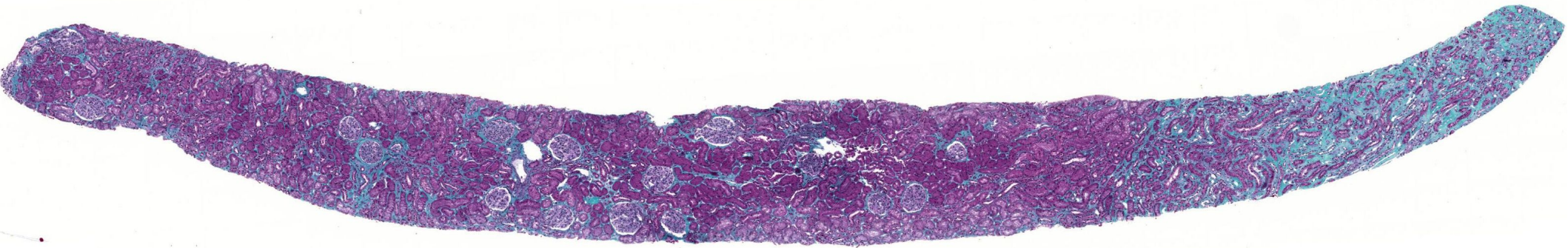
An approach for glomeruli detection in multisite digital pathology

Raphaël Marée*^, Stéphane Dallongeville^, Jean-Christophe Olivo-Marin^, Vannary Meas-Yedid^

^Bioimage Analysis Unit, Institut Pasteur, Paris, France

* Systems and Modeling, EE&CS, Univ. of Liège, Belgium

ISBI
2016
Prague



Large biomedical studies have to be multicentric

Benefits :

- Large number of participants
- Wider range of population groups (<> genetic, environmental, demographic factors)

→ Increase the generalizability of the study

Challenges of *multisite* studies in digital pathology

Big data management

- High-resolution images (100 K x 100 K)
- Large number of whole-slide images (hundreds...thousands…)
- Very large number of « objects » (e.g. cells,...) to detect, quantify, and store
- Distributed and poorly structured

Image variabilities require robust pattern recognition

- Tissue (healthy or not, age, ...)
- Preparation protocols (staining, ...)
- Acquisition setups (slide scanners, ...)

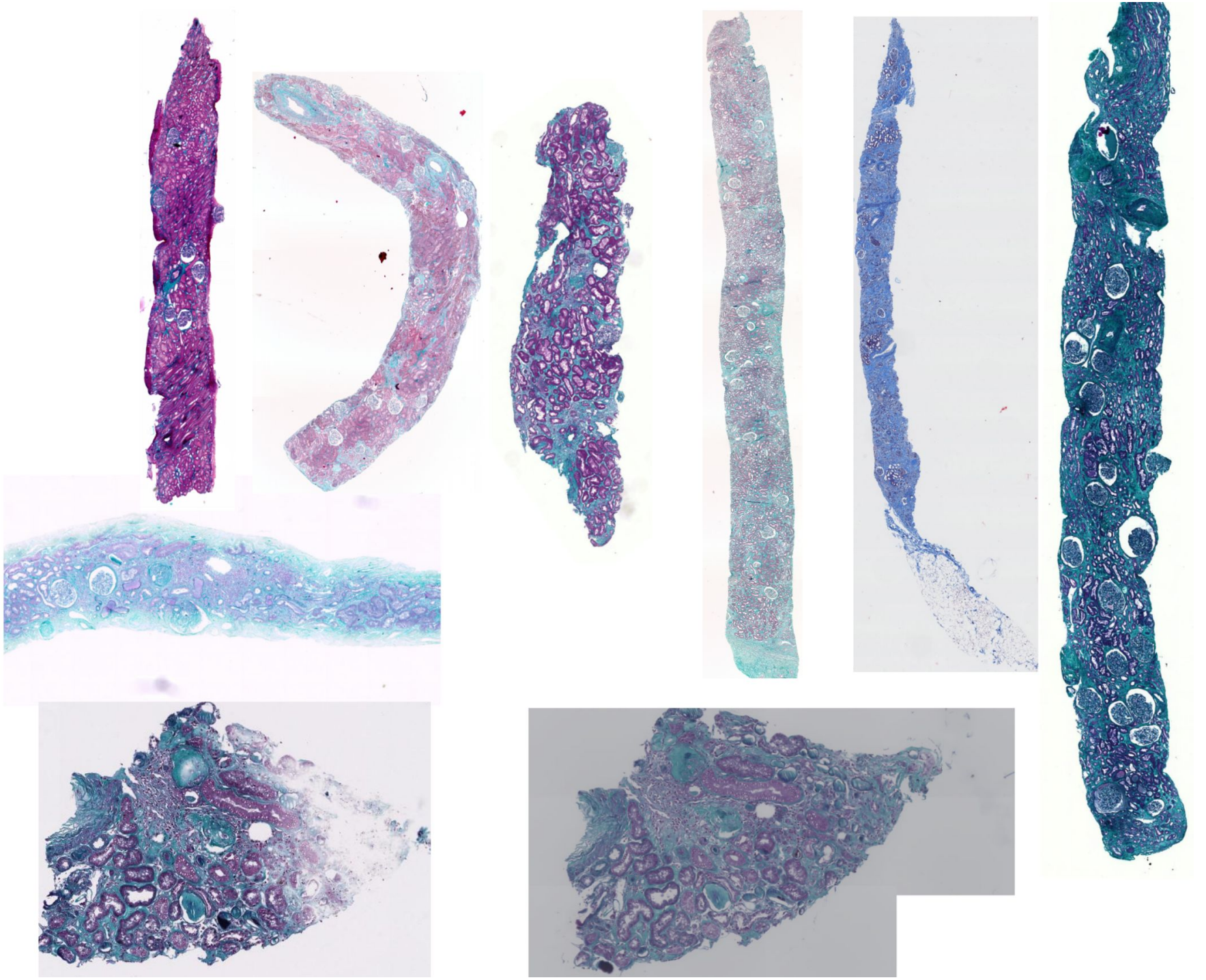
Collaborative software platforms:

