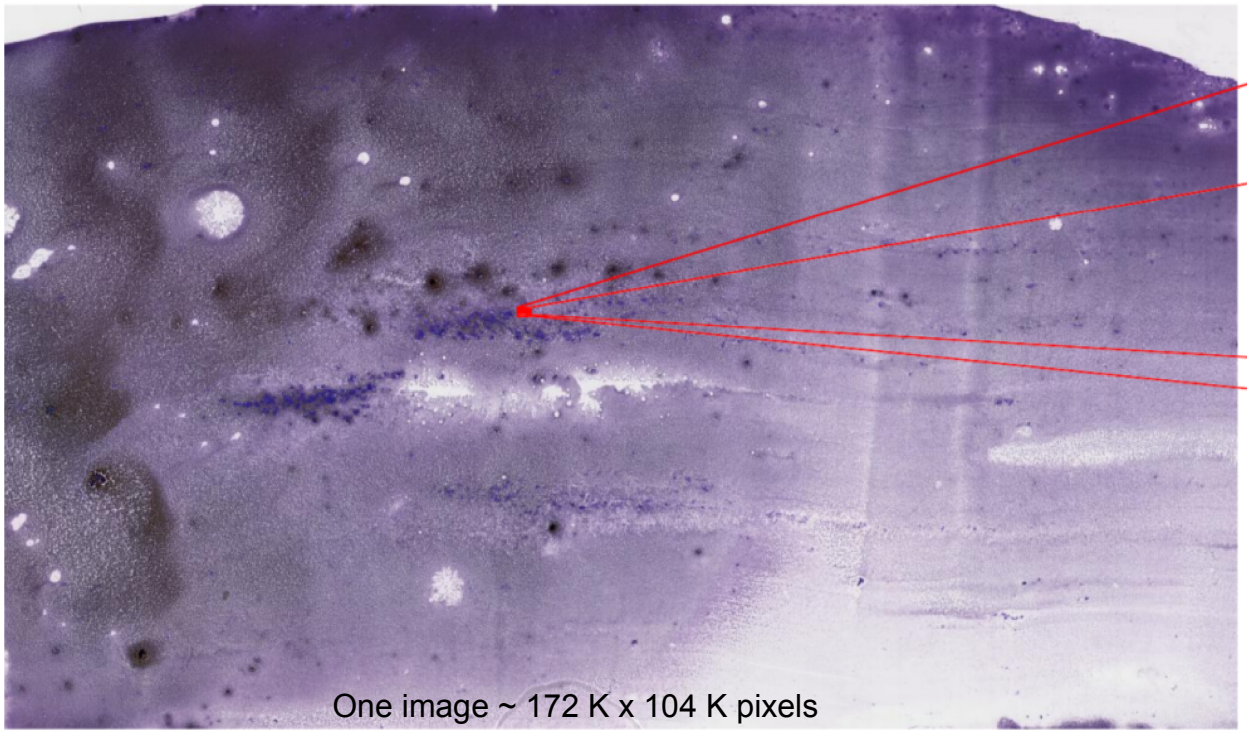
Object classification

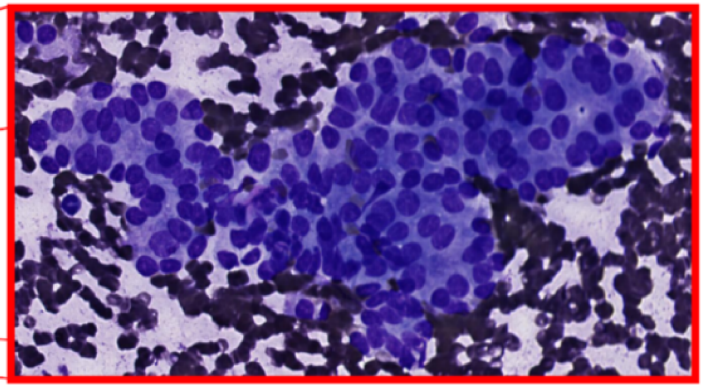
Marée et al. Pattern Recognition Letters 2016 ; ISBI 2016
Delga et al., 2014 ; Jeanray et al., PLoS ONE 2015 ;

