Development of AI-based diagnostic tools based on massive digital pathology datasets

LUMI AI Factory Launch @ Pikku-Finlandia
2.4.2025

Pekka Ruusuvuori, Associate Professor
Institute of Biomedicine, University of Turku
pekka.ruusuvuori@utu.fi
LinkedIN: @PekkaRuusuvuori

Mira Valkonen – Sartaj Salman - Masi Valkonen – Pekka Ruusuvuori (PI)
Hesam Hakimnejad – Muhammad Adnan – Hyder Abbas - Rimsha Kaokab – Henrique Hiram Libutti – Niloufar Rahimizadeh

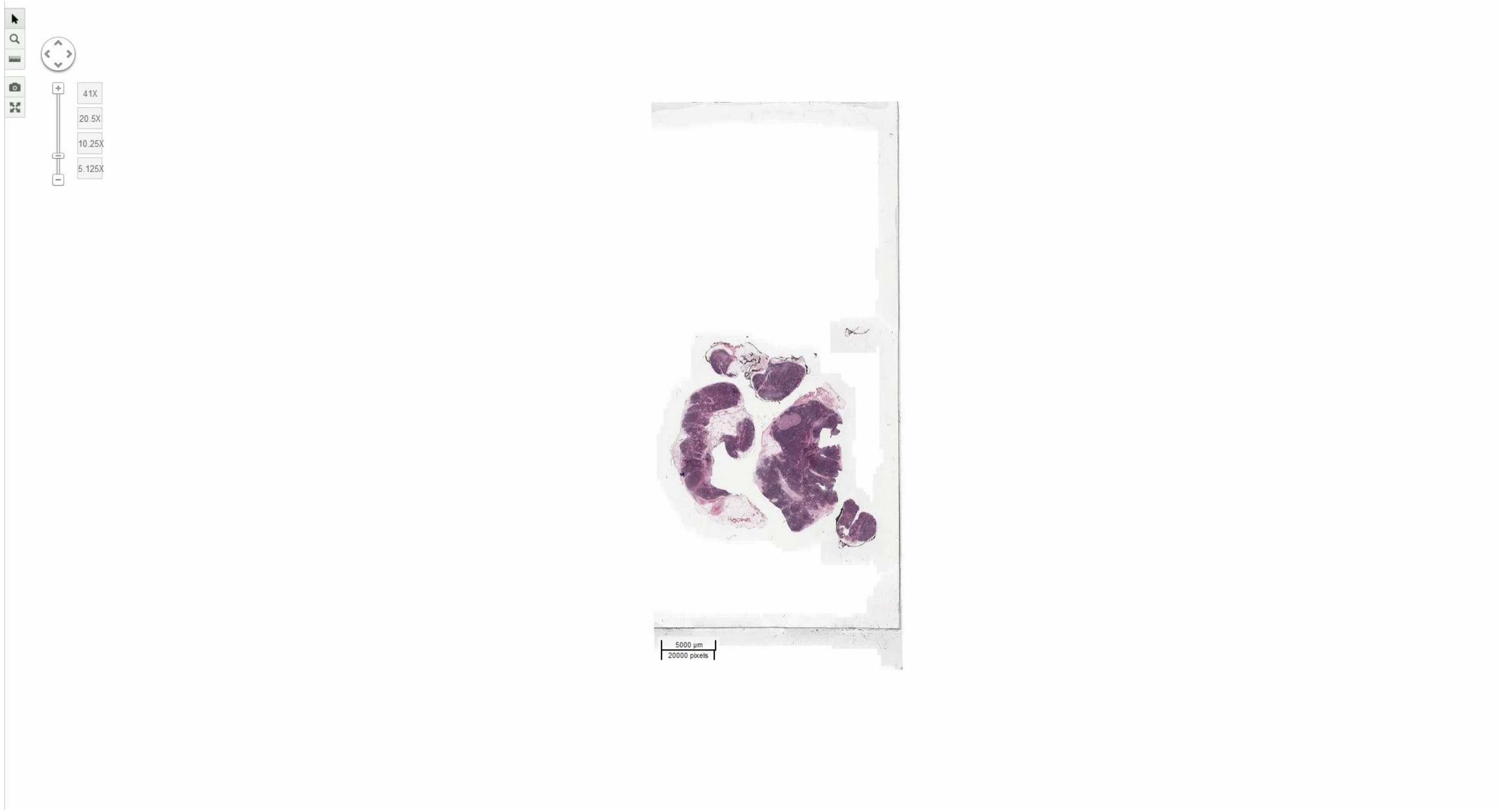
Bioimage informatics research group

We develop computational methods and tools using modern machine learning to quantify and extract information from biomedical images for diagnostics, decision making and for data-driven cancer research and medicine.

ACADEMY OF FINLAND
CSC
ERA PerMed
Syöpäsäätiö
Cancer Foundation Finland
SIGRID JUSELIUS STIFTELSE
EPIC MegaGrants RECIPIENT
InFLAMES
Solution is in Immunity

ComPatAI consortium

- ComPatAI:
  - Large scale AI-development for histopathology
  - Extreme scale access projects for LUMI-G (2023, 2024, 2025)
  - Research Council of Finland funded project ”Computationally intensive modeling of histopathology using generative and predictive AI”, 2024-2026 (Ruusuvuori & Latonen)
- People:
  - Assoc.Prof. Pekka Ruusuvuori, University of Turku, director
    - AI-development, generative AI, high-performance computing, computational pathology
  - Adj.Prof. Leena Latonen, University of Eastern Finland:
    - Virtual staining in histopathology, histopathology expertise, AI development
  - Adj. Prof. Teemu Tolonen, FIMLAB Laboratories & Tampere University Hospital
    - Pathology expertise, FIMLAB data access



One whole slide image: 97 000 x 220 000 pixels = **21.3 Gigapixels**

AI-based diagnostic systems with human expert level accuracy

Deep learning based detection and grading of cancer

UNIVERSITY OF TURKU

Can AI detect and grade prostate cancer?

