

CYTOMINE

A rich internet application for remote visualization, collaborative annotation, and automated analysis of whole slide images

Raphaël Marée

**GIGA Bioinformatics Core Facility
Systems and Modeling, Dept. EE&CS
University of Liège, Belgium**



3rd European Conference on Whole Slide Imaging and Analysis
BioQuant, TIGA center (Heidelberg), 30th November 2013



Our Cytomine software relies on...

- Whole-slide scanners to convert glass slides into images

+

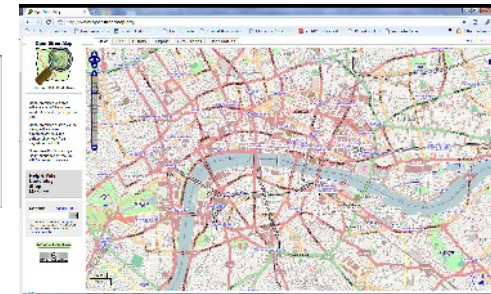


- Modern web development tools & open-source libraries



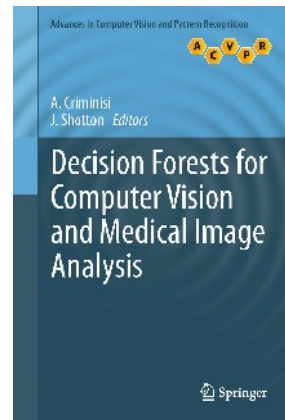
(+/- 500 person-years)

+



- Recent algorithms in machine learning and image analysis

+



- High-performance computing and mass storage equipments



Software 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, 3DHitech MRXS, Leica SCN, Roche BIF...)
- Users with **authentication** (e.g. LDAP), **access rights**, and **roles**
- **Specific ontologies** with user-defined, vocabulary terms

The screenshot displays a web application interface for managing histology projects and annotations. At the top, there is a navigation bar with filters for 'By name' (set to 'ULU'), 'Slides' (0 - 0), 'Images' (0 - 1599), 'Annotations' (0 - 29013), 'By discipline' (All), and 'By ontology' (All). Below the navigation bar, a grid of project cards is shown. Each card displays the project name, discipline, ontology, and various statistics such as 'Number of Images', 'Number of user annotations', 'Number of job annotations', and 'Number of validated annotations'. A 'New project' button is visible on the left. On the right side, a detailed ontology tree for 'ULG-LBTD-TISSUS' is shown, listing various tissue types and structures with corresponding color-coded boxes. Two arrows point from the 'ULG-LBTD-AGAR23' project card to a set of histology images and the ontology tree.

By name: ULU | Slides: 0 - 0 | Images: 0 - 1599 | Annotations: 0 - 29013 | By discipline: All | By ontology: All

ULG-LBTD-TRANSFERT-...
Name: UUG-LBTD-TRANSFERT-ADOPTIF
Discipline: HISTOLOGY
Ontology: ULG-LBTD-TISSUS
Number of Images: 18
Number of user annotations: 2
Number of job annotations: 1760
Number of validated annotations: 381

ULG-LBTD-AGAR15-FOIE
Name: UUG-LBTD-AGAR15-FOIE
Discipline: HISTOLOGY
Ontology: ULG-LBTD-AGAR15-FOIE
Number of Images: 12
Number of user annotations: 0
Number of job annotations: 0
Number of validated annotations: 0

ULG-LBTD-AGAR15-POUM...
Name: UUG-LBTD-AGAR15-POUMON
Discipline: HISTOLOGY
Ontology: ULG-LBTD-TISSUS
Number of Images: 74
Number of user annotations: 571
Number of job annotations: 5270
Number of validated annotations: 7647

ULG-LBTD-AGAR23
Name: UUG-LBTD-AGAR23
Discipline: HISTOLOGY
Ontology: ULG-LBTD-TISSUS
Number of Images: 76
Number of user annotations: 64
Number of job annotations: 12338
Number of validated annotations: 2597

ULG-LBTD-AGAR25
Name: UUG-LBTD-AGAR25
Discipline: HISTOLOGY
Ontology: ULG-LBTD-TISSUS
Number of Images: 40
Number of user annotations: 105
Number of job annotations: 2985
Number of validated annotations: 1097

ULG-LBTD-AGAR28
Name: UUG-LBTD-AGAR28
Discipline: HISTOLOGY
Ontology: ULG-LBTD-TISSUS
Number of Images: 70
Number of user annotations: 30
Number of job annotations: 2493
Number of validated annotations: 767

ULG-LBTD-AGG
Name: UUG-LBTD-AGG
Discipline: HISTOLOGY
Ontology: ULG-LBTD-TISSUS
Number of Images: 60
Number of user annotations: 478
Number of job annotations: 20910
Number of validated annotations: 3512

ULG-LBTD-BOYDEN
Name: UUG-LBTD-BOYDEN
Discipline: HISTOLOGY
Ontology: ULG-LBTD-BOYDEN
Number of Images: 1
Number of user annotations: 0
Number of job annotations: 0
Number of validated annotations: 0

ULG-LBTD-BREBIS-IC-F...
Name: UUG-LBTD-BREBIS-IC-F...
Discipline: HISTOLOGY
Ontology: UUG-LBTD-BREBIS-HE-F...
Number of Images: 0
Number of user annotations: 0
Number of job annotations: 0
Number of validated annotations: 0

ULG-LBTD-BREBIS-SONNET...
Name: UUG-LBTD-BREBIS-SONNET...
Discipline: CYTOLOGY
Ontology: UUG-LBTD-SONNET...
Number of Images: 20
Number of user annotations: 0
Number of job annotations: 0
Number of validated annotations: 0

ULG-LBTD-CD31
Name: UUG-LBTD-CD31
Discipline: IMMUNOHISTOCHEMISTRY
Ontology: UUG-LBTD-CD31
Number of Images: 0
Number of user annotations: 0
Number of job annotations: 0
Number of validated annotations: 0

ULG-LBTD-CD31
Name: UUG-LBTD-CD31
Discipline: CYTOLOGY
Ontology: UUG-LBTD-CD31
Number of Images: 19
Number of user annotations: 172
Number of job annotations: 0
Number of validated annotations: 0

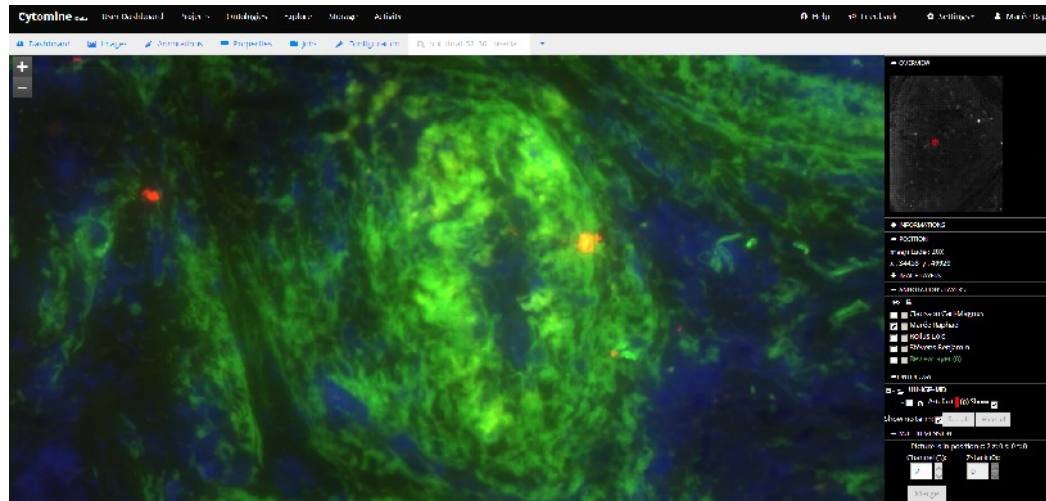
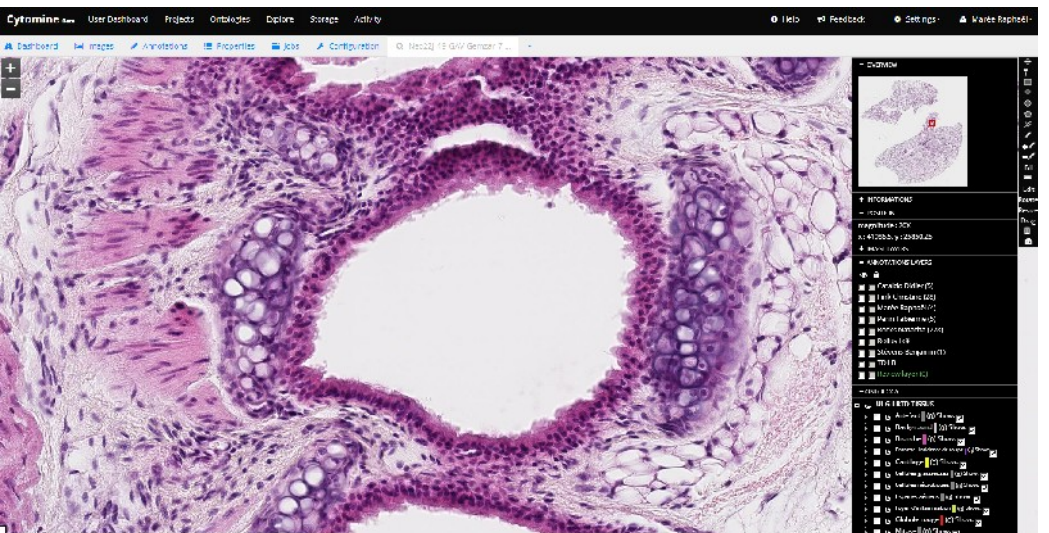
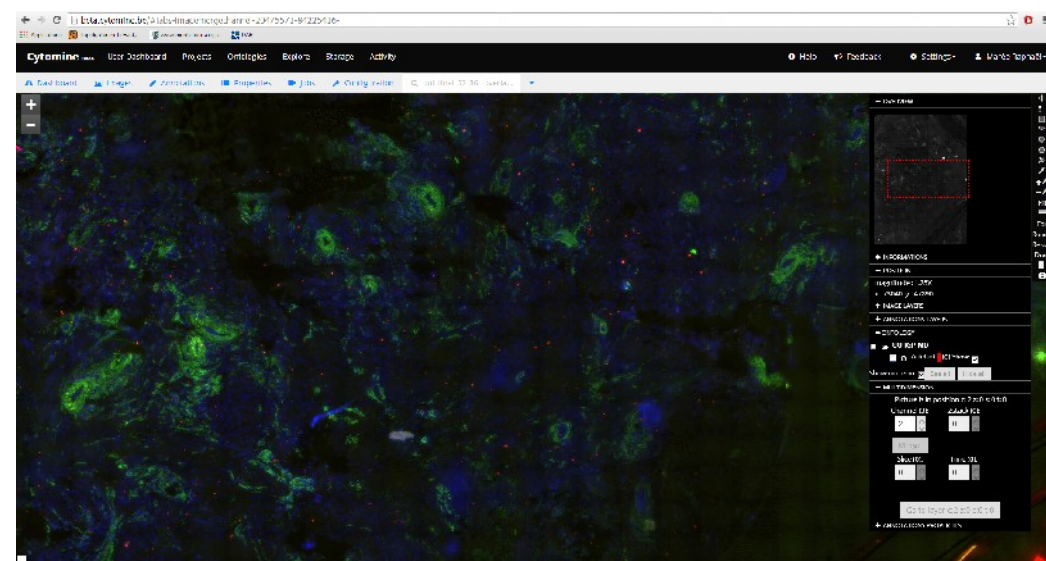
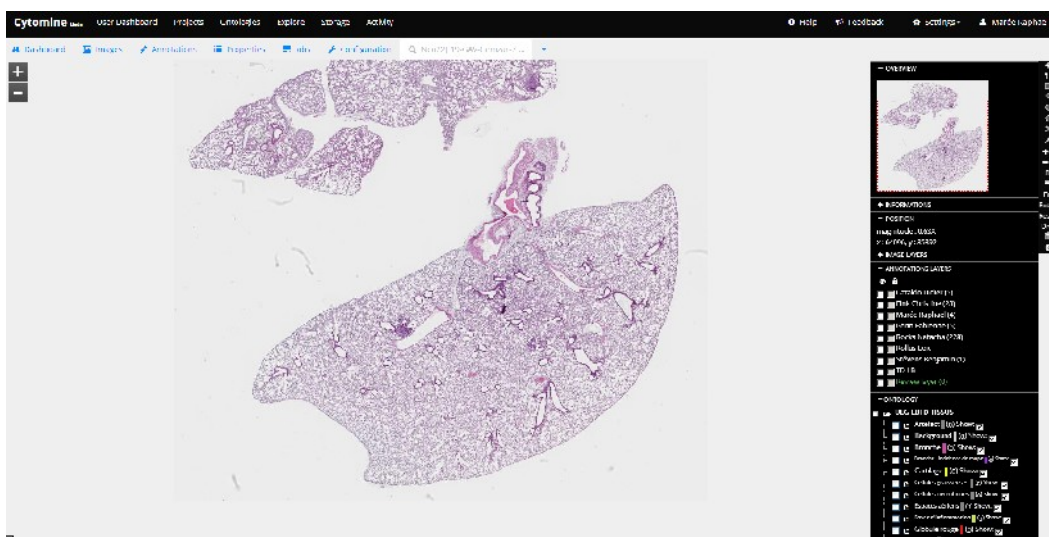
ULG-LBTD-CD31-HISTO
Name: UUG-LBTD-CD31-HISTO
Discipline: HISTOLOGY
Ontology: UUG-LBTD-TISSUS
Number of Images: 11
Number of user annotations: 1
Number of job annotations: 0
Number of validated annotations: 0