Object classification in multi-gigapixel images

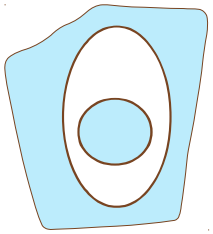


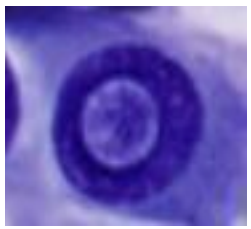


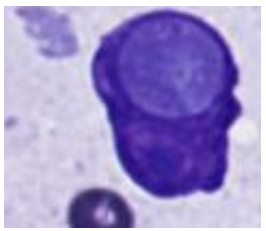


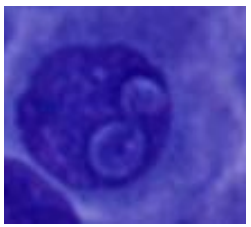


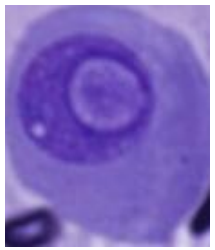
x130





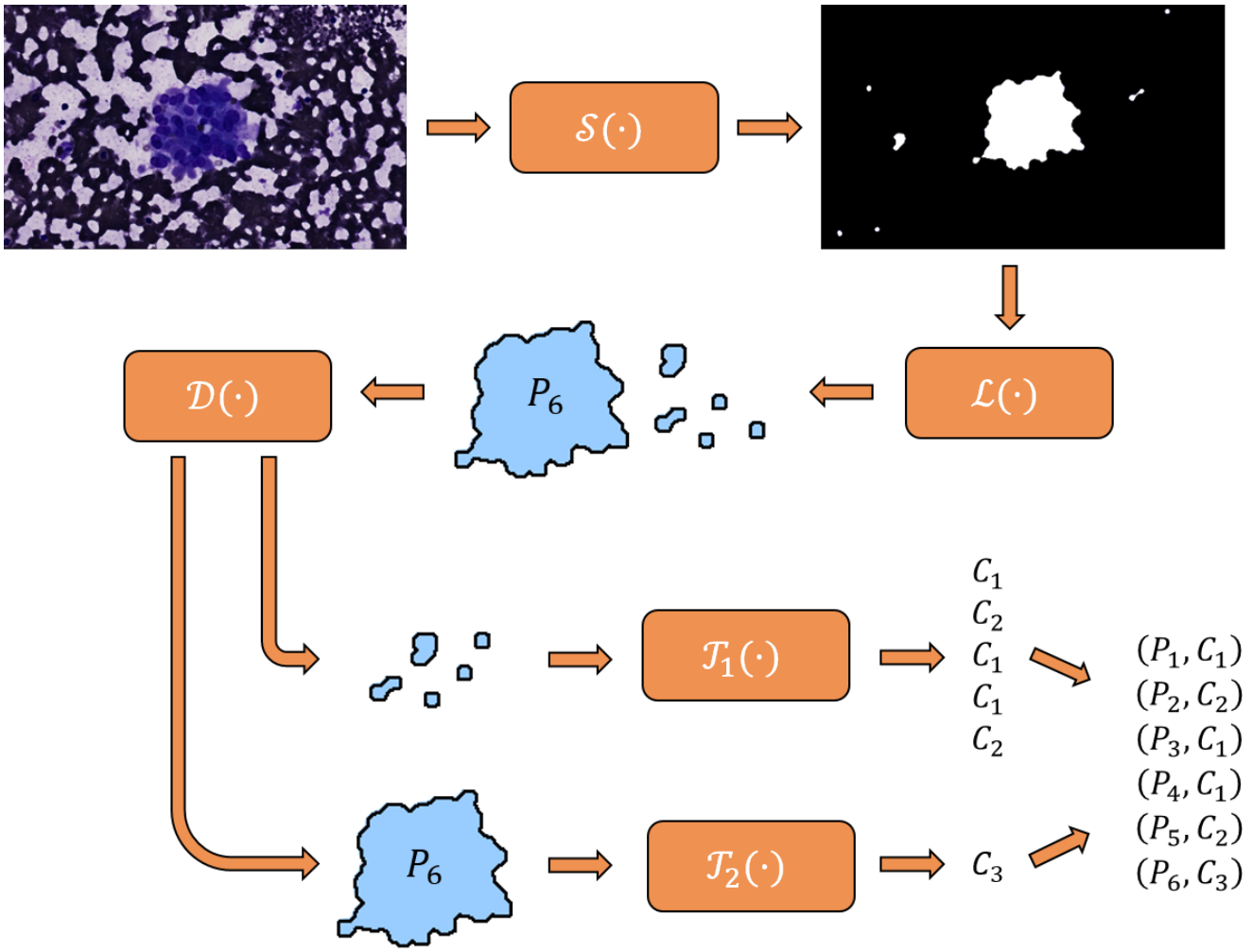






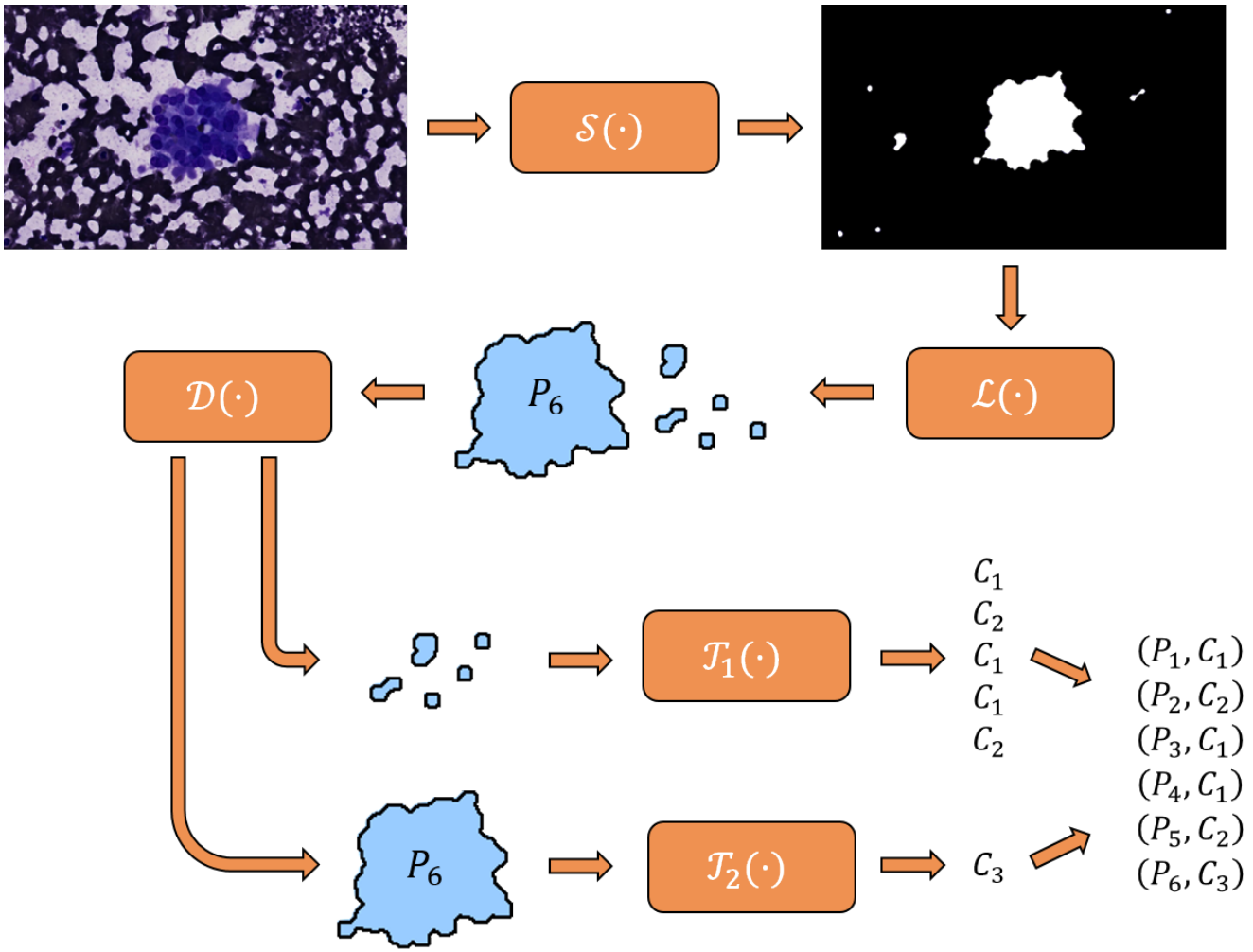
Thyroid FNAB project with ULB Erasme (Prof. I. Salmon)

SLDC : a workflow for object detection & classification in multi-gigapixel images



Segment : Produces binary masks
Locate : Extracts polygons representing the objects of interest from binary masks
Dispatch : Identifies the most appropriate classifier for processing each polygon
Classify : Produces a classification label (and probability estimates) for each polygon

SLDC : a workflow for object detection & classification in multi-gigapixel images