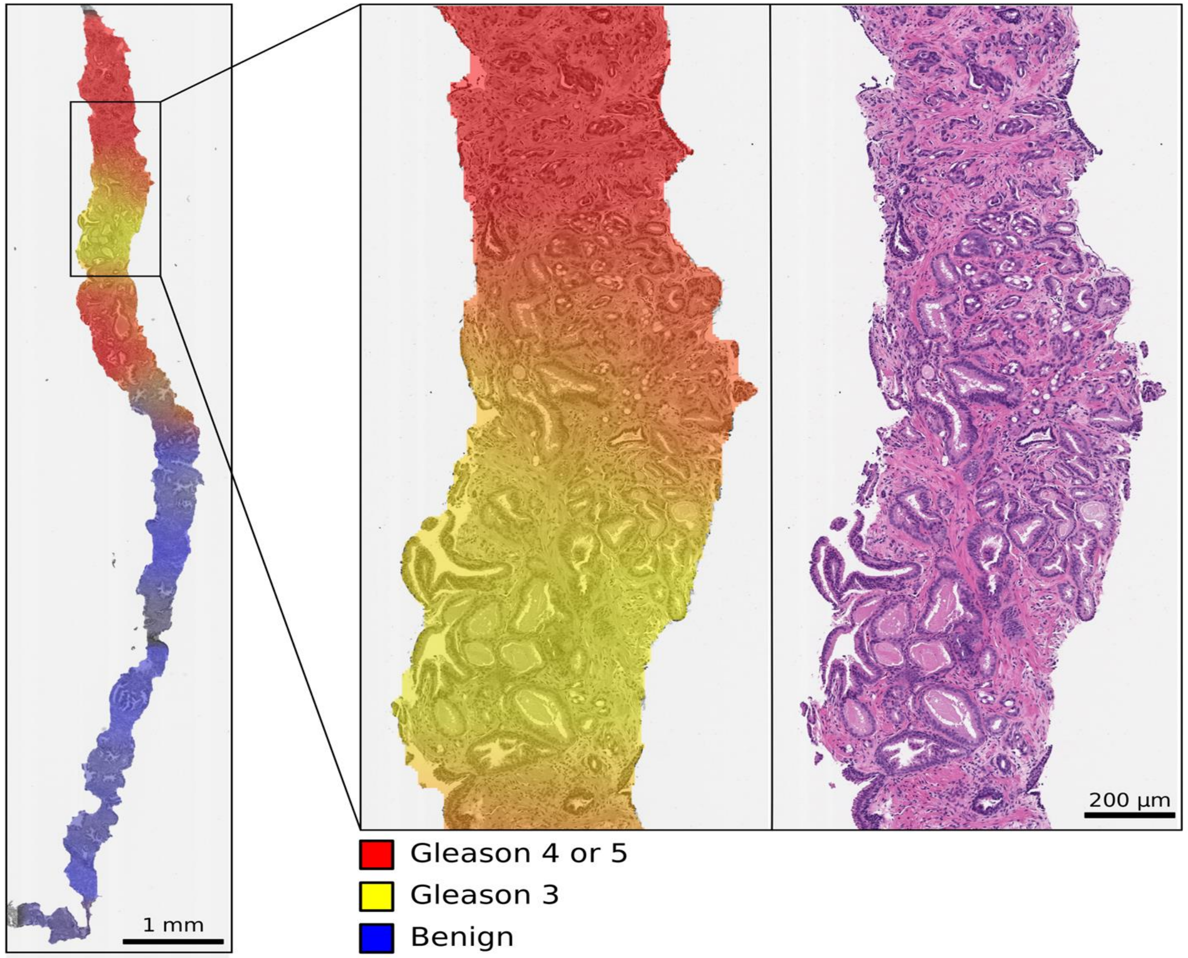

THE LANCET
Oncology

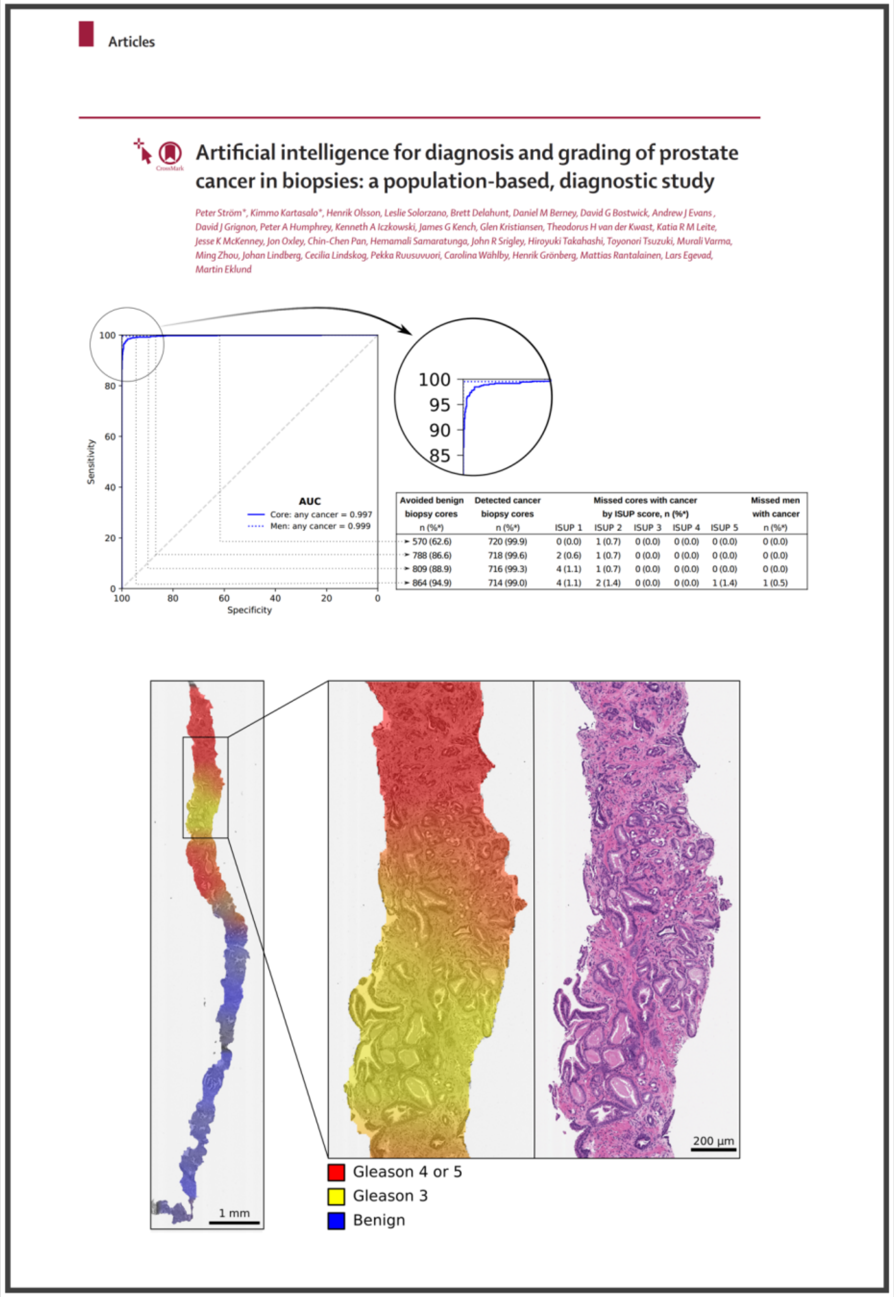
ARTICLES | VOLUME 21, ISSUE 2, P222-232, FEBRUARY 01, 2020

Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study

Peter Ström, MSc † • Kimmo Kartasalo, MSc † • Henrik Olsson, MSc • Leslie Solorzano, MSc • Prof Brett Delahunt, MD • Prof Daniel M Berney, MD • Prof David G Bostwick, MD • Andrew J Evans, MD • Prof David J Grignon, MD • Prof Peter A Humphrey, MD • Prof Kenneth A Iczkowski, MD • Prof James G Kench, MD • Prof Glen Kristiansen, MD • Prof Theodorus H van der Kwast, MD • Prof Katia R M Leite, MD • Jesse K McKenney, MD • Jon Oxley, MD • Chin-Chen Pan, MD • Prof Hemamali Samaratunga, MD • Prof John R Srigley, MD • Hiroyuki Takahashi, MD • Prof Toyonori Tsuzuki, MD • Murali Varma, MD • Prof Ming Zhou, MD • Johan Lindberg, PhD • Cecilia Lindskog, PhD • Pekka Ruusuvuori, PhD • Prof Carolina Wählby, PhD • Prof Henrik Grönberg, MD • Mattias Rantalainen, PhD • Prof Lars Egevad, MD • Martin Eklund, PhD • Show less • Show footnotes

Published: January 08, 2020 • DOI: https://doi.org/10.1016/S1470-2045(19)30738-7 • Check for updates