- Groups of pathologists
  - Ground-truth creation
  - Validation of algorithms

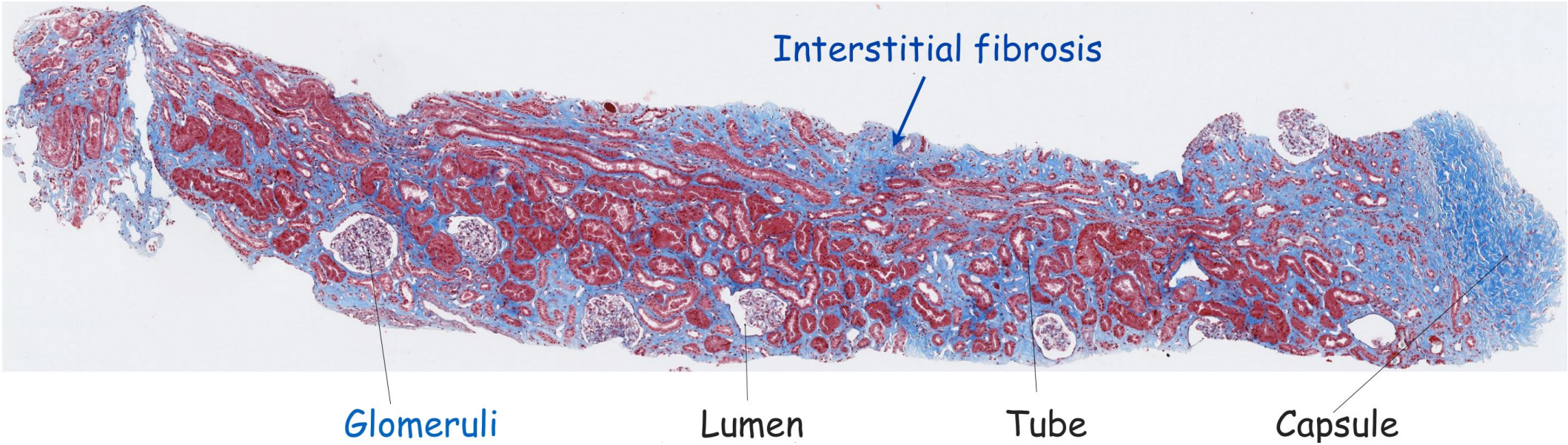
- Groups of image analysts
  - Image processing/analysis
  - Machine learning

...

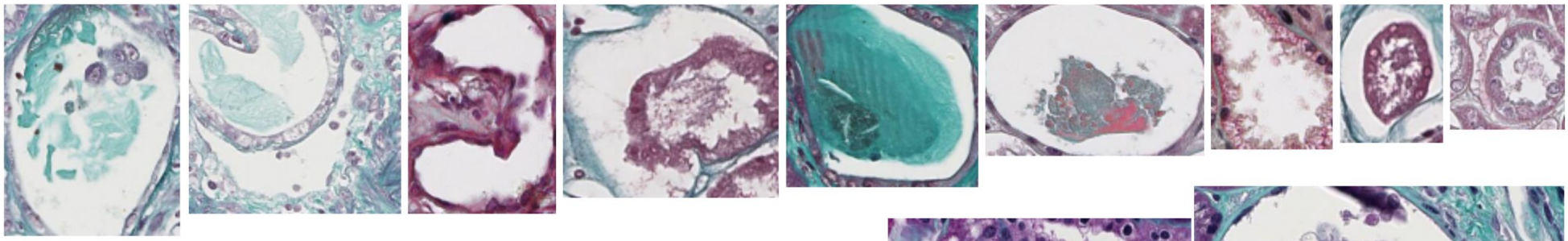
Challenges of glomeruli detection (1/2)



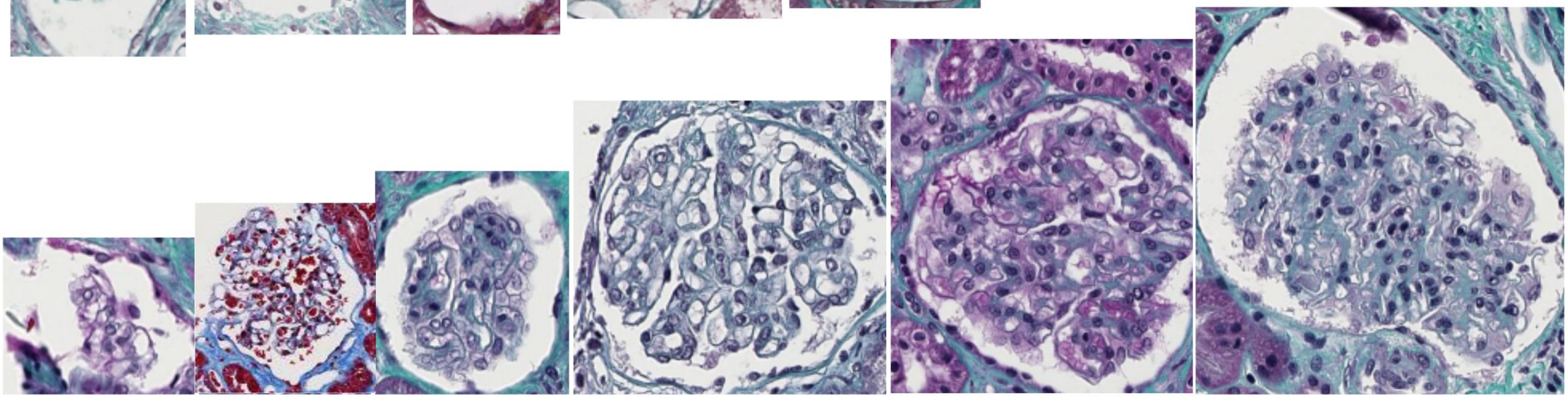
Challenges of glomeruli detection (2/2)



Non



Glom

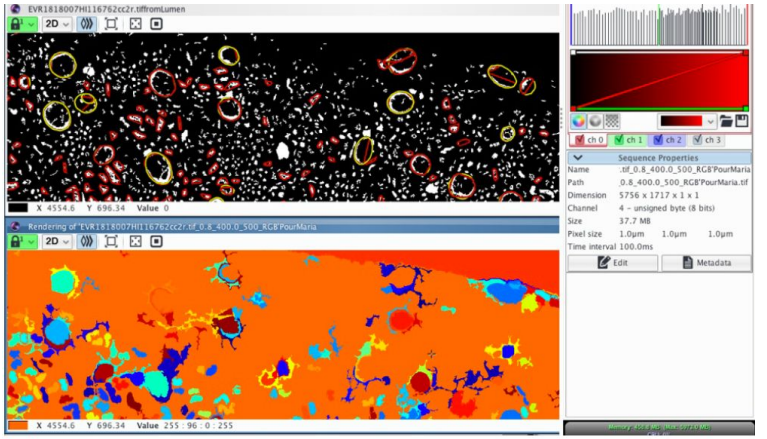


Our approach overview

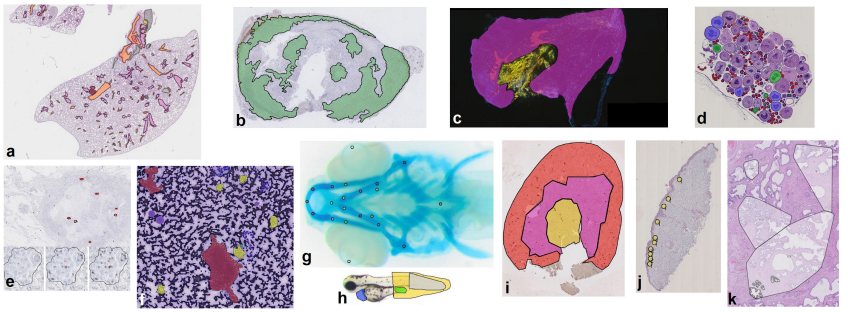
icy

(icy.bioimageanalysis.org ; Nature Methods, 2012) :