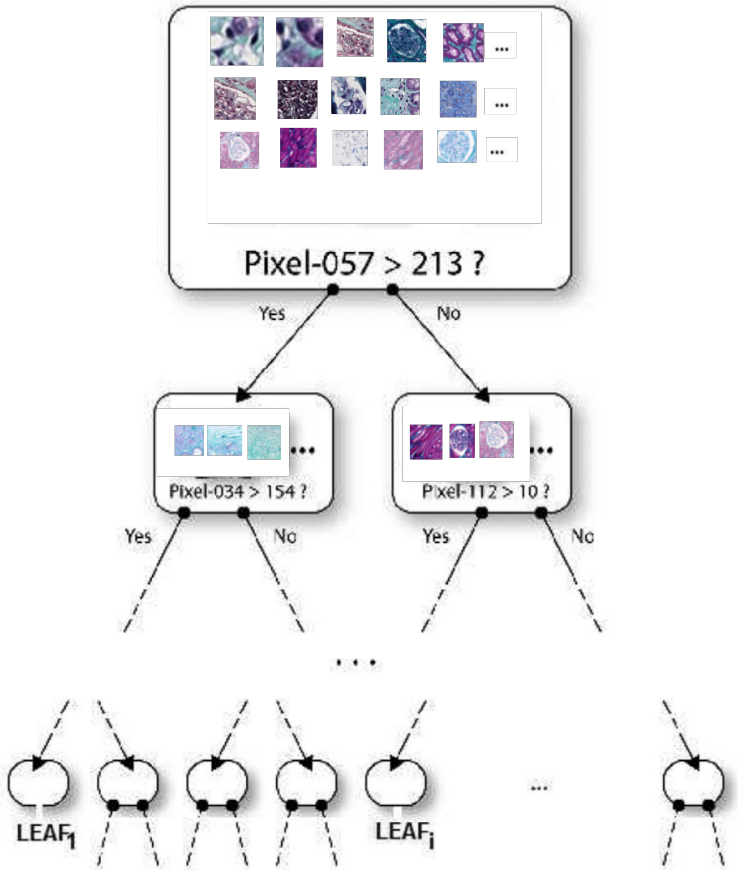


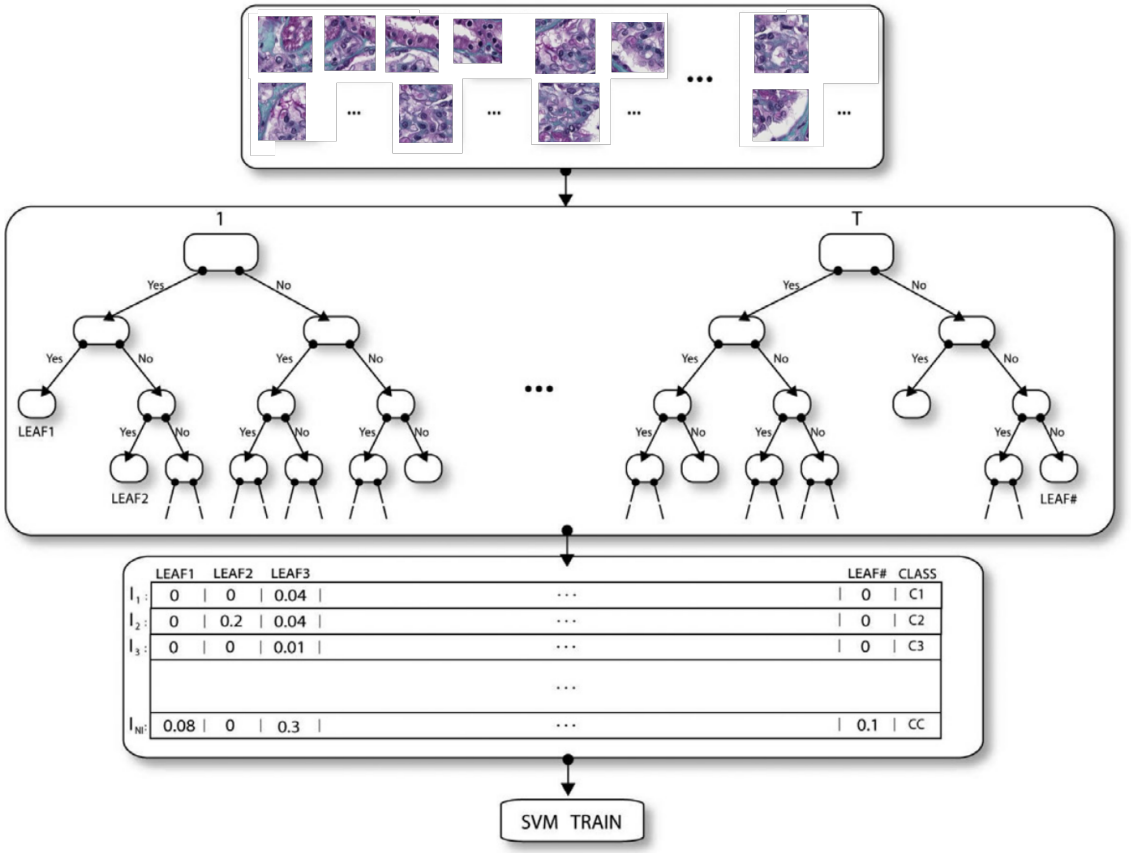
Segment : Color deconvolution (Ruifrok & Johnston, Anal Quant Cytol Histol., 2001)
Locate : Connected components + polygon merging (tiles)
Dispatch : Classifier* to sort Cells / Patterns / Artefacts
Classify : Classifier* Inclusion vs Normal ; Classifier Proliferative pattern vs Normal pattern
*Based on Random subwindows and Extra-Trees (Marée et al., Pattern Recognition Letters ; 2016)

Generic classifier

+

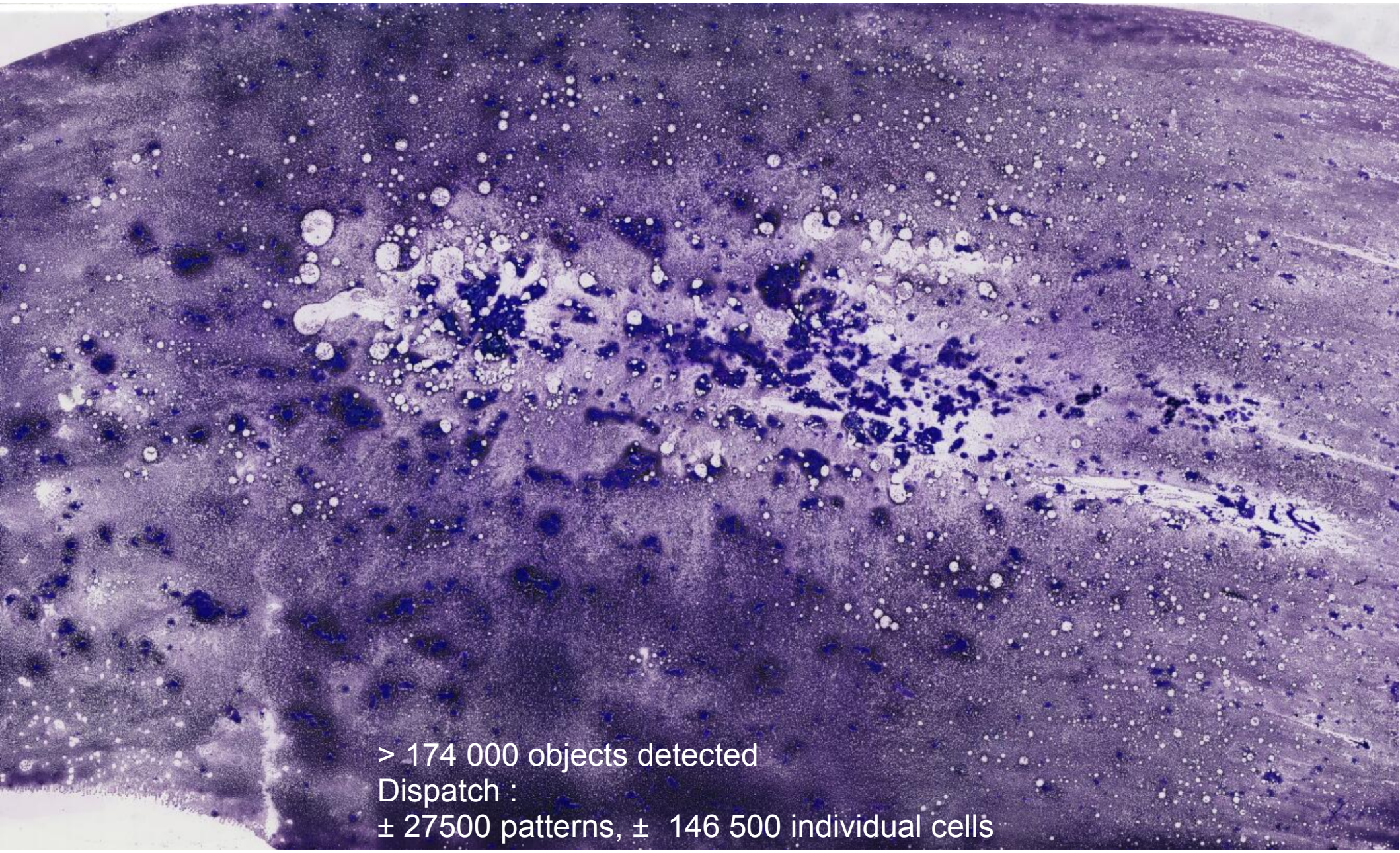
Random subwindows (patches) extracted in images
Described by raw pixel values
Extremely Randomized Trees for feature learning + Final linear SVM classifier