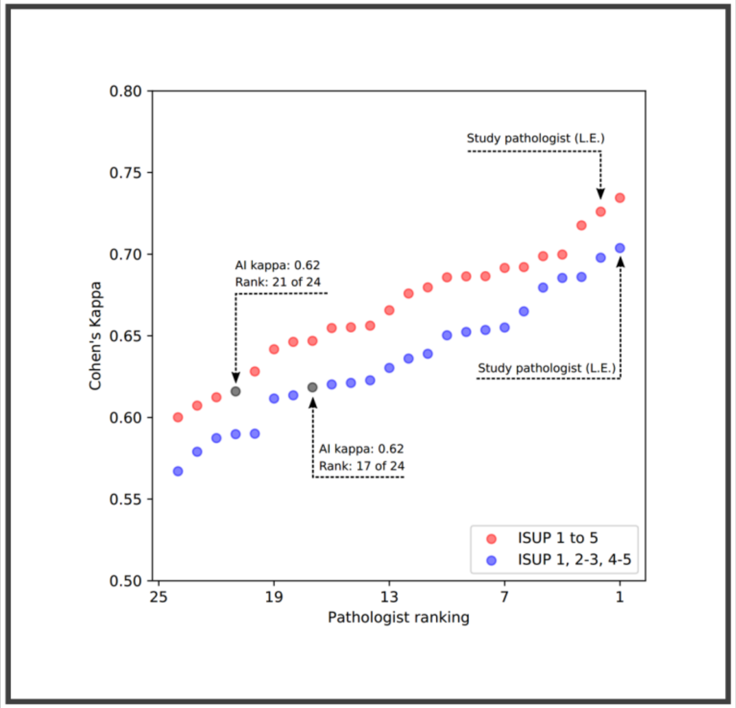


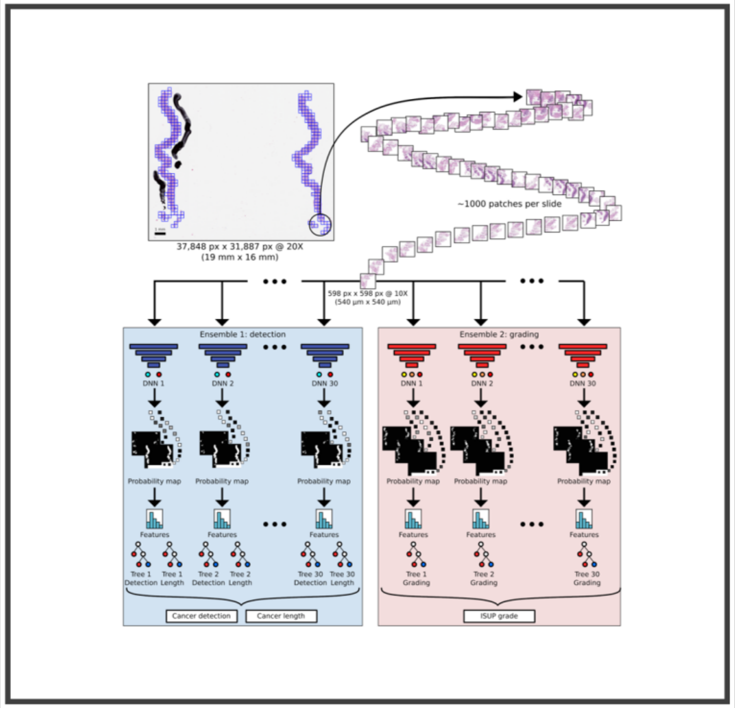
Articles

Artificial intelligence for diagnosis and grading of prostate cancer in biopsies: a population-based, diagnostic study

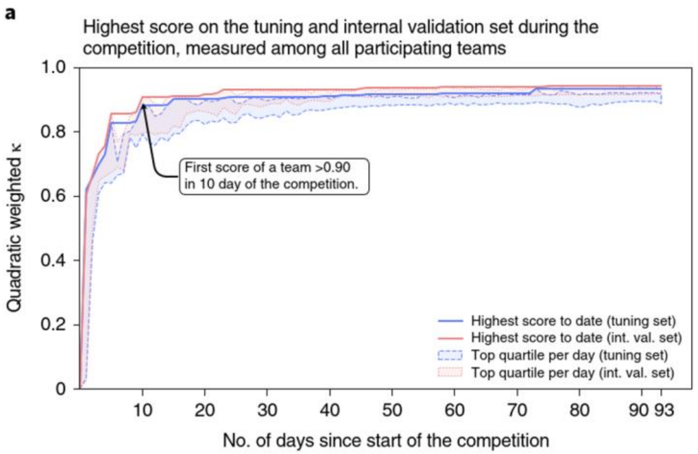
Peter Ström*, Kimmo Kartasalo*, Henrik Olsson, Leslie Solorzano, Brett Delahunt, Daniel M Berney, David G Bostwick, Andrew J Evans , David J Grignon, Peter A Humphrey, Kenneth A Iczkowski, James G Kench, Glen Kristiansen, Theodorus H van der Kwast, Katia R M Leite, Jesse K McKenney, Jon Oxley, Chin-Chen Pan, Hemamali Samaratunga, John R Srigley, Hiroyuki Takahashi, Toyonori Tsuzuki, Murali Varma, Ming Zhou, Johan Lindberg, Cecilia Lindskog, Pekka Ruusuvuori, Carolina Wählby, Henrik Grönberg, Mattias Rantalainen, Lars Egevad, Martin Eklund

Avoided benign biopsy cores n (%*)	Detected cancer biopsy cores n (%*)	Missed cores with cancer by ISUP score, n (%*) ISUP 1	ISUP 2	ISUP 3	ISUP 4	ISUP 5	Missed men with cancer n (%*)
570 (62.6)	720 (99.9)	0 (0.0)	1 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
788 (86.6)	718 (99.6)	2 (0.6)	1 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
809 (88.9)	716 (99.3)	4 (1.1)	1 (0.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
864 (94.9)	714 (99.0)	4 (1.1)	2 (1.4)	0 (0.0)	0 (0.0)	1 (1.4)	1 (0.5)





AI-system diagnoses prostate cancer from whole slide images on human expert level



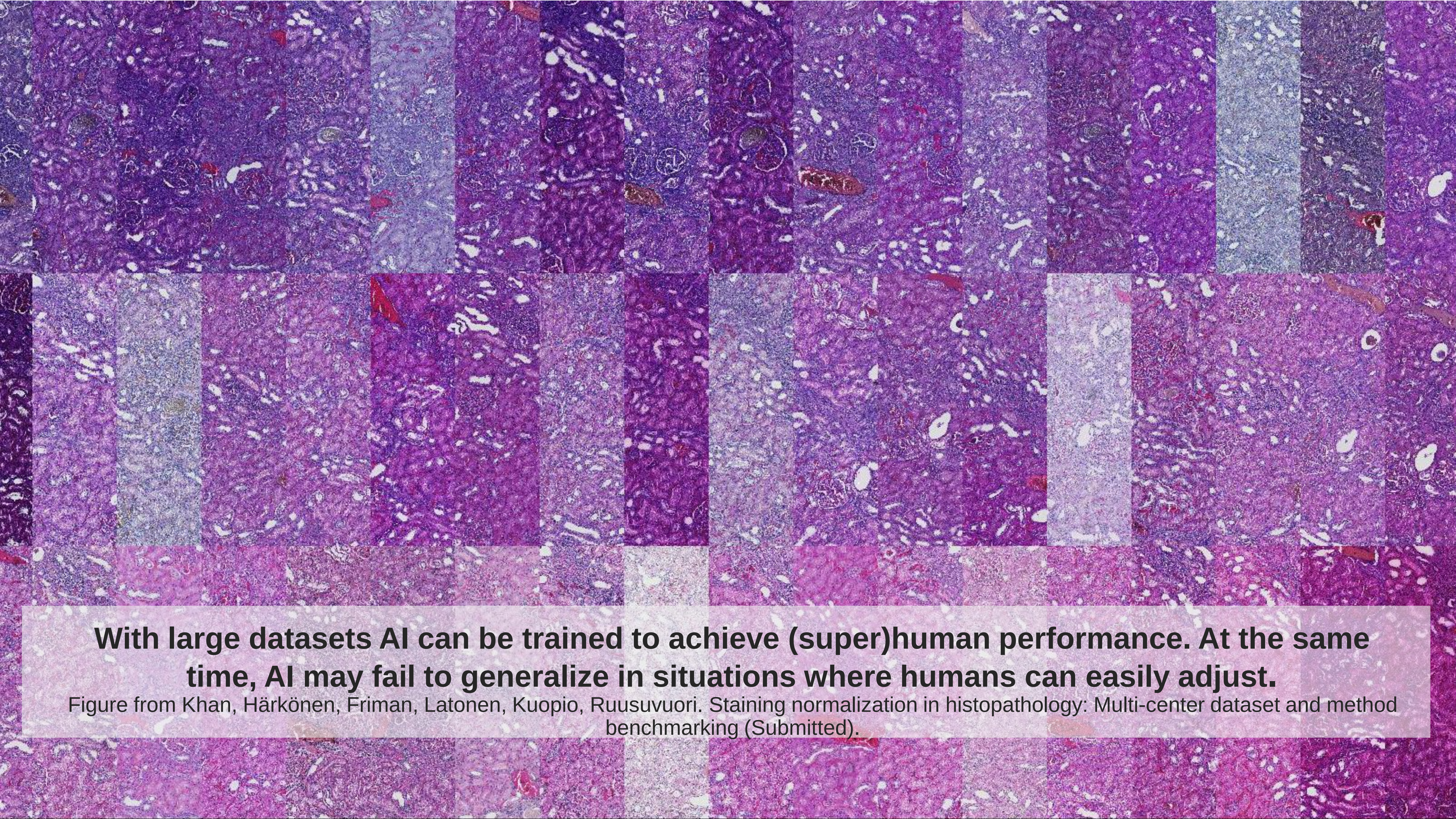
ARTICLES
https://doi.org/10.1038/s41591-021-01620-2
nature medicine
Check for updates
OPEN

Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge

Wouter Bulten[1,60], Kimmo Kartasalo[2,3,60], Po-Hsuan Cameron Chen[4,60], Peter Ström[2], Hans Pinckaers[1], Kunal Nagpal[4], Yuannan Cai[4], David F. Steiner[4], Hester van Boven[5], Robert Vink[6], Christina Hulsbergen-van de Kaa[6], Jeroen van der Laak[1,7], Mahul B. Amin[8], Andrew J. Evans[9], Theodorus van der Kwast[10], Robert Allan[11], Peter A. Humphrey[12], Henrik Grönberg[2,13], Hemamali Samaratunga[14], Brett Delahunt[15], Toyonori Tsuzuki[16], Tomi Häkkinen[3], Lars Egevad[17], Maggie Demkin[18], Sohier Dane[18], Fraser Tan[4], Masi Valkonen[19], Greg S. Corrado[4], Lily Peng[4], Craig H. Mermel[4], Pekka Ruusuvuori[3,19,61], Geert Litjens[1,61], Martin Eklund[2,61] and the PANDA challenge consortium*

PANDA
prostate cancer grade assessment

Crowdsourced AI development outperforms human experts in prostate cancer grading



With large datasets AI can be trained to achieve (super)human performance. At the same time, AI may fail to generalize in situations where humans can easily adjust.

Figure from Khan, Härkönen, Friman, Latonen, Kuopio, Ruusuvuori. Staining normalization in histopathology: Multi-center dataset and method benchmarking (Submitted).

Towards foundational AI models using population level digital pathology cohort

- Major Finnish hospitals have adopted fully digital workflow in pathology diagnostics
  - Turku & Tampere university hospitals use digital pathology in routine diagnostics:
    ~600 000 whole slide images annually
- Routine digitization of all biopsies leads to petabyte level data collections
- We currently have access to roughly ~1M WSI collection
- Can digital pathology data from routine diagnostics collected at hospitals be harnessed to build better diagnostic tools?

UNIVERSITY OF TURKU

Massive digital pathology data + high performance computing + AI

- Access to massive, high-quality digital pathology data
- State-of-the-art high-performance computing resources: LUMI supercomputer & CSC as a national service provider
- High level of computational data science & AI expertise
- Research Council of Finland and Cancer Foundation Finland fund the study

News

15.01.2024

New Finnish LUMI projects chosen: another step forward in cancer research, advances in multimodal AI across dozens of languages, and more

The fourth batch of Finnish research projects to utilize LUMI, Europe's fastest supercomputer, have been chosen, and will run on the system using its massive GPU capacity. The projects are, for example, continuing the development of computational pathology tools, advancing multimodal AI across dozens of languages, studying small-scale dynamos in stellar convection zones and much more.

< Back to search results

Granted Funding > Computationally intensive modeling of ...

Computationally intensive modeling of histopathology using generative and predictive AI

Acronym	ComPatAI
Description of the granted funding	Emergence of digital pathology has led to a leap in availability of digitalized whole slide images, providing a wealth of data for developing computational methods for interpreting the images. Realizing the full... Show more

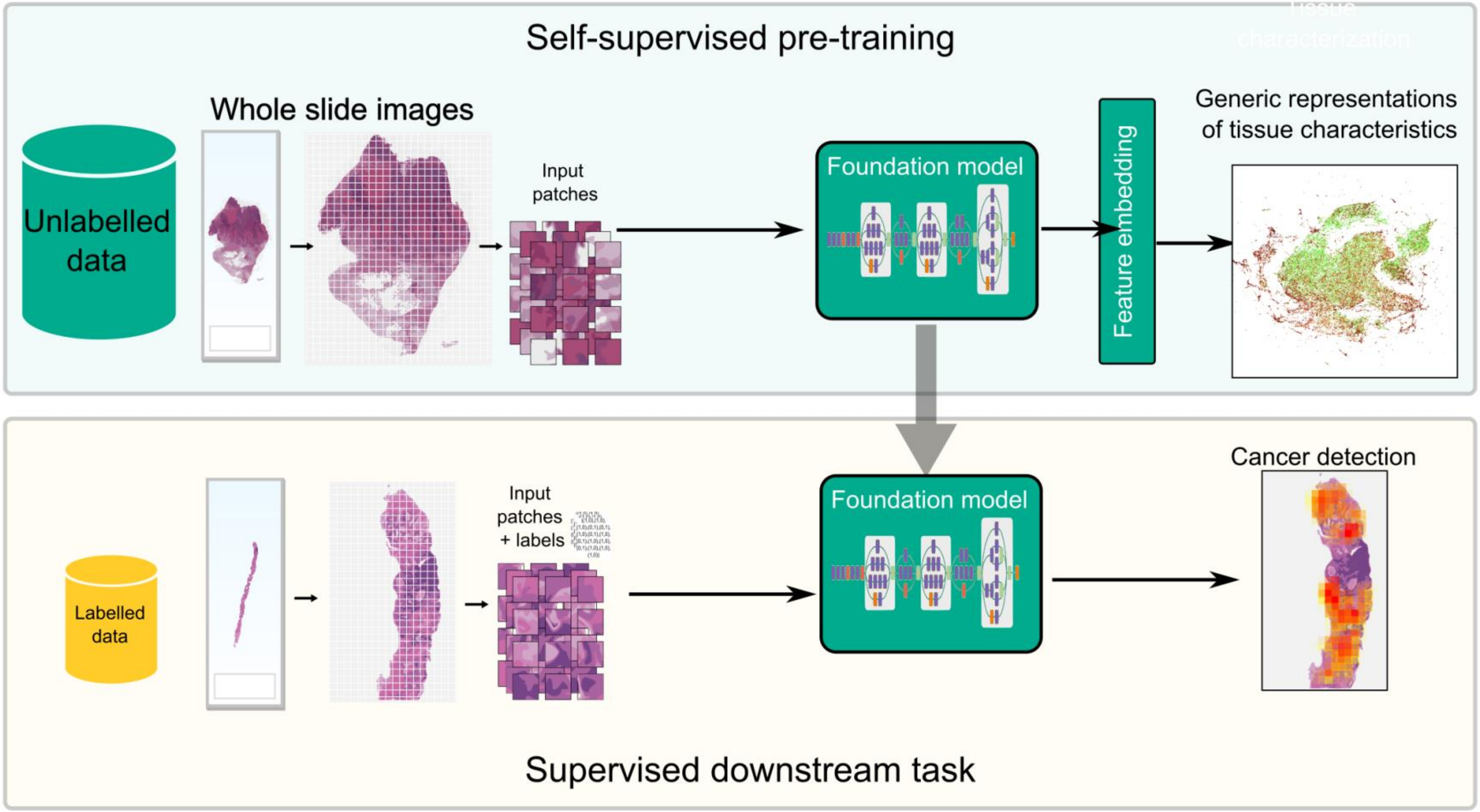
Syöpäsäätiö Cancer Foundation Finland

CSC

Research Council of Finland

UNIVERSITY OF TURKU

Towards generalizable AI for histopathology: foundation models using self-supervised learning (SSL)



Valkonen, M., Aho, A., Tolonen, T., Latonen, L., Kather, J., Ruusuvuori, P. Self-supervised learning architectures as the basis for foundation models in histopathology: A comparative study (Manuscript)

