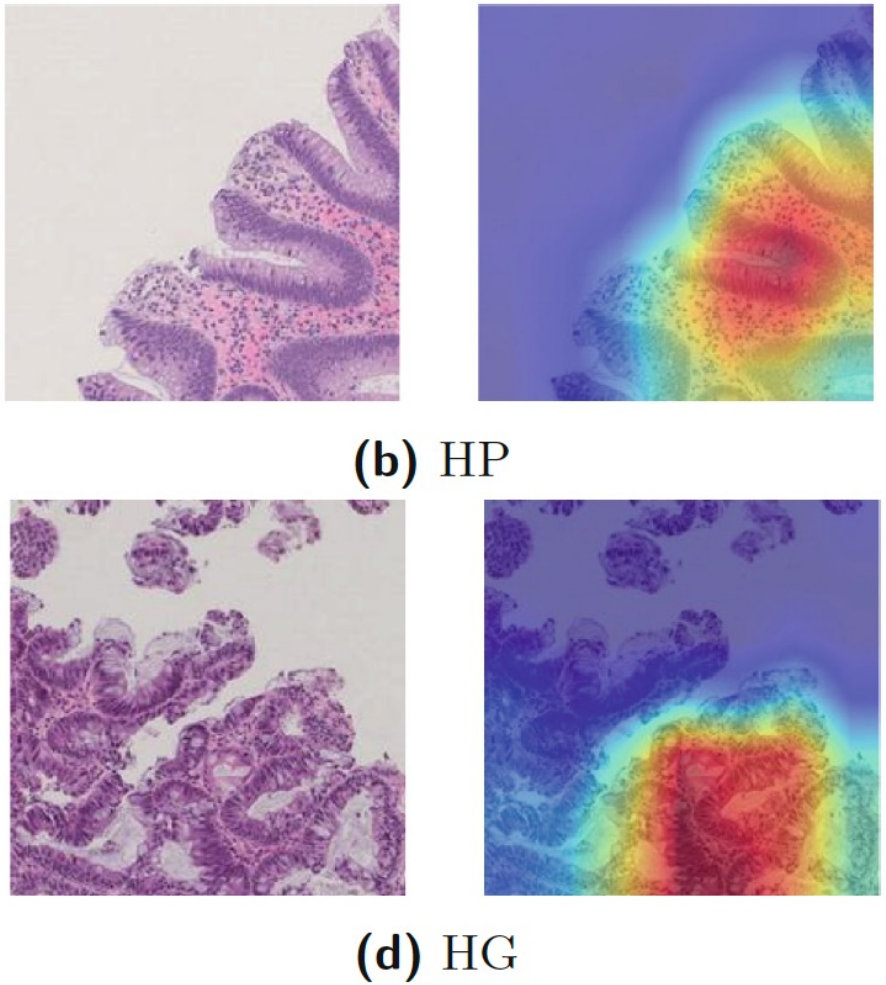
UNIVERSITÀ
DEGLI STUDI
DI TORINO

AZIENDA OSPEDALIERO - UNIVERSITARIA
Città della Salute e della Scienza di Torino

# L' intelligenza artificiale nella diagnostica istologica degli adenomi: l'esperienza Torinese




**L. BERTERO**

Div. of Pathology, Dept. Medical Sciences

University of Turin, Italy

17 dicembre 2021

nature medicine

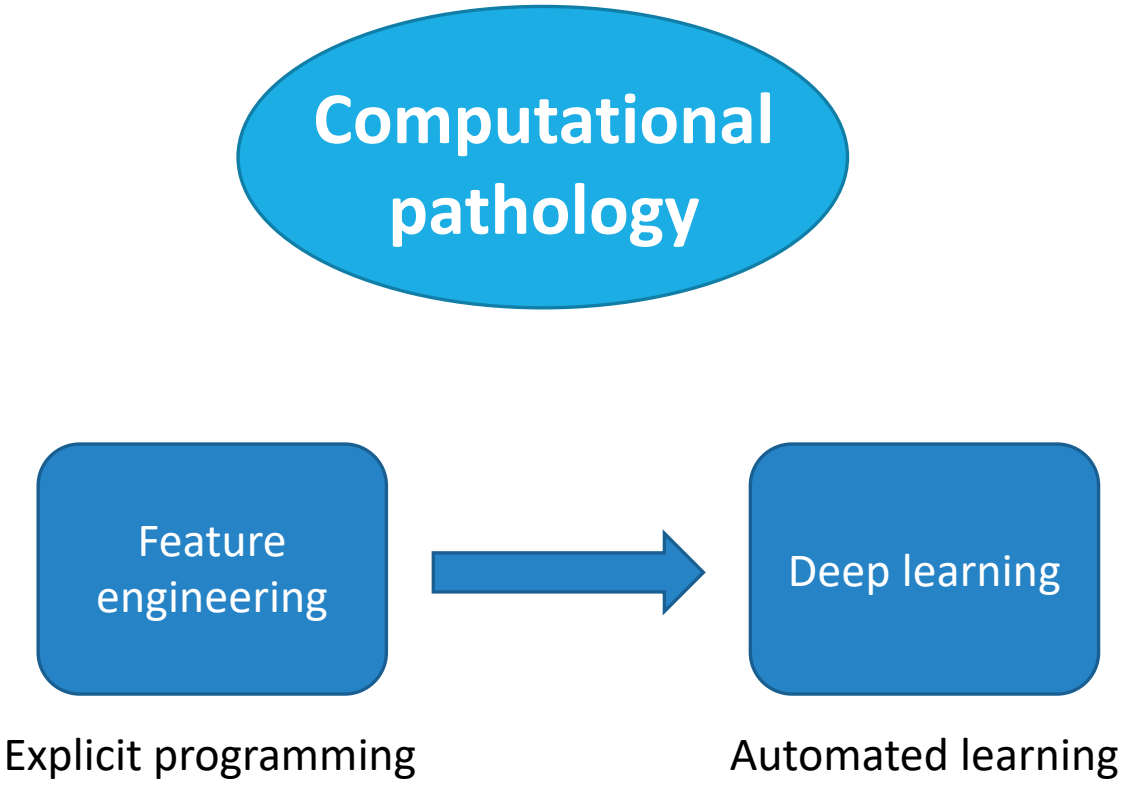
REVIEW ARTICLE

https://doi.org/10.1038/s41591-021-01343-4

Check for updates

# Deep learning in histopathology: the path to the clinic

Jeroen van der Laak[1,2] ✉, Geert Litjens[1] and Francesco Ciompi[1]




**Box 1 | Definitions**

**Deep learning**
A machine learning approach in which algorithms are trained for a specific task (or set of tasks) by exposing a multilayered artificial neural network to (typically a large amount of) training data, without the need for handcrafted engineering of features to be extracted from the data. The resulting algorithm has learned a hierarchical representation of the data that is subsequently used for tasks such as classification, detection or segmentation. The term deep refers to artificial neural networks built using many layers, in other words a deep neural network.

**Digital pathology**
The digitization of the traditional diagnostic process of analyzing cells and tissue with a microscope via whole-slide scanners and computer screens.

**Computational pathology**
The computational analysis of digital images obtained through scanning slides of cells and tissues.

**Radiomics/pathomics**
Techniques to extract a (usually very large) set of features from radiological or histopathological digital images, respectively, using computational algorithms of data analysis. These features are successively used to feed (usually supervised) prediction models targeting clinically relevant end points, such as prognosis.

**End-to-end training**
In the context of machine learning models, possibly consisting of a pipeline with multiple steps, end-to-end training refers to the procedure of learning the optimal value of all parameters of a model simultaneously rather than sequentially (that is, one step at a time).

**Whole-slide images**
Digital images obtained by digitizing complete histopathological glass slides using a high-resolution scanner.

**Convolutional neural networks**
Deep learning approach consisting of a series of convolutional layers to process data (usually bi-dimensional) from input to output. Each layer implements the convolution operation between the input data and a set of filters (that is, small matrices), whose numerical values are automatically learned in an end-to-end training fashion.

**Graphics processing units**
Microprocessor specifically designed to process many data samples simultaneously, such as parts of digital images or features extracted from images.

**Image segmentation**
The operation of decomposing the semantic content of an image into multiple segments, where each segment contains pixels belonging to the same semantic category (for example, the tumor region).

**U-Net models**
Deep learning models based on two convolutional neural networks, one that encodes the input image into a set of features, and one that decodes those features to produce a segmentation output. The name, introduced in 2015 by Ronneberger et al.[145], indicates the U shape that the two convolutional neural networks form, where the encoder and decoder are connected via skip connections.

