

Fundamentals of Computational Pathology: Part I

Nasir Rajpoot

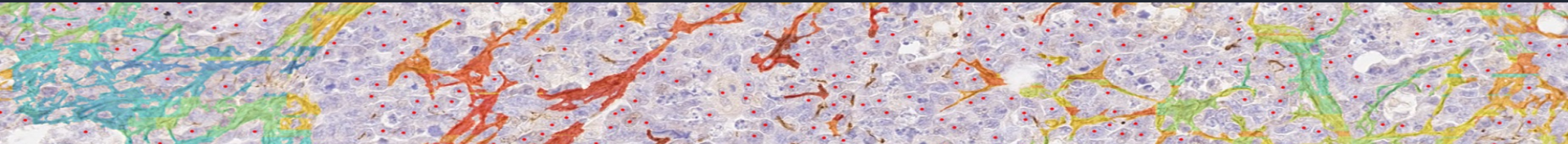
Professor of Computational Pathology

Wolfson Fellow, The Royal Society

Director, Tissue Image Analytics (TIA) Centre

University of Warwick, UK

Croucher ASI AI for Biomedicine (Fri 28 Nov 2025)



Computer Science at Warwick



TIA Centre Objectives



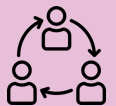
Research

- Hub for AI/ML research on tissue/multi-omic based cancer diagnostics, prognostics and therapeutics



Development

- Clinical-grade AI solutions
- Translation & commercialisation



Outreach & Collaboration

- Patient & Public / Industrial & Clinical outreach
- External seminar series, training workshops



The Practice of Medicine is Subjective



Practice of Medicine is Subjective

Medicine deals with making **judgements** and **predictions**, for the management of various aspects of patient's health (or disease)

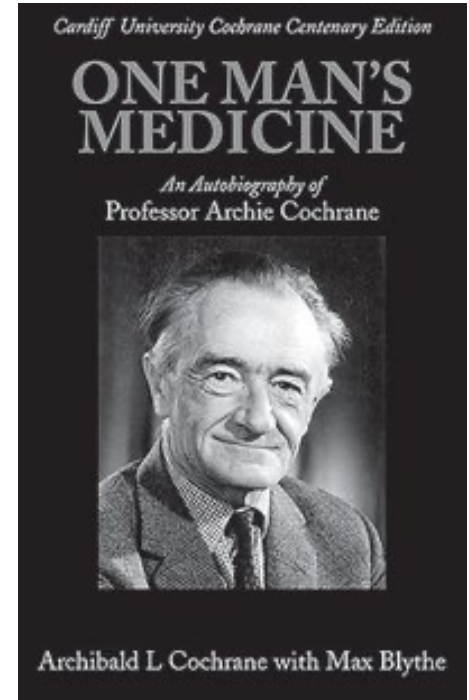
- **Diagnosis** → predict/guesstimate a subject's medical condition
- **Prognosis** → predict the prognosis of a cancer patient based on the stage of the disease and the patient's overall health
- **Chronic disease management** → predict future blood sugar levels and adjust medication and lifestyle recommendations
- **Treatment planning** → predict treatment outcomes
- **Preventive medicine** → predicting risk of disease for a given subject





Scientific Validation is a Must

- Physicians kept doing what they did because they had **always done it that way**
- They didn't need scientific validation. **They just knew!**
- Cochrane despised this attitude. He called it "**the God complex**"





Science Begins with Counting



Siddhartha Mukherjee
Pullitzer Prize Winner

“To understand a phenomenon, a scientist must **describe** it .. **objectively**. If cancer medicine was to be transformed into a rigorous science, then cancer would need to be counted somehow – **measured in some reliable, reproducible way.**”

The Emperor of All Maladies



22 June 2023



Histopathology Basics

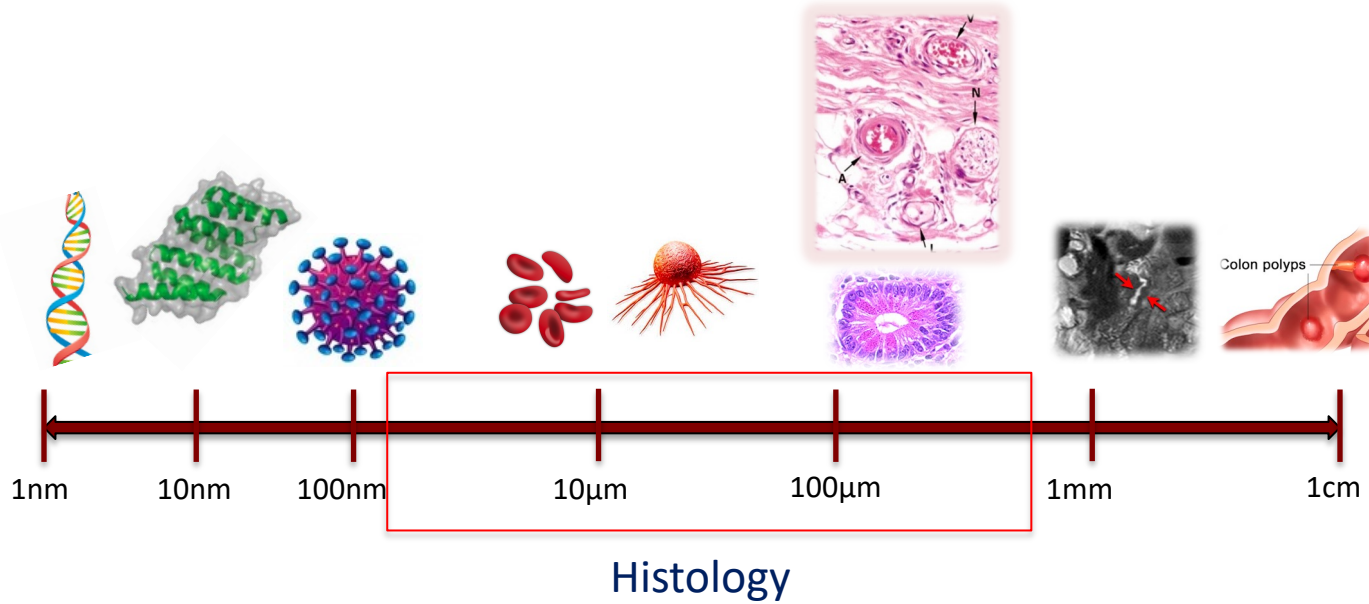
(Histo)Pathology Basics

- **Histopathology** → Science of studying disease by **analyzing tissue** at the **microscopic** scale
- This involves:
 - Tissue slide preparation
 - Slide examination
 - Diagnostic reporting