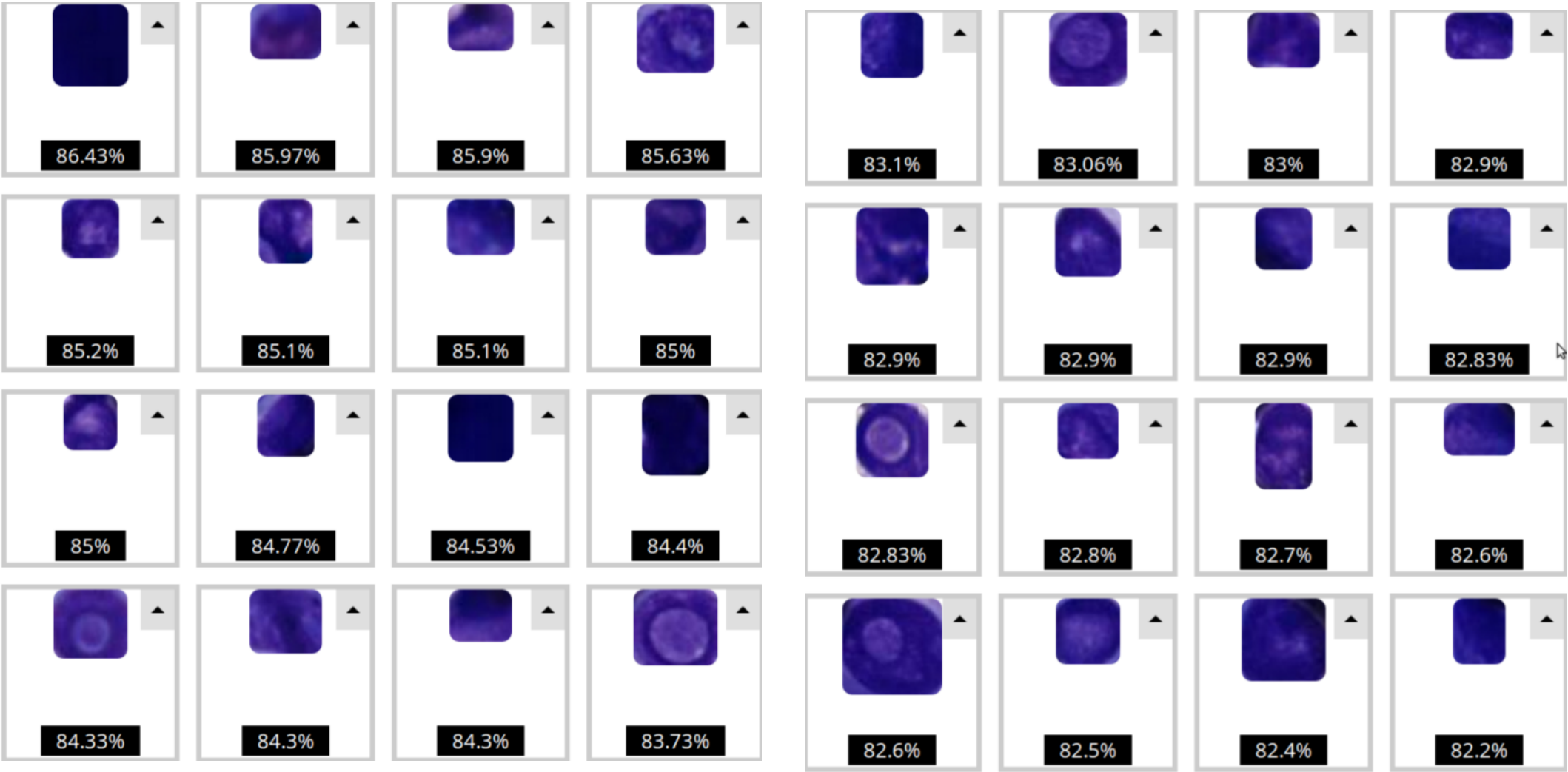


(Marée et al., Pattern Recognition Letters 2016)

Application on a whole-slide image

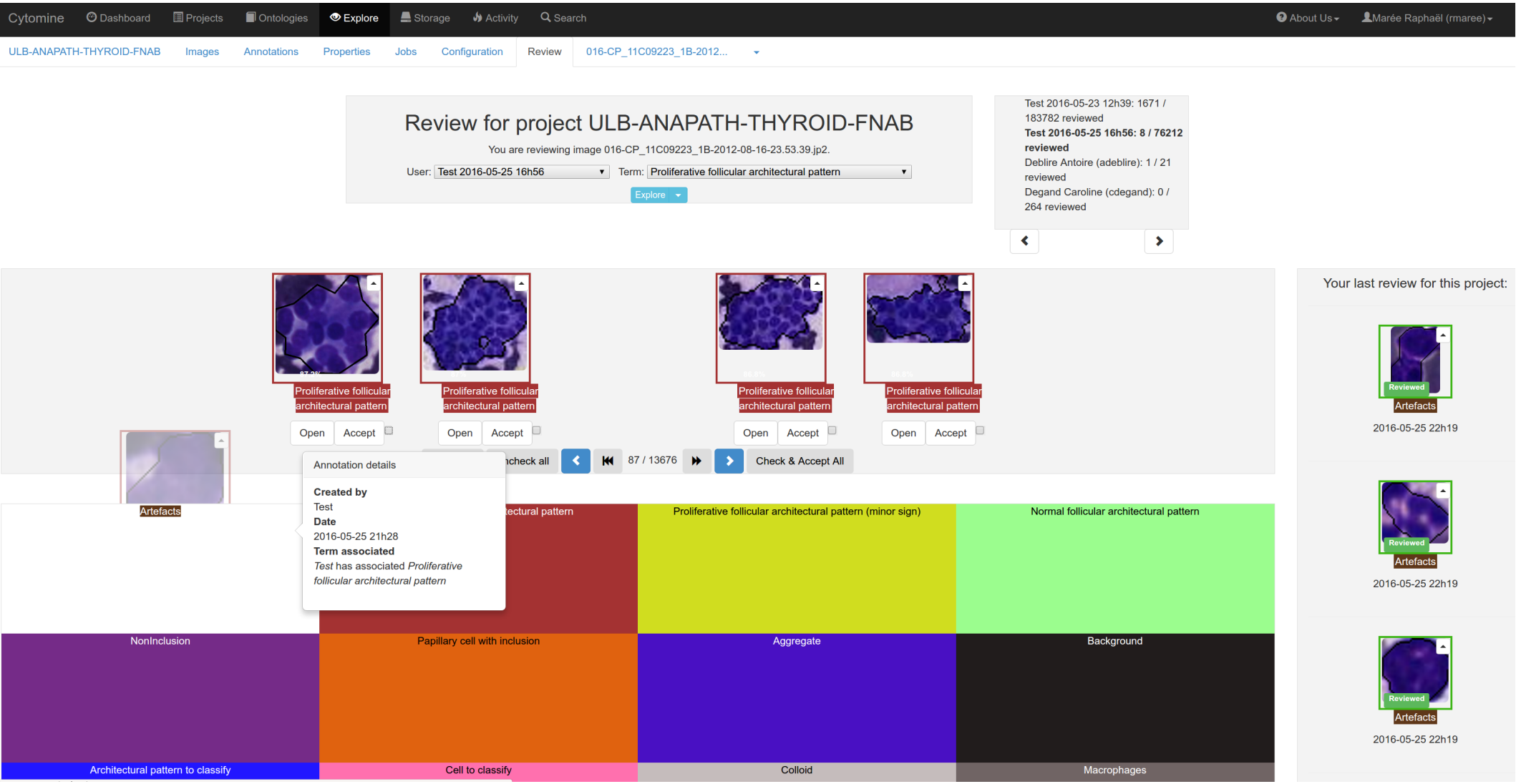


Application on a whole-slide image : Detected intranuclear inclusions

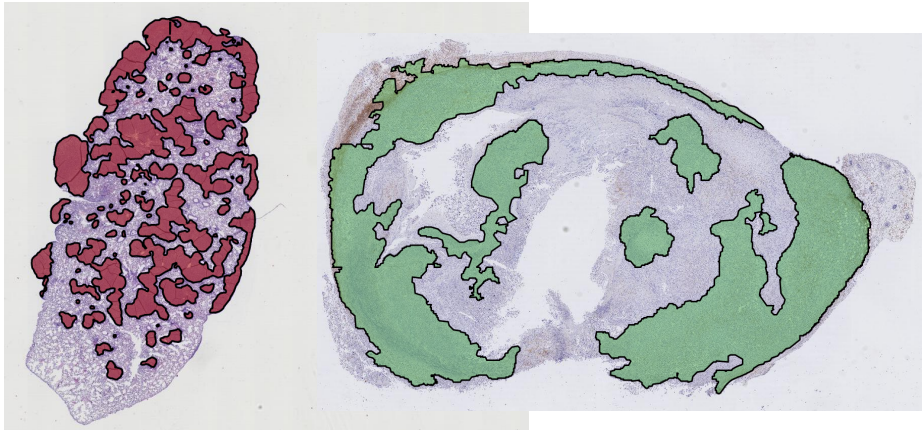


cytomine

cytomine features :
Algorithm proofreading : classification

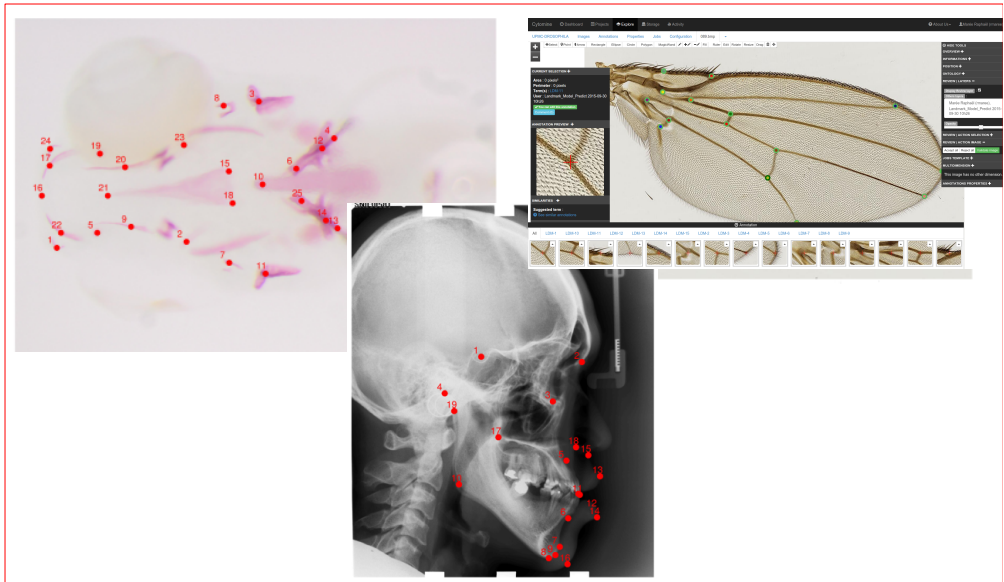


cytomine features : Semi-automated analysis using machine learning and proofreading