- Image analysis
  - Glomeruli candidate detection (Transplantation, 2011)
  - Color normalization/transfer



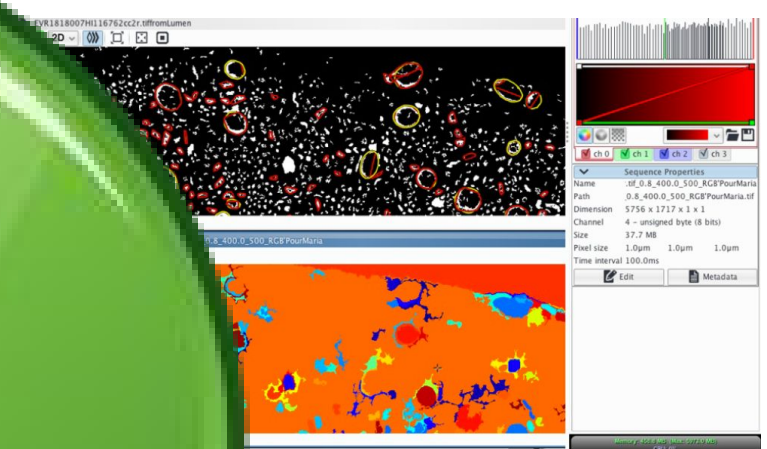
cytomine



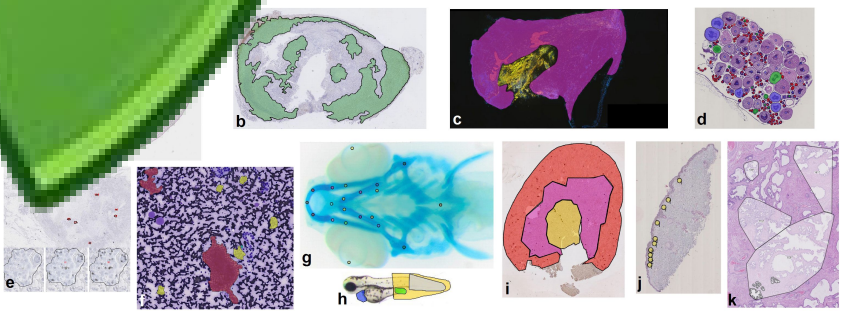
(www.cytomine.be ; Bioinformatics, 2016) :

- Database to organize whole-slide images (various formats) and semantic annotations
- A rich internet application : data sharing and collaborative modes (proofreading)
- Supervised machine learning algorithms for object recognition
- Open API and software templating mechanisms to add algorithms

Our approach overview



(icy.bioimageanalysis.org
- Image analysis
  - Glomeruli candidate splantation, z
  - Color normalization

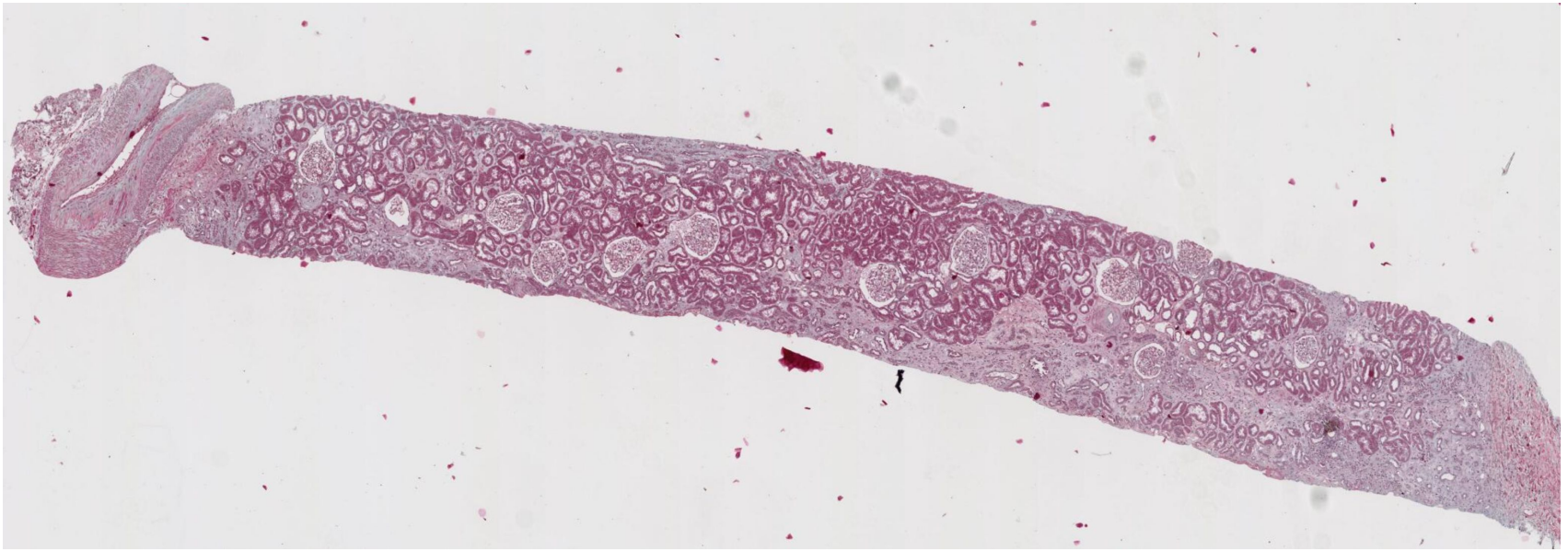


(www.cytomine.be ; Bioinformatics, 2016) :
- Database to organize whole-slide images (various formats) and semantic annotations
- A rich internet application : data sharing and collaborative modes (proofreading)
- Supervised machine learning algorithms for object recognition
- Open API and software templating mechanisms to add algorithms

Proposed approach :