**Data augmentation**
The operation of artificially modifying some properties of input data (for example, image contrast, orientation, color and so on) with the aim of feeding a computational model with multiple variations of the same piece of data.

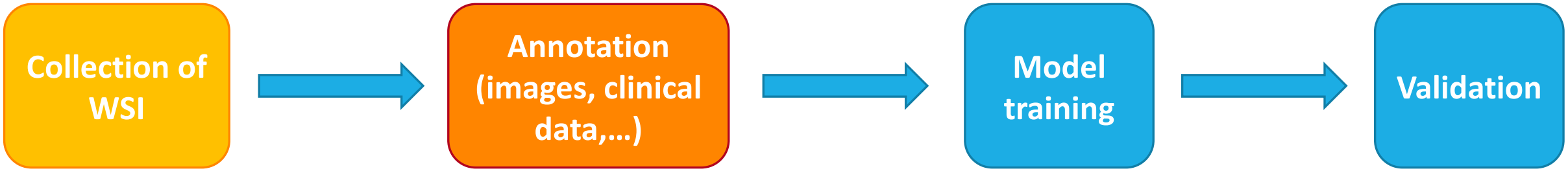
**Model regularization**
In machine learning, indicates the process of constraining a model's parameters to small values, discouraging complex models, therefore reducing the risk of overfitting the training data.

- Predetermined feature selection
- Multiple interactions pathologists/informaticians needed
- Time consuming

Feature engineering → Deep learning

- Automated learning
- Freely available source codes of effective neural network architectures
- Superior results in most cases

## Deep learning overall workflow:



**ImageNet Large Scale Visual Recognition Competition (ILSVRC)**

- Since ~2010
- Efficacy of CNN (convolutional neural networks)

**CAMELYON challenge**

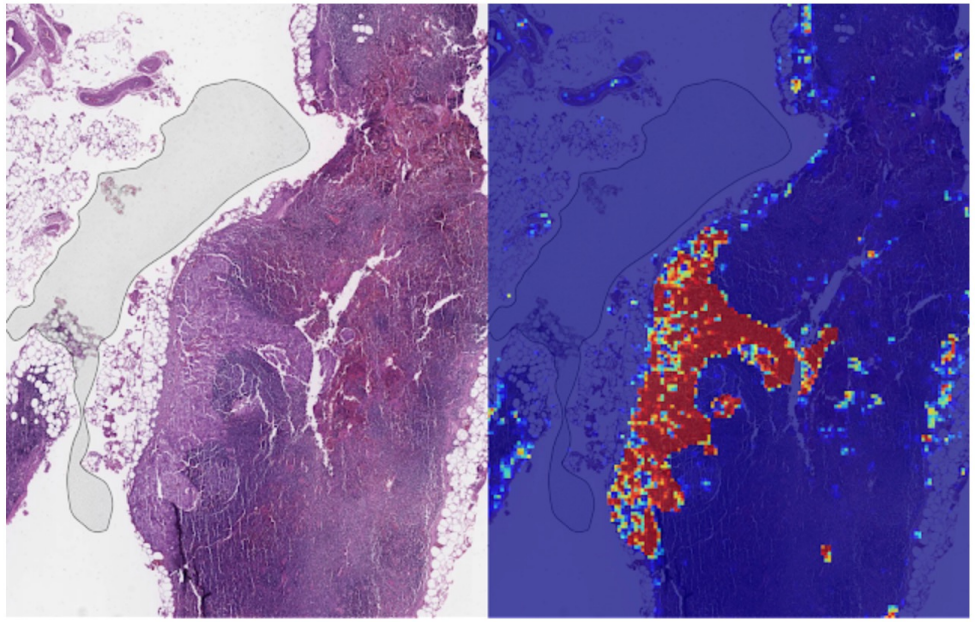
- Breast cancer metastases in sentinel lymph nodes
- Dataset of 1399 manually annotated WSI