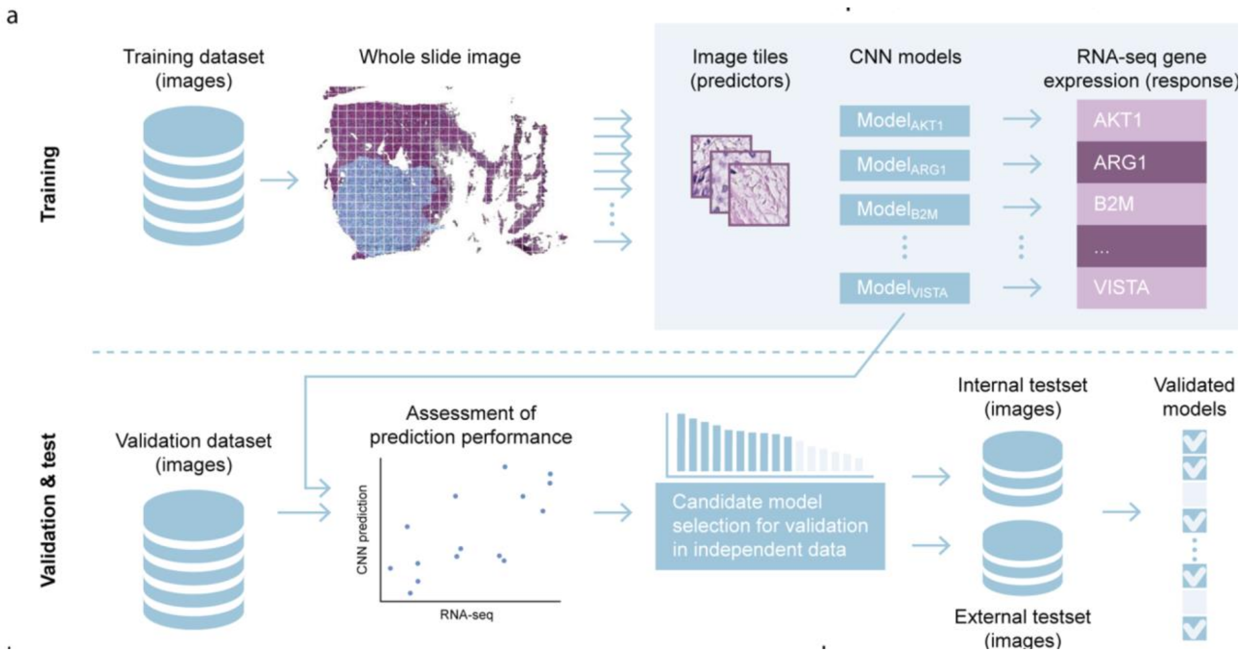
Combining pathology and genetics using AI

Prediction of gene expression from histopathology images

UNIVERSITY OF TURKU

Computational pathology beyond human vision: Gene expression prediction directly from histology

- CSC Grand Challenge Pilot Project
- Transcriptome-wide Expression-MOrphology (EMO) analysis in breast cancer
  - Individual deep convolutional neural networks (CNNs) were optimised to predict the expression of 17,695 genes from hematoxylin and eosin (HE) stained whole slide images (WSIs).
  - Ensembles of 30 CNN x 18 000 genes = huge computational effort



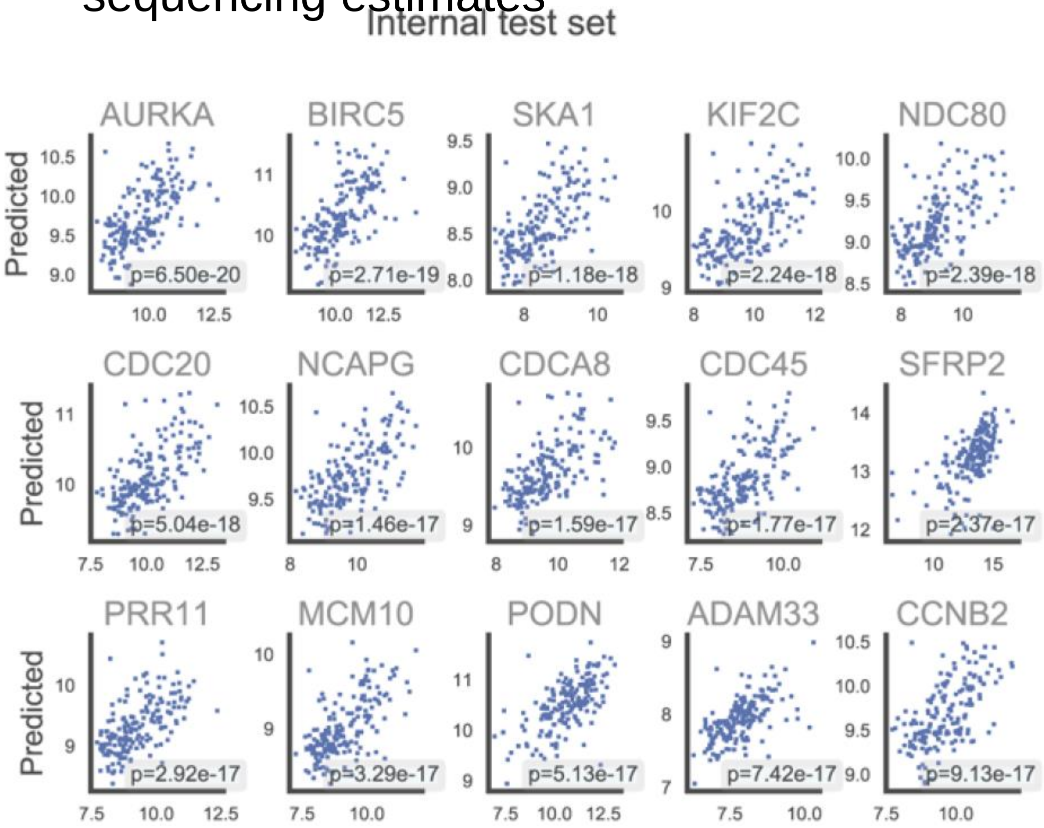

CANCER RESEARCH | CONVERGENCE AND TECHNOLOGIES

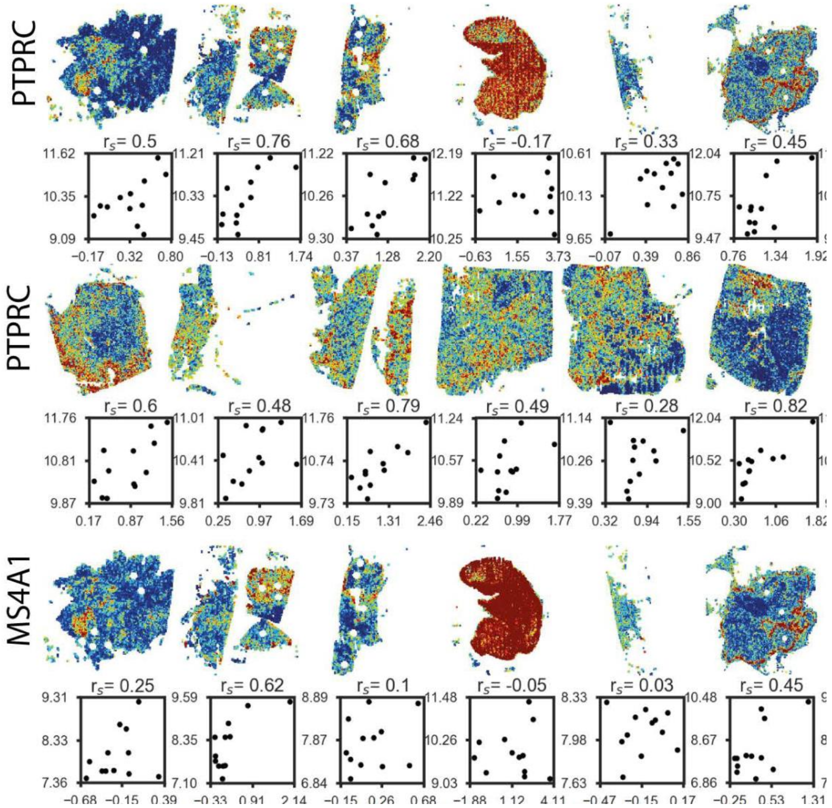
Predicting Molecular Phenotypes from Histopathology Images: A Transcriptome-Wide Expression–Morphology Analysis in Breast Cancer

Yinxi Wang[1], Kimmo Kartasalo[1,2], Philippe Weitz[1], Balázs Ács[3,4], Masi Valkonen[5], Christer Larsson[6], Pekka Ruusuvuori[2,5], Johan Hartman[3,4,7], and Mattias Rantalainen[1,7]

Computational pathology beyond human vision: Gene expression prediction directly from histology