ULG-LBTD-CD31-IPHA-...
Name: UUG-LBTD-CD31-IPHA-...
Discipline: IMMUNOHISTOCHEMISTRY
Ontology: UUG-LBTD-IPHA-...
Number of Images: 30
Number of user annotations: 0
Number of job annotations: 0
Number of validated annotations: 0

ULG-LBTD-TISSUS
Artefact
Bronche
Bronche - incidence de coupe
Cartilage
Cellules graisseuses
Cellules nécrotiques
Espaces aériens
Foyer d'inflammation
Globe rouge
Mitose
Muscle
Poumon
Poumon non insufflé
Unknown
Zone floue
Marquage
Alpha-smooth muscle actin
Collagen
Tumeurs
Adénocarcinomes
Cellules tumorales
Foyer hyperplasie nodulaire
Tumeurs épidermoïdes
Vaisseaux
Vaisseau lymphatique
Vaisseau sanguin

Software features : Visualize

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

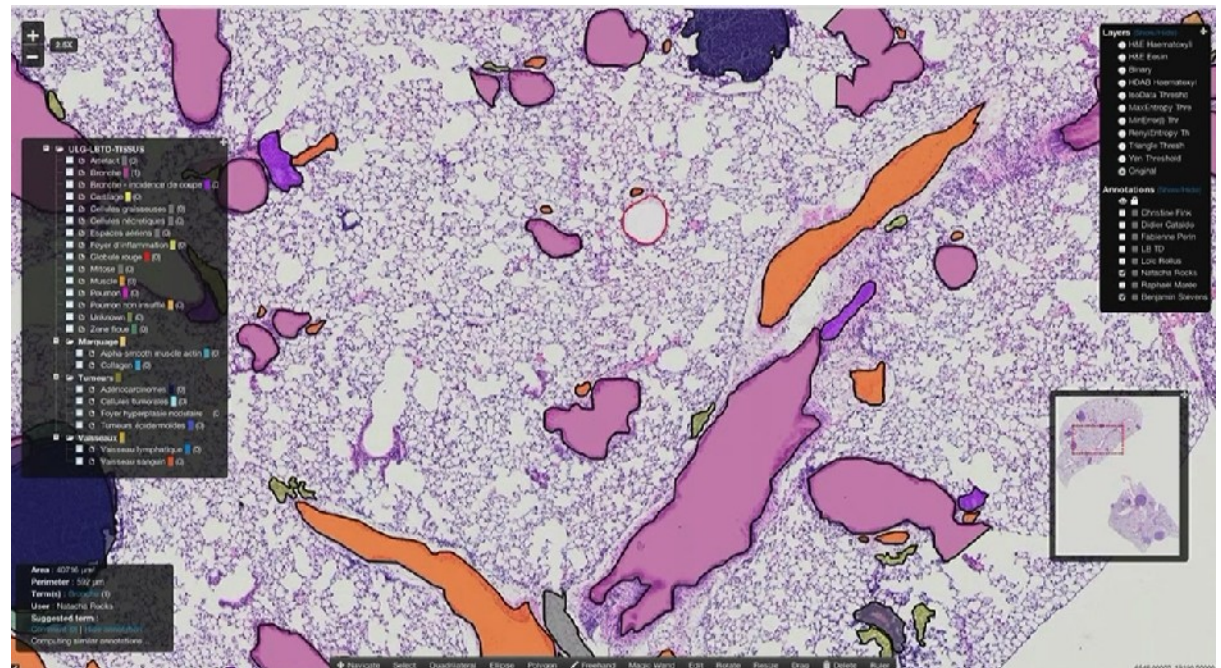
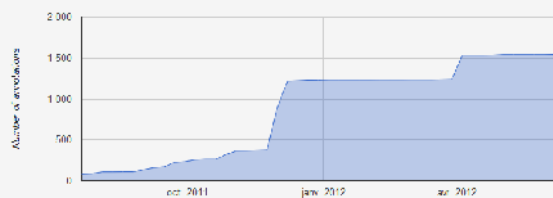
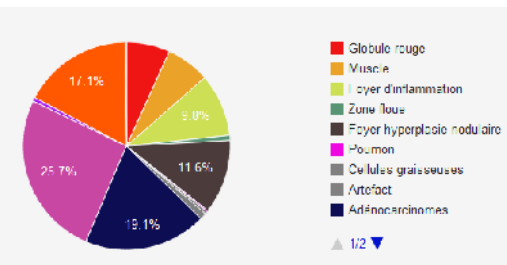
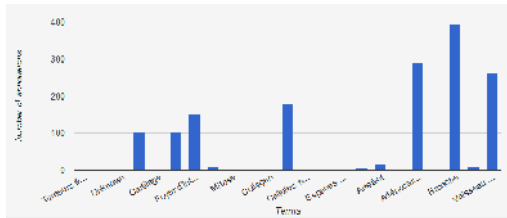
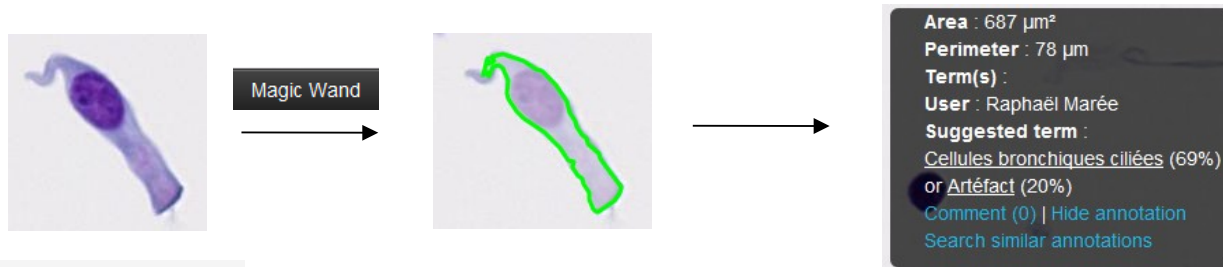


1 tissue slice = 35000 x 30000 pixels (0.23µm/pixel)

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

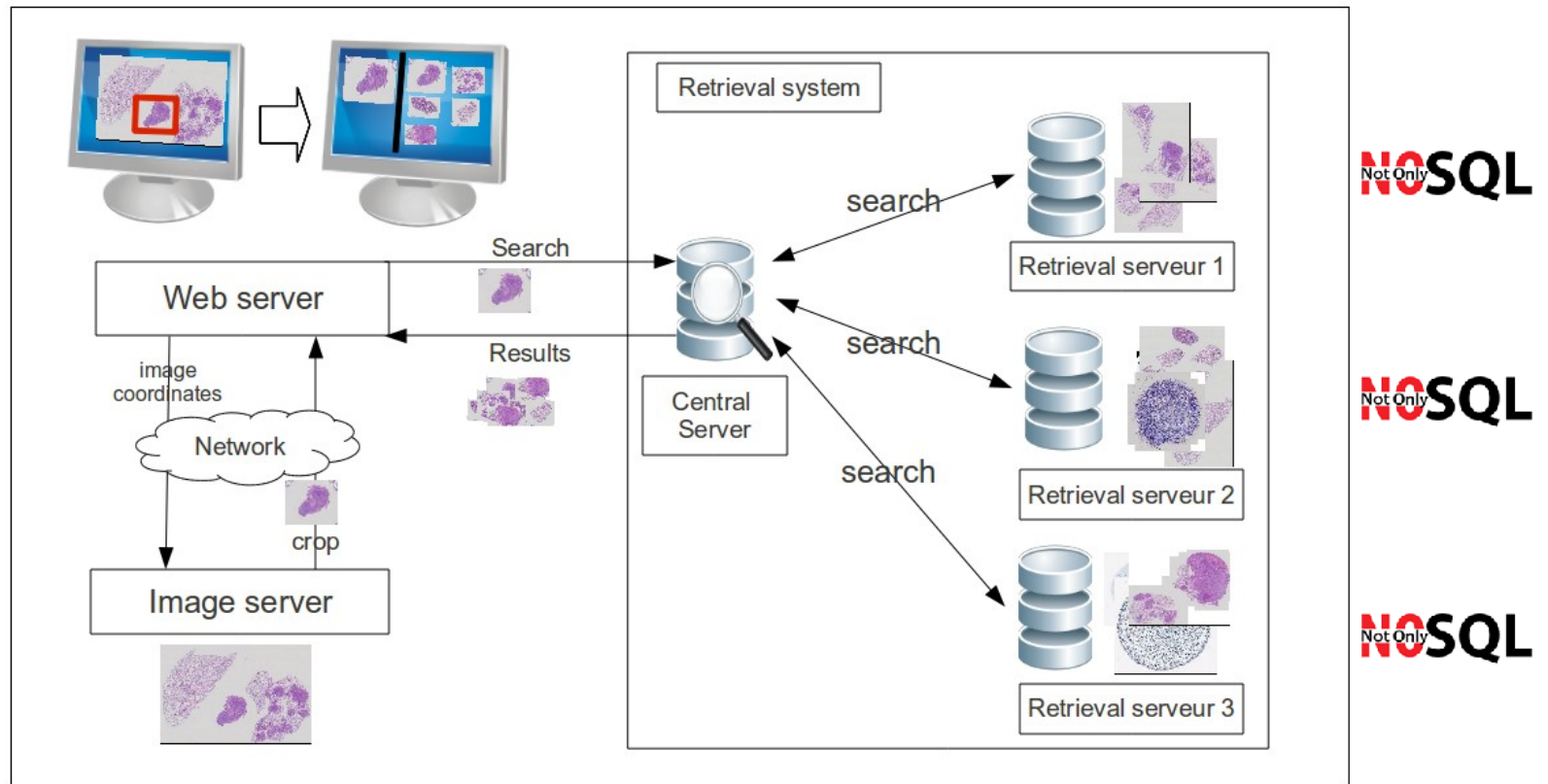
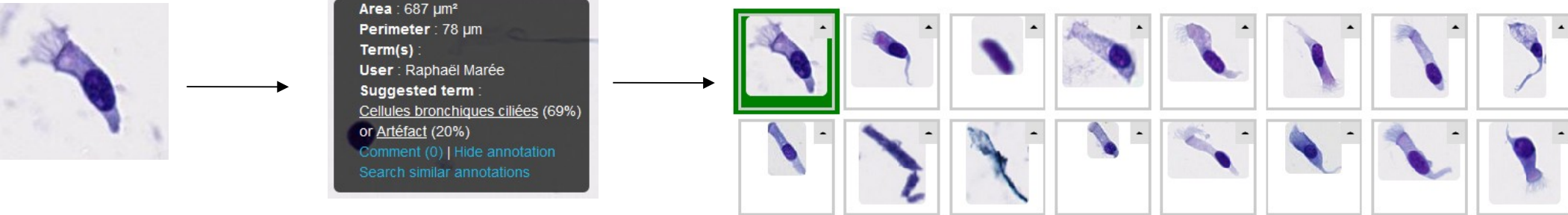
Software features : Annotate

- **Annotate** images using various **drawing tools**, with **user-specific layers**
- Describe ROIs with **ontology terms** (term suggestion using **CBIR**)
- Describe images and ROIs with any **key-value properties** or **text description**
- Build up pathology atlases and gather annotation statistics



