# Computational pathology: an increasingly growing field that requires training and interaction between pathologists and computer scientists


M. Tecilla[1], F. Romero-Palomo[1], C. Gámez Serna[2], F. Arcadu[2], Y. Cohen[3], V. Schumacher[1]

[1] Roche Pharma Research and Early Development (pRED), Pharmaceutical Sciences, BIOmics and Pathology - Roche Innovation Center Basel, Switzerland
[2] Roche Pharma Research and Early Development (pRED), Early Development Informatics - Roche Innovation Center Basel, Switzerland
[3] Roche Pharma Research and Early Development (pRED), Safety Informatics - Roche Innovation Center Basel, Switzerland


## QR-codes

Download

Contacts


## 1. Abstract

**Background.** Pathology is switching from an "analog" to a "digital" era. Digitalization is not limited to slide scanning; artificial intelligence (AI)-based tools and computer-aided diagnoses are increasingly used. In this context, pathologists will benefit from adjustments to training, enabling them to understand and use these technologies properly. Here we report on our preliminary implementation of a computational pathology (CP) training program at Roche. **Methods.** Veterinary Pathology (VP) and Computer Vision (CV) specialists worked together to improve joint knowledge in these fields. The work was divided into: 1) Review available resources, 2) Identify and outline the scientific questions to be addressed with AI, and 3) "Hands-on" assessment of the acquired knowledge by means of close VP/CV interaction in algorithm development or testing of commercially available software. **Results.** The review process led to the creation of a digital interactive AI digital glossary, with videos and links to external resources. During the "hands-on" experience with use case examples, python programming skills or in-depth knowledge of commercially available AI tools were acquired. **Conclusions.** The glossary was a useful resource for the pathology and informatics groups that will be extended to cover Unix and relevant commercial software. In our experience, Unix knowledge is a must-have in CP algorithm development. Python knowledge improves pathologist/computer scientist communication and helps in developing complex tools; nevertheless, commercial tools may offset programming basic daily-based tasks.


## 2. Background

Digital-pathology is the natural evolution of telepathology and started at the end of the 1990's by introducing the first Whole Slide Scanners (WSSs) on the market. Since then, technology has improved on to modern days with high-performance computers (HPCs) allowing computerized analysis on entire Whole Slide Images (WSI) and enabling the new era of computational pathology. Although changes have occurred over the past 20 years, in the early times the limited numbers of WSIs and HPCs have limited the wide adoption of these technologies. Fortunately, this is no longer the case and the technology and know-how is becoming accessible to a broader audience, moving pathology progressively from an "analog" to a "digital" era. Pathologists face a radical change in this context, moving from light microscopes to digital microscopes, WSS, and computer-aided diagnosis. This technological shift represents a double challenge for pathologists. First, the technological gap frequently requires changing mindset and routines acquired over years of training and practice. Second, pathologists' knowledge and abilities are challenged by new tools, perceived as a potential replacement for pathologists' work. For these reasons, pathologists need to understand and embrace this evolutionary change and will need the training to consider these technological change for future generations. This poster will present how we envisioned the preliminary implementation of computational pathology training in Roche.

## 3. Materials and Methods

A team of three pathologists and three computer vision scientists was created to try to define a general *modus operandi* to start with.

### Review of Digital Pathology AI resources

- Review last 6 years peer-reviewed literature
- Identify additional audio-video resources

### Outline questions to address

- Establish which pathology topics can be investigated using machine learning (ML) and AI

### «Hands-on» experience

Assess the acquired knowledge by:

- Close VP/CV interaction in algorithm development
- Testing of current commercially available software

## 4. Results

### Review of AI online resources

**Literature:** Among all the papers screened, a total of 34 (the complete list is available here) ranging from 2014 to 2020 were selected. The resulting literature review was gathered in a spread-sheet. The topics covered in these papers were considered as backbone for the glossary construction.

**Additional resources:** Additional resources were selected from YouTube, Online blogs, AI companies, University websites and Roche internal learning platform. All the resources were collected in an online interactive glossary shared within Roche.

**Glossary:** The glossary was created as a live document with the idea to expand or edit the topics included based on the feedbacks and literature updates. To make this tool available to all Roche's employees, the document was created using Google documents. Glossary layout is depicted in **Fig. 1** and topics are included in the **Table 1**.