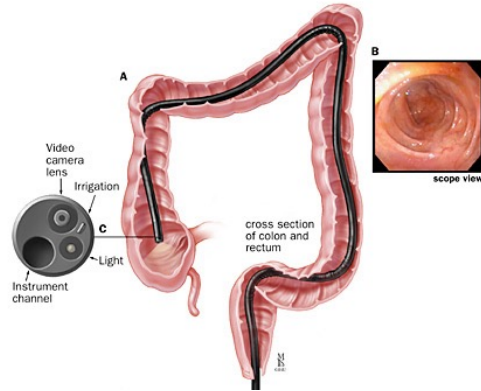




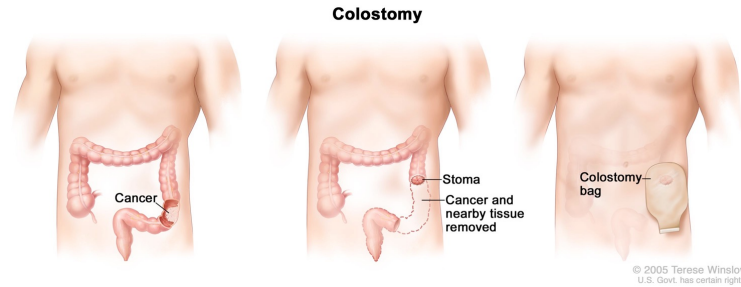
Oncology Length-Scales



Colon Tissue Extraction



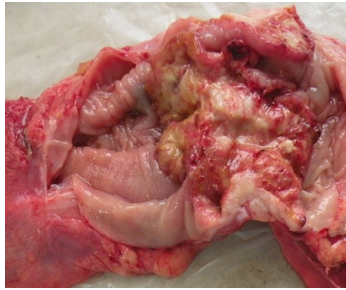
Biopsy



Surgical Resection

Tissue Slide Preparation

Gross tissue specimen



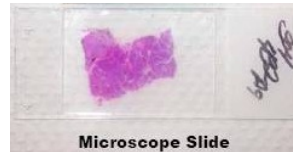
Fixation & embedding
of the tissue



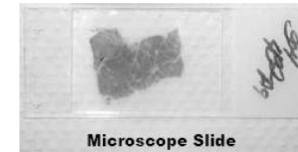
Cut thin ($3-5\mu\text{m}$)
sections



Send to the lab
for examination



Slide staining



Paste tissue onto the slide

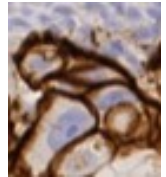
Histopathology Slide Examination

- The ***Gold standard*** for cancer diagnosis and grading of solid tumours
- Visual examination of the slide analyses tissue morphology and variation from normal cues of tissue morphology
- Morphology (architecture, structure, shape, texture) of the tissue at the microscopic resolution

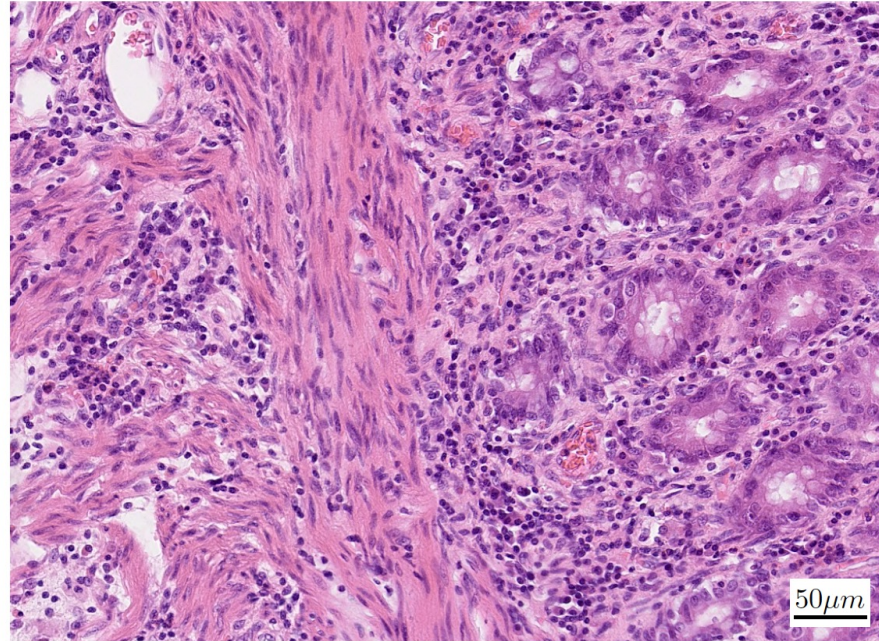


Slide Staining

- Cells and other extracellular material making up most tissue specimens are colorless
- Hematoxylin & Eosin (**H&E**) are the most commonly used stains:
 - **Hematoxylin** → chromatin rich nuclei as dark blue
 - **Eosin** → cytoplasm as pink
 - **Limitations**: H&E staining makes it difficult to distinguish between different kinds of nuclei or to assess the expression of various proteins/genes
- For that reason, immunohistochemical (**IHC**) stains are used



H&E Staining



Cell nuclei stained by Hematoxylin appear dark blue or purple. Cytoplasm regions of cells and collagen in connective tissue appear pink as they are stained by Eosin