Google AI Blog

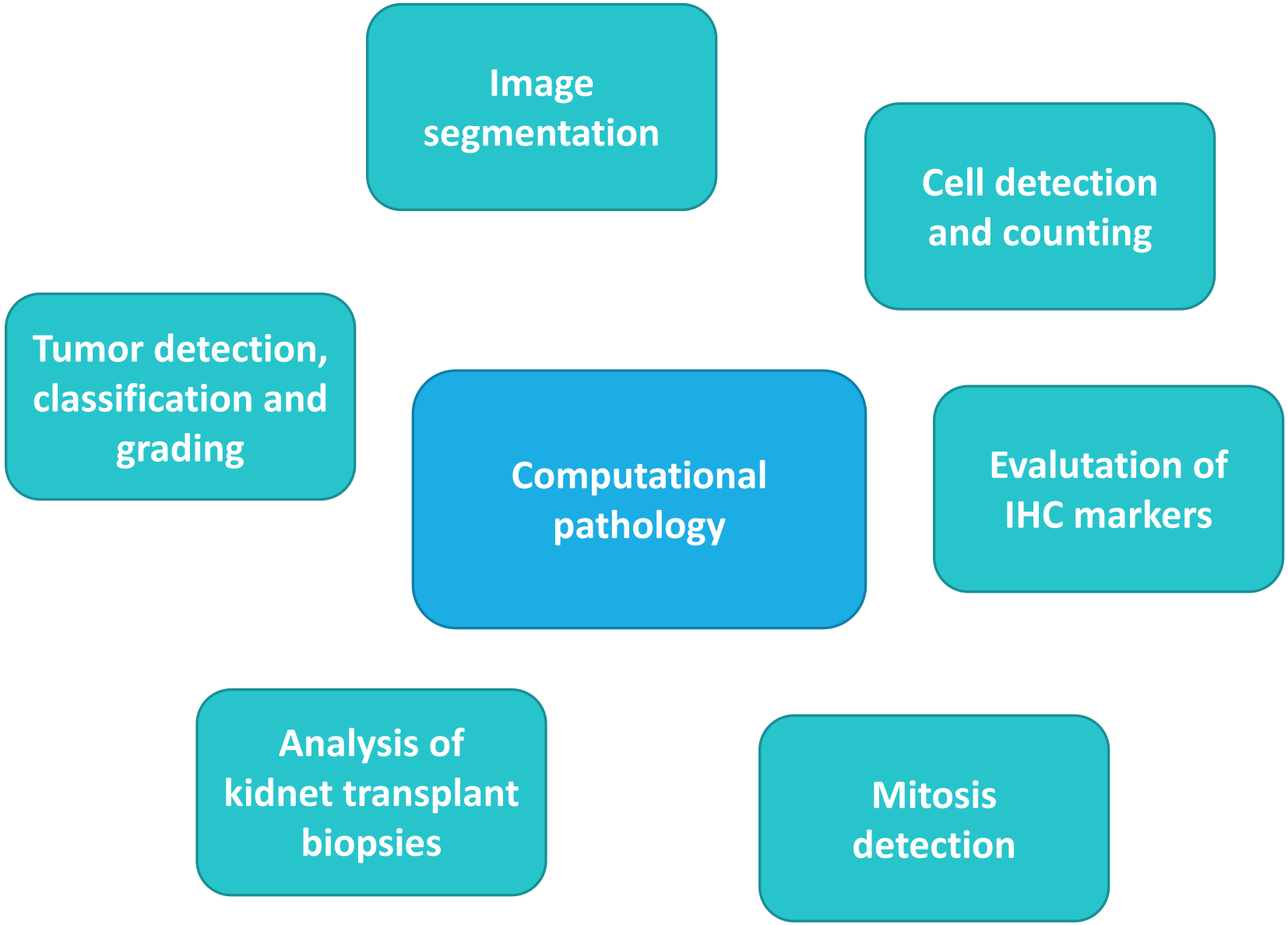
The latest from Google Research

Applying Deep Learning to Metastatic Breast Cancer Detection

Friday, October 12, 2018

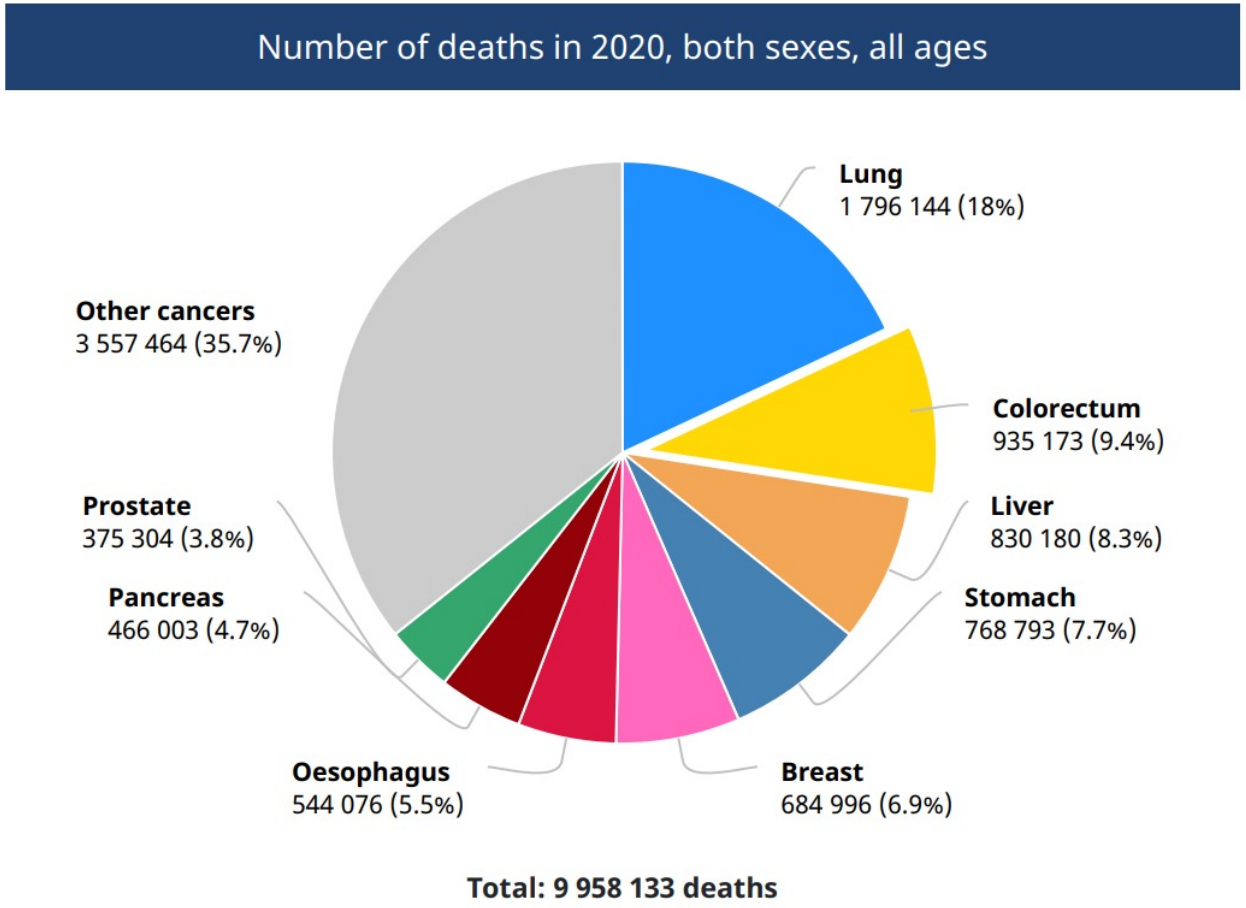


- **Reducing repetitive and time-consuming tasks**
- **Lower interobserver variability**



# Colorectal carcinoma

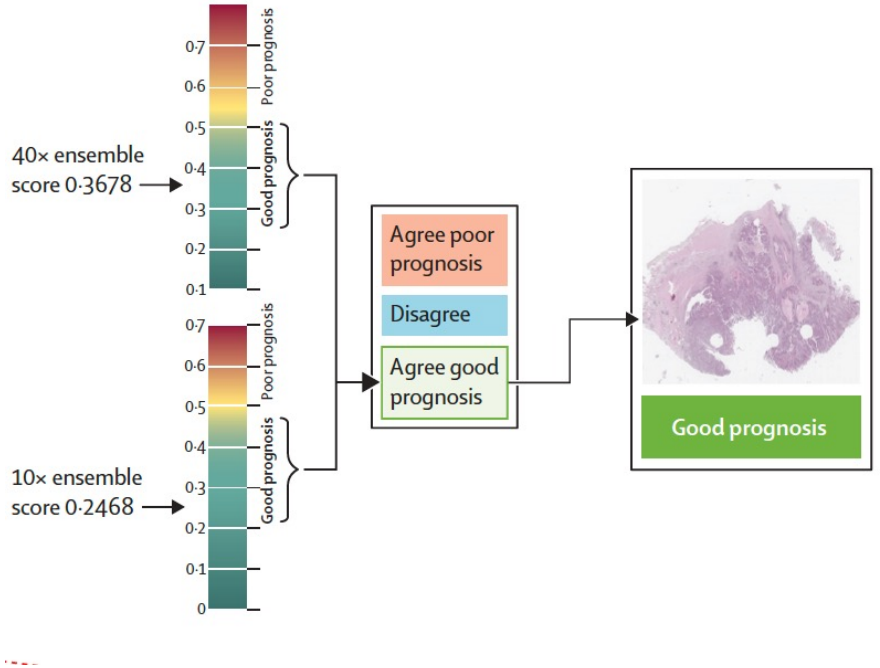
- Colorectal carcinoma (CRC) is the second most deadly and the third most common cancer (Globocan 2020)
- Colorectal cancer screening enables prompt detection of early CRC or preinvasive lesions, but represents a significant workload for both endoscopy and pathology units



# Digital pathology for colorectal carcinoma

- Distinction between tumor tissue and stroma (Kather JN et al. *Sci Rep* 2016)
- Outcome prediction (Bychkov D et al. *Sci Rep* 2018; Kather JN et al. *PLoS Med* 2019; Skrede O et al. *Lancet* 2020)
- Molecular profile prediction (Yamashita R et al., *Lancet Oncol* 2020; Sirinukunwattana K et al. *Gut* 2021; Bilal M et al. *Lancet Digit Health* 2021)

- Adenoma classification…



# Adenoma classification

## Deep Learning for Classification of Colorectal Polyps on Whole-slide Images

Bruno Korbar[1,2], Andrea M. Olofson[3], Allen P. Miraflor[3], Catherine M. Nicka[3], Matthew A. Suriawinata[3], Lorenzo Torresani[2], Arief A. Suriawinata[3], Saeed Hassanpour[1,2,4]

J Pathol Inform 2017, 1:30

**Table 1: Our dataset: The distribution of colorectal polyp types in crop images used in this work**

| Colorectal polyp type | Acronym | Number of image crops |
|---|---|---|
| Hyperplastic polyp | HP | 405 |
| Sessile serrated polyp | SSP | 612 |
| Traditional serrated adenoma | TSA | 258 |
| Tubular adenoma | TA | 360 |
| Tubulovillous/villous adenoma | TVA/V | 202 |
| Normal | - | 237 |
| Total | - | 2074 |

**Table 4: Whole-slide classification results: Results of our final model for classification of colorectal polyps on 239 whole-slide images in our test set**

| | HP ($n$=37) (%) | SSP ($n$=39) (%) | TSA ($n$=38) (%) | TA ($n$=39) (%) | TVA/V ($n$=38) (%) | Normal ($n$=48) (%) | Total ($n$=239) (%) |
|---|---|---|---|---|---|---|---|
| Accuracy | 89.8 (85.3-93.3) | 89.5 (85.0-93.1) | 94.7 (91.1-97.2) | 93.1 (89.2-96.0) | 95.8 (92.5-97.9) | 95.0 (91.5-97.4) | 93.0 (89.0-95.9) |
| Precision | 90.9 (86.6-94.2) | 86.11 (81.1-90.2) | 100.0 (98.5-100) | 83.3 (78.0-87.8) | 97.2 (94.3-98.9) | 80.7 (75.1-85.5) | 89.7 (85.2-93.2) |
| Recall | 81.1 (75.5-85.8) | 81.6 (76.1-86.3) | 89.5 (84.9-93.0) | 89.7 (85.2-93.3) | 92.1 (88.0-95.2) | 95.8 (92.5-98.0) | 88.3 (83.6-92.1) |
| F1 score | 85.7 (80.6-89.9) | 83.8 (78.5-88.2) | 94.4 (90.8-97.0) | 86.4 (81.4-90.5) | 94.6 (90.9-97.1) | 87.6 (82.8-91.5) | 88.8 (84.1-92.5) |

HP: Hyperplastic polyp, SSP: Sessile serrated polyp, TSA: Traditional serrated adenoma, TA: Tubular adenoma, TVA/V: Tubulovillous/villous adenoma