1. Segmentation of candidates within whole slides using image analysis

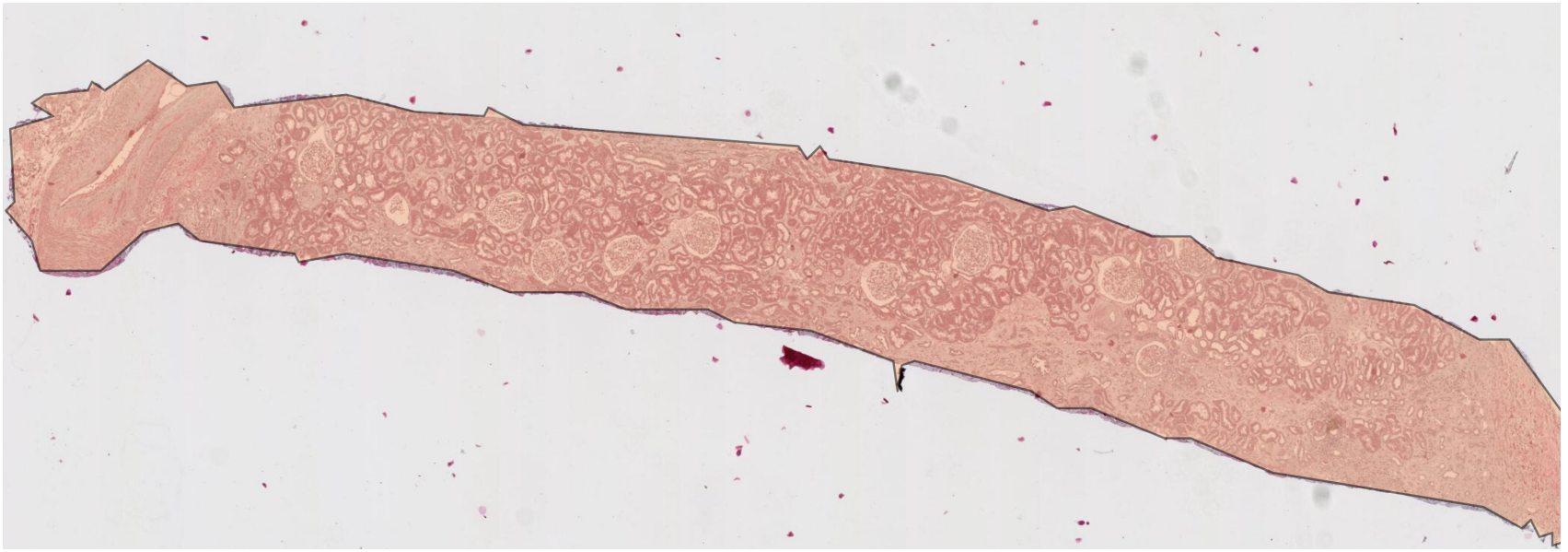
Section detection : Otsu thresholding

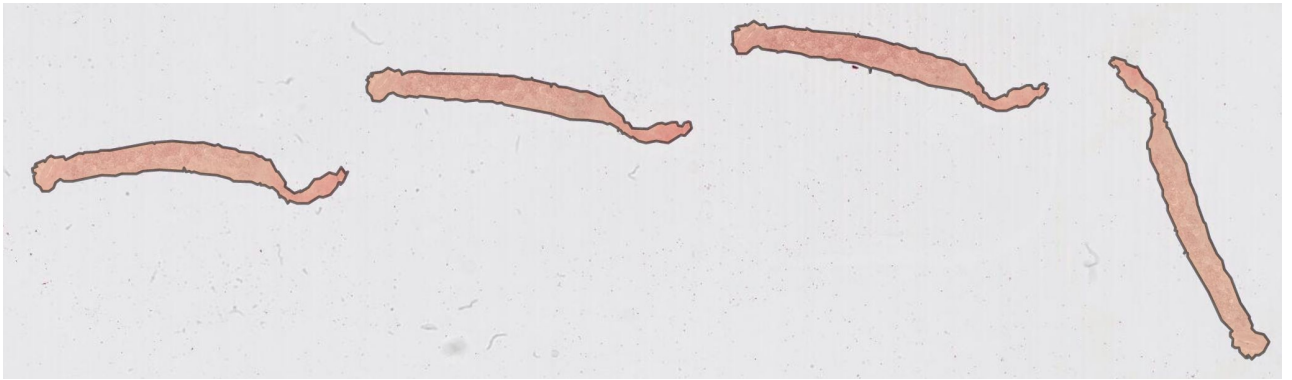


Proposed approach :

1. Segmentation of candidates within whole slides using image analysis

Section detection : Otsu thresholding

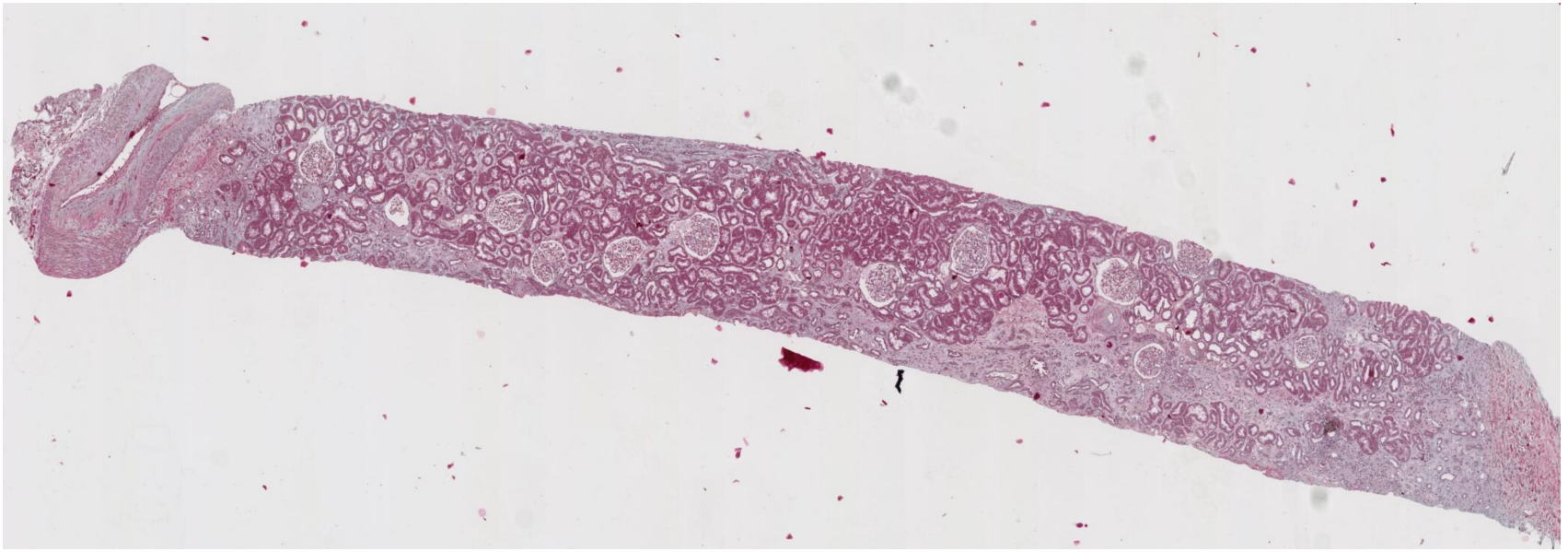




Proposed approach :

1. Segmentation of candidates within whole slides using image analysis

Candidate glomeruli detection : Lumen extraction (contour + convex Hull)
+ Ellipse detection (Fitzgibbon, 1999)
+ Superpixel segmentation (Felzenszwalb, 2004)