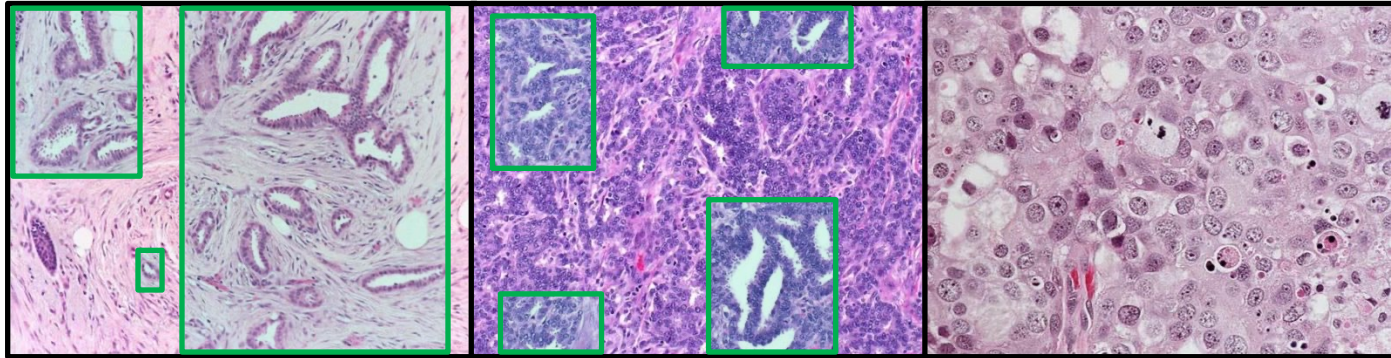
Breast Cancer Grading

Criteria	Score	Description
Gland Formation	1	more than 75% of the tumour forms gland
	2	10–75% of the tumour forms gland
	3	less than 10% of the tumour forms gland
Nuclear Atypia	1	small, regular and uniform nuclei
	2	moderate increase in size and variability
	3	marked variation
Mitosis Count	1	0–9 mitosis in 10 HPF
	2	10–19 mitosis in 10 HPF
	3	greater than 19 mitosis in 10 HPF

The Nottingham BCa Grading System

Grade I (score 3-5), Grade II (score 6-7), Grade III (score 8-9)

Nottingham BCa Grading (1)

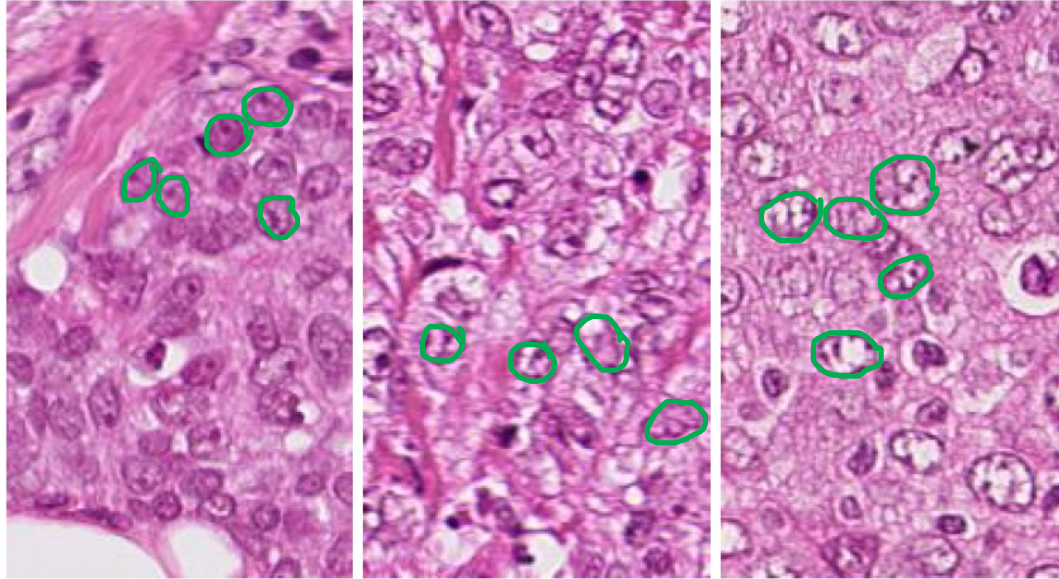


Source: <https://pathology.jhu.edu/breast/staging-grade/>

Gland formation

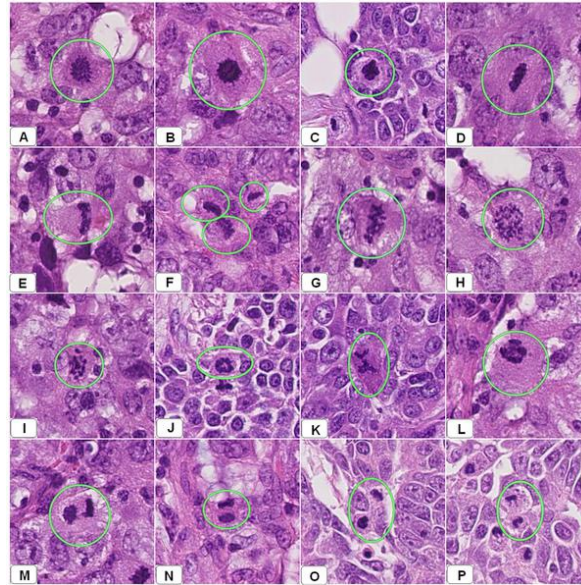
Left to right: low-grade, medium-grade and high-grade tumour examples

Nottingham BCa Grading (2)