# Adenoma classification

JAMA Network Open

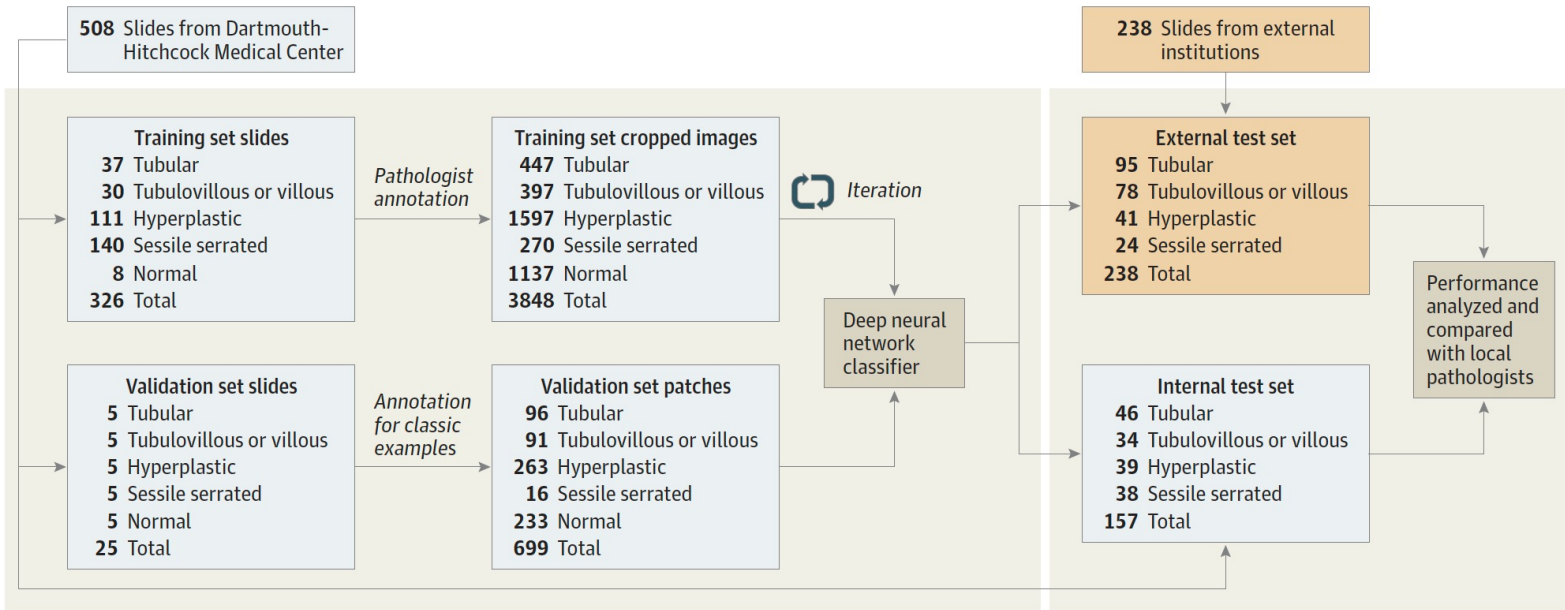
Original Investigation | Health Informatics

## Evaluation of a Deep Neural Network for Automated Classification of Colorectal Polyps on Histopathologic Slides

Jason W. Wei, BA; Arief A. Suriawinata, MD; Louis J. Vaickus, MD, PhD; Bing Ren, MD, PhD; Xiaoying Liu, MD; Mikhail Lisovsky, MD, PhD; Naofumi Tomita, MS; Behnaz Abdollahi, PhD; Adam S. Kim, MD; Dale C. Snover, MD; John A. Baron, MD; Elizabeth L. Barry, PhD; Saeed Hassanpour, PhD

*JAMA Network Open*. 2020;3(4):e203398. doi:10.1001/jamanetworkopen.2020.3398

Figure 1. Data Flow Diagram for the Study



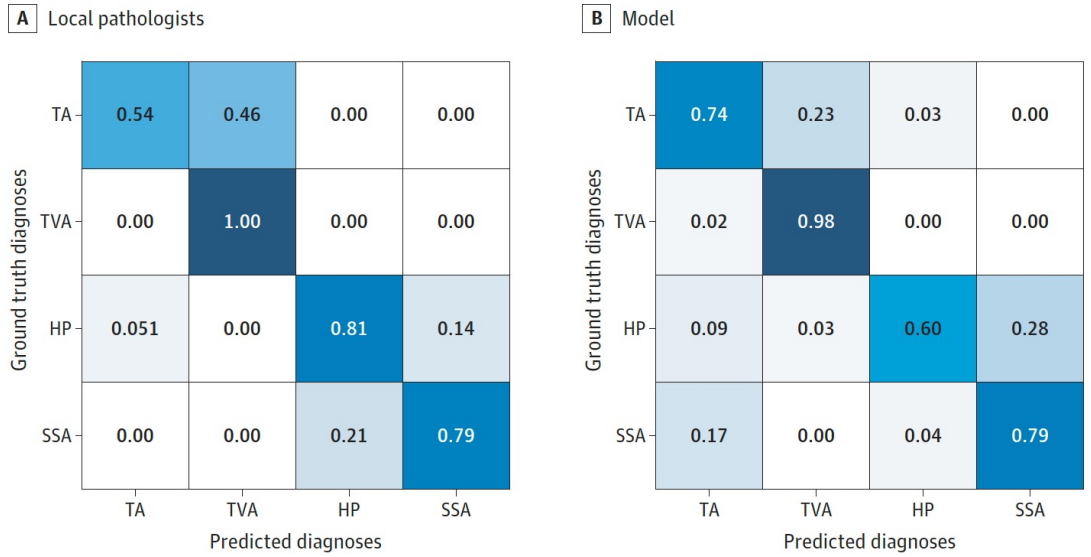
# Adenoma classification

Table. Per-Class Comparison Between Local Pathologists and the Deep Neural Network Model in Classifying Colorectal Polyps on Internal and External Test Sets

| | Internal test set (n = 157) | | | | | | External test set (n = 238) | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Local pathologists | | | Deep neural network | | | Local pathologists | | | Deep neural network | | |
| Polyp type | Accuracy, % | Sensitivity, % | Specificity, % | Accuracy, % | Sensitivity, % | Specificity, % | Accuracy, % | Sensitivity, % | Specificity, % | Accuracy, % | Sensitivity, % | Specificity, % |
| TA | 89.8 | 76.1 | 95.5 | 93.0 | 89.1 | 94.6 | 79.8 | 53.7 | 97.2 | 84.5 | 73.7 | 91.6 |
| TVA | 94.3 | 88.2 | 95.8 | 95.5 | 97.1 | 95.1 | 81.5 | 100 | 77.7 | 89.5 | 97.6 | 87.8 |
| HP | 89.8 | 76.9 | 94.1 | 92.4 | 82.1 | 95.8 | 91.6 | 80.8 | 96.8 | 85.3 | 60.3 | 97.5 |
| SSA | 91.7 | 81.6 | 95.0 | 93.0 | 78.9 | 97.5 | 93.3 | 79.2 | 94.8 | 88.7 | 79.2 | 89.7 |
| Mean | 91.4 | 80.7 | 95.1 | 93.5 | 86.8 | 95.7 | 86.6 | 78.4 | 91.6 | 87.0 | 77.7 | 91.6 |

## Limitations:

- Lack of dysplasia grading
- Lack of normal tissue
- Lower performance during external testing




*JAMA Network Open*. 2020;3(4):e203398. doi:10.1001/jamanetworkopen.2020.3398

# High Performance Computing boosting Biomedical Applications

DEEPHEALTH

## PARTNERS

**22 partners** from 9 European Countries

### Research Organisations

Universitat Politècnica de València
Università degli Studi di Torino
Barcelona Supercomputing Center – Centro Nacional de Supercomputación
list
UNIMORE – Università degli Studi di Modena e Reggio Emilia
CRS4 – Ideas become life
Otto von Guericke Universität Magdeburg
EPFL
Karolinska Institutet
security tech Stelar law research

### Health Organisations

Azienda Ospedaliero-Universitaria Città della Salute e della Scienza di Torino
CHUV
Fundación para el Fomento de la Investigación Sanitaria y Biomédica de la Comunitat Valenciana
Spitalul Clinic Prof. Dr. Theodor Burghele București
Stockholms läns landsting