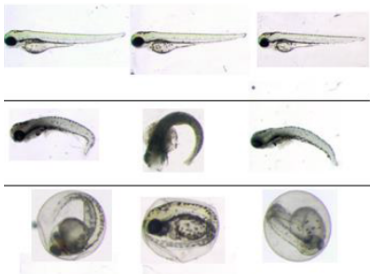
Tissue segmentation

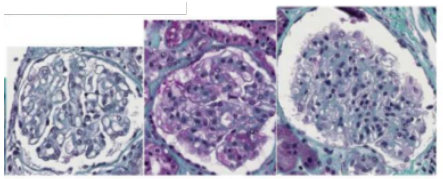
Dumont et al.,VISAPP 2009
Marée et al., ISBI 2014
Leroi et al., Oncotarget 2015



Landmark detection

Vandaele et al., Submitted
Huang et al., IEEE TMI 2015

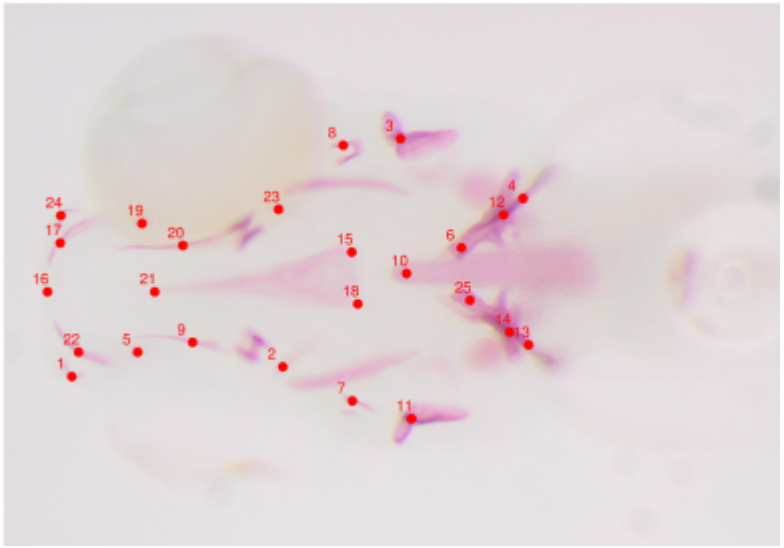




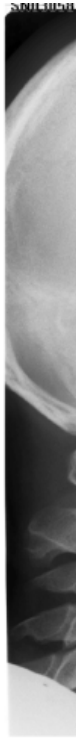
Object classification

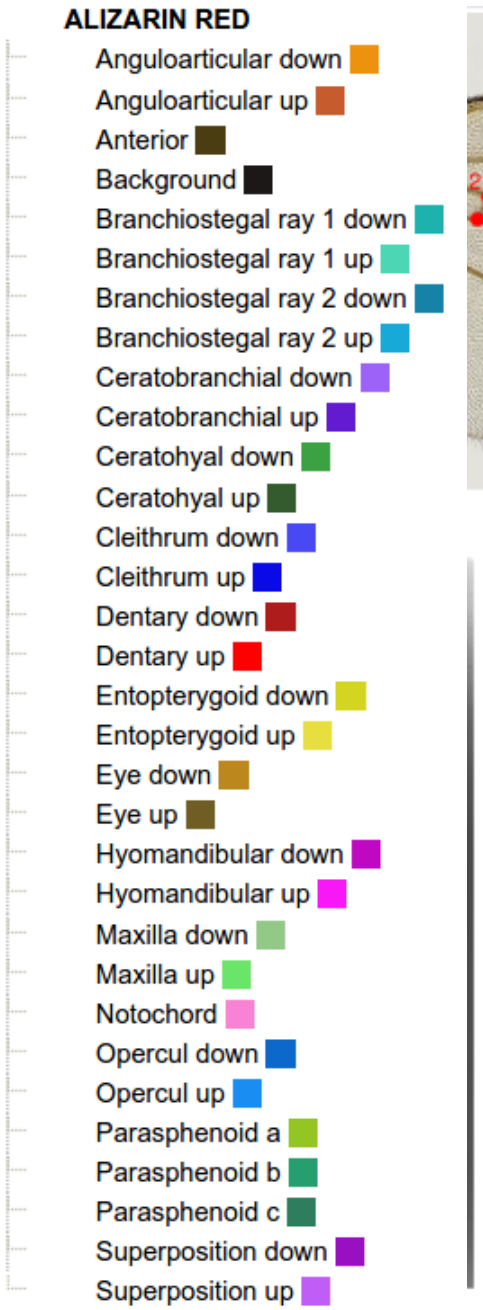
Marée et al. Pattern Recognition Letters 2016 ; ISBI 2016
Delga et al., 2014 ; Jeanray et al., PLoS ONE 2015 ;

Supervised landmark detection

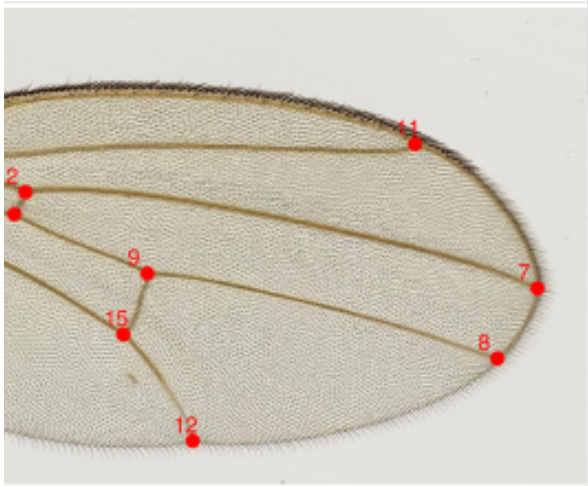


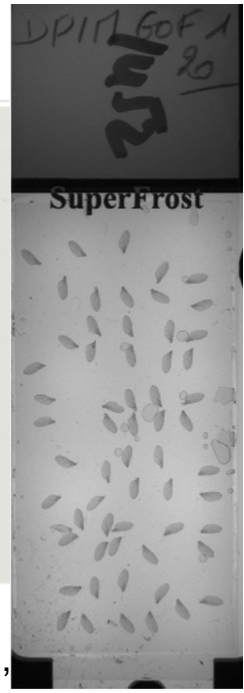
ZEBRA
(M.Muller's lab, GIGA)





(C-T Huang's lab, NTUST)





DROSO
(F.Peronnet's lab, UPMC)

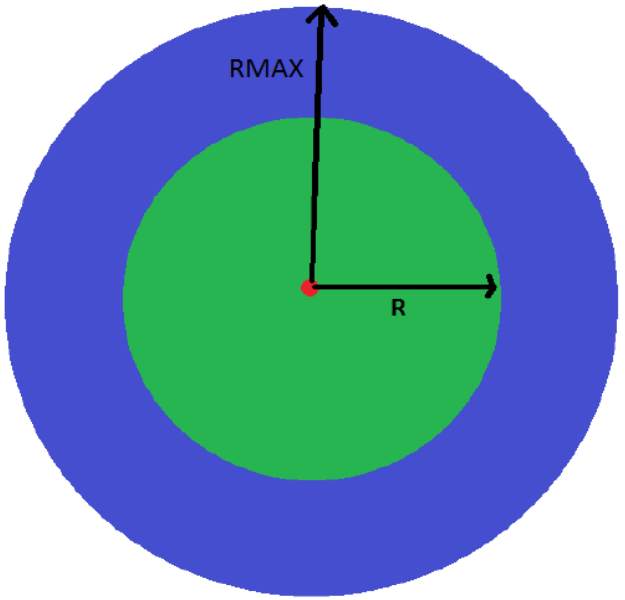
Vandaele et al., Landmark detection in 2D images : a tree-based approach. To be submitted