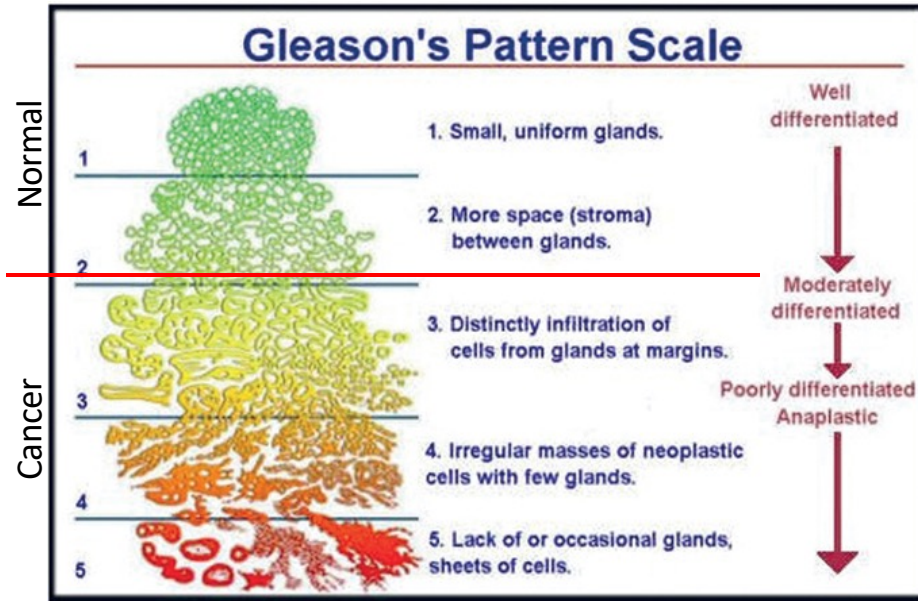
Left to Right: Nuclear atypia/pleomorphism score 1-3

Nottingham BCa Grading (3)



Mitotic cells

Prostate Cancer Grading



Source: John Murtagh, Jill Rosenblatt, Justin Coleman, Clare Murtagh: *John Murtagh's General Practice*, 7e
Copyright © McGraw-Hill Education. All rights reserved.

Gleason Score = Primary Score (largest tumor area) + Secondary Score (2nd largest tumour area)

Prostate Cancer Grading

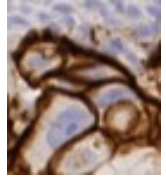
Gleason Score	Grade Group	Characteristics
6	Grade Group 1	• Less aggressive • Very slow growing • Low risk
3 + 4 = 7	Grade Group 2	• Slightly aggressive • Slow growing • Low to Intermediate risk
4 + 3 = 7	Grade Group 3	• Moderately aggressive • Fast growing • Intermediate to High risk
8	Grade Group 4	• Aggressive • Rapidly growing • High risk
9-10	Grade Group 5	• Highly aggressive • Rapidly growing • High risk

$$3+4 \neq 4+3$$

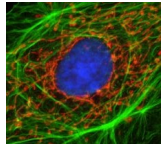
Gleason Score = Primary Score (largest tumor area) + **Secondary Score** (2nd largest tumour area)

IHC Staining

- **H&E Limitations**: It is difficult to distinguish between different kinds of nuclei or to assess the expression of various proteins/genes
- **IHC/IF** markers (antibodies) are used to identify the presence and concentration of certain proteins in different parts of the tissue
 - Antibody reaction is visualized using an additional chemical, typically 3,3'-Diaminobenzidine (**DAB**), which stains the reactive area a shade of **brown**
 - Usually, a **Hematoxylin** counterstain is also used to simultaneously identify cell nuclei as a shade of **blue**



IHC



IF

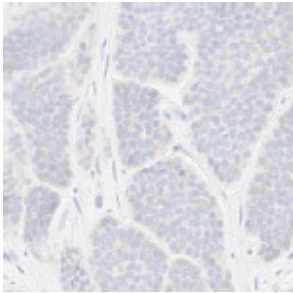
IHC Staining – Binding Locations

- IHC markers may be expressed in particular cell **organelles**, such as the nucleus, cytoplasm, cell membrane or a combination of those
- **Major Advantage of IHC Staining:** Allows the observer a greater insight into the chemical makeup of tissue
 - Composition of tumor/immune cells
 - Used directly and indirectly for determining the degree of malignancy

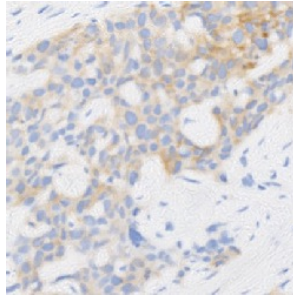


IHC Example – HER2 in BCa

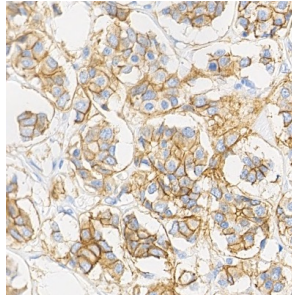
- Human epidermal growth factor receptor 2 (HER2) is a protein that mainly influences the growth of malignant epithelial cells in breast cancer
- Used as predictive test for targeted anti-HER2 therapy such as Herceptin (*Trastuzumab*)