**Fig. 1** Glossary general layout

| Table 1: Index of relevant topics included in the glossary: |
|---|
| ***AI/ML/DL resources – General:*** |
| - What is artificial intelligence? Machine learning? Deep learning? |
| - Unsupervised vs Supervised learning |
| - How classifiers, CNNs work? |
| - What is overfitting and how to avoid it? |
| - What metrics to use to monitor the training process and how to interpret them? |
| ***AI/ML/DL resources - Pathology-oriented:*** |
| - Webinars, seminars, presentations of deep learning applied to digital pathology |
| ***Glossary:*** |
| - List of definitions on the topics of pathology, computational pathology, and machine learning. |
| ***Additional Links:*** |
| - Links to software and tools for digital pathology. |

### "Hands-on": Python programming

The pathologist with basic Python knowledge was asked to learn how to mask WSIs in a trial and error fashion. In order to solve the task the pathologist needed to: learn how to interact with the local HPC, open and process WSIs, and work with images to create and apply the final mask. Although the local HPC implements a primary, web-based user interface to interact with, basic knowledge of bash commands was crucial to execute different basic operations with files and folders, and install Python libraries. Code was written and executed using Jupyter notebook, installed on the local HPC. To open and interact with WSIs, OpenSlide, and Large-image python libraries were used. Large-image library ended up being incredibly user friendly, allowing to integrate basic routines from OpenSlide with some advanced features like tiles extraction. For image processing, CV2, PIL, and Scikit-image libraries were used. Os library was used to interact with the operating system, Numpy for arrays operation, and Matplotlib for displaying images. The pathologist assigned 5% of his working time to learn and practice for this task. Five months were required to obtain preliminary results (**Fig. 2**), but the learning experience was useful to improve interaction with CVs.

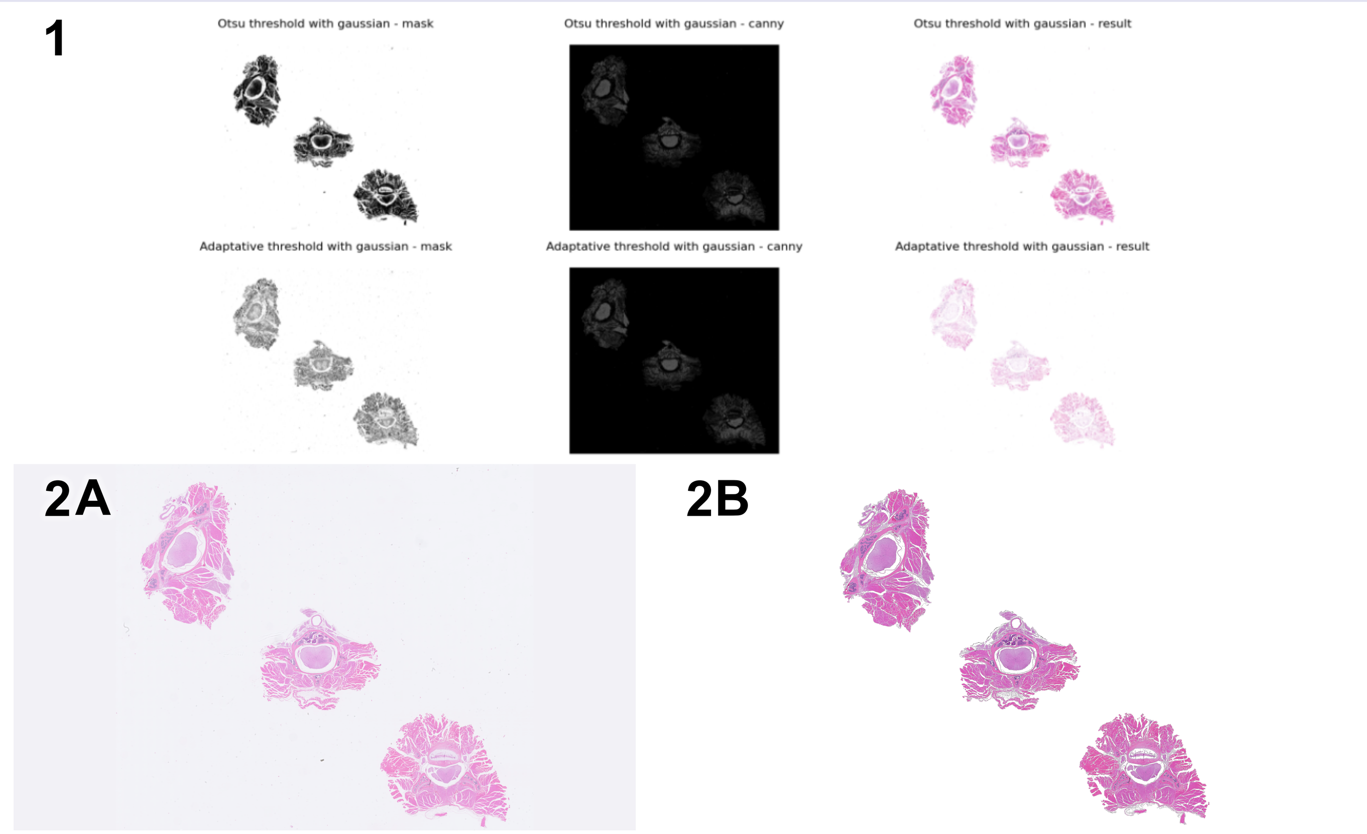


**Fig. 2** 1) Examples of masking using different type of thresholding. 2) A: Image before masking, B: Resulting pictures once background has been removed, using the created mask.

