FGAI4H-D-044

Shanghai, 2-5 April 2019

Source: TG-Histo driver

Title: TG-Histo (AI for histopathological diagnostics) update

Purpose: Discussion

Contact: Frederick Klauschen E-mail: frederick.klauschen@charite.de
Charite Berlin, Germany

AI for histopathological diagnostics

Machine learning-based profiling of tumor-infiltrating lymphocytes in breast cancer

Shanghai (via remote) 3. 4. 2019, FGAI4H-C-018

ITU-T Focus Group on AI for Health
Frederick Klauschen
Charité Berlin

Histological slide

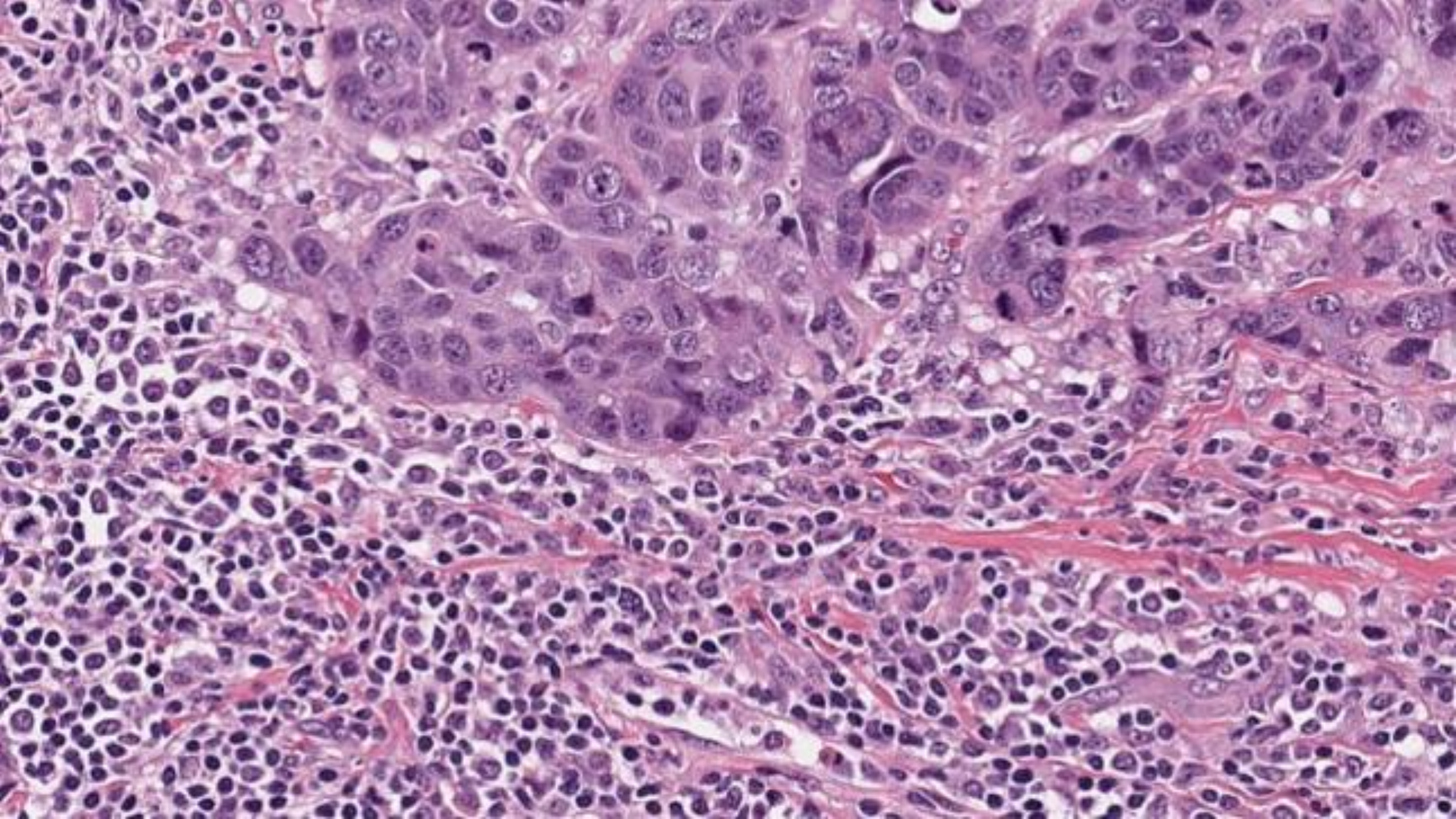
Microscopic diagnostics

Manual evaluation

Example:
Cancer diagnostics for Immuno-Oncology

Cancer present?