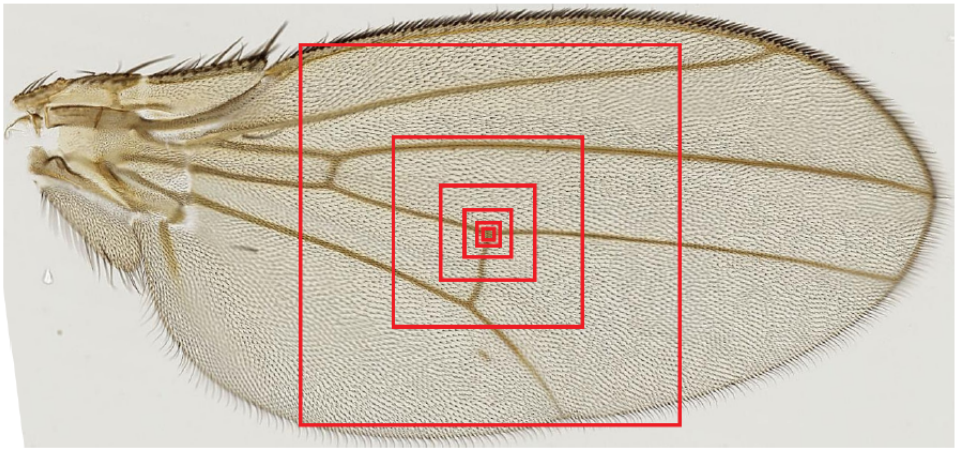
Training

Binary classification problems using Extremely Randomized Trees classifier

Training dataset for each landmark:

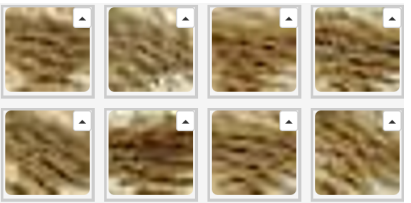
Positive examples : n pixels in the neighboorhood ((x,y) < R) of the landmark from each training images.





Each pixel is described by
6 multiresolution subwindows (16x16)

Scale 0 :



Negative examples : p*n pixels beyond the neighboorhood but within a certain distance to balance the training set (R < (x,y) < RMAX)

Prediction

For each landmark :

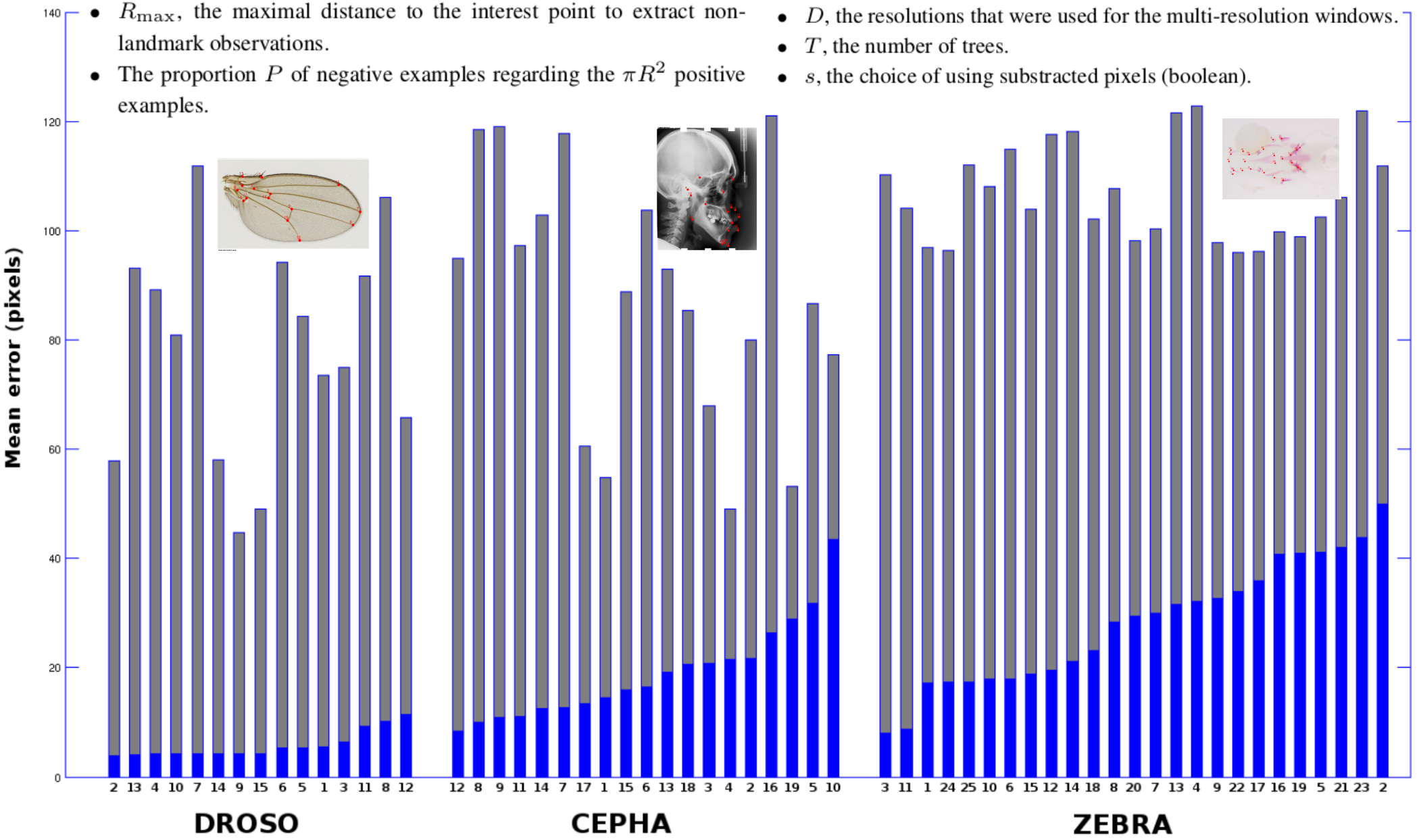
Subsample NP candidate pixels (multivariate normal distribution, centered on observed mean position in the training set), described by multiresolution features.

Classify each candidate pixel into positive/negative using the Extra-Trees classifier

Predicted landmark position is the median position of the pixels predicted as positive with the highest probability