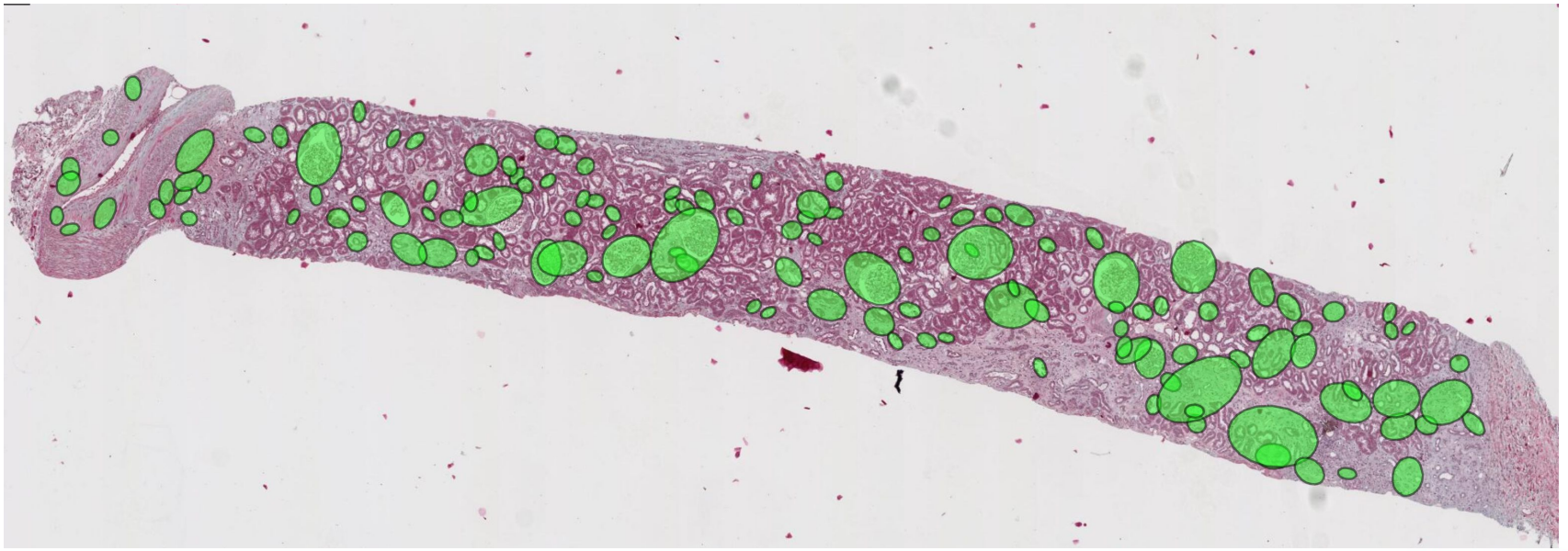


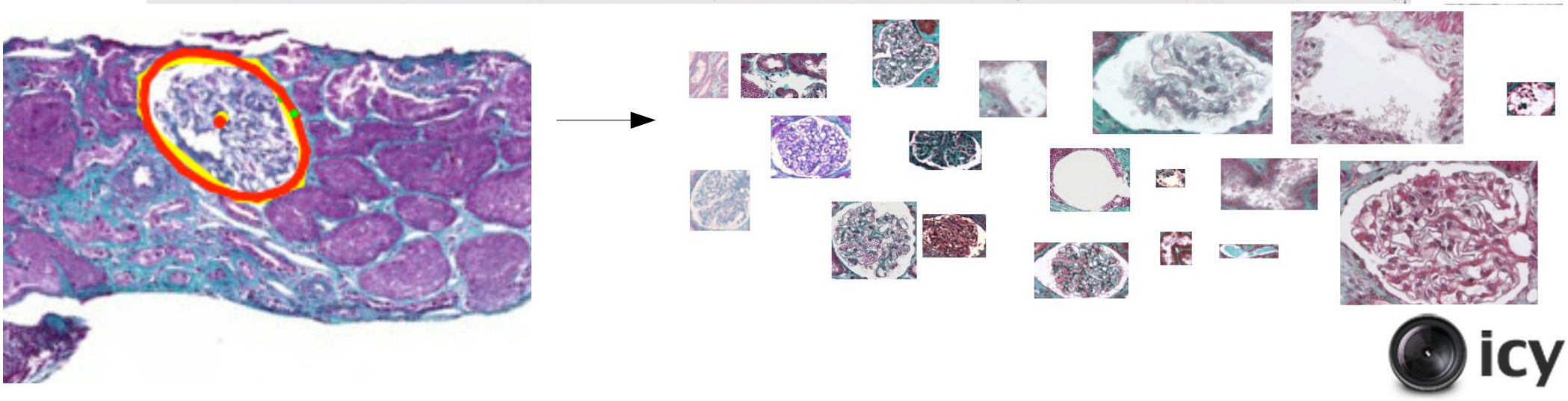
icy

Proposed approach :

1. Segmentation of candidates within whole slides using image analysis

Candidate glomeruli detection : Lumen extraction (contour + convex Hull)
+ Ellipse detection (Fitzgibbon, 1999)
+ Superpixel segmentation (Felzenszwalb, 2004)