### Large Industrial Partners

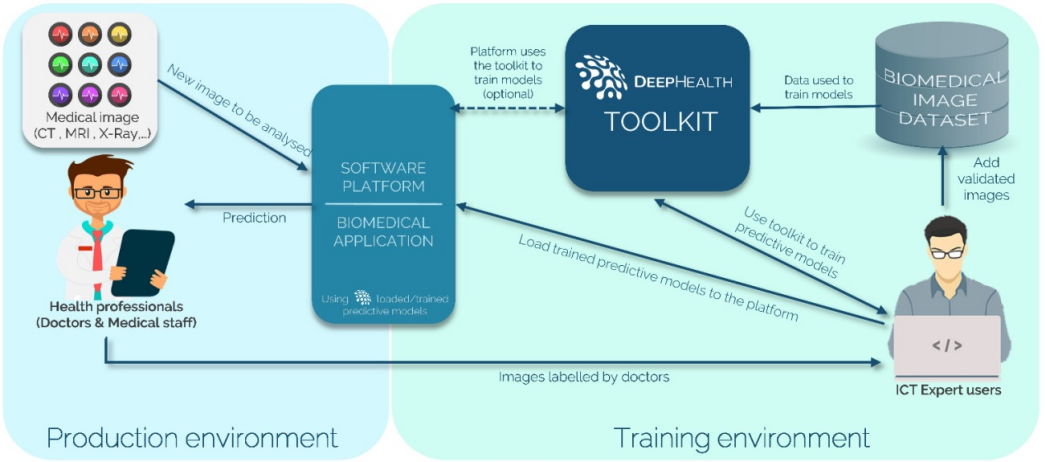
SIMAVI – Software Imagination & Vision
everis – an NTT DATA Company
PHILIPS
THALES

### SME Industrial Partners

WINGS ICT Solutions
prodesign electronic gmbh
TREE Technology

## Aim

Provide **High Performance Computing (HPC)** power at the service of biomedical applications; and apply **Deep Learning** (DL) and **Computer Vision** (CV) techniques on large and complex biomedical datasets to support new and more efficient ways of diagnosis, monitoring and treatment of diseases




This project has received funding from the European Union's Horizon 2020 research innovation programme under grant agreement No. 825111

DEEPHEALTH

UNIVERSITÀ DEGLI STUDI DI TORINO

UniTOPatho

di.unito.it

DIPARTIMENTO DI INFORMATICA

## Use Cases

**14 pilot test-beds in 3 areas:**

**Neurological diseases**

- Migraine and Seizures prediction
- Major Depression
- Dementia
- Study of structural changes in lumbar spine pathology
- Population model for Alzheimer's Disease
- Epileptic seizures detection
- Objective fatigue assessment for multiple sclerosis patients

**Tumor detection and early cancer prediction**

- Chest cancer detection
- Prostate tumor diagnosis
- Skin cancer melanoma detection

**Digital pathology and automated image annotation**

- Classification of whole-slide histological images of colorectal biopsy samples
- CT brain perfusion maps synthesis
- Deep Image annotation
- Image Analysis and prediction for Urology

# Colon cancer diagnosis

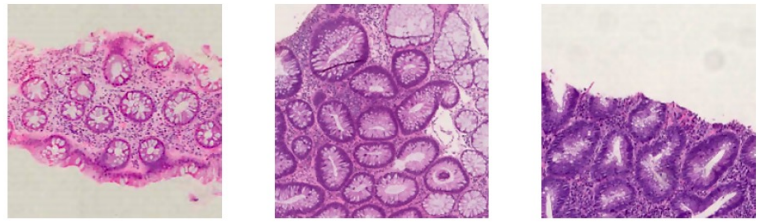
DeepHealth

**Colon cancer is one of the most frequent causes of death. Screening programs can enable prompt diagnosis and treatment of this aggressive disease, but they also lead to higher caseloads and costs for the already strained European healthcare services. DeepHealth can help streamline pathological diagnosis of colon biopsies.**

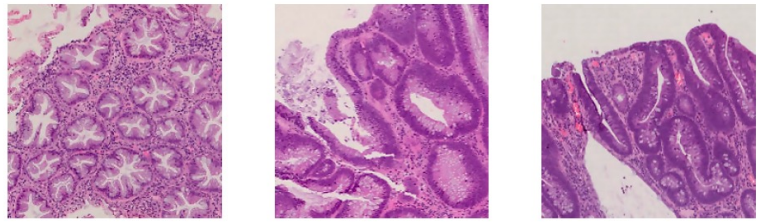
# Dataset (WSI images)

| | HP | NORM | TA.HG | TA.LG | TVA.HG | TVA.LG | Total |
|---|---|---|---|---|---|---|---|
| Slides | 62 | 30 | 34 | 232 | 44 | 55 | **457** |
| $R_t$ | 158 | 112 | 145 | 777 | 264 | 245 | **1701** |
| $A_t$ $[\mathrm{cm}^2]$ | 9.91 | 18.38 | 7.94 | 71.74 | 60.45 | 41.86 | **210.29** |

- H&E slide acquired on the Hamamatsu Nanozoomer S210 scanner (200X)
- Manual annotation according to 6 classes:
  - NORM: normal tissue
  - HP: hyperplastic polyp
  - TA.LG: tubular adenoma, low-grade dysplasia
  - TA.HG: tubular adenoma, high-grade dyplasia
  - TVA.LG: tubulo-villous adenoma, low-grade dysplasia
  - TVA.HG: tubulo-villous adenoma, high-grade dysplasia



(a) NORM (b) TA.LG (c) TA.HG



(d) HP (e) TVA.LG (f) TVA.HG

Perlo D. et al. MICAD 2021

- CNN: ResNet-18
- Pre-training on the ImageNet classification task
- Data augmentation: one random operation between rotation, equalization, solarization, inversion and contrast enhancing

Patches normalization: relevant features are not embed in color, but in image texture and signal strenght