### "Hands-on": Commercial AI software

One of the pathologists learned how to use a commercial software for image analysis and AI. Consequently, he developed a tool for beta amyloid plaques detection in H&E stained slides from a mouse model of Alzheimer's disease. In this approach, β-amyloid (βAMY)-stained slides from FFPE sections were used as ground truth to support the annotation of amyloid plaques in H&E stained consecutive sections. The tissue samples were subdivided into groups for training, testing and validating the algorithm. The following steps summarize the workflow done (**Fig. 3**) which required about 3 months at 5% working time:

1. Registration of WSIs aligning βAMY and H&E stains.
2. Train Random forest classifier for automatic βAMY tissue detection and annotations.
3. Export of annotations from βAMY-IHC into H&E and manual adaptation/correction to the contour of the amyloid plaque.
4. Train a CNN (DenseNet) to identify β amyloid plaques on H&E sections.

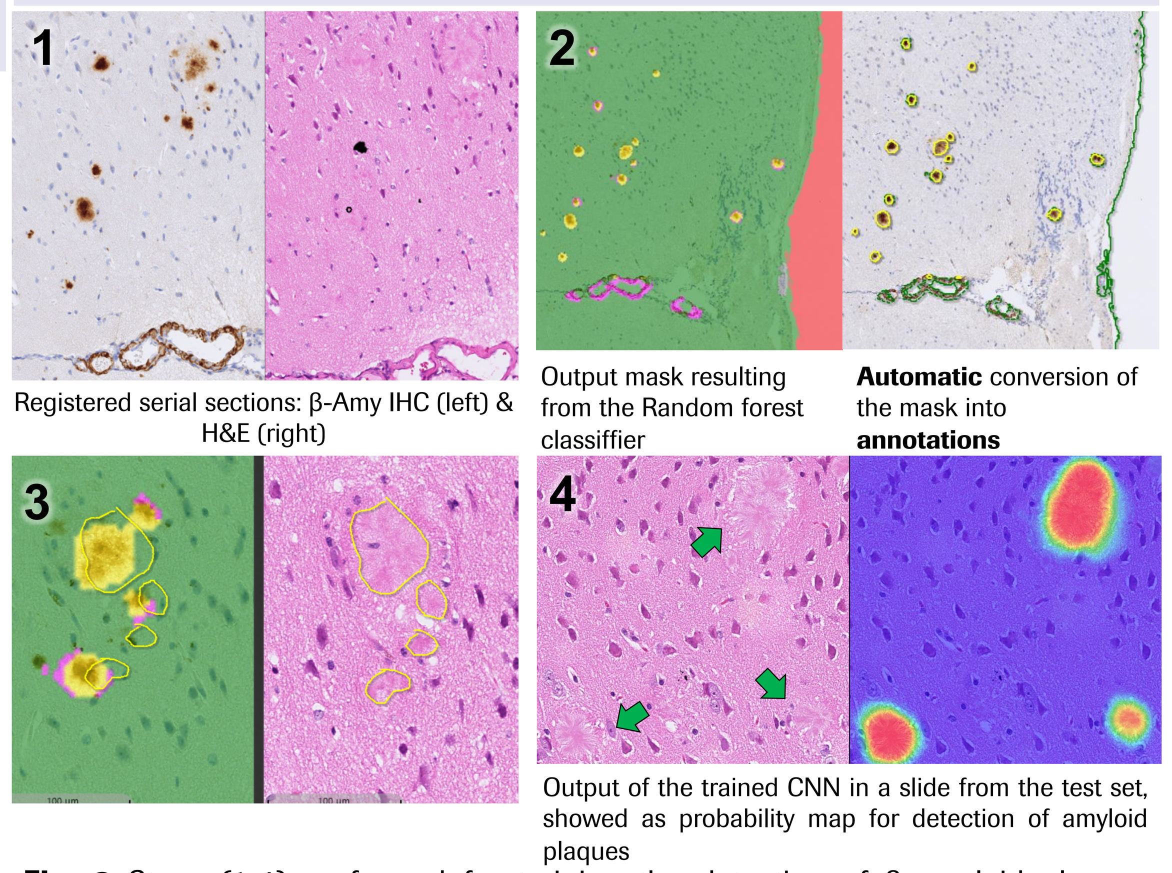


**Fig. 3** Steps (1-4) performed for training the detection of β amyloid plaques algorithm from mouse Alzheimer diseases using a commercial Image analysis software with deployment of CNNs.

## 5. Conclusions

The glossary was a useful resource for the pathology and informatics groups that will be extended to cover Unix and relevant commercial software. In our experience, good knowledge of topics included in the General section of the Glossary (AI/ML/DL resources – General) is essential to allow CP and VP to acquire a common language, and improve collaboration. Unix knowledge is a must-have in CP algorithm development to appropriately interact with Linux based servers. Python knowledge CP/VP communication and helps in developing complex tools. Nevertheless, commercial tools may offset basic programming for daily-based tasks..

*Doing now what patients need next*