Quantify immune cells!



Identify
Cancer!

Quantify
Immune cells!

ML-based TiL profiling in breast cancer

- The use case was introduced in document B-014 that contains background information.

- The present document FGAI4H-C-018 explains how
  - the histopathology images are annotated
  - data would be provided
  - machine learning models can be benchmarked

2 Annotation of the histopathology images: What?

Specifications:

- Digitized standard Hematoxylin & Eosin (H&E) stained histological slides
- List of the tissue components/classes:

cancer tissue
- multiple subtypes
  - focus on NST (no-special-type) and invasive-lobular breast cancer

normal tissue
- normal breast gland and duct epithelium
- connective tissue (fibers, cells)
- fatty tissue
- bone tissue
- blood and lymphatic vessels
- nerves

immune system
- lymphocytes
- granulocytes
- monocytes/macrophages
- plasma cells

necrotic tissue

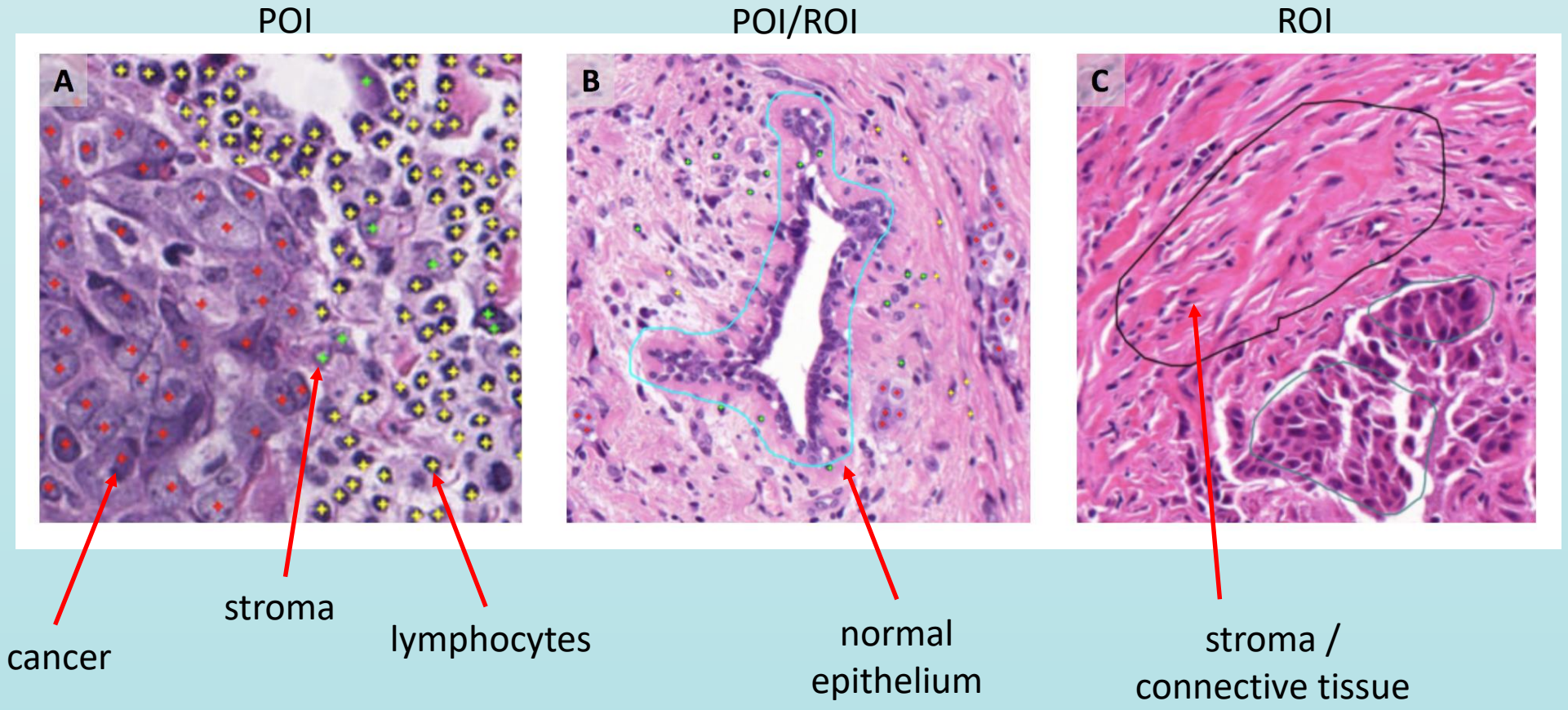
artifacts

background

2 Annotation of the histopathology images: How?

- Annotations should be flexibly reusable with different patch sizes extractable from annotation coordinates (saved in xml-format)
- Annotation procedure (single cell “point” vs. area “region” annotation)
  - positive annotations
    - point annotations (POI): cell nuclei are marked, relevant for heterogeneous tissues (e. g. individual lymphocytes between cancer cells)
    - region annotations (ROI): regions containing at least 95% cells of respective class
  - negative annotations
    - region annotations (ROI): regions negative of a certain class, i. e. region may contain any cells, but none of the respective

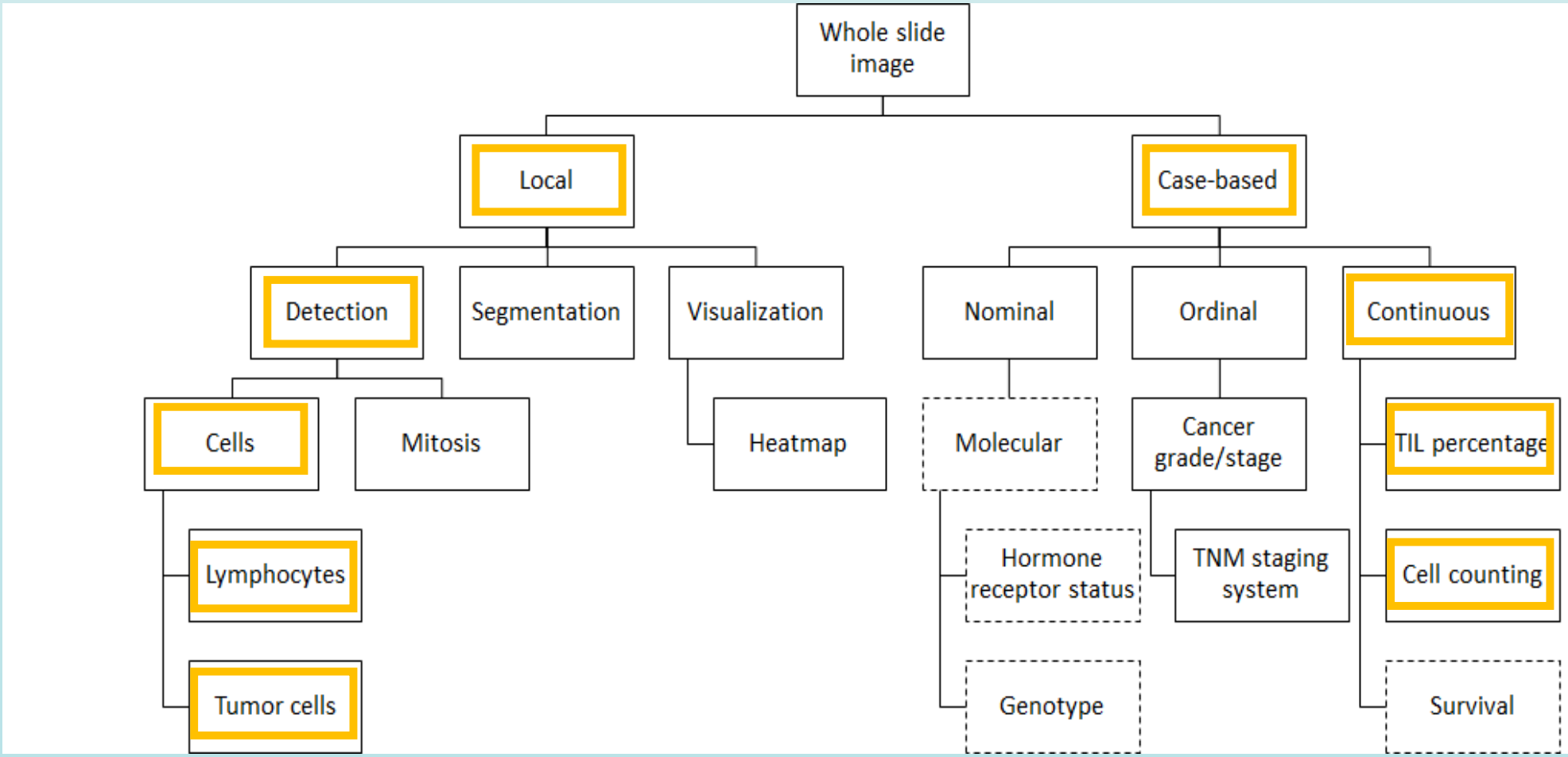
2 Annotation of the histopathology images