Software features : Search

– Visual search of regions of interest



Software features : Share

- **Share images** through simple URLs

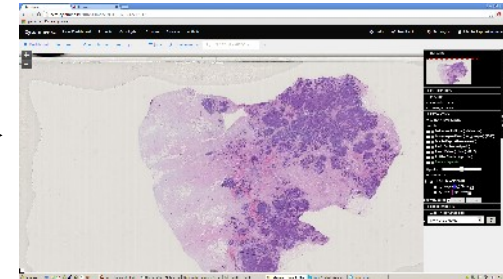
<http://beta.cytomine.be/#tabs-image-83151073-86503947-> →

Sign in to Cytomine

Username

Password

Remember me



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

<http://beta.cytomine.be/#share-annotation/92024416>

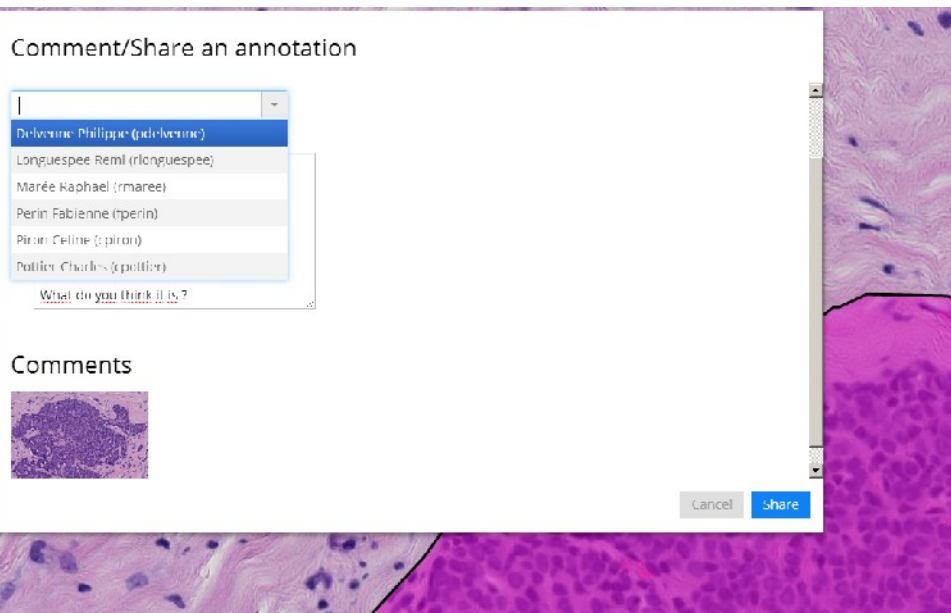
From: cytomine.ajp@gmail.com To: marée.marée@guigard.fr
Subject: Cytomine - Marée Raphaël (mareae) shared an annotation with you Date: Fri, 29 Nov 2013 22:12:05 +0100 (EET)

Dear Marée Raphaël (mareae),

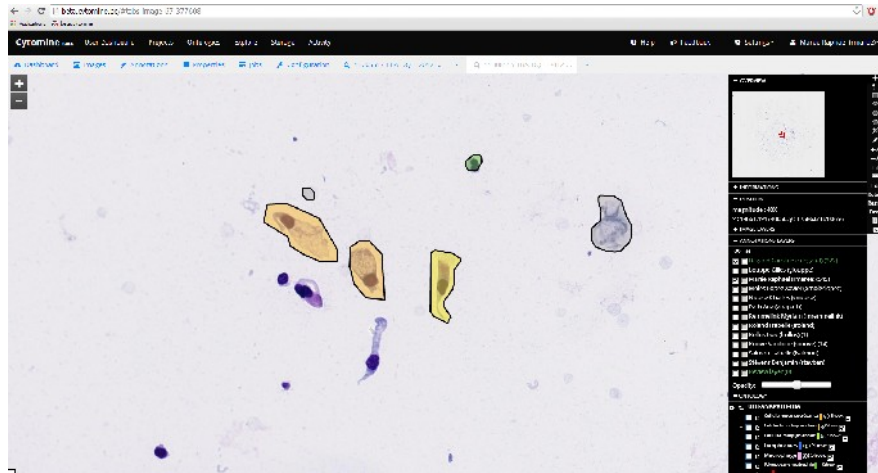
Marée Raphaël (mareae) shared an annotation with you and said :

Dear colleague, Do you think it is grade C ?

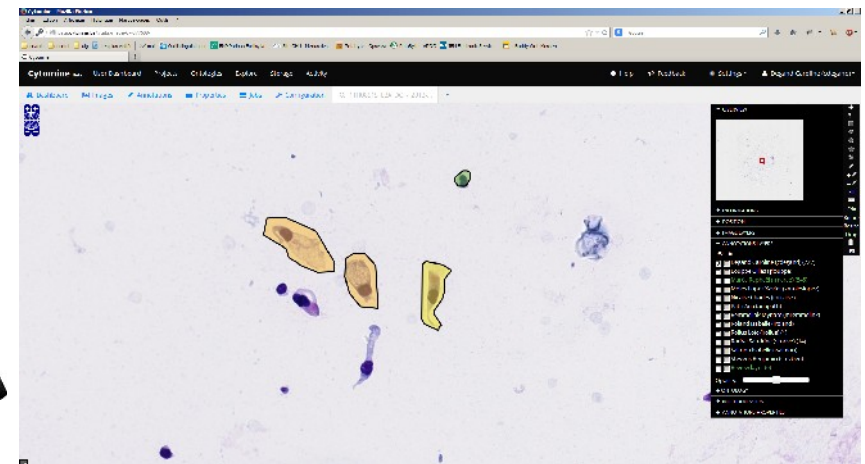
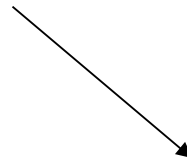
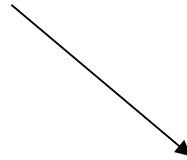
Navigate to <http://beta.cytomine.be/#share-annotation/92024416> in order to reply.
Navigate to <http://beta.cytomine.be/#tabs-image-83151073-86503947-92024416> in order to view the annotation within its context, or click on the thumbnail.



Software features : Live broadcast



- Degand Caroline (cdegand) (222)
- Louppe Gilles (gloupe)
- Marée Raphaël (rmaree) (545)
- Moles Lopez Xavier (xmoleslopez)
- Nicaise Charles (cnicaise)
- Path Ana (anapath)
- Rimmelink Myriam (mremmelink)
- Roland Isabelle (iroland)
- Rollus Loic (lrollus) (1)
- Rorive Sandrine (srorive) (14)
- Salmon Isabelle (isalmon)
- Stévens Benjamin (stevben)

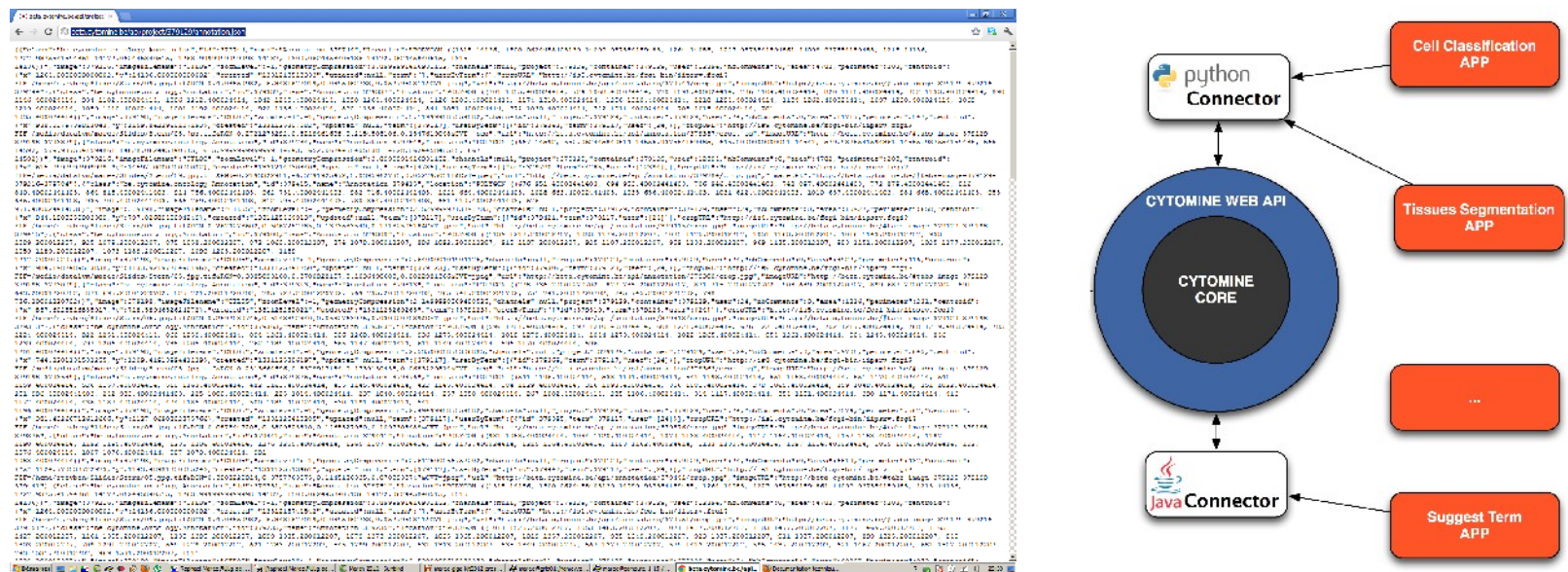


Start following Marée Raphaël (rmaree)

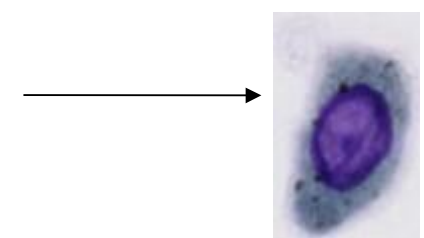


Software features : Extend and reproduce

- Integrate third-party softwares through **web services** with HTTP requests and **import/export data** through JSON messages
 - <http://beta.cytomine.be/api/project.json>
 - <http://beta.cytomine.be/api/annotation.json?&project=60&term=4777&users=14,16>



- <http://beta.cytomine.be/api/annotation/75499.json>
- <http://beta.cytomine.be/api/annotation/75499/crop.jpg?zoom=0>
- <http://beta.cytomine.be/#tabs-image-67-58147-75499>

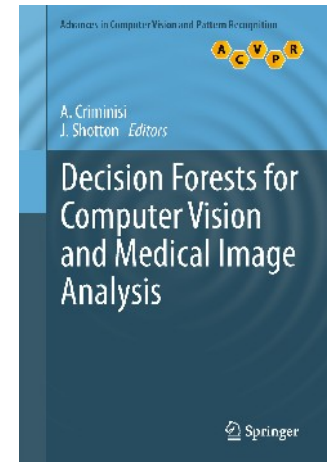


- Software parameters and results are recorded in the centralized database to ease traceability and reproducibility

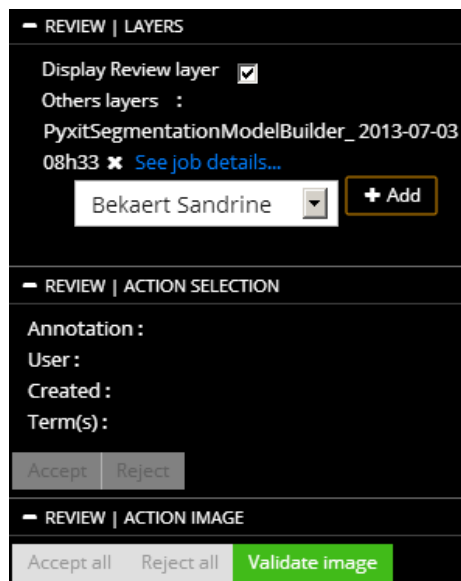
Software features : Analyze and proofread

- **Generic machine-learning** based image recognition (without user-defined rules nor explicit features)

Marée et al. (2013). *Extremely Randomized Trees and Random Subwindows for Image Classification, Annotation, and Retrieval*. Invited chapter in A., Criminisi & J., Shotton (Eds.), *Decision Forests in Computer Vision and Medical Image Analysis*, Advances in Computer Vision and Pattern Recognition, pp. 125-142. Springer.