- Prediction of both tumour average gene expression and intra-tumour spatial expression is possible directly from tissue morphology
- Predicted expressions in 9,334 (52.75%) genes were significantly associated with RNA sequencing estimates

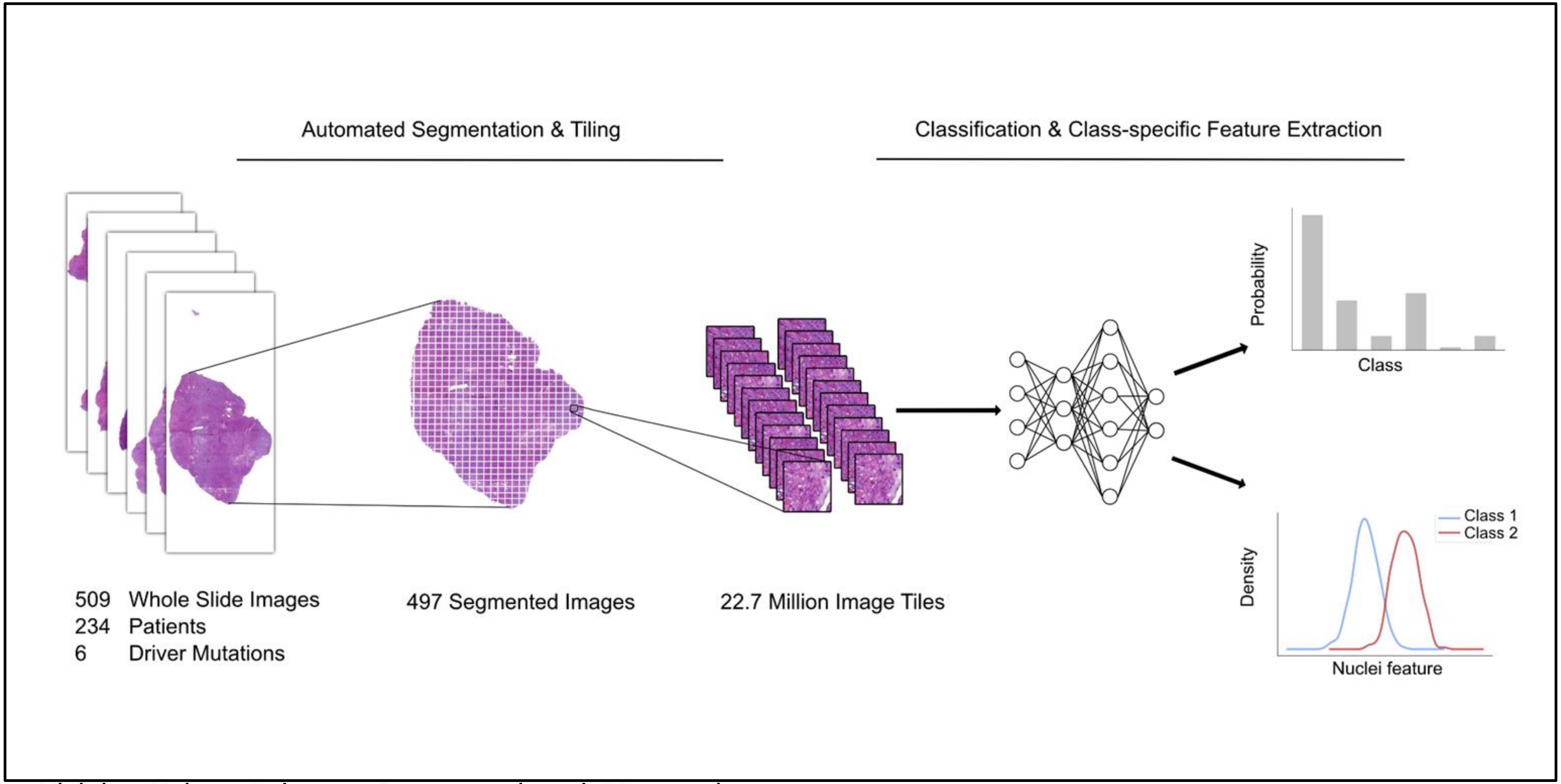




Computational pathology beyond human vision

Prediction of mutations from histopathology images

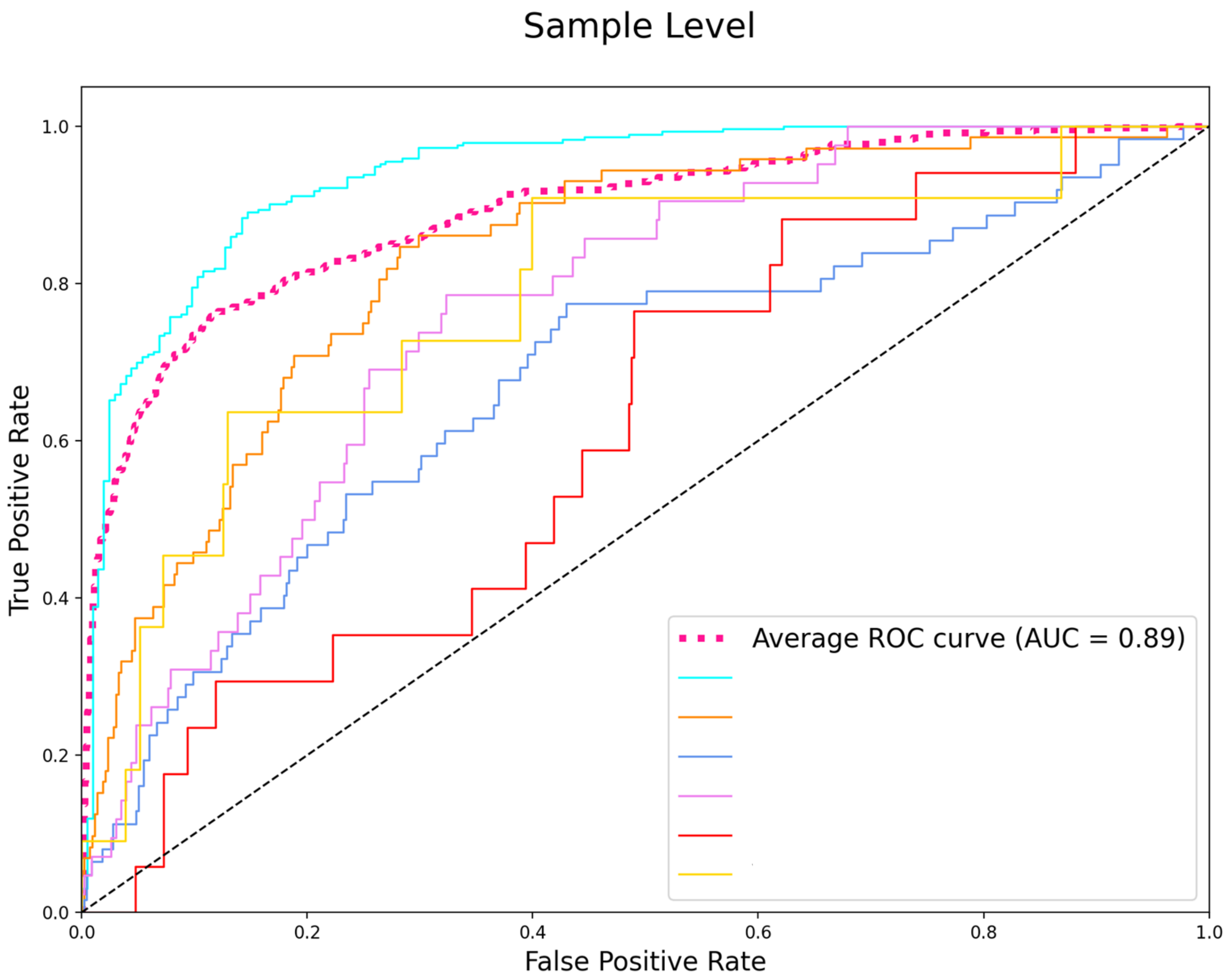
UNIVERSITY OF TURKU



Kiviaho, Nykter, Aaltonen, Ruusuvuori et al. Manuscript.