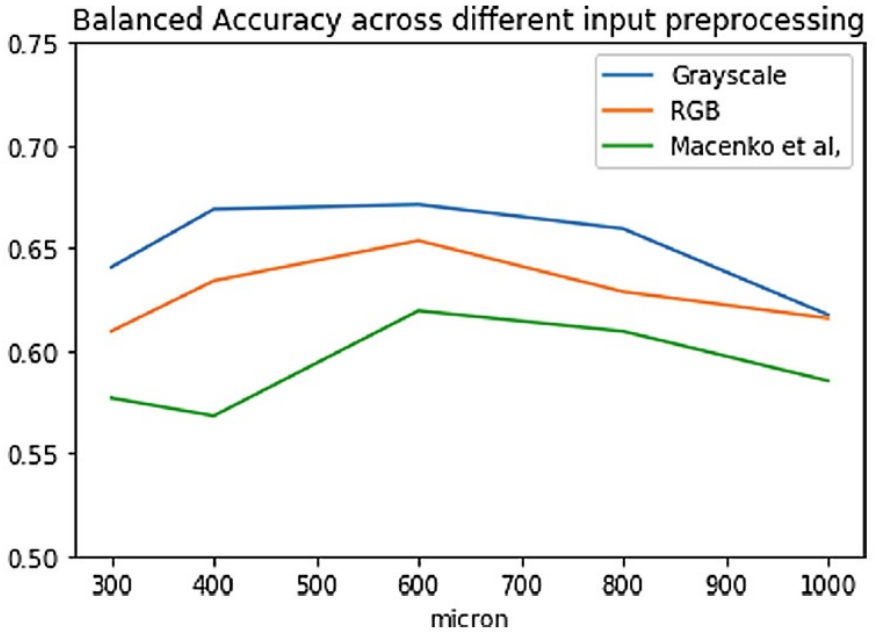



**Fig. 2.** Patches classification performance.

Perlo D. et al. MICAD 2021

# Patches resolution:

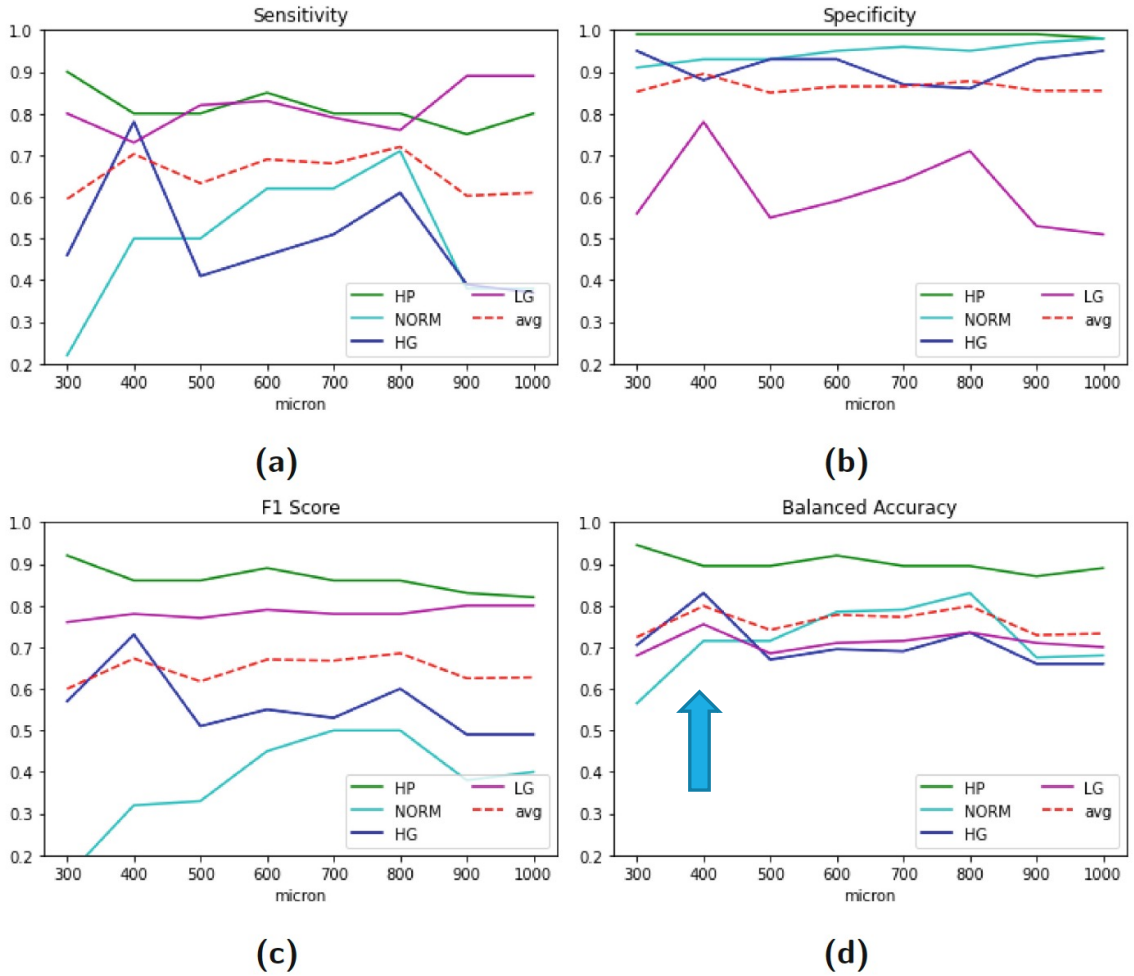



**Table 3.** Human dysplasia diagnostic performance comparison

| | | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|
| Hyperplastic | Our (400 µm) | 0.90 | 0.80 | **0.99** |
| | Our (600 µm) | **0.92** | **0.85** | **0.99** |
| | Pathologist [8] | 0.79 | 0.30 | 0.97 |
| Low grade | Our (400 µm) | **0.76** | 0.73 | **0.78** |
| | Our (600 µm) | 0.71 | **0.83** | 0.59 |
| | Pathologist [8] | 0.66 | 0.57 | 0.69 |
| High grade | Our (400 µm) | **0.83** | 0.78 | 0.88 |
| | Our (600 µm) | 0.70 | 0.46 | **0.93** |
| | Pathologist [8] | **0.83** | **0.81** | 0.84 |

- Achieved results are similar to those reported by Denis B et al. (*Eur J of Gastroenterol Hepatol* 2009)

Perlo D. et al. MICAD 2021

# Dysplasia grading

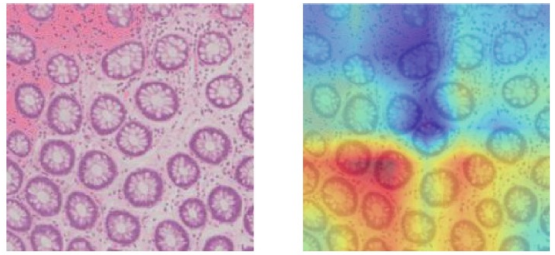
**Table 4.** WSI inferences: confusion matrices.

**(a)** $\varphi$ =600 µm, gray-scale

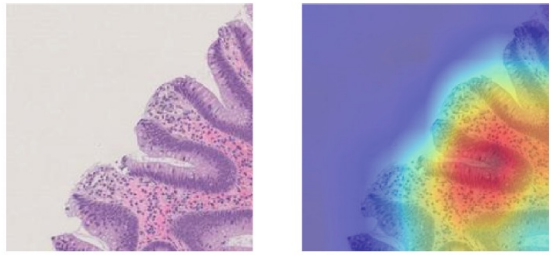
| Gr. truth \ Predicted | HP | NORM | HG | LG |
|---|---|---|---|---|
| HP | **0.85** | 0 | 0.05 | 0.1 |
| NORM | 0.12 | **0.75** | 0 | 0.12 |
| HG | 0.02 | 0 | **0.63** | 0.35 |
| LG | 0.03 | 0.09 | 0.18 | **0.7** |

**(b)** $\varphi$ =600 µm, RGB

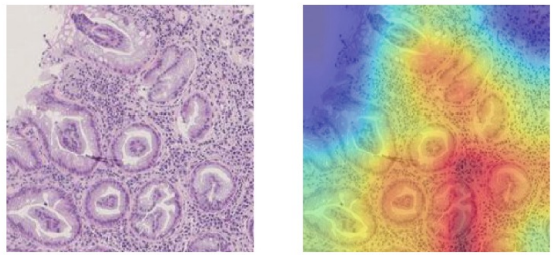
| Gr. truth \ Predicted | HP | NORM | HG | LG |
|---|---|---|---|---|
| HP | **0.75** | 0.05 | 0 | 0.2 |
| NORM | 0 | **0.62** | 0 | 0.38 |
| HG | 0 | 0.02 | **0.61** | 0.37 |
| LG | 0.03 | 0.06 | 0.15 | **0.76** |