- Built-in interfaces for **algorithm evaluation** and **collaborative proofreading**



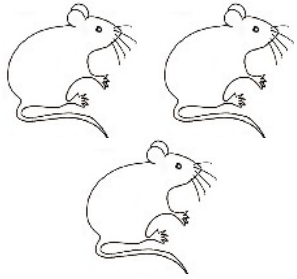
Marée et al. (2014). *A hybrid human-computer approach for large-scale image based measurements using web services and machine learning*. To appear in Proc. IEEE International Symposium on Biomedical Imaging (ISBI)

Applications

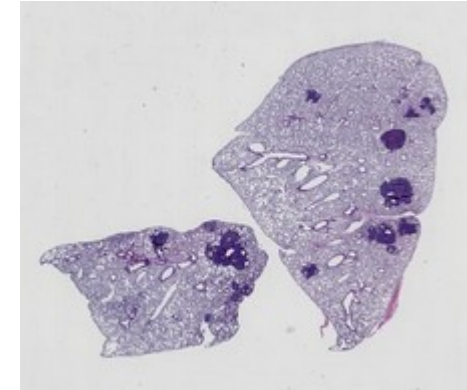
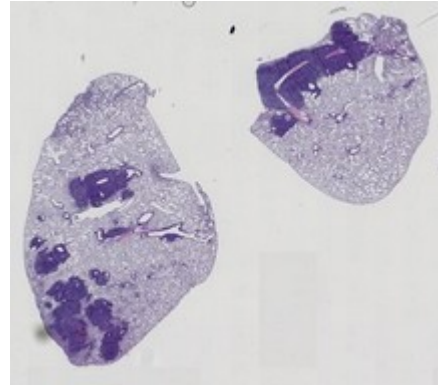
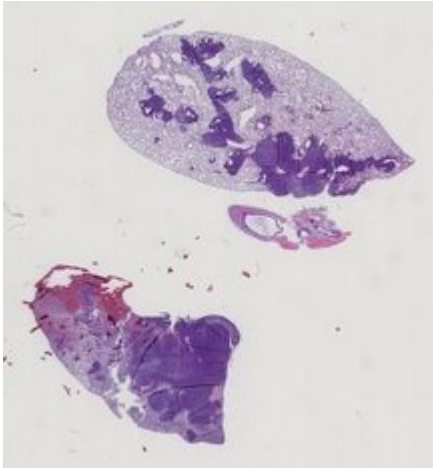
LUNG tumor tissue quantification

(ongoing collaboration with D. Cataldo, N. Rocks, at LBTD, GIGA)

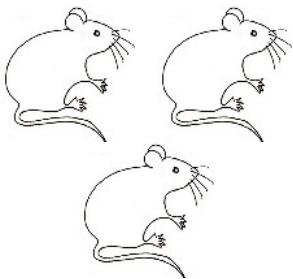
What is the impact of condition X/Y/... on lung tumor onset and progression ?



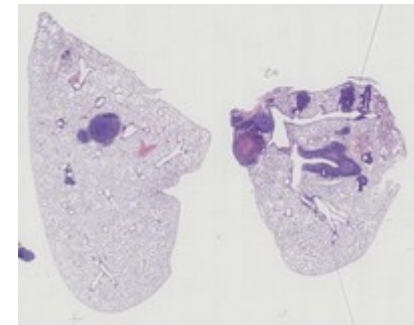
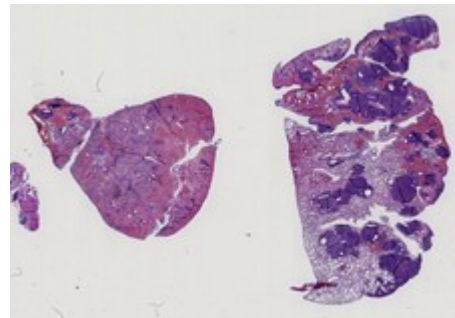
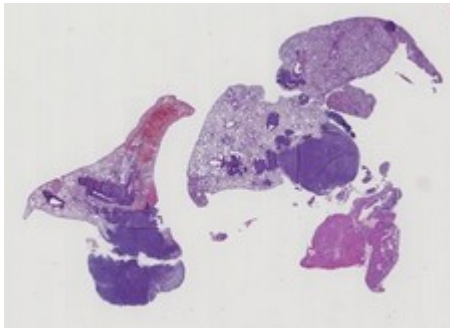
Condition X



...



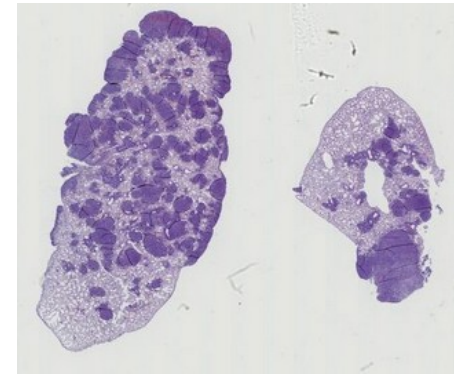
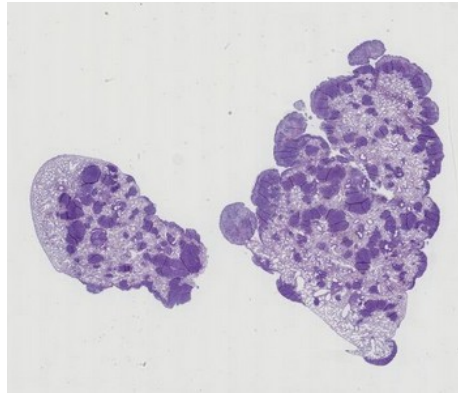
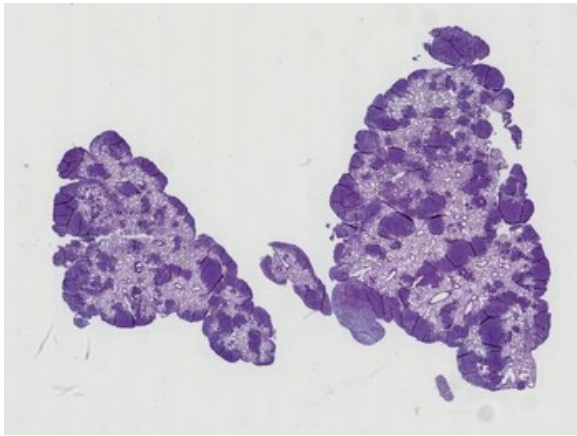
Condition Y



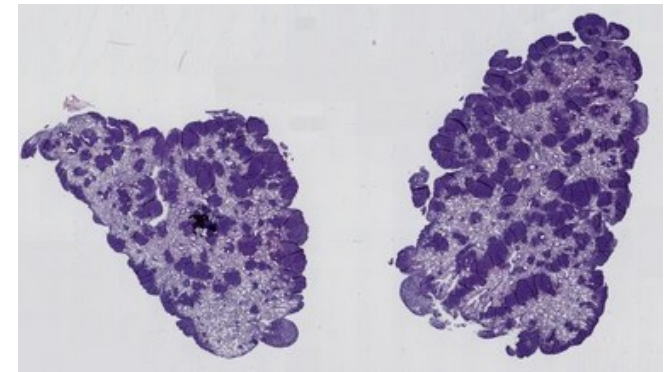
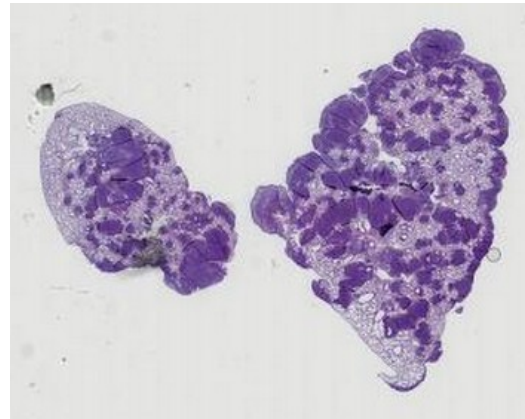
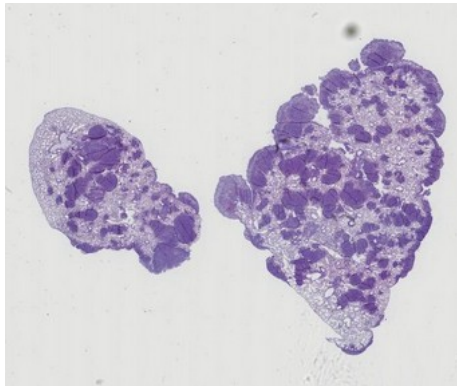
...

■
■
■

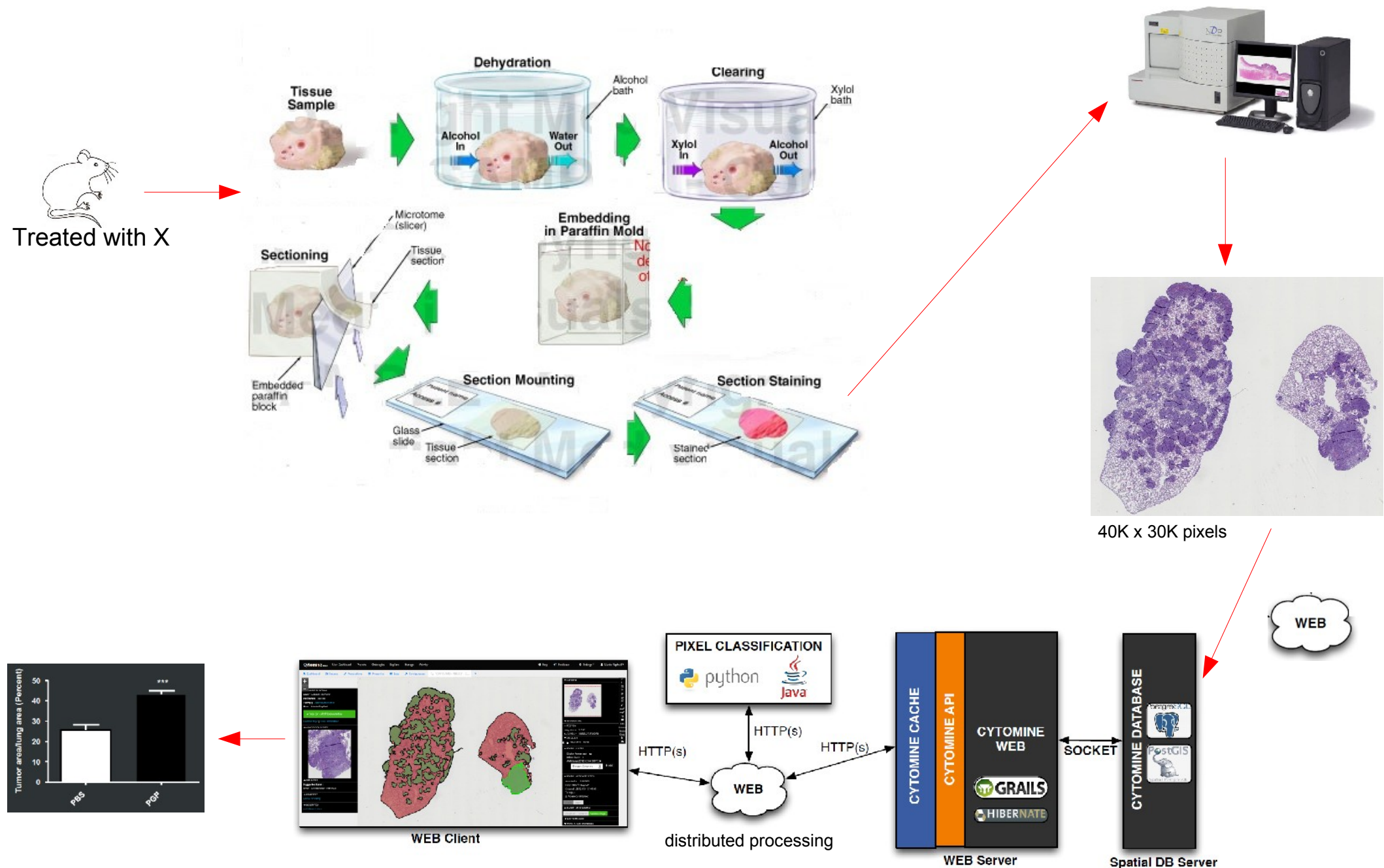
Tens or hundreds of glass slides to be quantified per study...



...