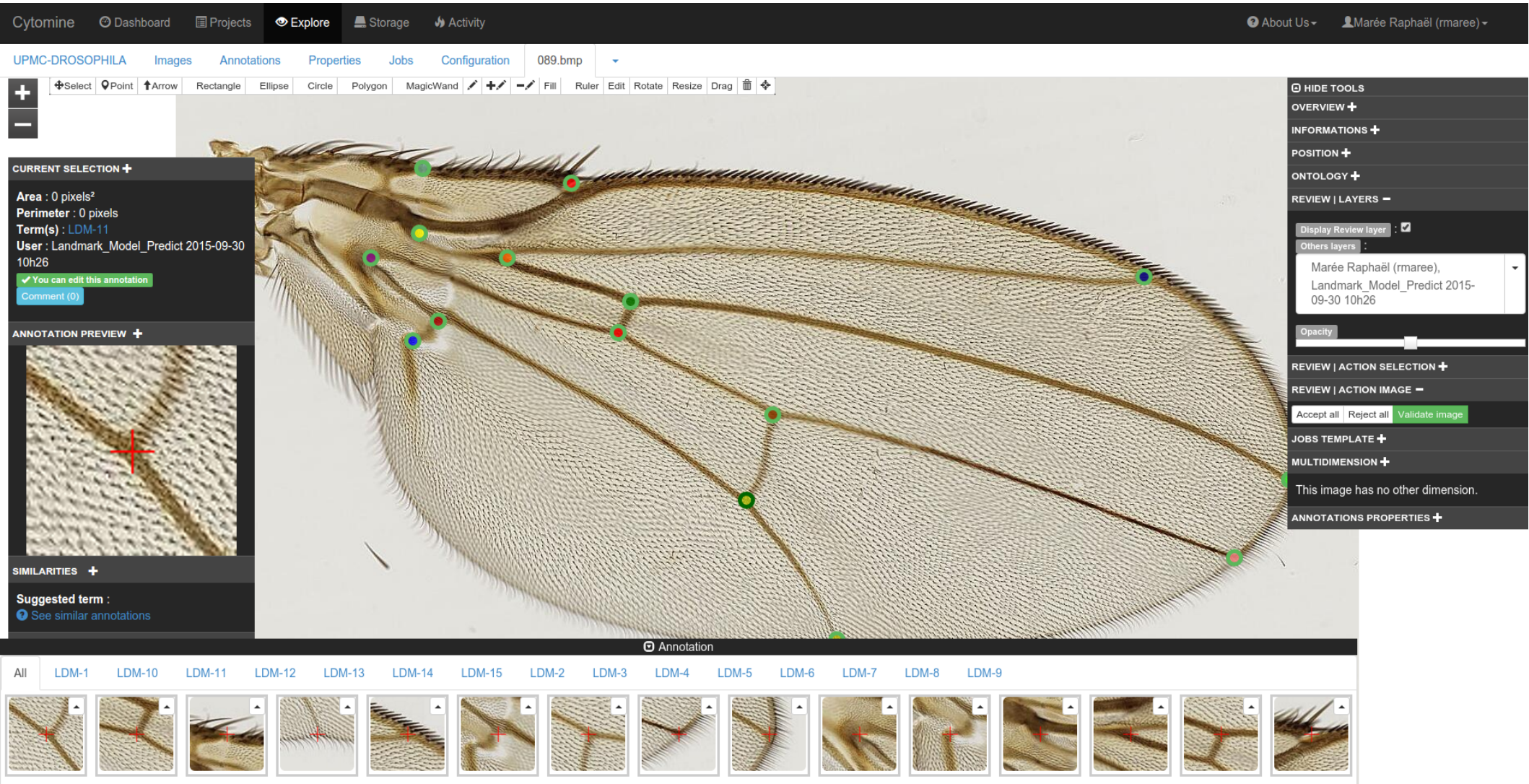
Empirical evaluation

- W, the size of the multi-resolution windows.
- R, the distance to the landmark position determining the training pixel output class.
- $R_{\max}$, the maximal distance to the interest point to extract non-landmark observations.
- The proportion P of negative examples regarding the πR^2 positive examples.
- N_p, the number of pixels randomly extracted during prediction.
- N_r, the number of repetitions with rotation to add to the dataset.
- α the maximal angle of the rotations (in degree).
- D, the resolutions that were used for the multi-resolution windows.
- T, the number of trees.
- s, the choice of using substracted pixels (boolean).

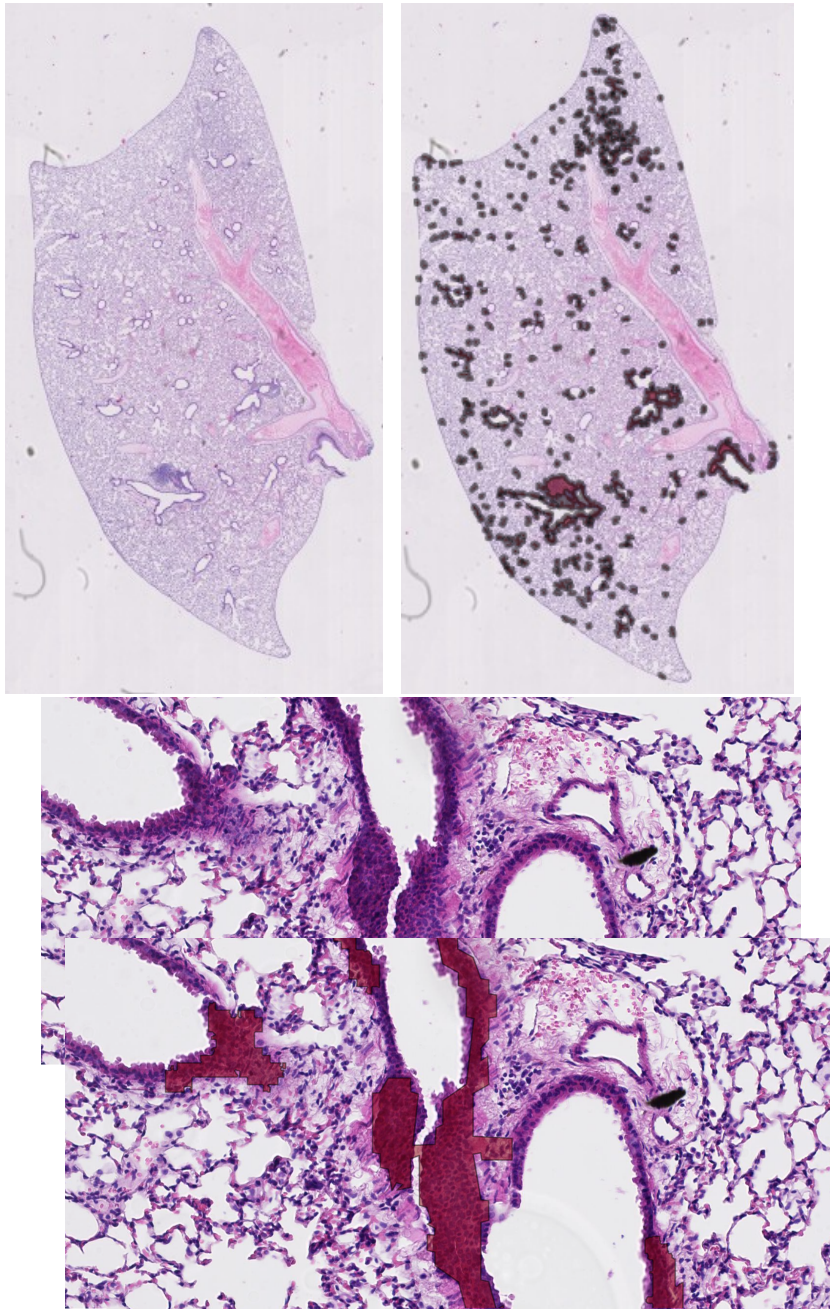


Vandaele et al., Landmark detection in 2D images : a tree-based approach.To be submitted

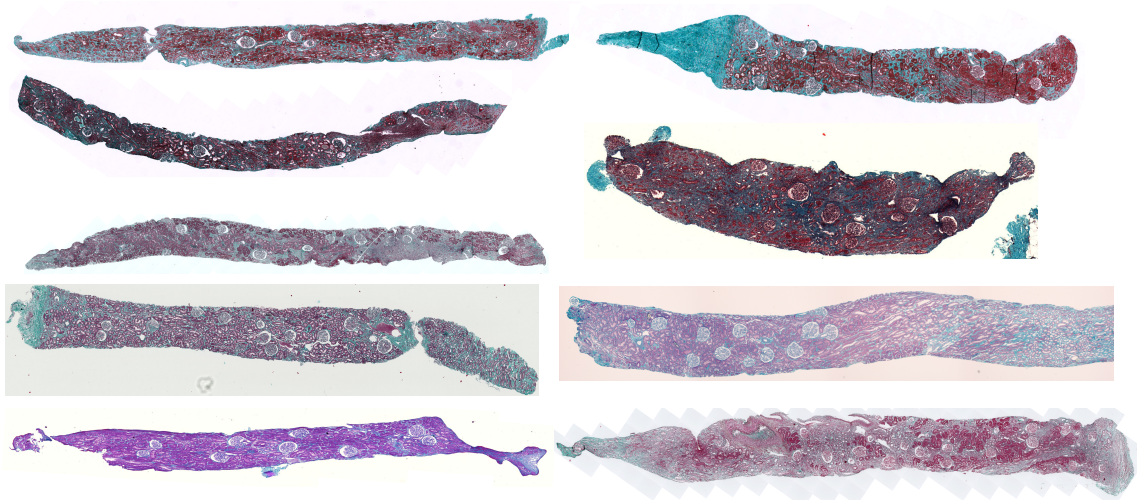
cytomine features :
Algorithm proofreading : interest point detection