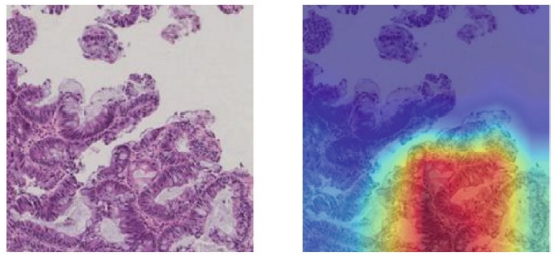
**(a)** NORM



**(b)** HP



**(c)** LG



**(d)** HG

- Poor results in distinguishing TA versus TVA/VA

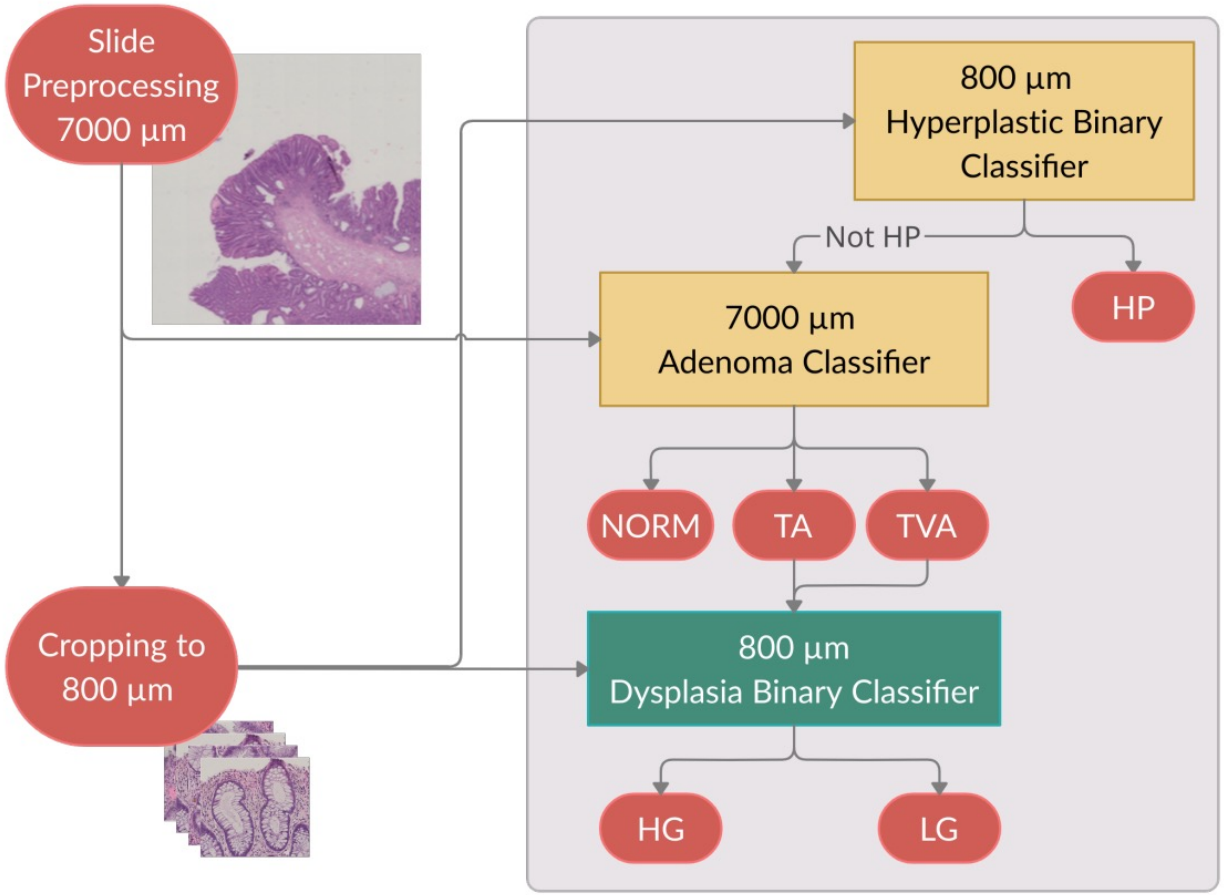
Perlo D et al. MICAD 2021

# Multi-resolution analysis

| | Patch scale $\sigma$ [µm] | | | | | |
|---|---|---|---|---|---|---|
| Type | 100 | 800 | 1500 | 4000 | 7000 | 8000 |
| BA (6-class) | 0.40 | 0.45 | **0.46** | 0.41 | 0.37 | 0.38 |
| NORM | 0.70 | 0.66 | 0.72 | 0.76 | 0.78 | 0.71 |
| HP | 0.81 | **0.92** | 0.85 | 0.70 | 0.60 | 0.69 |
| TA (HG+LG) | 0.65 | 0.66 | 0.65 | 0.71 | **0.76** | 0.70 |
| TVA (HG+LG) | 0.64 | 0.67 | 0.68 | 0.74 | **0.84** | 0.76 |

**Table 2**: Preliminary experiments: overall BA for all of the six classes (first row) and BA for each polyp type, plus normal tissue.

- Adenoma type and dysplasia grade are best classified at different scales




Barbano CA et al. IEEE ICIP 2021

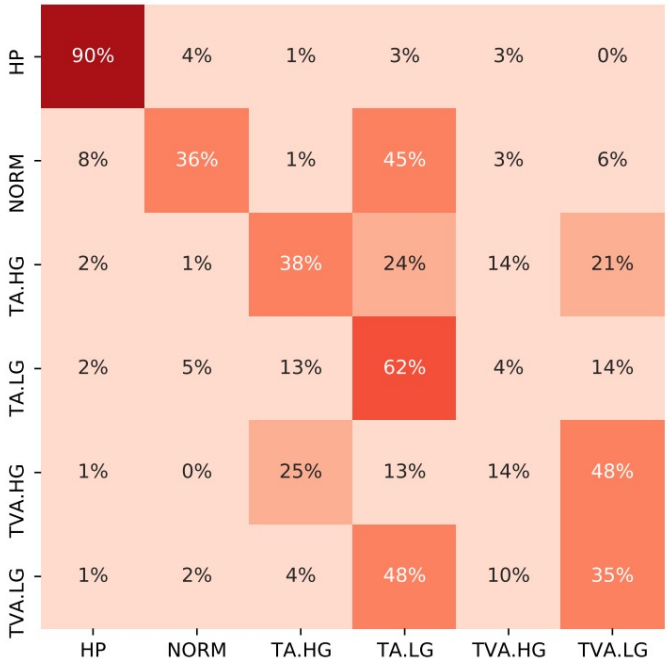
# Multi-resolution analysis

| | HP | NORM | TA | | TVA | |
|---|---|---|---|---|---|---|
| | | | HG | LG | HG | LG |
| Sensitivity | 0.86 | 0.79 | 0.60 | 0.50 | 0.78 | 0.52 |
| Specificity | 0.93 | 0.87 | 0.92 | 0.94 | 0.96 | 0.92 |
| BA | 0.89 | 0.83 | 0.76 | 0.72 | 0.87 | 0.72 |

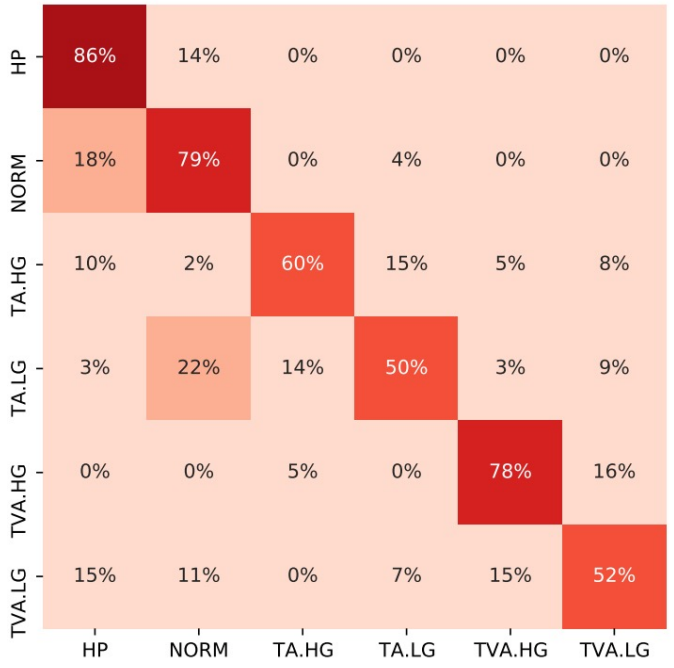
**Table 3**: Sensitivity, Specificity and BA per class.

| | $\sigma$ | HP | NORM | TA | TVA |
|---|---|---|---|---|---|
| Baseline | 800 | **0.92** | 0.66 | 0.66 | 0.67 |
| Baseline | 1500 | 0.85 | 0.72 | 0.65 | 0.68 |
| Baseline | 7000 | 0.60 | 0.78 | 0.76 | 0.84 |
| Multi-resolution | - | 0.89 | **0.83** | **0.81** | **0.87** |

**Table 4**: Comparison of the class BA between the baseline and the proposed multi-resolution approach.




(a) Baseline




(b) Multi-resolution Ensemble

Limitations:

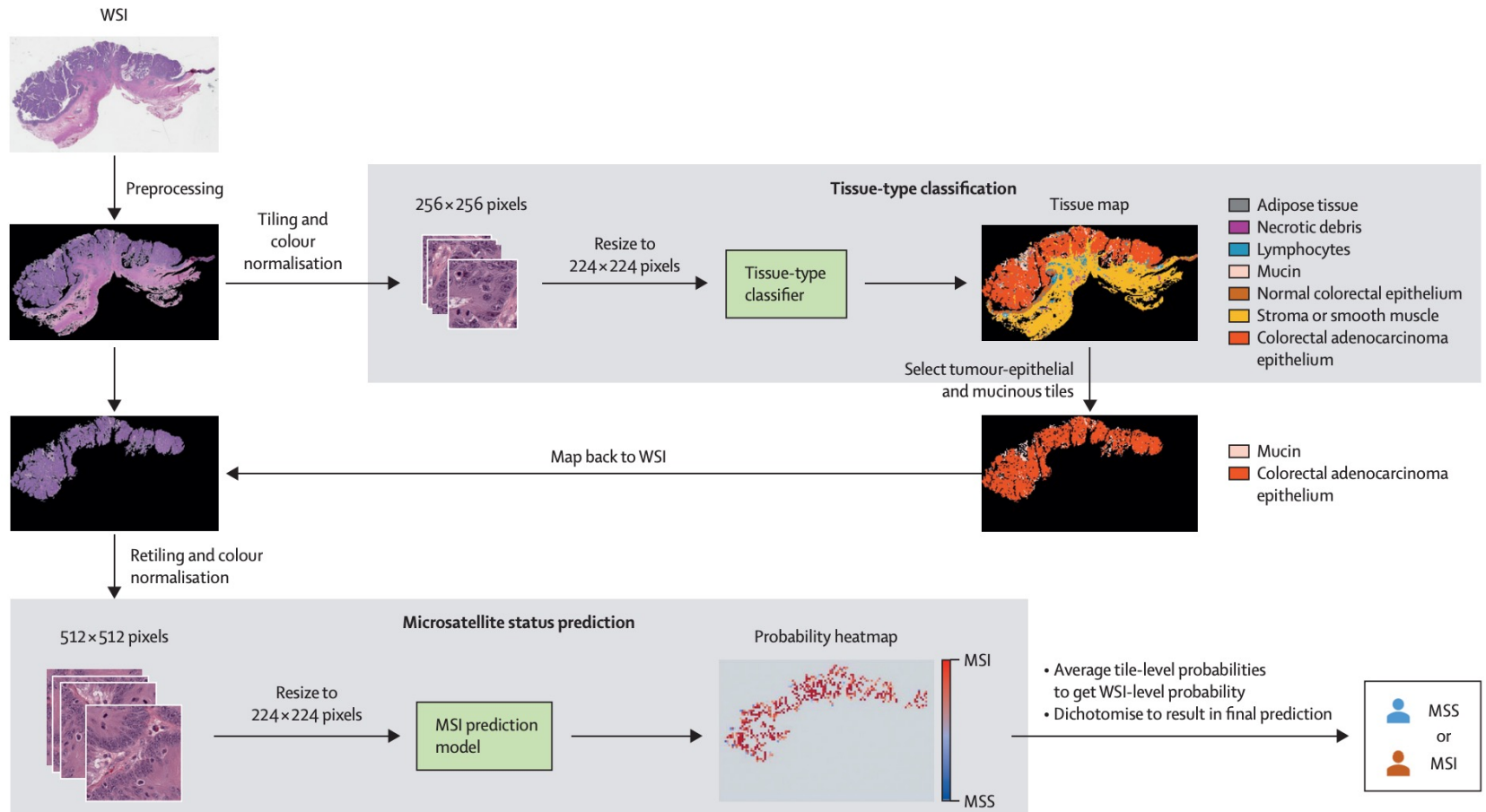
- Some entities missing (serrated adenomas, invasive adenocarcinomas,…)
- Larger dataset is warranted
- Lack of external validation

Barbano CA et al. IEEE ICIP 2021

# Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study

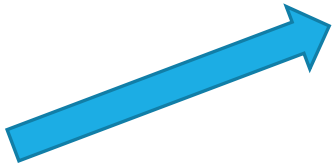
*Rikiya Yamashita, Jin Long, Teri Longacre, Lan Peng, Gerald Berry, Brock Martin, John Higgins, Daniel L Rubin*, Jeanne Shen**

*Lancet Oncol* 2020; 22: 132–41



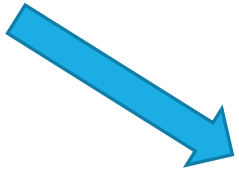
# Challenges

- Collect large-scale annotated datasets (images and clinical annotations)



**Combined H&E and IHC for WSI annotation**

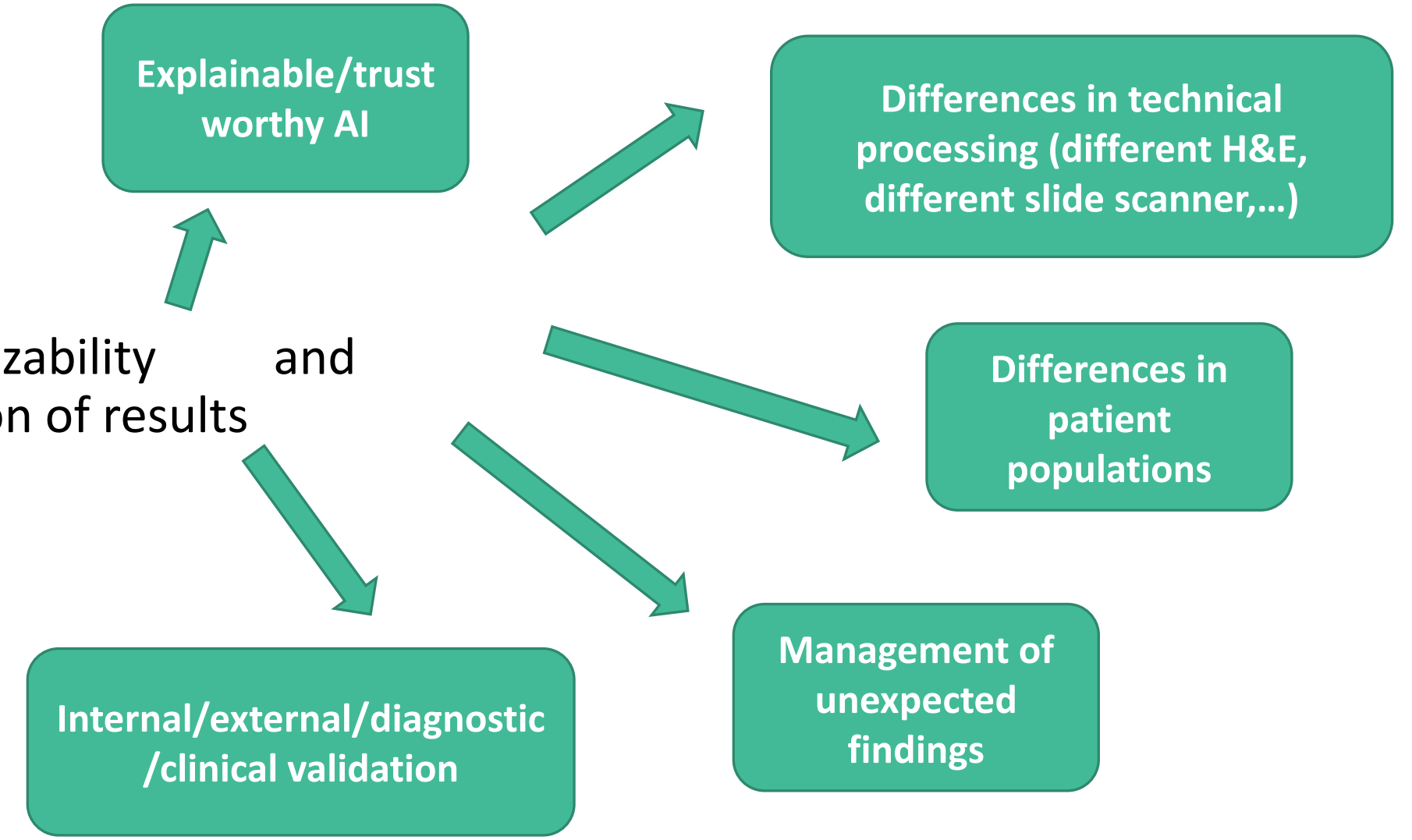
**Weakly supervised learning**



**Scailing up input to whole WSI**

# Challenges

- Generalizability and validation of results




Data augmentation or normalization

# Thank you!

DEEPHEALTH

- Prof.ssa P. Cassoni
- Dr. L. Bertero
- Dr. A. Gambella
- Dr. E. Bottasso

- Prof. M. Grangetto
- Dr. E. Tartaglione
- Dr. D. Perlo
- Dr. C. A. Barbano
- Dr. A. Fiandrotti

- NTT Data Spain
- Universitat Politecnica de Valencia (UPV)
- Philips Medical Systems Netherland BV
- Software Imagination & Vision SRL (SIMAVI)
- Wings ICT Solutions Information & Communication Technologies IKE
- Thales Six GTS France SAS
- Commissariat a l'energie atomique et aux energies alternatives
- Barcelona Supercomputing Center
- Pro Design Electronic GmbH
- Karolinska Institutet
- Fundacion para el fomento de la investigacion sanitaria y biomedica de la comunitat valenciana (FISABIO)
- Università degli Studi di Modena e Reggio Emilia (UNIMORE)
- Stockholms lans landsting
- AOU Città della Salute e della Scienza di Torino
- Ecole Polytechnique Federale de Lausanne
- Centre Hospitalier Universitarie Vaudois
- Tree Technology SA
- Otto Von Guericke Universitat Magdeburg
- Stelar Security Technology Law Research
- Spitalul Clinic Prof. Dr. Theodor Burghele
- Centro di Ricerca, Sviluppo e Studi Superiori in Sardegna SRL (CRS4)