Hybrid human-computer workflow



Hybrid human-computer workflow

1. Manual region contouring and labelling to provide training examples

The image illustrates a hybrid human-computer workflow for histological image analysis. The central component is the Cytomine software interface, which displays a large histological image with various regions manually contoured and labeled. The interface includes a sidebar with a hierarchical list of labels and their counts, such as:

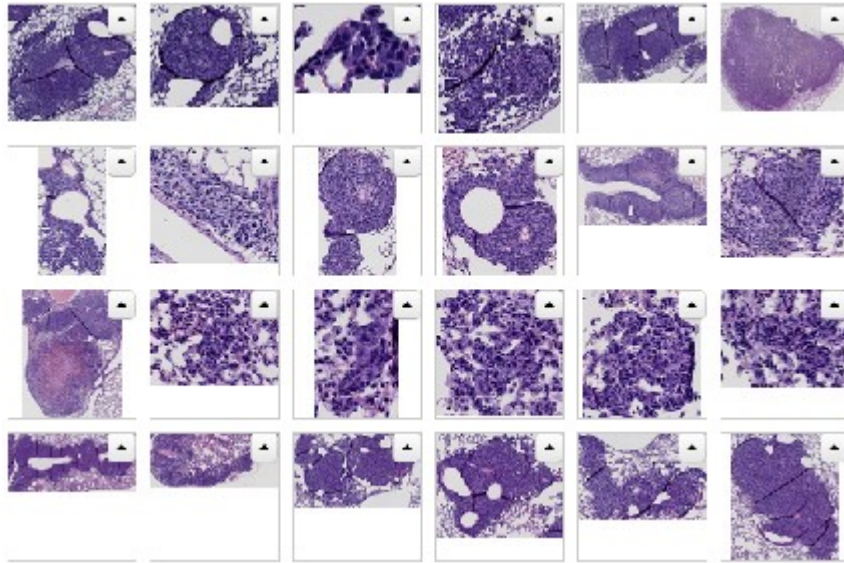
- III C: ADI-TISSUS
 - U Artificiel (0)
 - U Background (0)
 - U Branche (0)
 - U Cartilage (0)
 - R Cellules grasses (0)
 - R Cellules nécrotiques (0)
 - R Espaces aereux (0)
 - R Foyer d'inflammation (0)
 - R Globule rouge (0)
 - R Mitose (0)
 - R Muscle (0)
 - D Poumon (0)
 - D Poumon non insufflé (0)
 - D Unknown (0)
 - D Zone floue (0)
- Marquage
 - Alphasmooth muscle actin
 - Collagen (0)
- Tumeurs
 - Adenocarcinomes (0)
 - Cellules tumorales (0)
 - Foyer hyperplasie nodulaire
 - Nécrose tumorale (0)
 - Tumeurs Apicomorales (0)
 - Vaisseaux tumoraux (0)
- Vaisseaux
 - Vaisseaux lymphatiques (0)
 - Vaisseaux sanguins (0)

Surrounding the central window are several zoomed-in crop images of specific regions, each with a label:

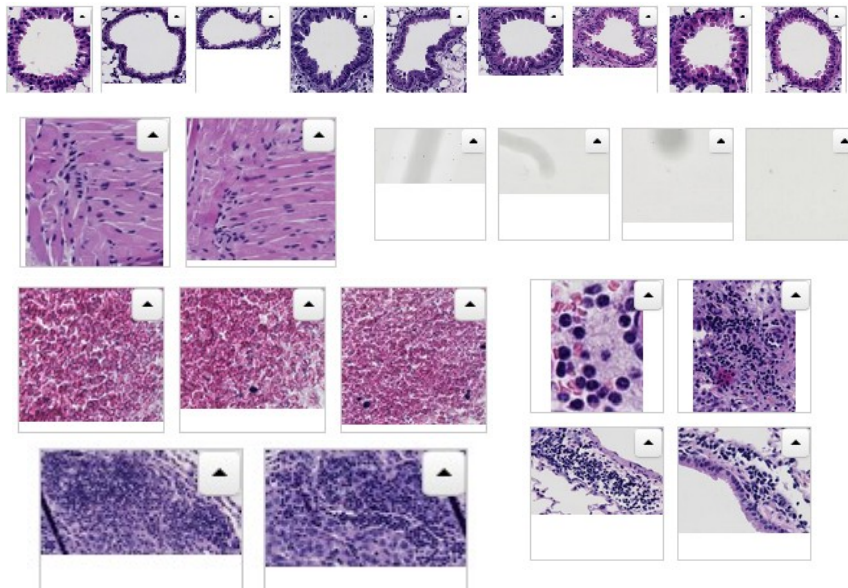
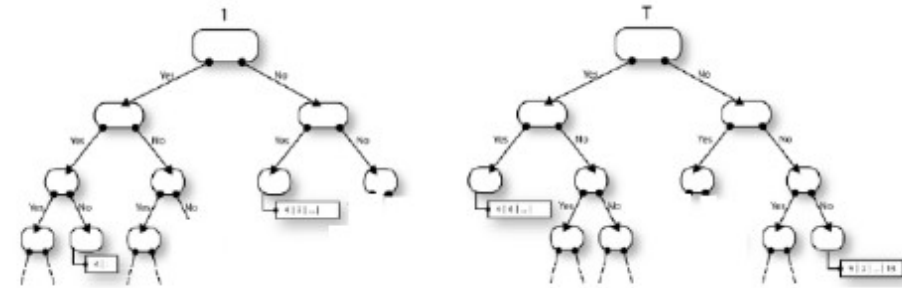
- necrosis
- inflammatory cells
- cartilage
- bronchus
- adenocarcinoma
- blood vessel
- red-blood

Hybrid human-computer workflow

2. Automatic training of image recognition model based on training examples













VS



Hybrid human-computer workflow

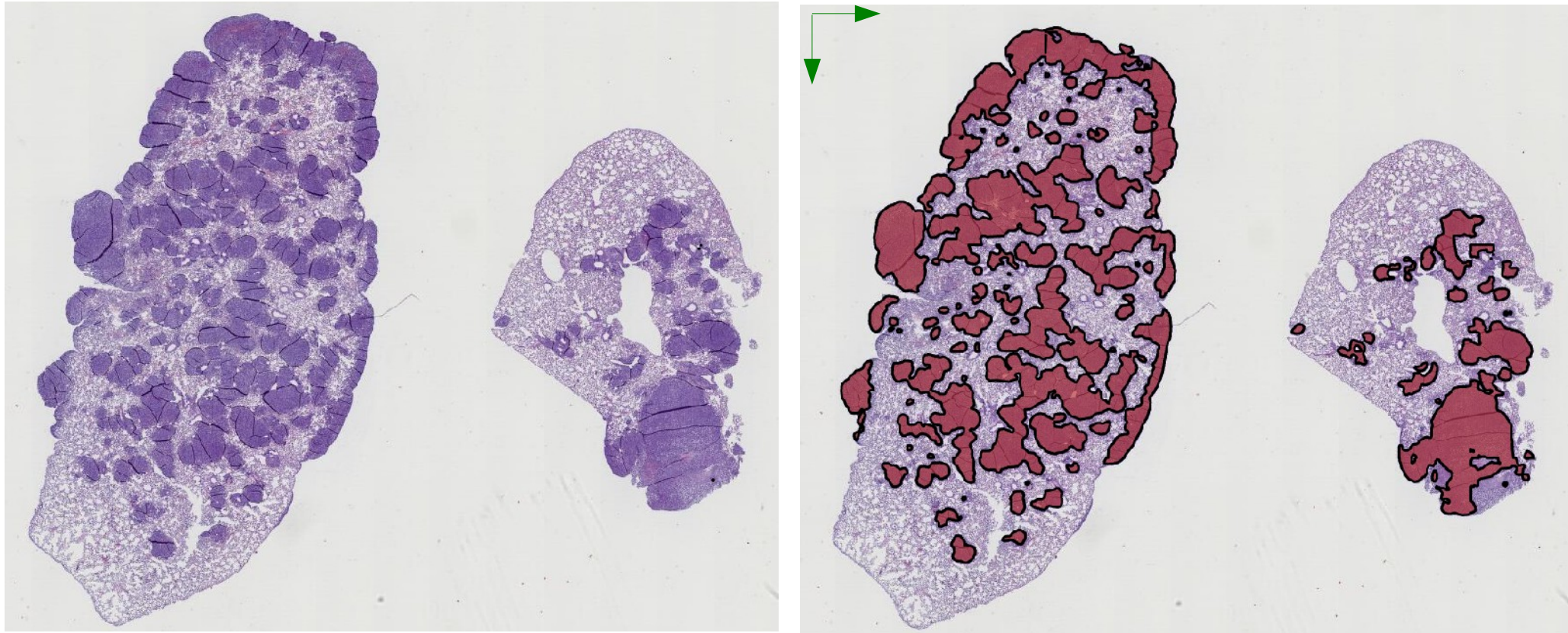
3. Batch processing of slides Reviewing slides

Filter by name

ID	Preview	Name	Width	Height	X	Resolution	User an.	Algo an.	Valid an.	Mime	Created	Status	Action
7889452		PGP POUMON PBS 3 21 - 2012-08-08 11.07.24.jp2	34816	29440	40	0.650µm/pixel	8	614	100	jp2	2012-10-18 10h50	Reviewed	Explore ▾
7889358		PGP POUMON PGP7 11 - 2012-08-07 12.07.00.jp2	36864	28416	40	0.650µm/pixel	2	130	119	jp2	2012-10-18 10h50	Reviewed	Explore ▾
7889229		PGP POUMON PGP7 41 - 2012-08-07 12.13.50.jp2	36864	28416	40	0.650µm/pixel	0	142	142	jp2	2012-10-18 10h50	Reviewed	Explore ▾
7889190		PGP POUMON PGP7 21 - 2012-08-07 12.09.19.jp2	34816	28928	40	0.650µm/pixel	5	122	124	jp2	2012-10-18 10h50	Reviewed	Explore ▾
7889180		PGP POUMON PGP7 31 - 2012-08-07 12.11.34.jp2	36864	26624	40	0.650µm/pixel	7	271	213	jp2	2012-10-18 10h49	Reviewed	Explore ▾
7889170		PGP POUMON PGP8 11 - 2012-08-07 12.18.57.jp2	38912	32256	40	0.650µm/pixel	3	169	151	jp2	2012-10-18 10h49	Reviewed	Explore ▾
7889160		PGP POUMON PGP8 31 - 2012-08-07 12.24.39.jp2	40960	32256	40	0.650µm/pixel	3	158	145	jp2	2012-10-18 10h49	In review	Explore ▾
7889150		PGP POUMON PGP8 41 - 2012-08-07 12.27.33.jp2	40960	33024	40	0.650µm/pixel	2	846	782	jp2	2012-10-18 10h49	In review	Explore ▾
7889140		PGP POUMON PGP8 21 - 2012-08-07 12.21.46.jp2	40960	32256	40	0.650µm/pixel	8	161	146	jp2	2012-10-18 10h48	In review	Explore ▾
7889130		PGP POUMON PGP10 41 - 2012-08-07 12.55.37.jp2	36864	28928	40	0.650µm/pixel	0	175	175	jp2	2012-10-18 10h48	In review	Explore ▾

Hybrid human-computer workflow

3. **Automatic** segmentation of tumors in new slide images



One image ~ 40 000 x 30 000 pixels
Tile-based pixel classification (tumor / nontumor) + contour processing

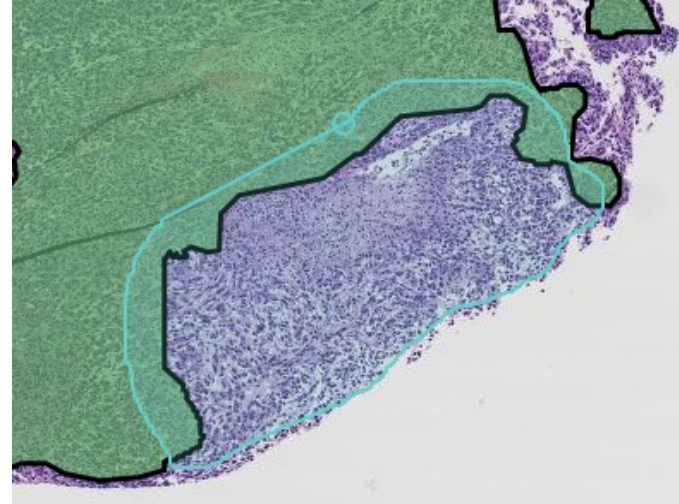
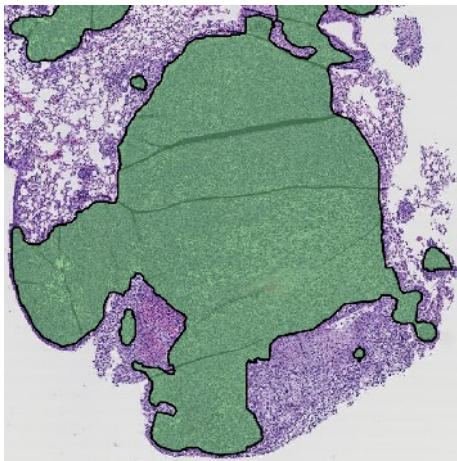
Hybrid human-computer workflow