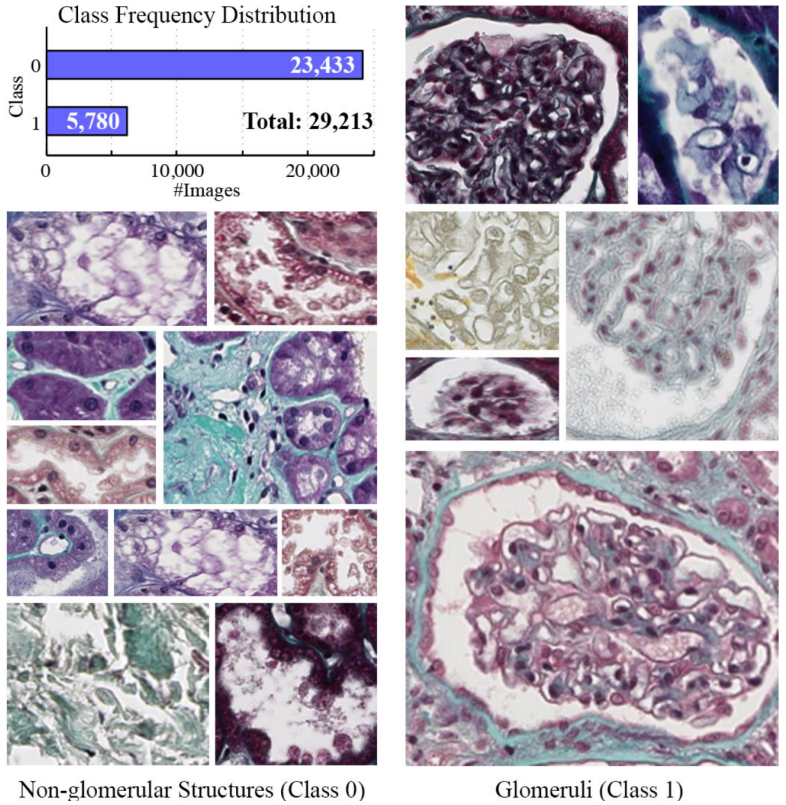


Current challenges: algorithm robustness



Ongoing evaluation on > 1000 slides (one lab)





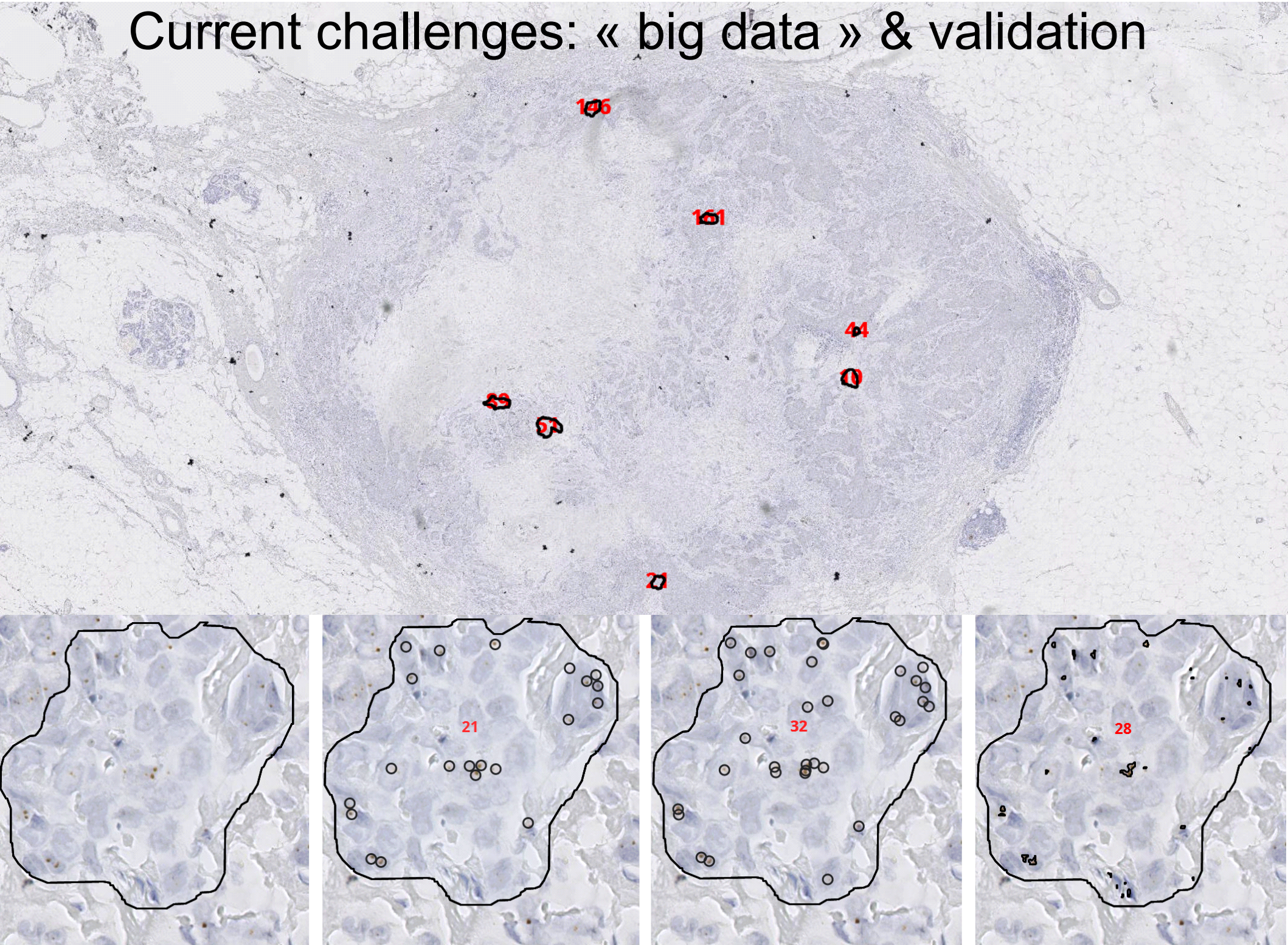
Non-glomerular Structures (Class 0)

Glomeruli (Class 1)

(Marée et al., ISBI 2016)

Institut Pasteur

Current challenges: « big data » & validation



cytomine summary :

- **Free** and **Open-source**
- **Easy installation procedure**

(check www.cytomine.be)

- **General-purpose** using generic software design, web services, user-defined ontologies and machine learning for semantic annotation

- Fosters **collaboration** between life scientists, pathologists, teachers, and computer scientists. Eases big image management for image analysts.

- Not universal but…
  - Many potential direct applications
    (J. Pathol, Oncotarget, BMC Cancer, Am J Transl Res, PLoS ONE,…)
  - **Fully extensible…**
  - **Please contribute: http://github.com/cytomine/**

- Permissive licence + cooperative with social economy goals

Acknowledgments

- Systems and Modeling (Montefiore Institute / GIGA-Research @ ULG):
  - Cytomine core development : Renaud Hoyoux, Loïc Rollus, Benjamin Stévens
  - Algorithms : Gilles Louppe, Jean-Michel Begon, Rémy Vandaele, Romain Mormont Pierre Geurts, Louis Wehenkel
- Other collaborators :
  - GIGA/ULG : Didier Cataldo, Natacha Rocks, Fabienne Perin, Christine Fink, Valérie Defaweux, Pascale Quatresooz
  - Students : Julien Confetti, Pierre Ansen, Olivier Caubo, Antoine Deblire
  - IFRES : Grégoire Vincke, Dominique Verpoorten
  - CRIFA : Brigitte Denis, Céline Snoeck, Christophe Laduron
  - ULB : Isabelle Salmon, Caroline Degand, Xavier Moles Lopez, Nicky D'Haene.
  - Pasteur Institute : Vannary Meas-Yedid, Jean-Christophe Olivo-Marin.
  - Medical University Graz : Philippe Kainz
  - UNamur : Patsy Renard, Eric Depiereux

Research grants of the Wallonia (DGO6) :
  - **CYTOMINE** (2010-2016) n° 1017072
  - **SMASH** (2012-2014) n° 1217606
  - **HISTOWEB** (2014-2017) n° 1017072

info@cytomine.be
www.cytomine.be

cytomine

MONTEFIORE INSTITUTE
Department of Electrical Engineering and Computer Science

GIGA

FEDER

Wallonie

belspo

Join us for the 2016 bioinformatics and computational pathology summit

What is it?
Automated analysis of such large-scale multi-gigapixel datasets is challenging and their combination with omics data is not trivial. This workshop wants to facilitate bridging opportunities between the bioinformatics, digital pathology, and tissue image analysis communities.

What topics are covered?

- Machine learning, image analysis
- Spatial or graph-based methods for high-level feature extraction
- Software and tools to explore large tissue images
- Software and tools to integrate tissue images with omics data
- Novel molecular imaging techniques on tissues
- Next generation and systems pathology
- Mass spectrometry imaging of formalin-fixed paraffin-embedded tissues
- Mass cytometry, and other spectroscopy techniques
- Non-human model research applications

Date and location:
World Forum, Den Hague, Netherlands
Saturday September 3, 2016

Confirmed keynote speakers:

- Jeffrey Fine, USA
- Jeroen Van der Laak, Netherlands

How to contribute?

- Posters, talks, proceedings
- Deadline for submission: May 31

Learn more?
Visit our website at http://cpw.pathomation.com

Sponsored by: PATHOMATION indica labs informed pathology