Predicting mutational status in Uterine Leiomyoma (Fibroids)

- Benign tumors of the uterine wall
- Affect one in four women
- Characterized by few driver mutations
- Good performance on mutations with reasonable sample size
- Less accurate for mutations with less samples
- Overall AUC 0.89



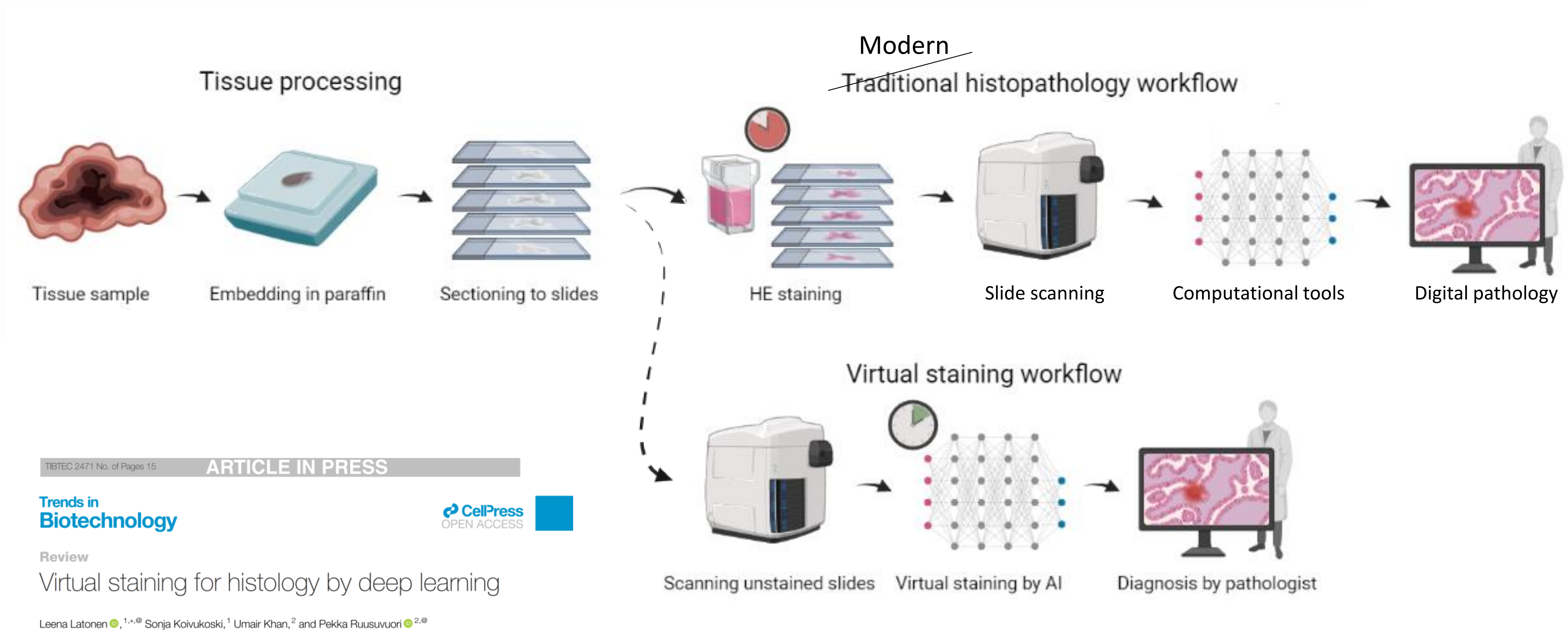
Kiviaho, Nykter (TUNI), Aaltonen (UH), Ruusuvuori et al. Manuscript.

Computational pathology beyond human vision

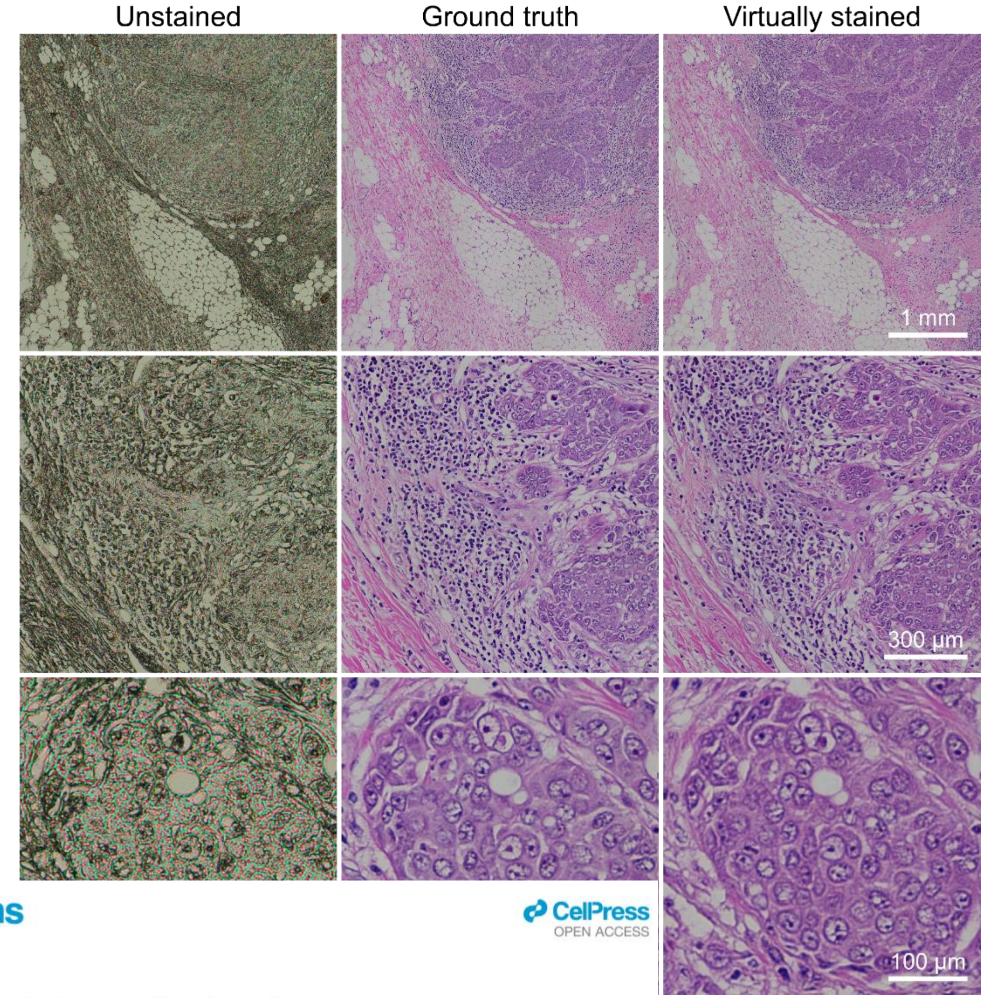
Virtual staining of unstained tissue using generative AI