3. Proofreading automatic segmentations

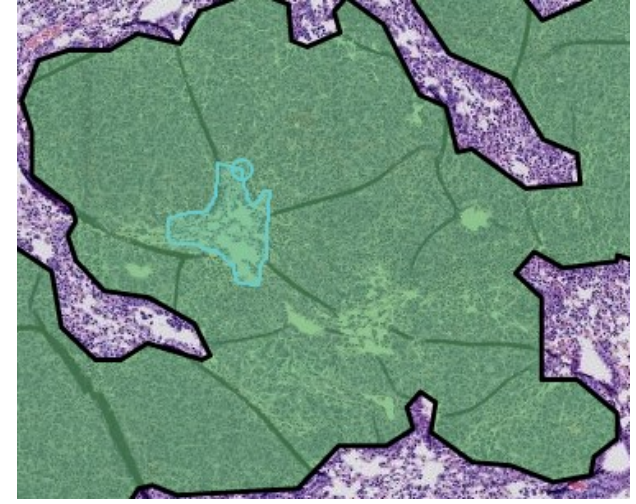
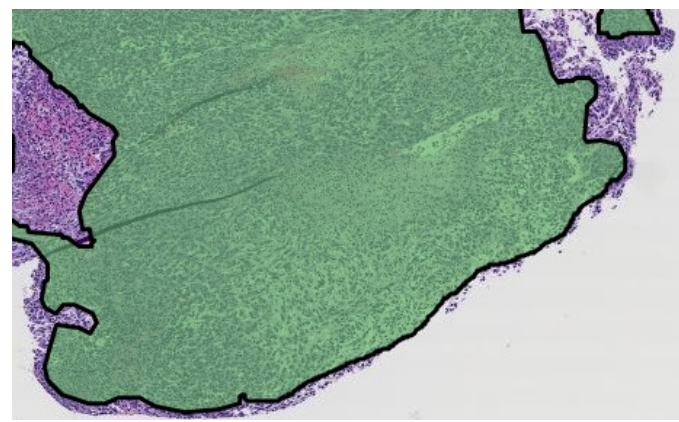
Edit Rotate Resize Drag Fill + Extend - Subtract Delete