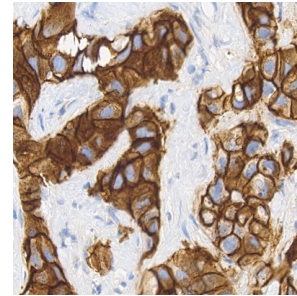
Grade 0



Grade 1



Grade 2

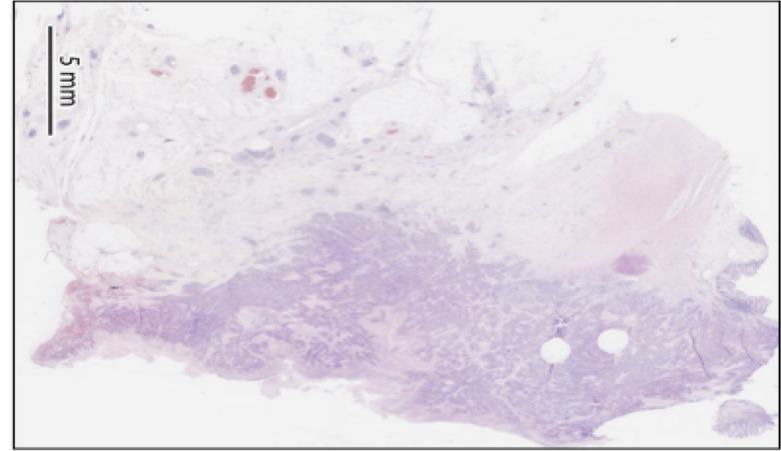


Grade 3

Introduction to Digital & Computational Pathology



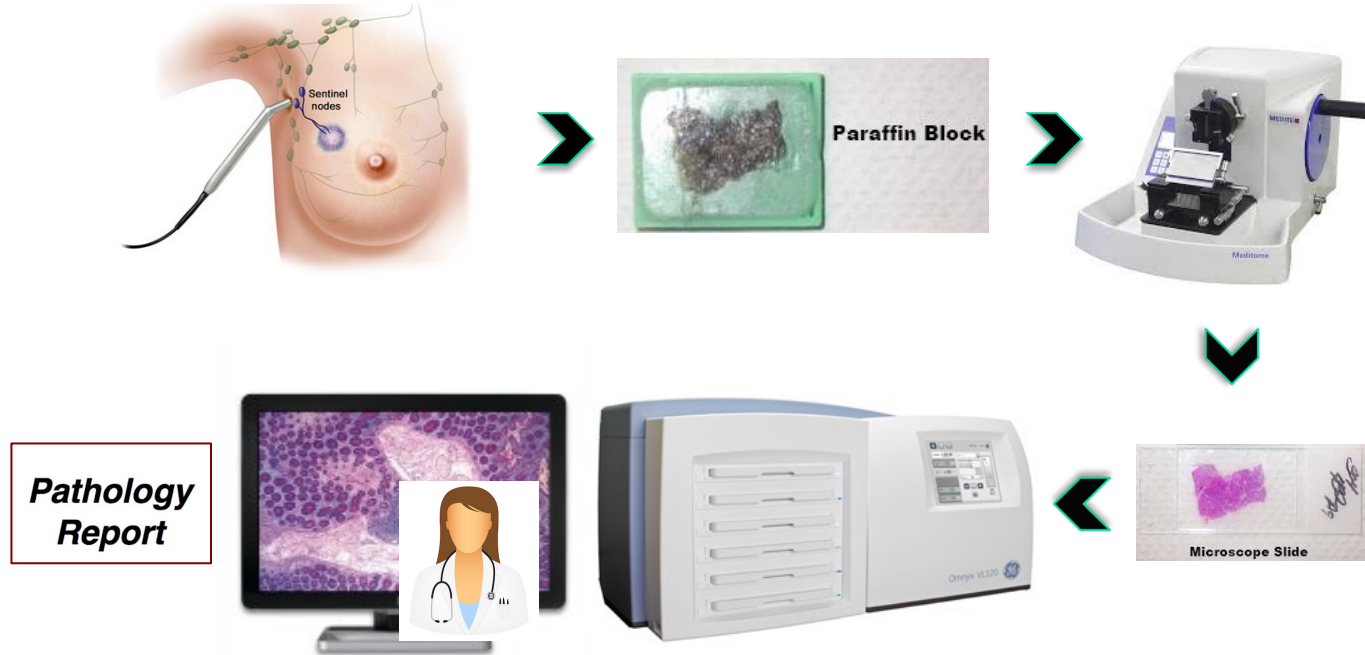
Cellular Pathology Goes/Going Digital



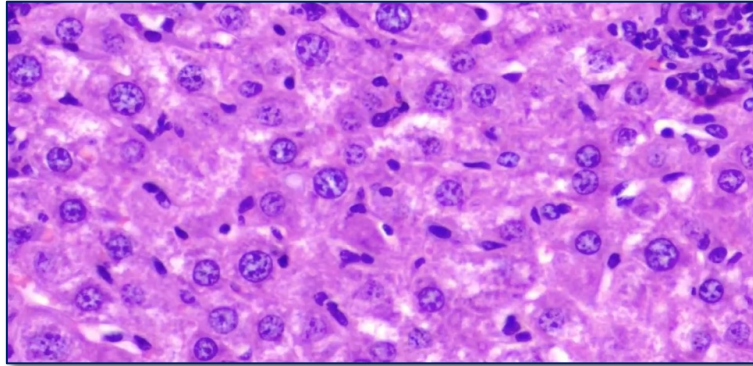
Typically, 25mm x 75 mm slide
 $0.25\mu\text{m}$ / pixel $\rightarrow$ Whole-Slide Image (WSI)