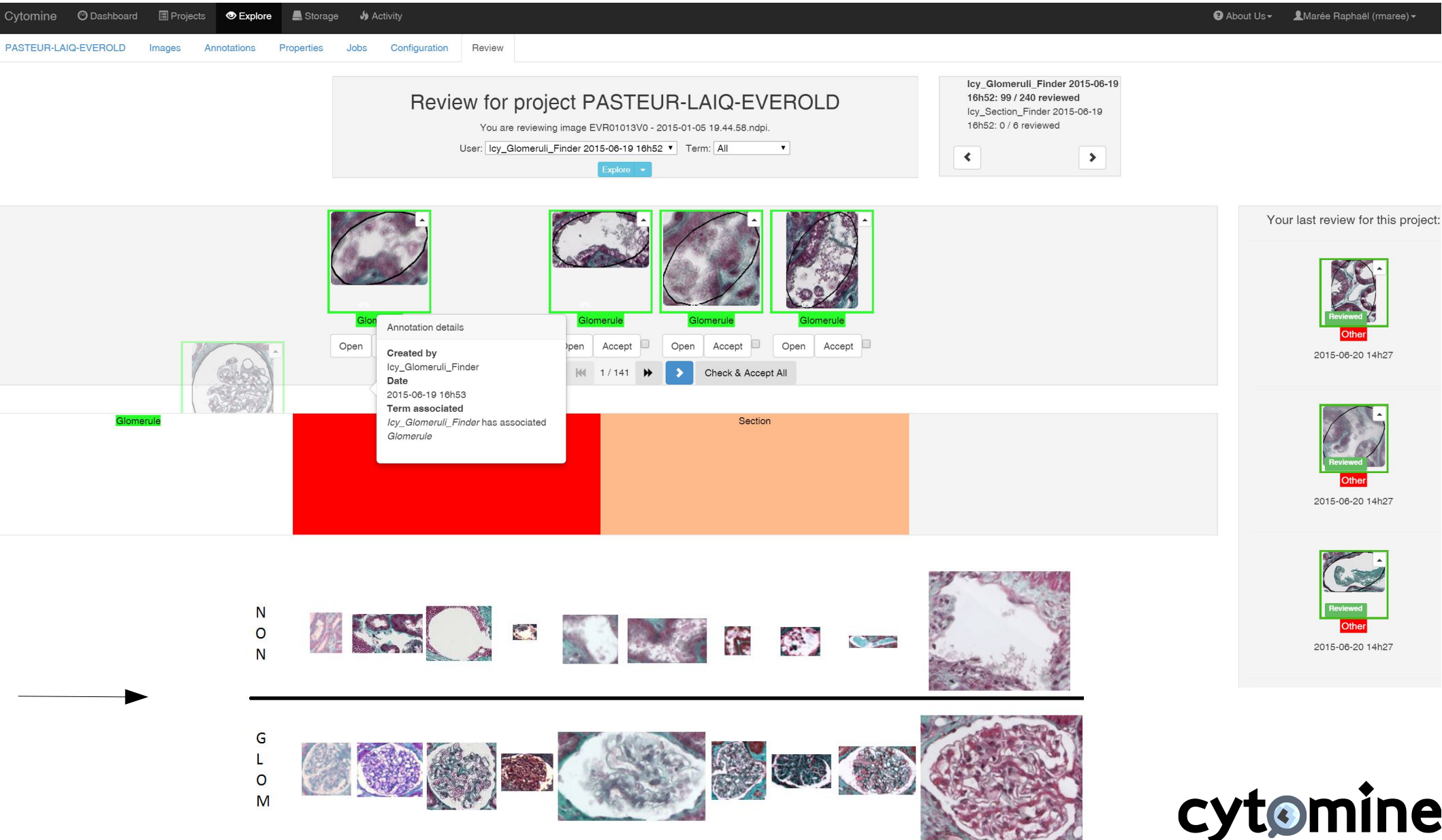




Proposed approach :

2. Classification of candidates using supervised algorithms

Semi-automated ground-truth creation: candidates generated by Step 1 are classified manually using Cytomine web proofreading UI → training set

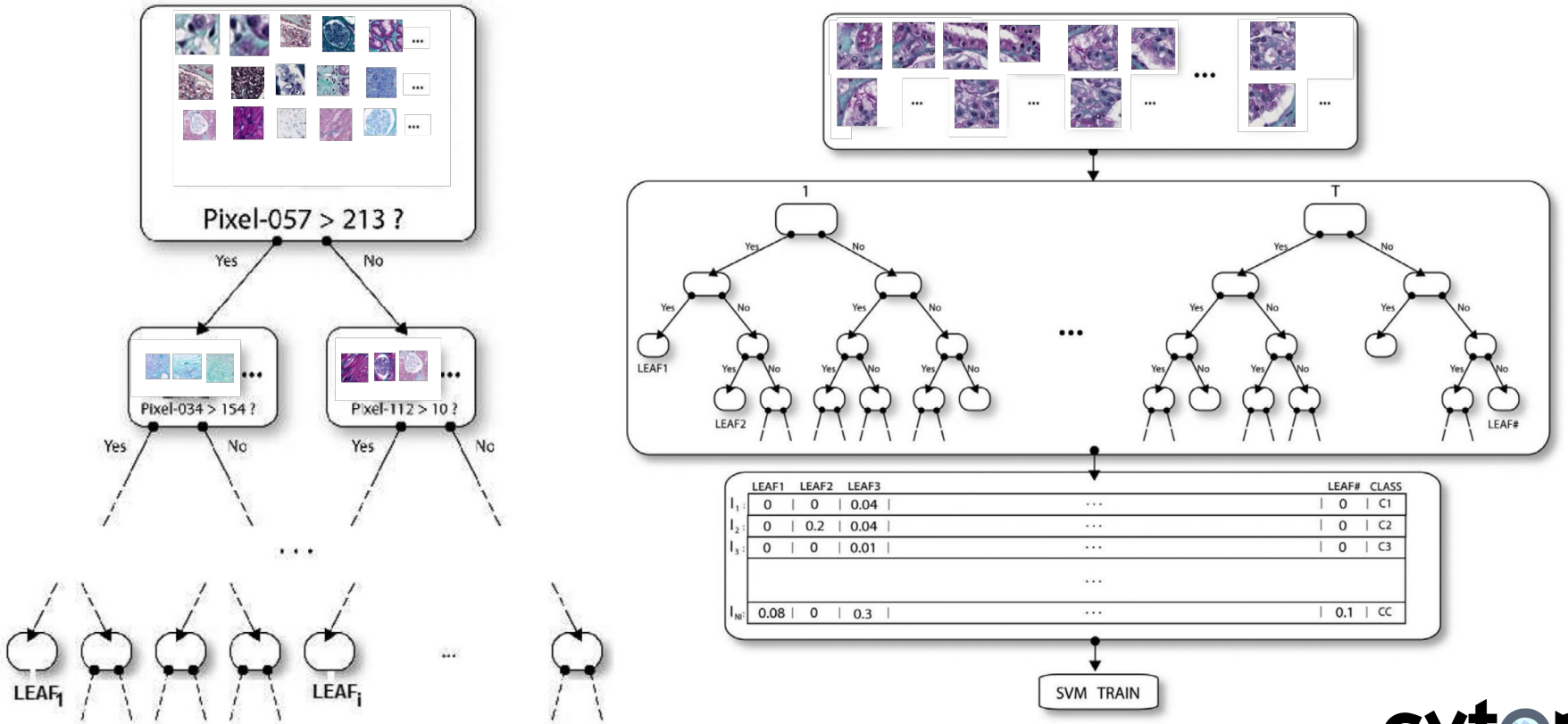


Proposed approach :

2. Classification of candidates using supervised algorithms

Generic image classification algorithm
(Marée et al., Pattern Recognition Letters 2016)**:**

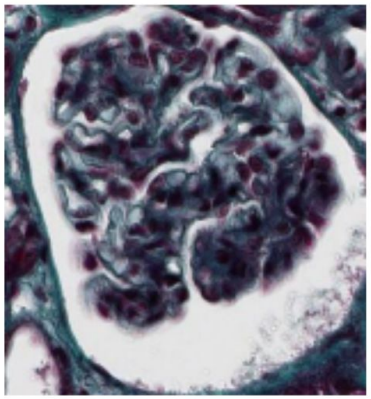
Random subwindows (patches) extracted in glom/non-glom images
Described by raw pixel values
Extremely Randomized Trees for feature learning
Final linear SVM classifier



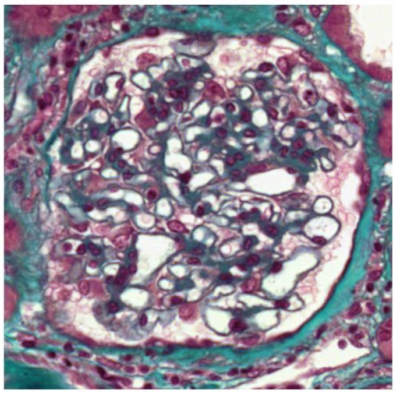
cytomine

Proposed approach :

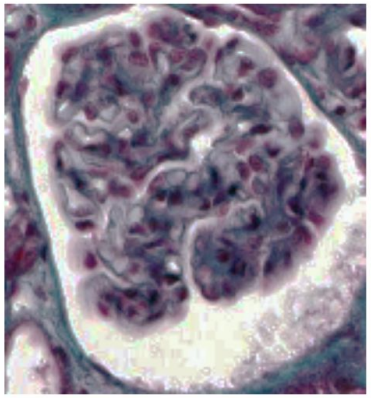
2. Classification … + color normalization