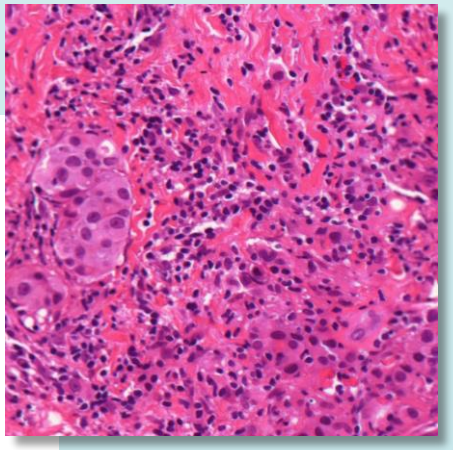


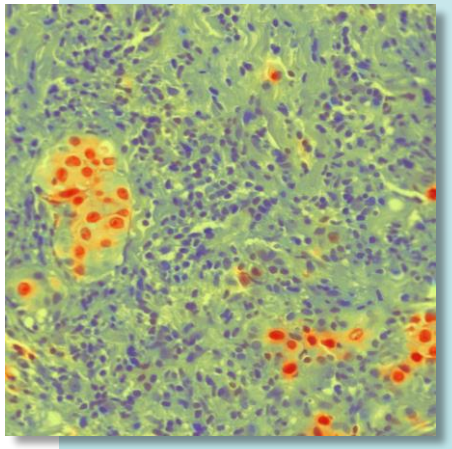
2 Provision of test and benchmarking images

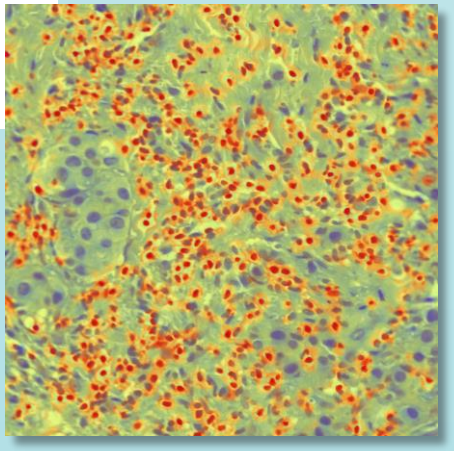
- Description of data and data structures/format:
  - benchmarking data will be selected from data sets available at participating institutions based on consensus by the pathologists to cover a representative spectrum of histological patterns
  - data provided as 1000x1000 images at 400x uncompressed
  - annotations in coordinates of POI and ROI stored in xml-format
  - 5-10 exemplary images made available with test annotations to provide overview of data
  - 25-50 densely annotated images will be available in an undisclosed fashion for benchmarking which will be performed on WHO/ITU servers and provide results according to section 3.

3 Benchmarking









Metrics

Level	Prediction task	Metric
Local	Segmentation	Jaccard index, Soerensen-Dice index, ROC/AuC
	Detection	***Accuracy, precision, recall, ROC/AuC***
	Visualization	Pixel flipping
Case-based	Nominal target variable	Accuracy, precision, recall, ROC/AuC
	Ordinal target variable	Accuracy, precision, recall, ROC/AuC, mean squared/absolute error
	Continuous target variable	Kaplan-Meier estimator Comparison to Cox regression ***Concordance index*** ***Mean squared/absolute error*** k-way accuracy (e.g. long/medium/short survival)

Current steps

- Ongoing discussion with the FDA on cooperation/standards
- Recruitment of additional board-certified pathologists to redundantly annotate data
- Contact with pathological societies about cooperation

High-throughput Truthing - Year 2

by Brandon D. Gallas

Year 2: High-throughput truthing of microscope slides to validate artificial intelligence algorithms analyzing digital scans of pathology slides: data (images + annotations) as an FDA-qualified medical device development tool (MDDT).