Digital Pathology

Digital Pathology → Digitization of tissue slides into whole-slide images

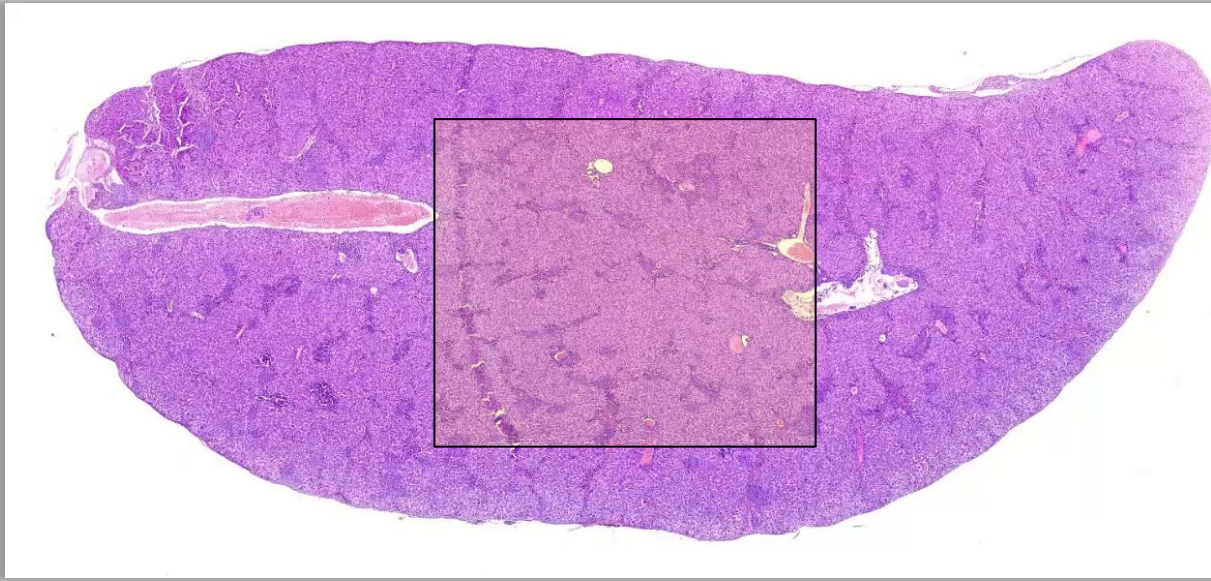


The Promise of Digital Pathology



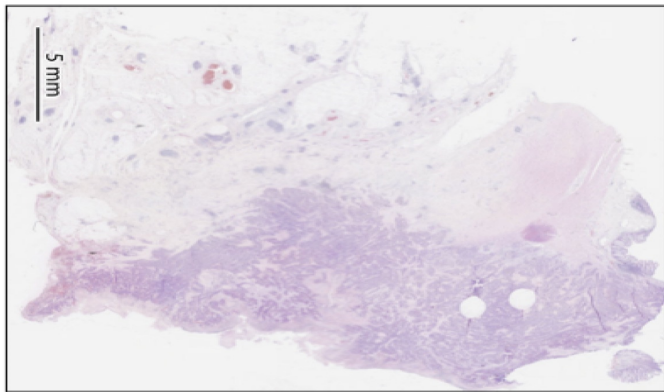
- As well as allowing the pathologists to provide **2nd opinion** and do **remote diagnostics** in a seamless manner than the conventional glass-slide based diagnostics, it allows analysis of digitised images by **computer algorithms**
- Digital pathology provides **spatial** (*aka* positional) **information** at microscopic resolution, important in the context of tumour heterogeneity

From Digital to Computational Pathology



Computational Pathology → Science of studying disease by analyzing digital images of tissue slides

Prime Time for Computational Pathology



25mm x 75 mm slide

0.25 μ m / pixel $\rightarrow$ 100K x 300K pixels $\rightarrow$ 90GB

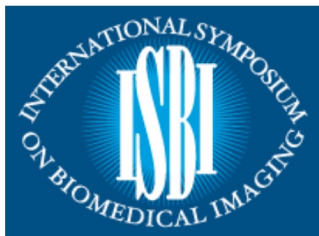
85,000+ chest X rays (1,024x1,024 pixels)

12K slides $\rightarrow$ >1PB raw pixel data



Opportunity for AI algorithms for
Faster Diagnosis & Personalized Treatment

CHIP @ ISBI'2008



ISBI'2008 Special Session on
Computational HistoPathology: Advances and New Challenges
(CHIP@ISBI2008)

14:00-16:00 on Thursday, 15 May 2008

(Paris, France)

NEW **Final Program**

Session Type: Special Oral; Location: La Seine B

Organisers:

Nasir Rajpoot

(University of Warwick, UK)

Tim Nattkemper

(Bielefeld University,
Germany)

Metin Gurcan

(Ohio State University, USA)

http://www2.warwick.ac.uk/fac/sci/dcs/research/combi/chip_isbi08/

Basic Problems (and Solutions) in Computational Pathology

Blur Detection & Estimation

Focus Issues

- Modern scanners are pretty good at producing good quality, high resolution WSIs
- However, the scanning process can produce images that are out of focus in some areas:
 - Many algorithms perform poorly on such images

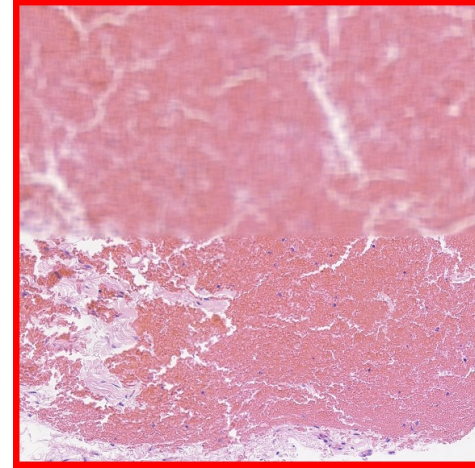
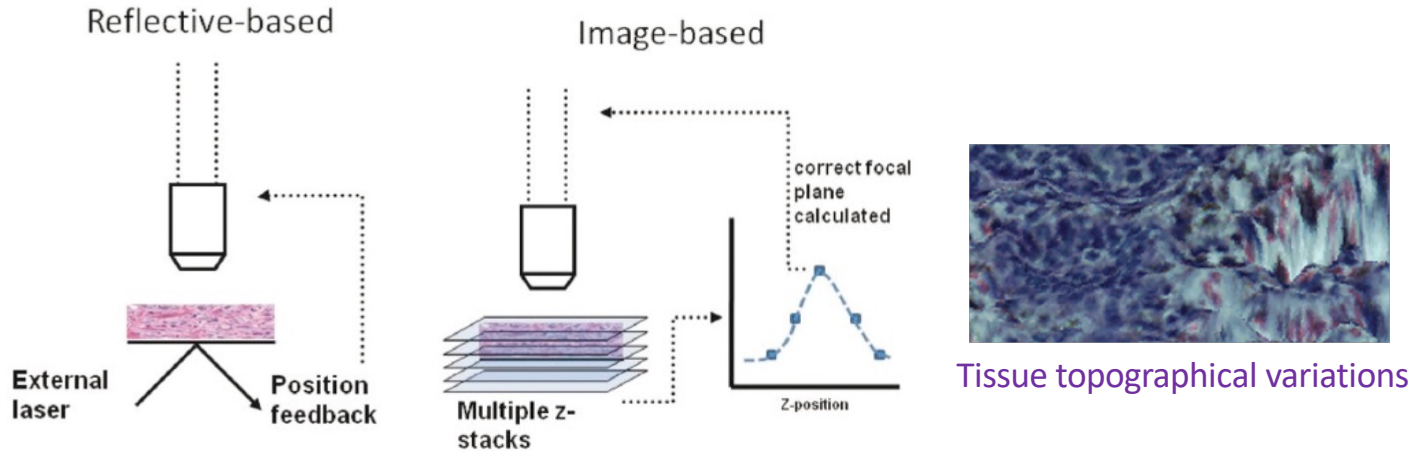