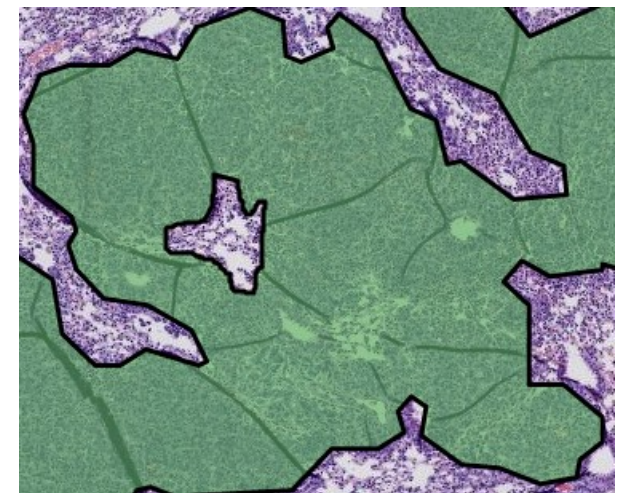
Fill



+ Extend

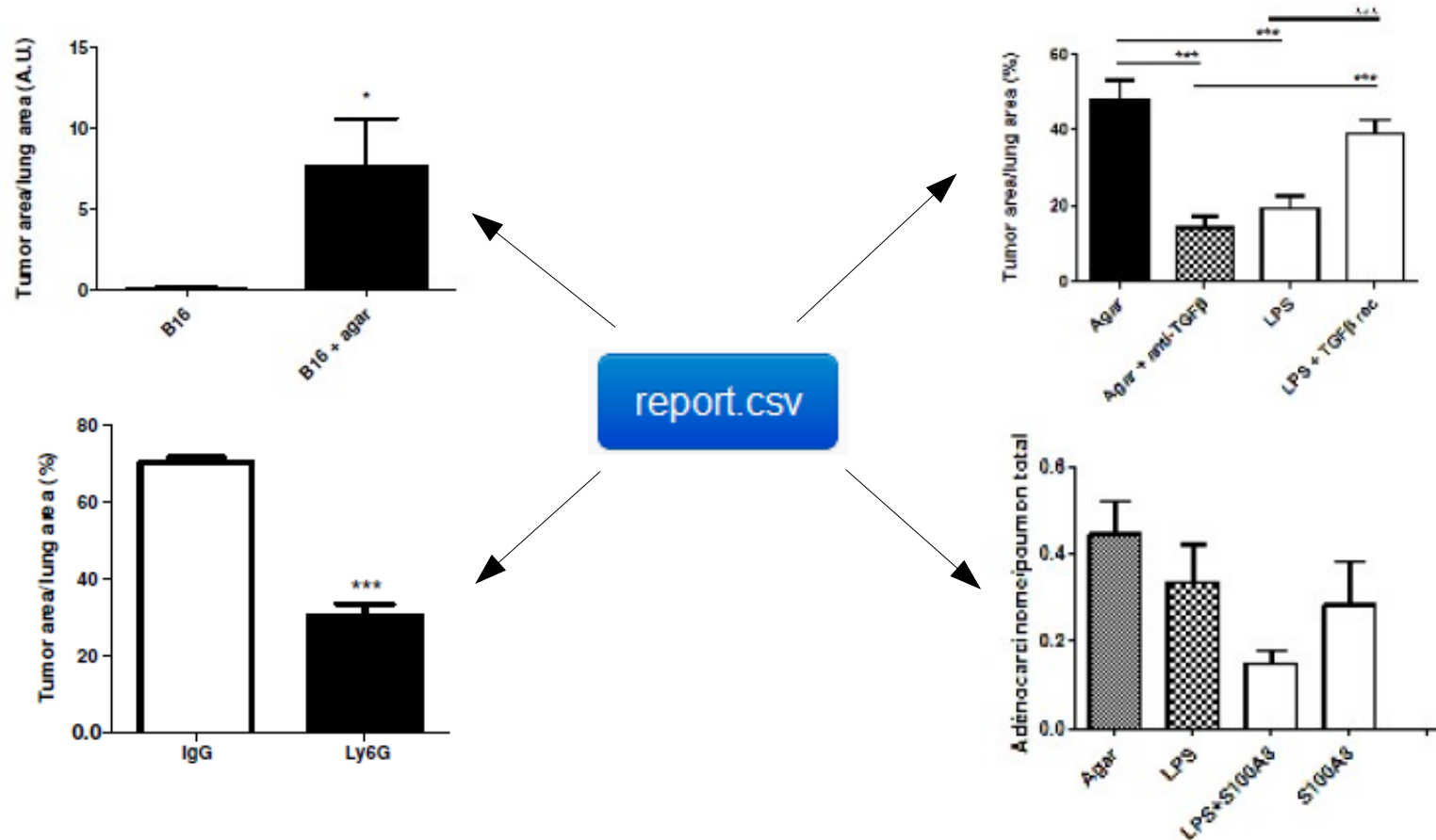


- Subtract



Hybrid human-computer workflow

4. Export statistics



> 500 whole-slide images analyzed with > 20 000 validated tumoral islets

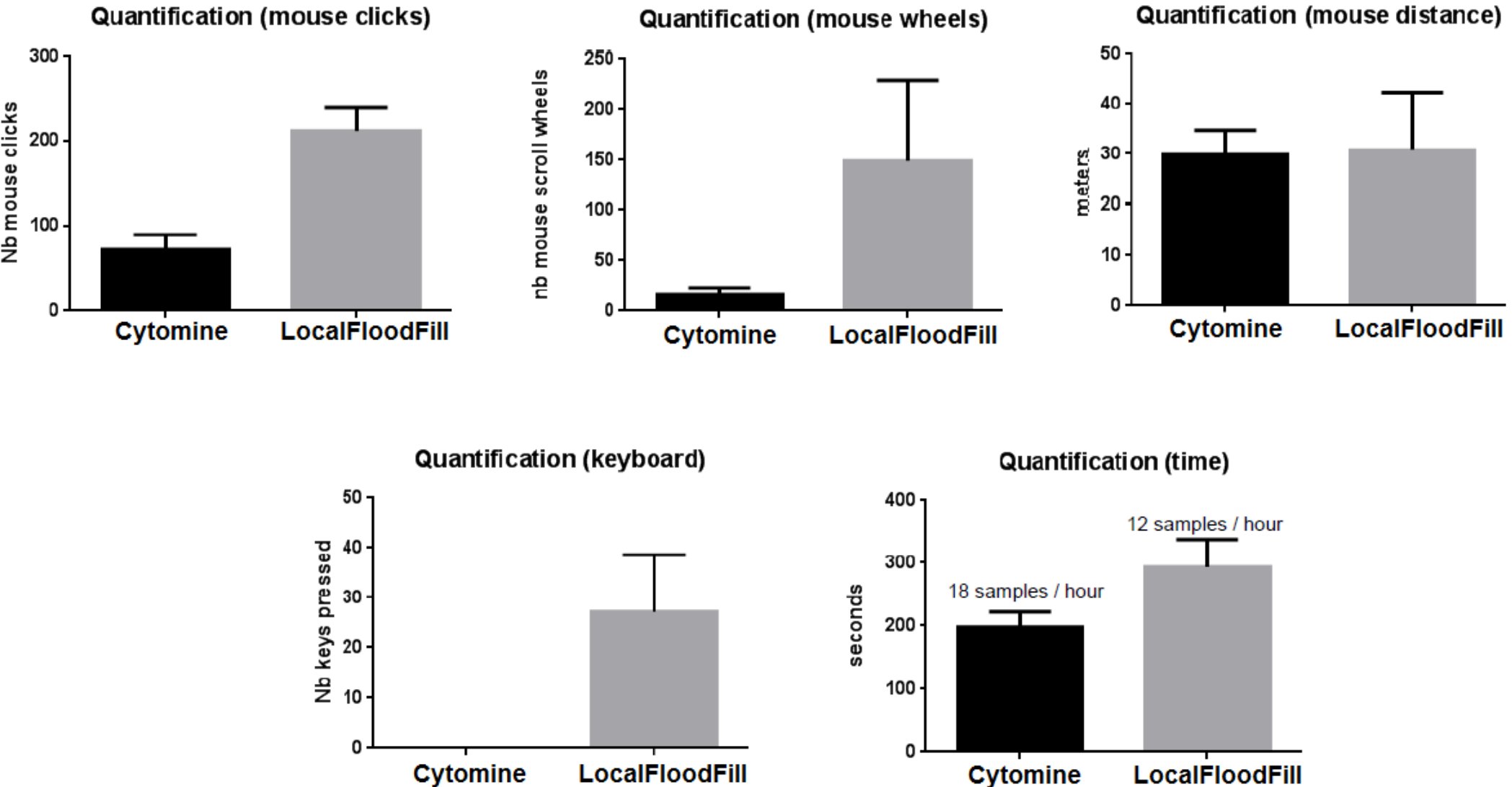
Hybrid human-computer workflow

4. Recognition performances : biologist's metrics : what is the impact on daily workload ?

Proofreading algorithm through WiFi connection

vs

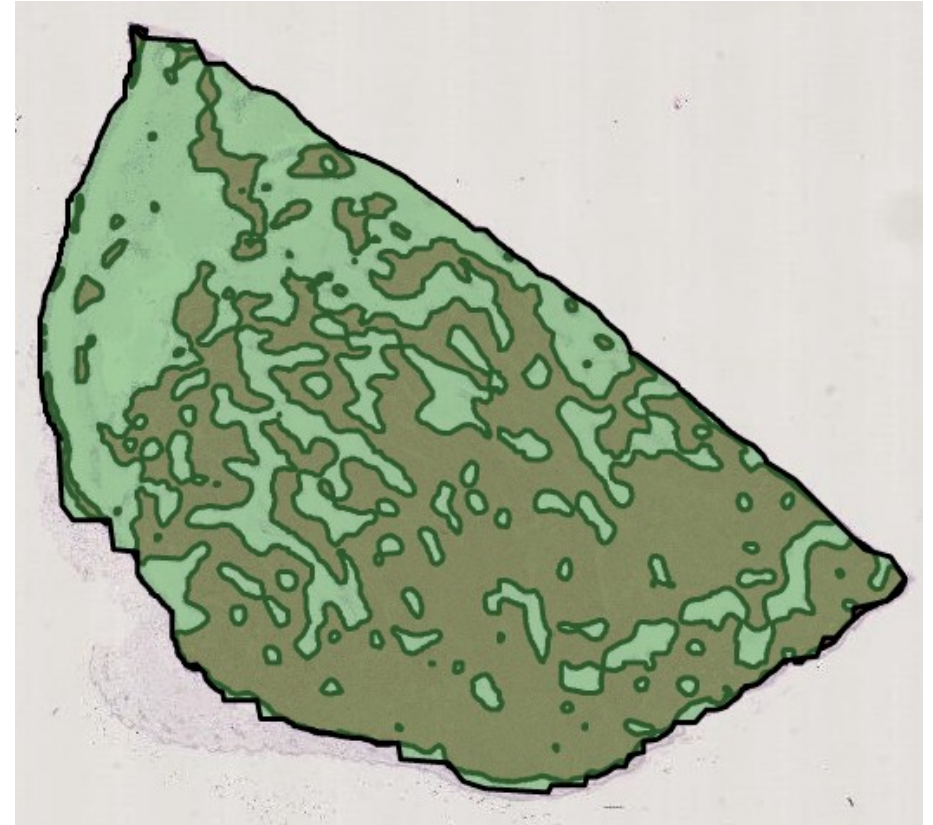
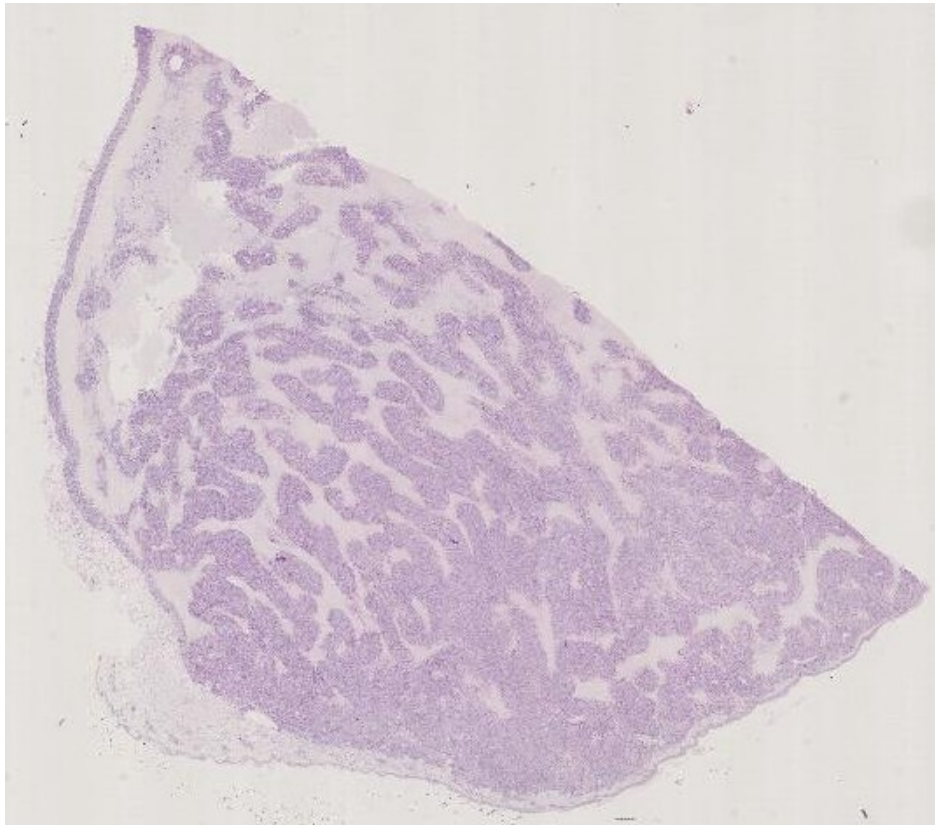
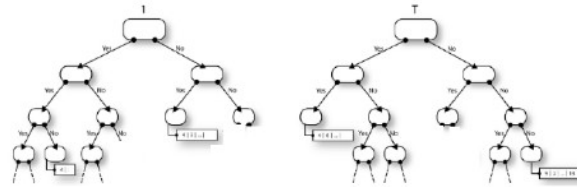
Flood fill algorithm on local computer



(statistics obtained for 5 slides using WinOMeter)

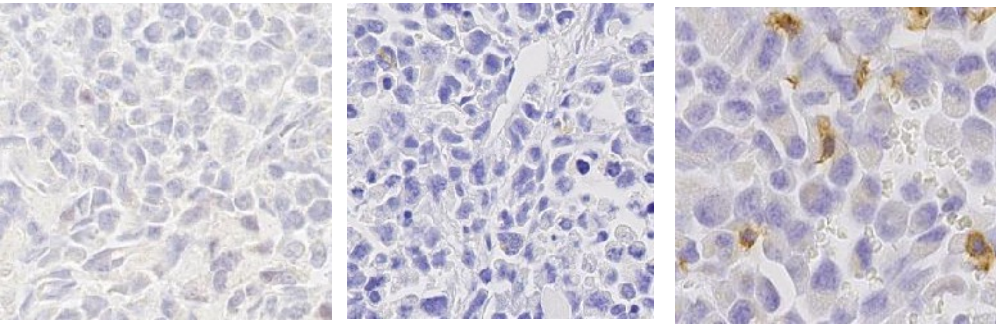
Other applications : tumor/necrosis (H&E)

(ongoing work with C. Pequeux at LBTD, GIGA)

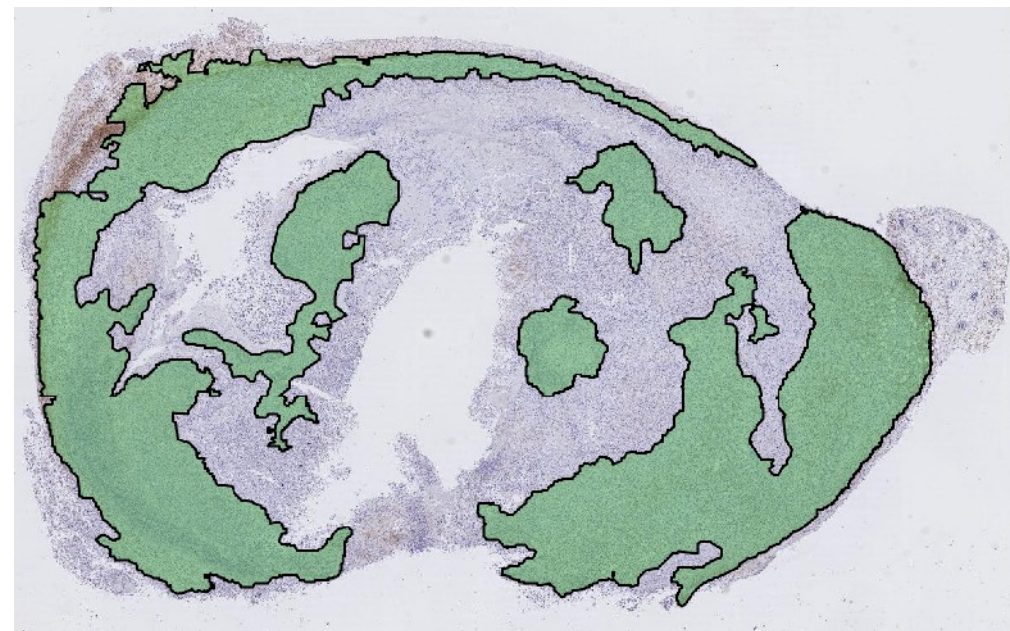
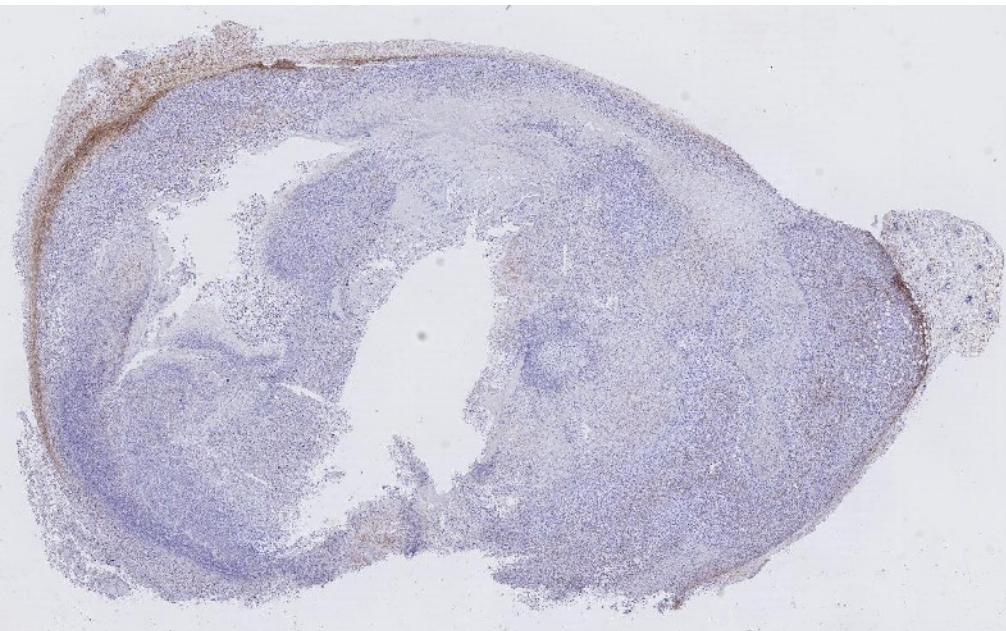
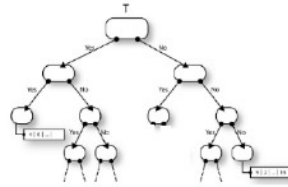
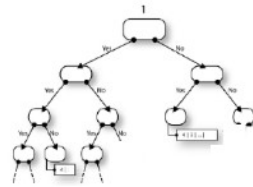
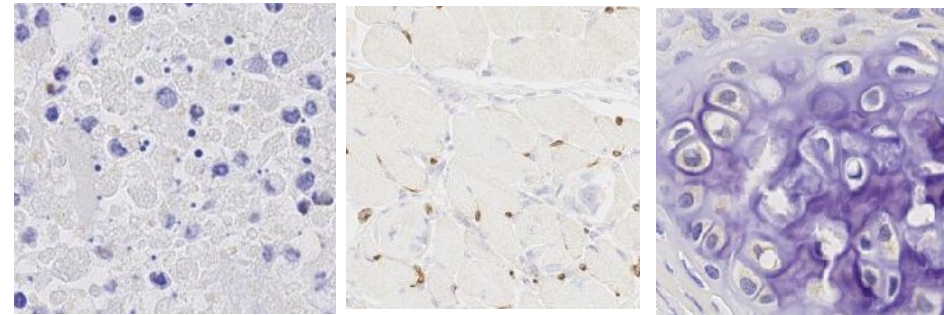


Other applications : tumor/necrosis (IHC)

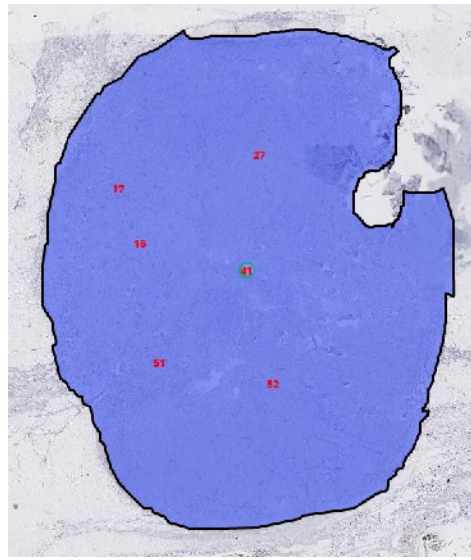
(ongoing work with Ph. Martinive, N. Leroi at LBTD, GIGA)



