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
  HTML  CSS  GRAILS  jQuery  OpenLayers™  Backbone.js  Hibernate  OpenSlide  PostGIS
  (+/- 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, 3DHistech MRXS, Leica SCN, Roche BIF...)
- Users with **authentication** (e.g. LDAP), access rights, and roles
- **Specific ontologies** with user-defined, vocabulary terms
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 flou channels 83000 x 100 000 pixels = 4 x 16GB image
Software features:

- **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

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

- **Share images** through simple URLs

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

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

  http://beta.cytomine.be/#share-annotation/92024416
Software features: Live broadcast
- 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

- 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)


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

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?
Tens or hundreds of glass slides to be quantified per study...
Hybrid human-computer workflow

Treated with X

40K x 30K pixels

To appear in Proc. IEEE ISBI 2014
Hybrid human-computer workflow

1. Manual region contouring and labelling to provide training examples
Hybrid human-computer workflow

2. Automatic training of image recognition model based on training examples
Hybrid human-computer workflow

3. Batch processing of slides
   Reviewing slides

<table>
<thead>
<tr>
<th>ID</th>
<th>Preview</th>
<th>Name</th>
<th>Width</th>
<th>Height</th>
<th>X</th>
<th>Resolution</th>
<th>User an.</th>
<th>Algo an.</th>
<th>Valid an.</th>
<th>Mime</th>
<th>Created</th>
<th>Status</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>7889452</td>
<td></td>
<td>PGP POU MON PBS 3 21 - 2012-08-08 11.07.24.jp2</td>
<td>34816</td>
<td>29440</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>8</td>
<td>614</td>
<td>100</td>
<td>jpg2</td>
<td>2012-10-18 10h50</td>
<td>Reviewed</td>
<td>Explore</td>
</tr>
<tr>
<td>7889358</td>
<td></td>
<td>PGP POU MON PGF7 11 - 2012-08-07 12.07.00.jp2</td>
<td>36864</td>
<td>28416</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>2</td>
<td>130</td>
<td>119</td>
<td>jpg2</td>
<td>2012-10-18 10h50</td>
<td>Reviewed</td>
<td>Explore</td>
</tr>
<tr>
<td>7889229</td>
<td></td>
<td>PGP POU MON PGF7 41 - 2012-08-07 12.13.50.jp2</td>
<td>36864</td>
<td>28416</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>0</td>
<td>142</td>
<td>142</td>
<td>jpg2</td>
<td>2012-10-18 10h50</td>
<td>Reviewed</td>
<td>Explore</td>
</tr>
<tr>
<td>7889190</td>
<td></td>
<td>PGP POU MON PGF7 21 - 2012-08-07 12.09.19.jp2</td>
<td>34816</td>
<td>28928</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>5</td>
<td>122</td>
<td>124</td>
<td>jpg2</td>
<td>2012-10-18 10h50</td>
<td>Reviewed</td>
<td>Explore</td>
</tr>
<tr>
<td>7889180</td>
<td></td>
<td>PGP POU MON PGF7 31 - 2012-08-07 12.11.34.jp2</td>
<td>36864</td>
<td>26624</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>7</td>
<td>271</td>
<td>213</td>
<td>jpg2</td>
<td>2012-10-18 10h49</td>
<td>Reviewed</td>
<td>Explore</td>
</tr>
<tr>
<td>7889170</td>
<td></td>
<td>PGP POU MON PGF8 11 - 2012-08-07 12.18.57.jp2</td>
<td>38912</td>
<td>32256</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>3</td>
<td>169</td>
<td>151</td>
<td>jpg2</td>
<td>2012-10-18 10h49</td>
<td>Reviewed</td>
<td>Explore</td>
</tr>
<tr>
<td>7889160</td>
<td></td>
<td>PGP POU MON PGF8 31 - 2012-08-07 12.24.39.jp2</td>
<td>40960</td>
<td>32256</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>3</td>
<td>158</td>
<td>145</td>
<td>jpg2</td>
<td>2012-10-18 10h49</td>
<td>In review</td>
<td>Explore</td>
</tr>
<tr>
<td>7889150</td>
<td></td>
<td>PGP POU MON PGF8 41 - 2012-08-07 12.27.33.jp2</td>
<td>40960</td>
<td>33024</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>2</td>
<td>846</td>
<td>782</td>
<td>jpg2</td>
<td>2012-10-18 10h49</td>
<td>In review</td>
<td>Explore</td>
</tr>
<tr>
<td>7889140</td>
<td></td>
<td>PGP POU MON PGF8 21 - 2012-08-07 12.21.46.jp2</td>
<td>40960</td>
<td>32256</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>8</td>
<td>161</td>
<td>146</td>
<td>jpg2</td>
<td>2012-10-18 10h48</td>
<td>In review</td>
<td>Explore</td>
</tr>
<tr>
<td>7889130</td>
<td></td>
<td>PGP POU MON PGP10 41 - 2012-08-07 12.55.37.jp2</td>
<td>36864</td>
<td>28928</td>
<td>40</td>
<td>0.650µm/pixel</td>
<td>0</td>
<td>175</td>
<td>175</td>
<td>jpg2</td>
<td>2012-10-18 10h48</td>
<td>In review</td>
<td>Explore</td>
</tr>
</tbody>
</table>
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
Hybrid human-computer workflow

4. Export statistics

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

Roles of polarized neutrophils on lung tumour development in an orthotopic lung tumour mouse model
Rocks et al., European Respiratory Society Annual Congress, 2013
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

To appear in Proc. IEEE ISBI 2014

Quantification (mouse clicks)

- Cytomine: 100
- LocalFloodFill: 200

Quantification (mouse wheels)

- Cytomine: 150
- LocalFloodFill: 150

Quantification (mouse distance)

- Cytomine: 30 meters
- LocalFloodFill: 30 meters

Quantification (keyboard)

- Cytomine: 30 keys
- LocalFloodFill: 40 keys

Quantification (time)

- Cytomine: 200 seconds
- LocalFloodFill: 400 seconds

(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)
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)
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


- 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