Image Quality – Autofocus



Source: Montalto et al., J Path Inf 2011

Blur Estimation

- Principle:
 - In-focus images have clearly defined edges
 - Out-of-focus images have no hard edges
- The ***blur effect**** method
 - Artificially blur the image
 - Blur measure → Compare the intensity of each pixel with its neighbors before and after blurring the image

* Crete et al., SPIE Electronic Imaging, 2007



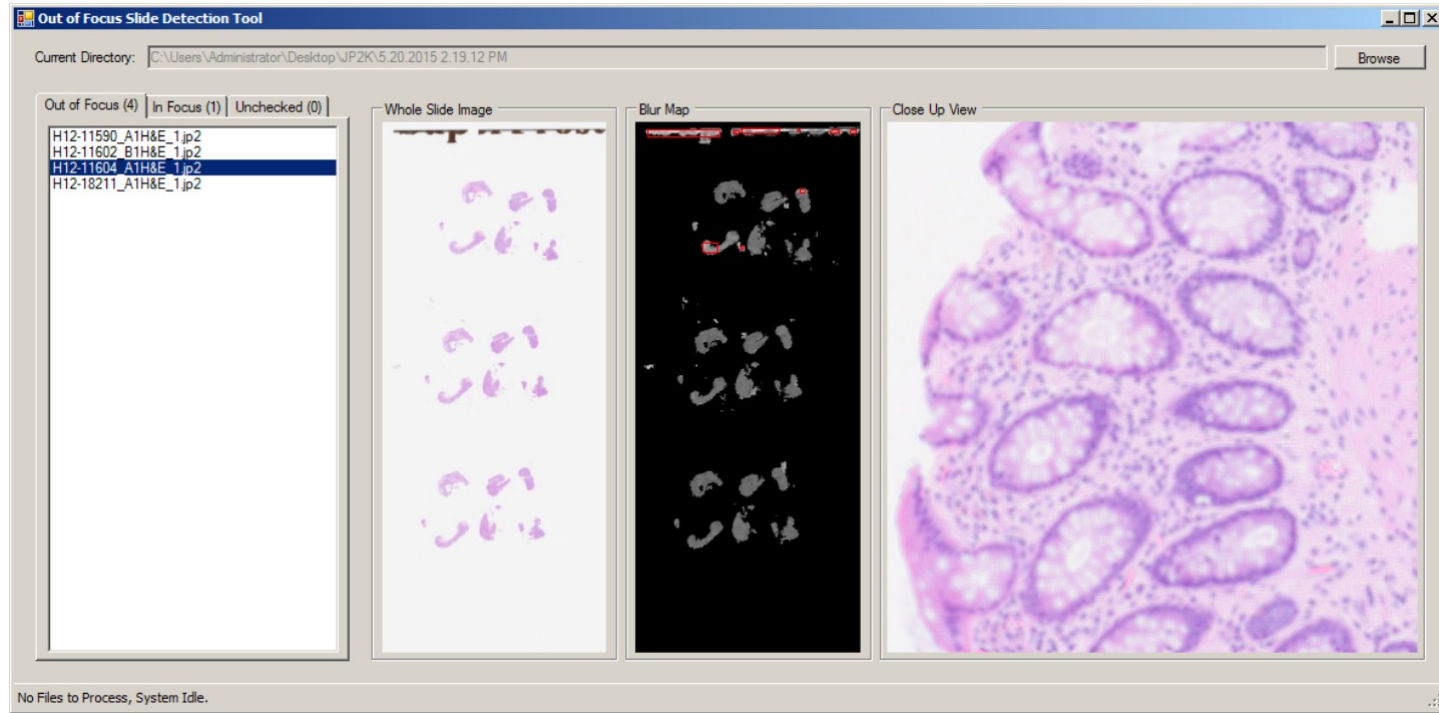
The *Blur Effect* – The Main Idea



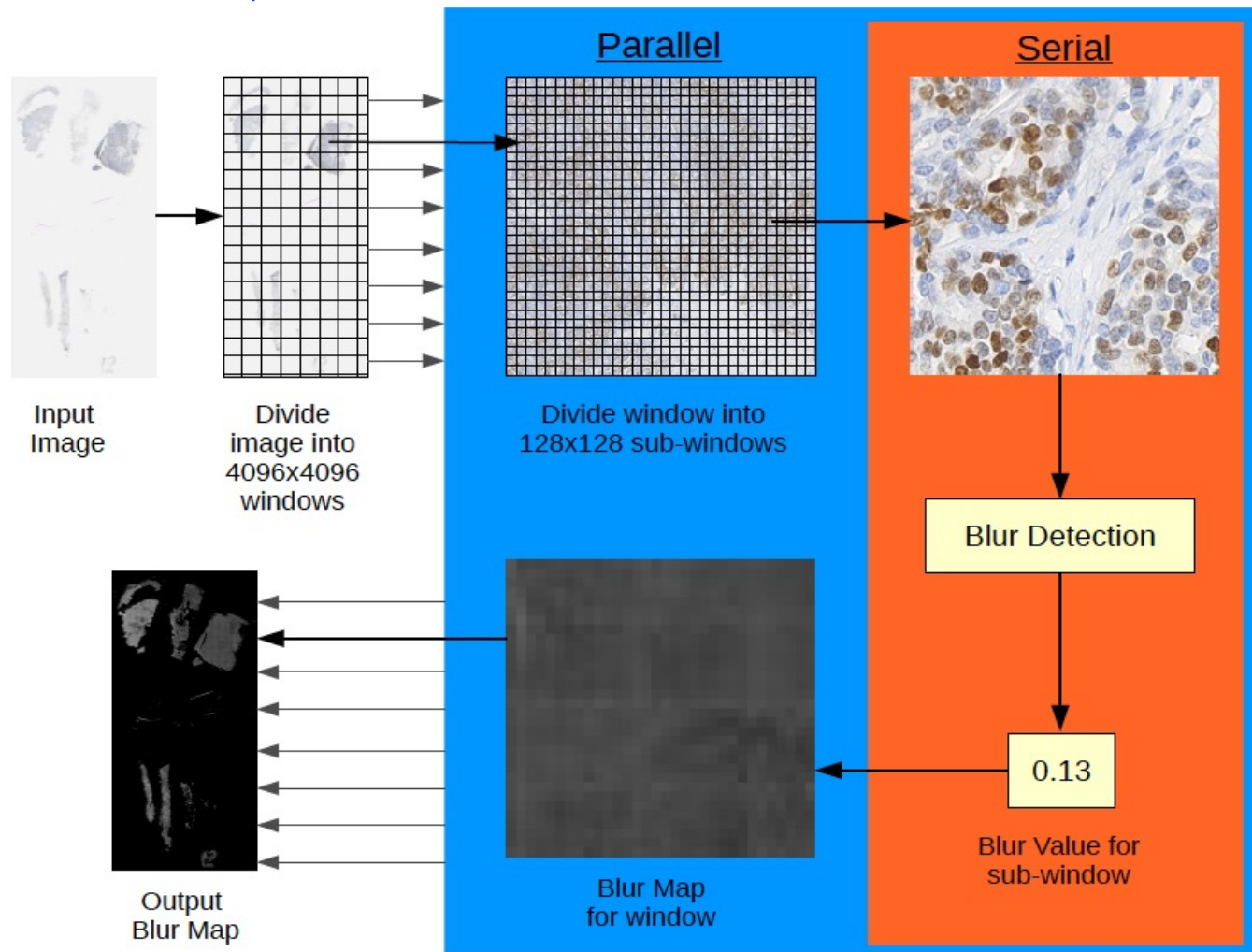
Figure 1: from left to right: original picture, original picture blurred with an averaging filter, blurred picture re-blurred with the same averaging filter.