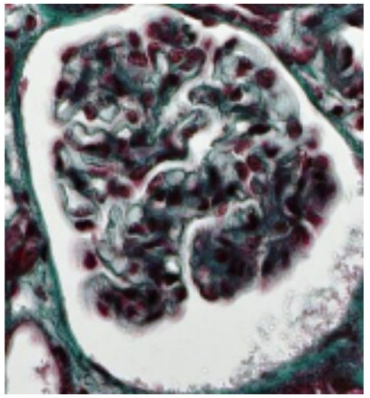
Source



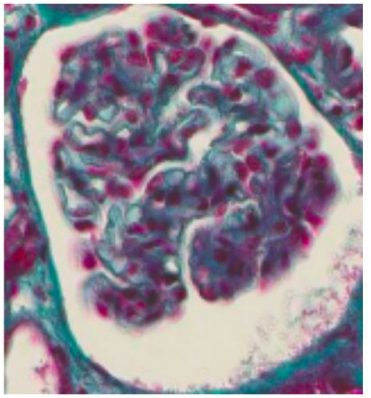
Target



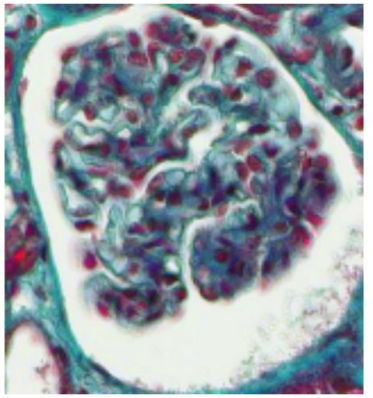
HS



ACE



RH



MM

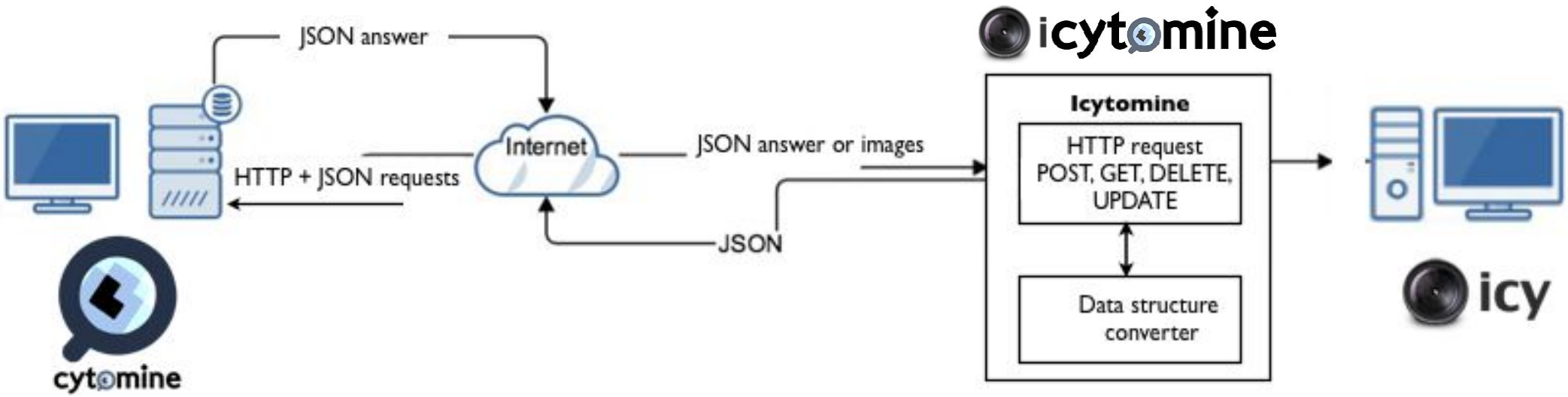
HS : Histogram equalization

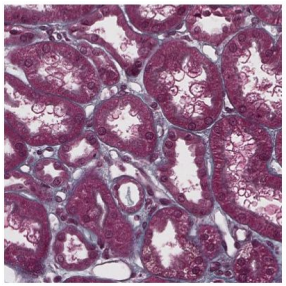
ACE : Automatic color enhancement (Getreuer, IPOL 2012)

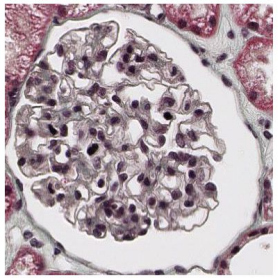
RH: Color transfer (per-pixel basis, equalization of μ and σ in the perceptual color space) (Reinhard et al., IEEE Computer Graphics and Applications 2011)

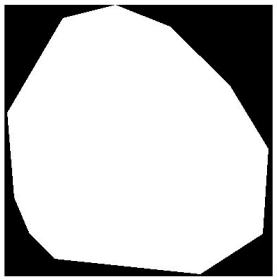
MM : SVD based approach to estimate color deconvolution matrices (Macenko et al., ISBI 2009)