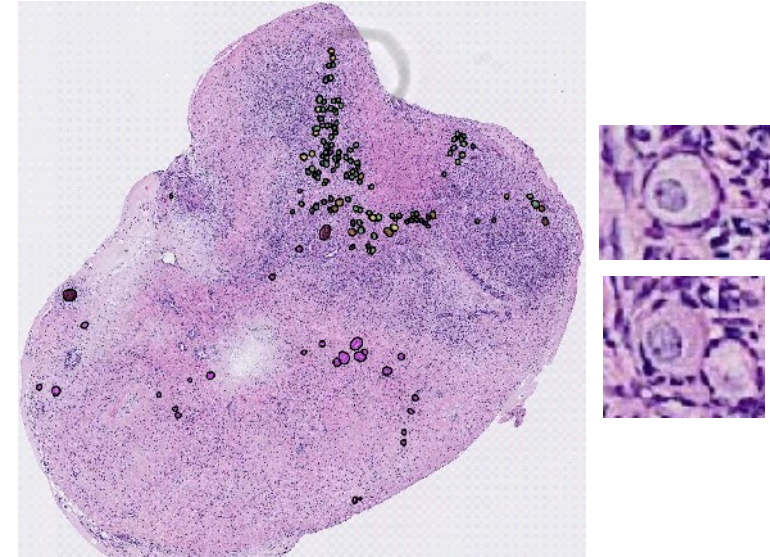
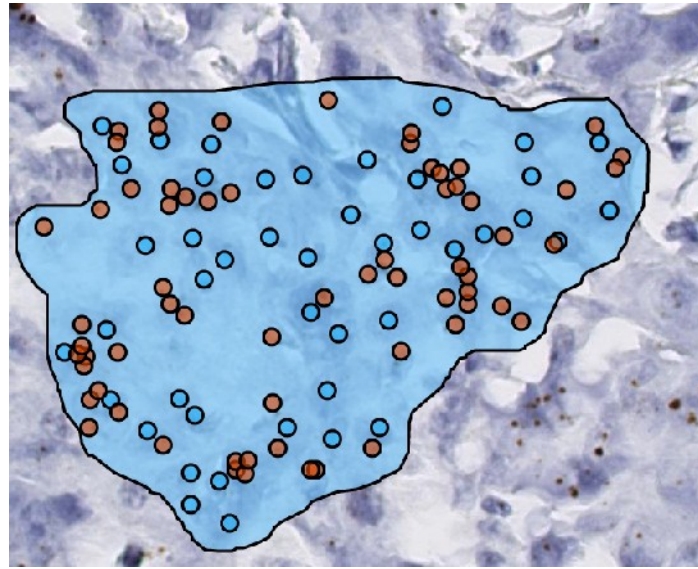
VS



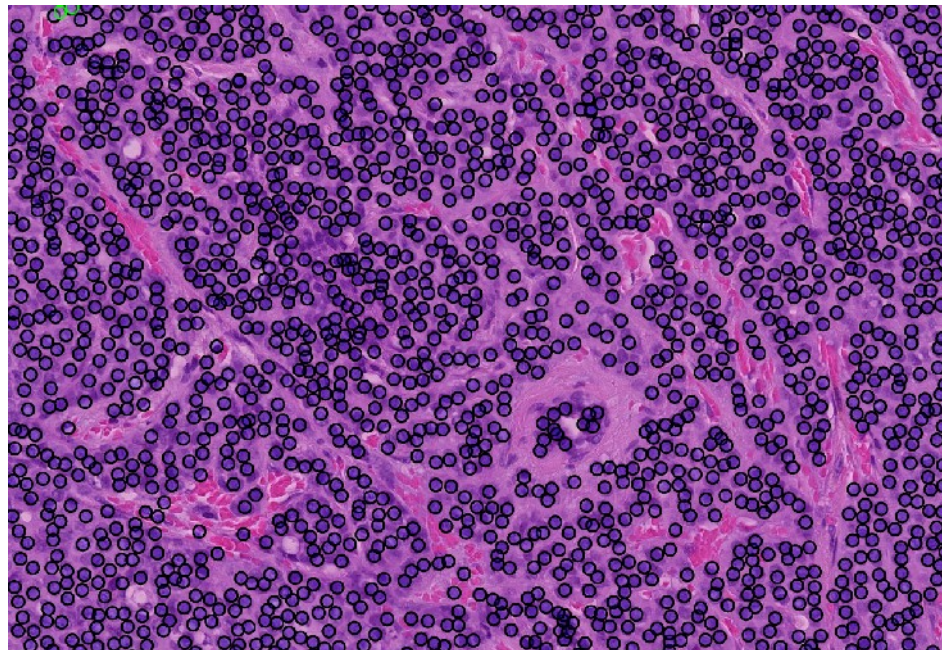
Other applications : counting



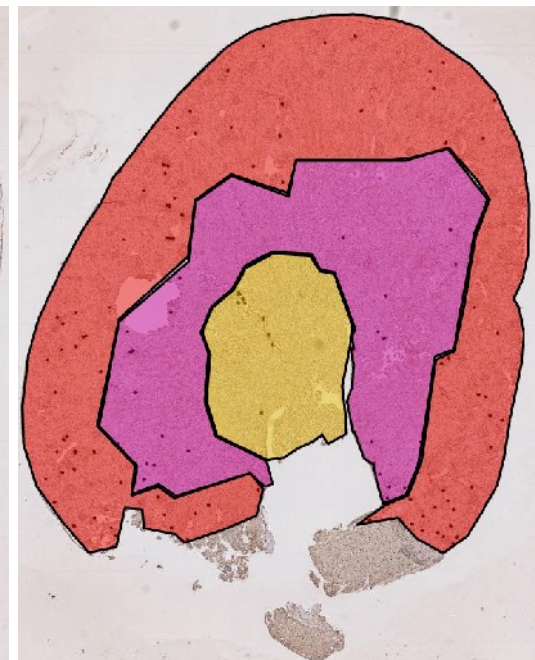
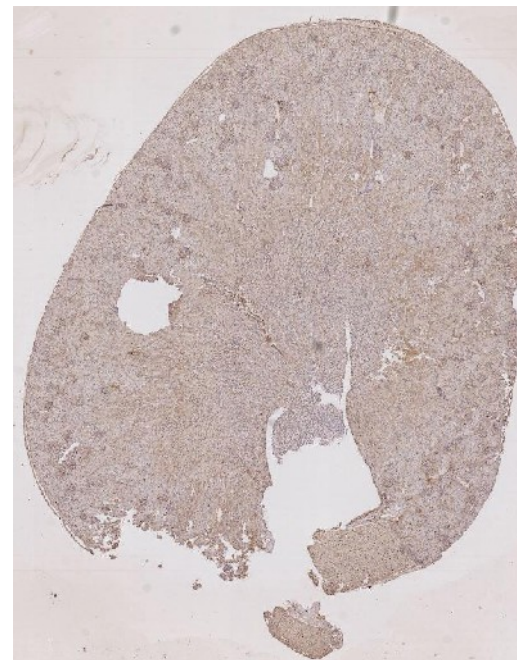
RNAscope spot counting, breast tumors (C. Josse, GIGA)



Follicle counting, ovarian (C. Munaut, GIGA)



H&E nucleus counting (R. Longuespée, GIGA)

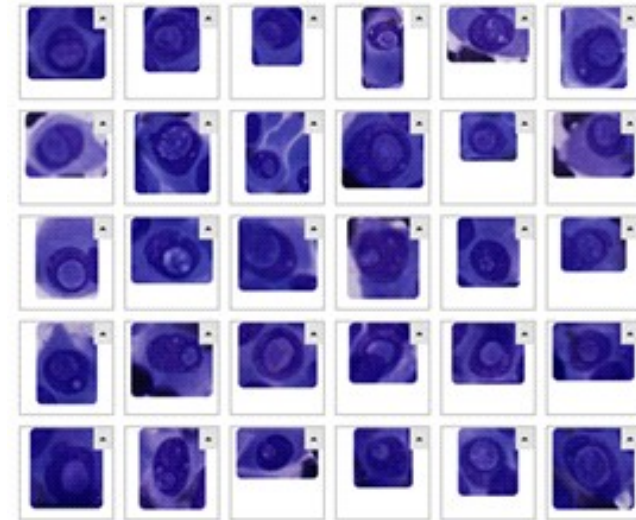


IHC positive cell counting, nephrology (F. Jouret, GIGA)

Other applications : diagnostic cytology

(fine-needle aspiration of the thyroid, ongoing work with I.Salmon at ULB)

The screenshot displays the Cytomine software interface. The main window shows a histological image of a fine-needle aspiration (FNA) of a thyroid. Several circular annotations are placed on the image, with lines connecting them to magnified inset images. The software's sidebar on the right shows a hierarchical tree of annotations for the project 'ULB-ANAPATH-THYROID-FNAB'. The tree includes categories like 'Architectural patterns', 'Nuclear features', and 'Others'. The main window also shows a toolbar with various tools for image manipulation and analysis.



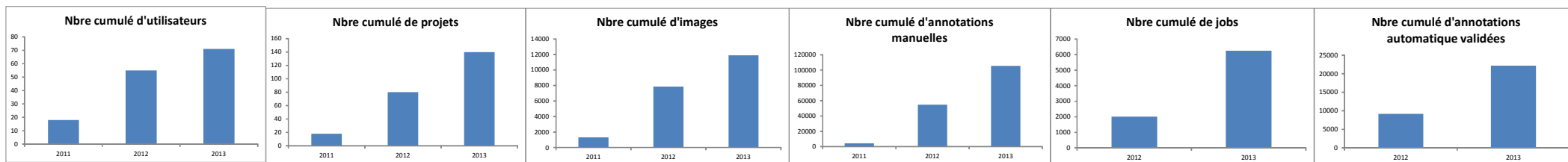
The screenshot displays the Cytomine software interface showing a review screen for the project 'ULB-ANAPATH-THYROID-FNAB'. The screen displays a grid of five images with classification labels and confidence scores. Below the grid is a legend with various classification categories and their corresponding colors.

Classification	Color
Normal follicular architectural pattern	Light Green
Proliferative follicular architectural pattern (minor sign)	Yellow
Proliferative follicular architectural pattern	Red
Aggregate	Purple
Architectural pattern to classify	Blue
Architectural patterns	Black
Artefacts	Brown
Background	Dark Grey
Cell to classify	Pink
Colloid	Light Grey
Macrophages	Dark Grey
Noninclusion	Purple
Normal follicular cell with pseudo-nucleolus	Light Green
Normal follicular cells	Green
Nuclear features	Black
Others	Red
PN (polyuclear)	Light Grey
Papillary cell NOS	Blue

Summary

CYTOMINE : a rich internet application

- Uses generic software design, web services, and machine learning
- Fosters collaboration between pathologists, life scientists, and computer scientists
 - Eases sharing of whole slides and annotations
 - Speeds up large-scale image quantifications
 - Offers mechanisms to integrate novel algorithms / image formats
- ~ 100 users, 150 projects, > 12K images, > 125K annotations



Future work

- Improve algorithm robustness and further speedup workflows
- Development for histology/anatomopathology training courses

Future work

- Improve algorithm robustness and further speedup workflows
- Development for histology/anatomopathology training courses

- Working together ?



Acknowledgments

- Systems and Modeling (GIGA-Research / Montefiore Institute): Loïc Rollus, Benjamin Stévens, Gilles Louppe, Olivier Stern, Nathalie Jeanray, Vincent Botta, Pierre Geurts, Louis Wehenkel
- CYTOMINE software beta-testing, etc. : Didier Cataldo, Natacha Rocks, Fabienne Perin, Christine Fink, Sandrine Bekaert, Myriam Remmelink, Caroline Degand, Isabelle Salmon, Sandrine Rorive, Audrey Voncken, Jessica Aceto, Yoann Curé, Benoist Pruvot, Marc Muller, Natacha Leroi, ...

Raphaël Marée is funded by **GIGA FEDER** grant and the **CYTOMINE** (2010-2014) research grant n° 1017072 of the Wallonia (DGO6). Benjamin Stévens is funded by **SMASH** spin-off grant n° 1217606 of the Wallonia

www.montefiore.ulg.ac.be/~maree/

www.cytomine.be



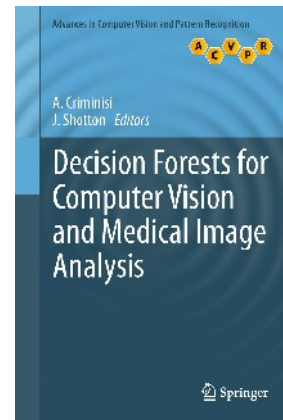
Related publications

- Marée et al., "A rich internet application for remote visualization and collaborative annotation of digital slide images in histology and cytology". BMC Diagnostic Pathology, 8(Suppl 1):S26, 30th September 2013



BioMed Central
DIAGNOSTIC PATHOLOGY

- Marée et al. (2013). *Extremely Randomized Trees and Random Subwindows for Image Classification, Annotation, and Retrieval*. Invited chapter in A., Criminisi & J., Shotton (Eds.), *Decision Forests in Computer Vision and Medical Image Analysis*, Advances in Computer Vision and Pattern Recognition, pp. 125-142. Springer.



- Marée et al. (2014). *A hybrid human-computer approach for large-scale image based measurements using web services and machine learning*. To appear in Proc. IEEE International Symposium on Biomedical Imaging (ISBI)