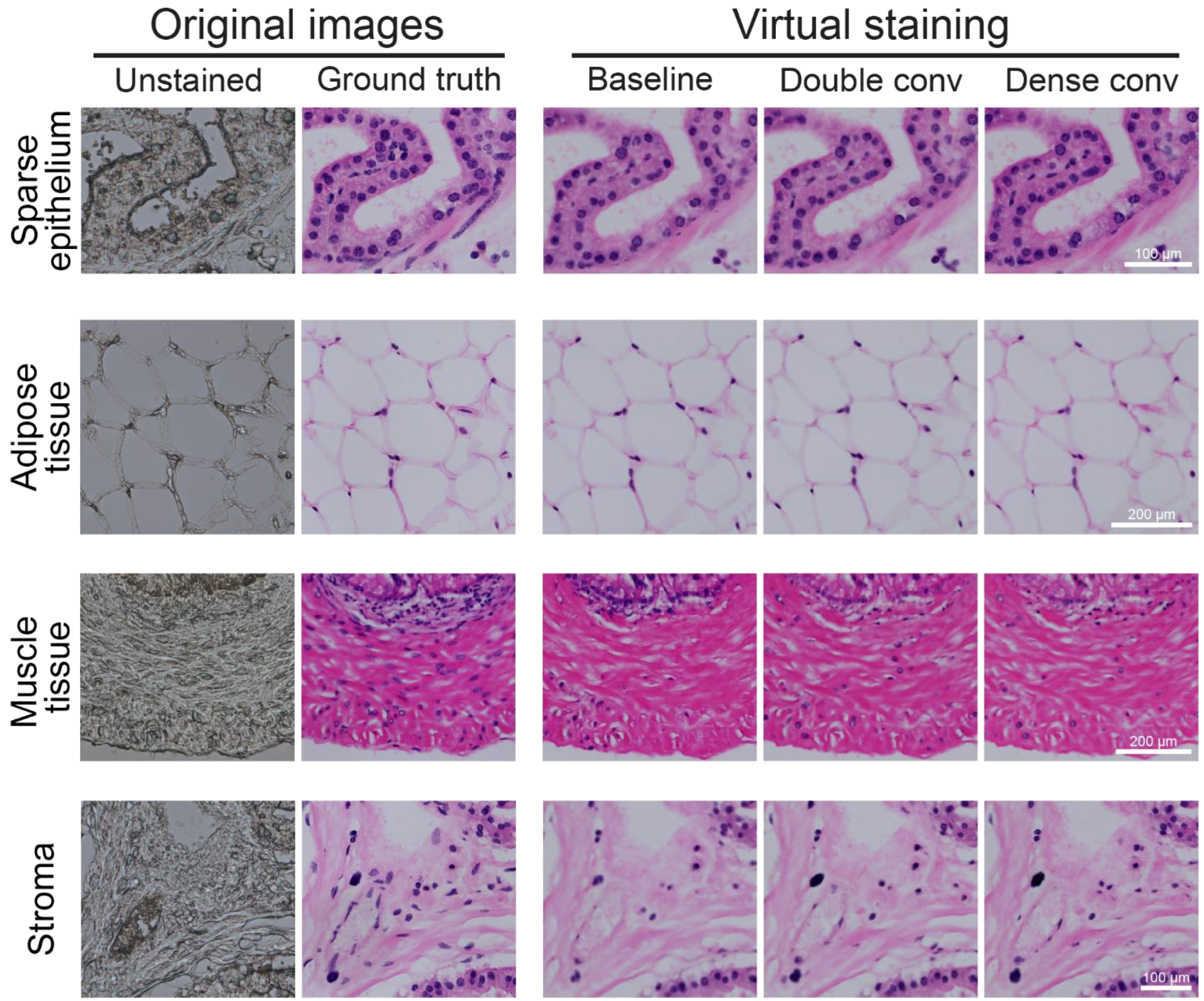
UNIVERSITY OF TURKU

Histopathological workflow - how much can be replaced virtually?



Virtual staining of unstained tissue using generative AI





Patterns

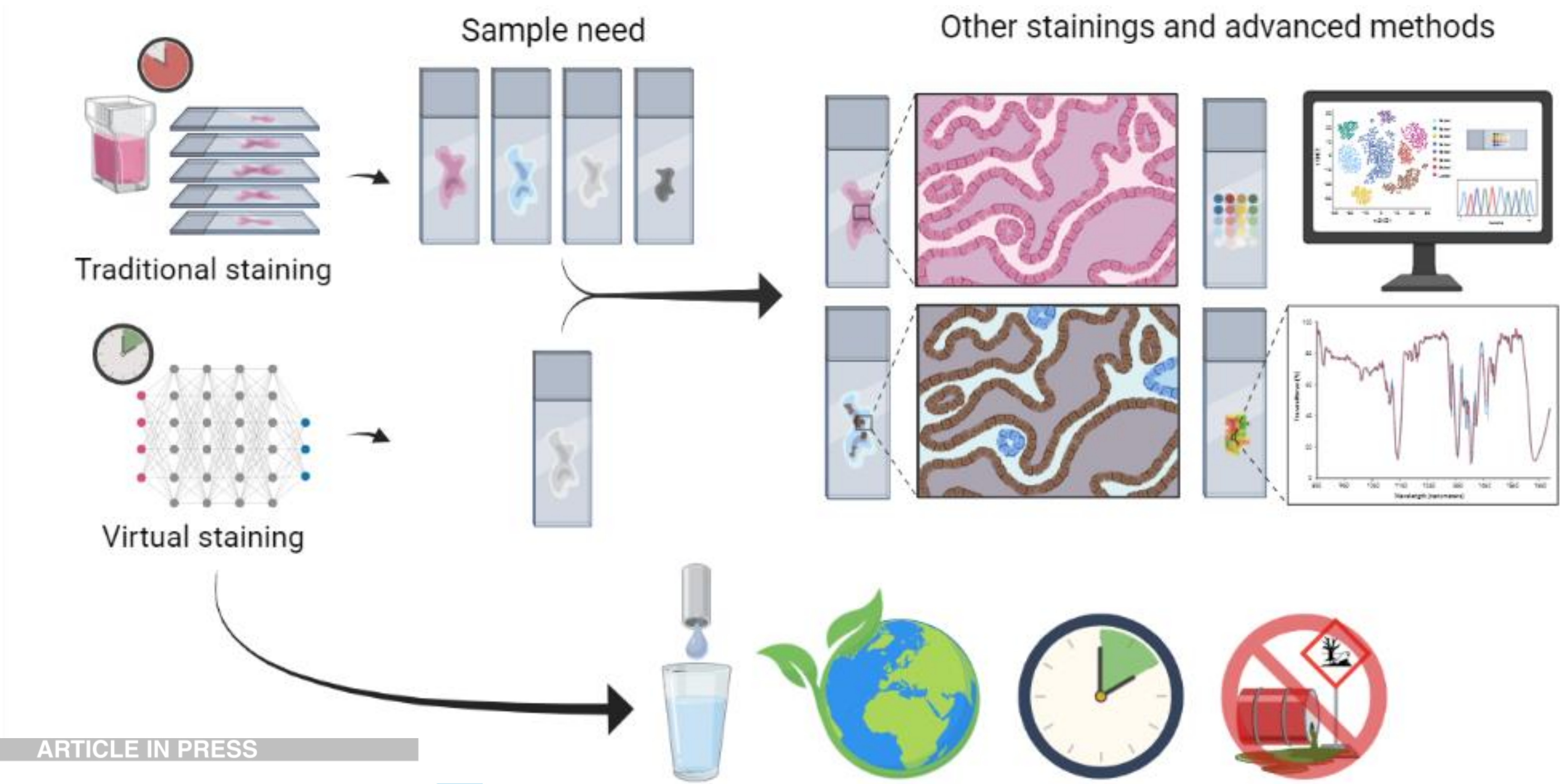
CelPress
OPEN ACCESS

Article

The effect of neural network architecture on virtual H&E staining: Systematic assessment of histological feasibility

Umair Khan,[1] Sonja Koivukoski,[2] Mira Valkonen,[3] Leena Latonen,[2,4] and Pekka Ruusuvuori[1,3,5,6,*]

Potential advantages of virtual staining



TIBTEC 2471 No. of Pages 15
ARTICLE IN PRESS
Trends in Biotechnology
CellPress OPEN ACCESS
Review

Virtual staining for histology by deep learning

Leena Latonen[1,*,@], Sonja Koivukoski,[1] Umair Khan,[2] and Pekka Ruusuvuori[2,@]

Translating research results for the benefit of society and economy

From innovations to start-ups, from novel methods to clinical trials

UNIVERSITY OF TURKU

Societal impact

- Two start-ups
  - Quva oy est. 2010, Louhi Health Data est. 2023
- IPR:
  - Two international patents, several (~10) invention reports
- Participating in four ongoing clinical trials where AI-systems are built for diagnostics, patient characterization and treatment outcome prediction

LUMI AI Factory

Speeding up development of AI-based decision support systems for biomedical imaging

UNIVERSITY OF TURKU

LUMI AI Factory – our wishlist for speeding up development cycle

- Massive scale means massive data transfer and computing – it takes time
  - Extreme-scale access projects for longer periods
  - Continued funding support
- Regulation and data access
  - Speed-up in handling ethical/data access/tissue processing permissions
  - For example, our timeline for tackling permissions >2yrs
  - Storage of diagnostic data – 2 years? Longer – who covers the cost?

UNIVERSITY OF TURKU

Thank you!

UNIVERSITY OF TURKU

pekka.ruusuvuori@utu.fi

@PekkaRuusuvuori