Implementation based on web services







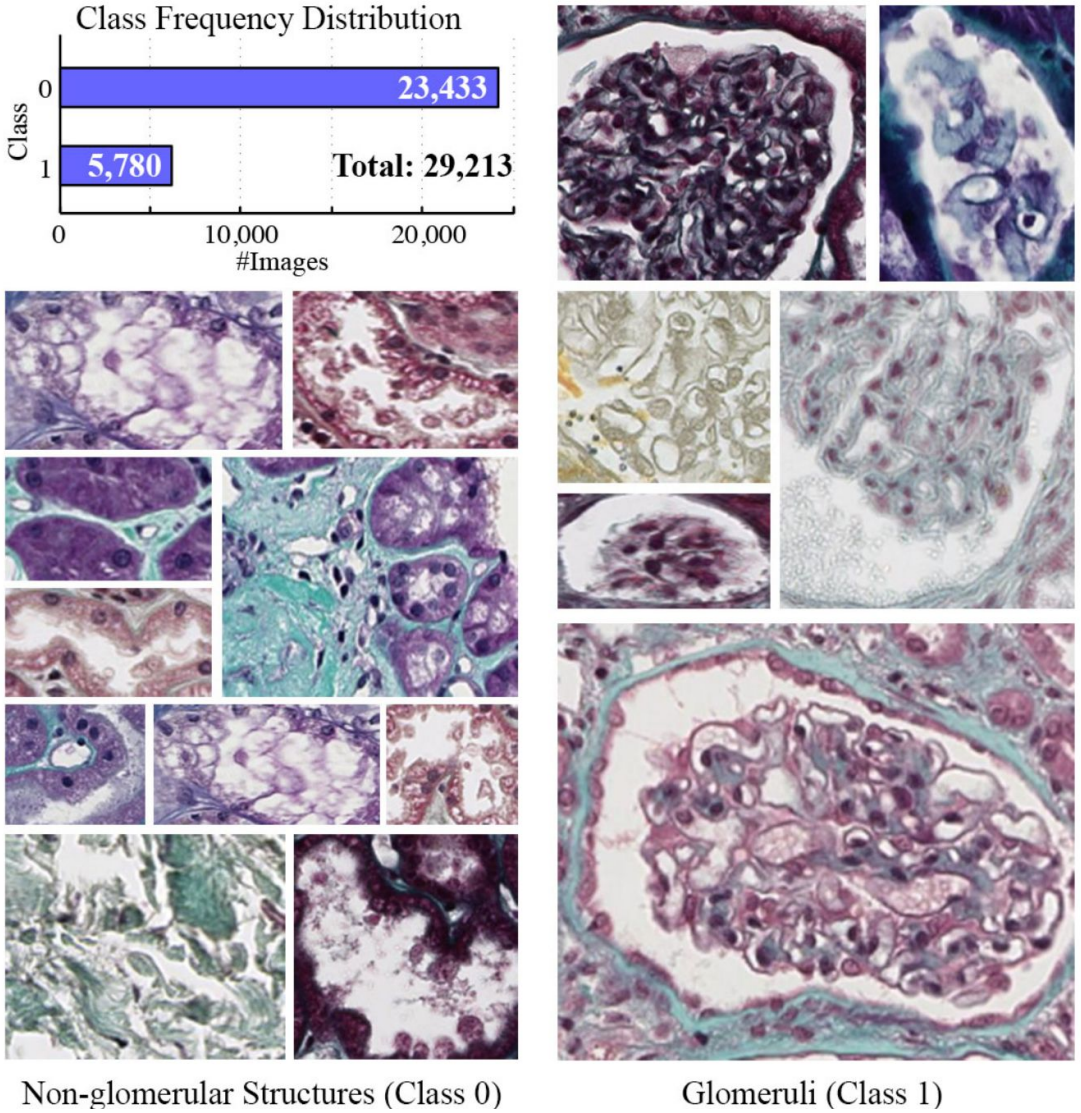


{"class":"be.cytomine.ontology.UserAnnotation","id":95189319,"created":"1390233213025","location":"POLYGON ((83409.75 96490.375, 83412.75 96492.375, 83412.75 96481.375, 83411.75 96476.375, 83408.75 96473.375 … 96489.375, 83409.75 96490.375))", "image":95062074,"project":94152784,"container":94152784, "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

Dataset and evaluation protocol

Dataset : 200 whole-slide images, 29213 annotated objects (ground-truth)



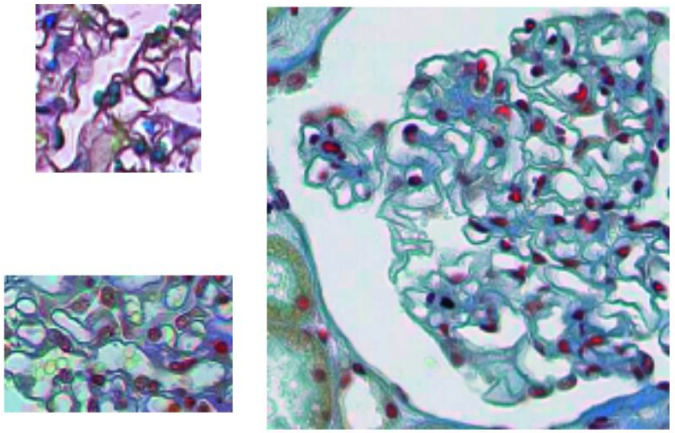
Training set: Objects from 100 slides (2927 glomeruli + 13648 non-glom)

Testing set: Objects from 100 unseen slides (2853 glomeruli + 9785 non-glom)

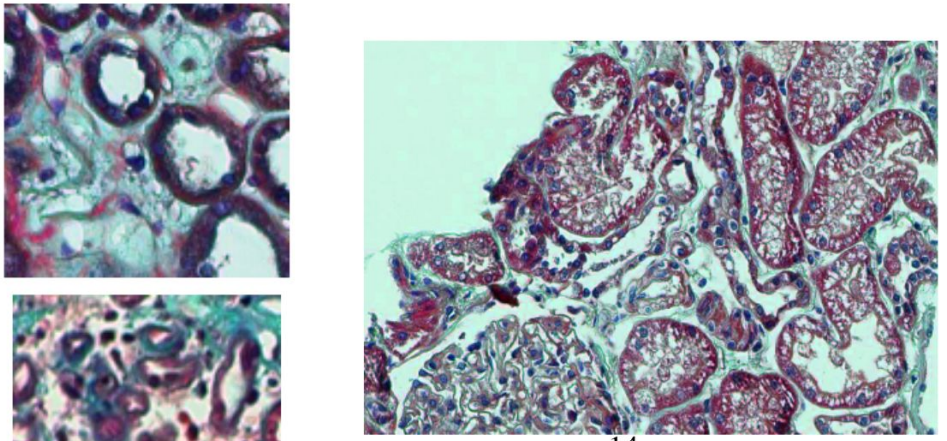
Glomeruli classification results

Method	Avg. acc.	NonGlom acc.	Glomeruli acc.
WND-CHARM GRAY	73.71	81.38	47.39
ET-FL GRAY	87.17	96.87	53.91
ET-FL HS	89.01	98.85	55.29
ET-FL COLOR	91.67	98.61	67.89
ET-FL ACE	91.04	98.71	64.73
ET-FL RH	92.62	98.95	70.93
ET-FL MM	**93.78**	**98.97**	**75.98**

Misclassification



NonGlom



Glom

Summary

- A novel approach combining Icy and Cytomine algorithms (open-source)

- It enables multisite and collaborative digital pathology

- Application for glomeruli detection in kidney human biopsies on a unprecedented dataset
  - > 90 % object recognition rate but...
  - Current results are not yet satisfactory for fully automated routine use

- Ongoing and future work :
- Improving results
  - Other color normalization techniques
  - Data augmentation and balancing
  - Other classification approaches (convnets...)
  - Consider other evaluation/optimization criteria more related to the final application

- Comprehensive analysis
  - Recognition of other kidney tissue components and their spatial relations

Acknowledgments

- Unité d'Analyse d'Images Biologiques (Pasteur) :
Osvaldo Fuica, Fabrice de Chaumont, Alexandre Dufour.

- Systems and Modeling (Montefiore Institute / GIGA-Research, University of Liège):
Renaud Hoyoux, Loïc Rollus, Gilles Louppe, Pierre Geurts, Louis Wehenkel

- Nephrologists and Pathologists : Eric Thervet (HEGP), Laure-Hélène Noël (Necker), Emmanuel Morelon (Hospices Civils de Lyon)

Institut Pasteur Carnot Maladies Infectieuses
Research grants of the Wallonia (DGO6) :
CYTOMINE (2010-2016) n° 1017072
HISTOWEB (2014-2017) n° 1318185

---

icy.bioimageanalysis.org
www.cytomine.be

INSTITUT CARNOT Pasteur MI

Institut Pasteur

MONTEFIORE INSTITUTE
Department of Electrical Engineering and Computer Science

FEDER

Wallonie