- Link to full proposal submitted 10/19/2018. Decisions expected in March 2019.
- Here is an executive summary (four slides) of the project with two new exciting deliverables.
  - 20190318-HTTpitchToUSCAP-5.pdf (169 KB, uploaded by Brandon D. Gallas 4 weeks ago)
  - Here is a project overview presentation given Nov.-Dec. 2018 to FDA/CDRO/OSEL management, the www.TILsinbreastcancer.org working group, project collaborators, and others.
    - 20181130-HTToverviewToCollabs.pdf (2 MB, uploaded by Brandon D. Gallas 4 months 1 day ago)
  - Link to list of collaborators
  - Link to updates

Project Overview

Pitch: We are launching a project to crowdsource pathologists and collect data (images + pathologist annotations) that can be qualified by the FDA/CDRH medical device development tool program (MDDT). The MDDT qualified data would be available to any algorithm developer to be used to validate their algorithm performance in a submission to the FDA/CDRH.

Notice, the year 2 title changed to emphasize, "data (images + annotations) as an FDA-qualified medical device development tool (MDDT)". If we can "qualify" a data set via FDA/CDRH MDDT program, it will be available to developers to use as their pivotal validation data in a submission to the FDA. That's the primary aim of year 2. In the lead up to the year 2 submission is the recruitment of partners to help. Check out the letters of support in the submission! We plan to organize data-collection events at meetings where all the pathologists go and at dedicated workshops at collaborating sites.

This project is generally open to new participants.

We will use the eeDAPstudies NCIPhub group to coordinate communications. So if you are a member, you will receive related communications about that project in addition to communications about the eeDAP MDDT. If you are not a member, sign up or check for updates here and in the blog. Updates will also be provided to the WSI working group on a less frequent basis.

Current Research Led By FDA

High-throughput truthing of microscope slides to validate artificial intelligence algorithms analyzing digital scans of same slides

- Partnership with academia, clinicians, and industry through Medical Device Innovation Consortium (MDIC)
- Focus on truth by pathologists, the microscope and TILs in breast cancer
- Status: Creating project structure, workgroups and leadership

- Key Deliverables:
  1. FDA qualified dataset for algorithm validation
  2. MDDT or 510(k) open or mock submission for
     - WSI viewer
     - TILs in breast cancer algorithm

Work to be done in the public domain.

MDDT:
Medical Device Development Tool
https://www.fda.gov/medicaldevices/scienceandresearch/medicaldevicedevelopmenttoolsmddt/

510(k):
Premarket submission for Class II medical devices
https://www.fda.gov/medicaldevices/deviceregulationandguidance/howtomarketyourdevice/premarketsubmissions/premarketnotification510k/default.htm

Dr. Alexander Arnold
Institute of Pathology, Charité Universitätsmedizin Berlin
Berlin Institute of Health

Miriam Hägele & Philipp Seegerer & Markus Wenzel
Fraunhofer HHI & Technische Universität Berlin
& Berlin Center for Machine Learning (BZML)

Prof. Dr. Carsten Denkert
University of Marburg
Reference pathology center German Breast Group (GBG)

Prof. Dr. Alexander Binder
Singapore University of Technology and Design (SUTD)
Singapore

Prof. Dr. Roberto Salgado
Chair, International TIL working group
Breast Cancer Translational Research Laboratory, Institut Jules Bordet, Université Libre de Bruxelles (U.L.B.), Brussels, Belgium;

Dr. Till Braunschweig
Deputy Director, Institute of Pathology, University Hospital, TU Aachen, Germany

Dr. Brandon Gallas
Division of Imaging, Diagnostics, and Software Reliability (DIDSR)
Office of Science and Engineering Laboratories (OSEL)
Center for Devices and Radiological Health (CDRH)
U.S. FDA
Silver Spring, MD, USA