Crete *et al.*, SPIE Electronic Imaging, 2007

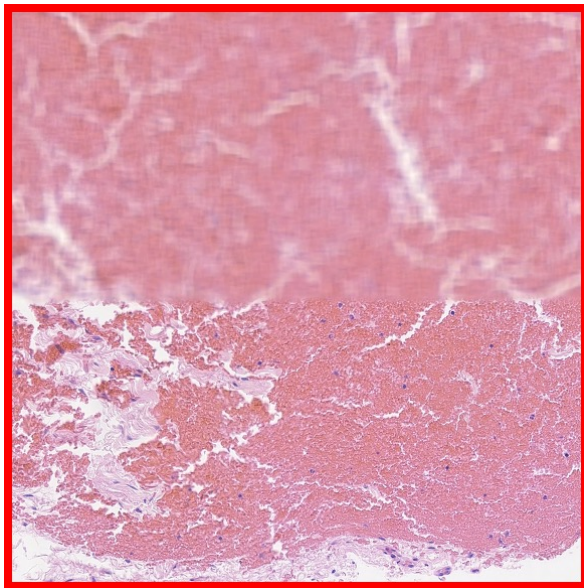
Blur Detection Tool



Source: Nick Trahearn, PhD thesis (University of Warwick, 2017)

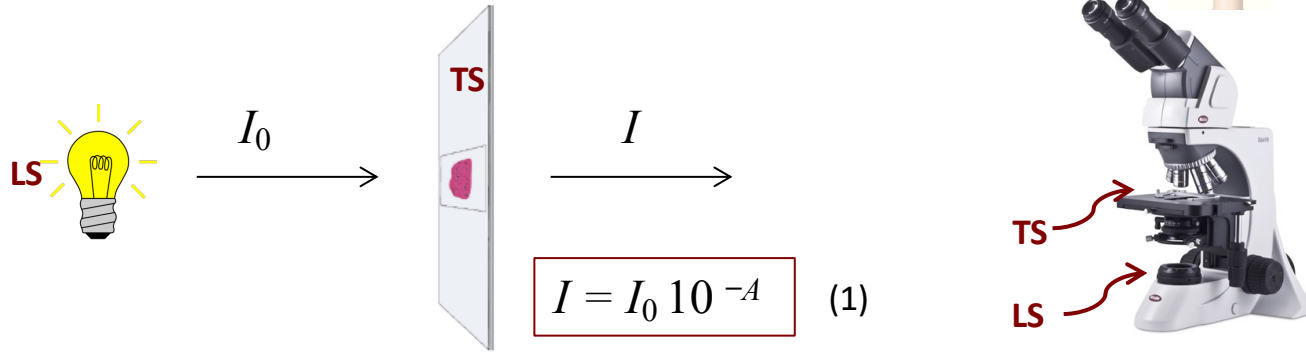


Blur Estimation Results



Stain Separation

The Beer-Lambert Law



- Beer's law:
 - Light absorbed $\propto$ concentration c of the attenuating species in the material (and path length)
- Lambert's law:

$A = \log_{10} (I_0 / I) = \alpha c x$

 (2)
 - Light absorbed $\propto$ its thickness x

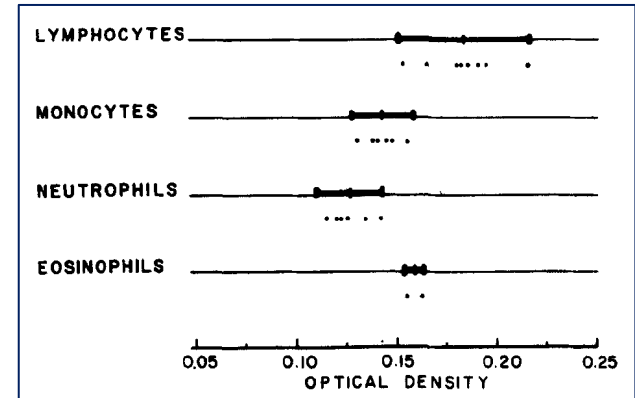
The Optical Density of a Pixel

- According to the Beer-Lambert law,

$$I(i,j) = I_0 10^{-\alpha c(i,j) x} \quad (3)$$

- The local **optical density (OD)** or Absorption A is then defined as follows:

$$\begin{aligned} d(i,j) &= \alpha c(i,j) x \\ &= \log_{10} (I_0 / I(i,j)) \end{aligned} \quad (4)$$



Prewitt's OD based cell classification (1965)

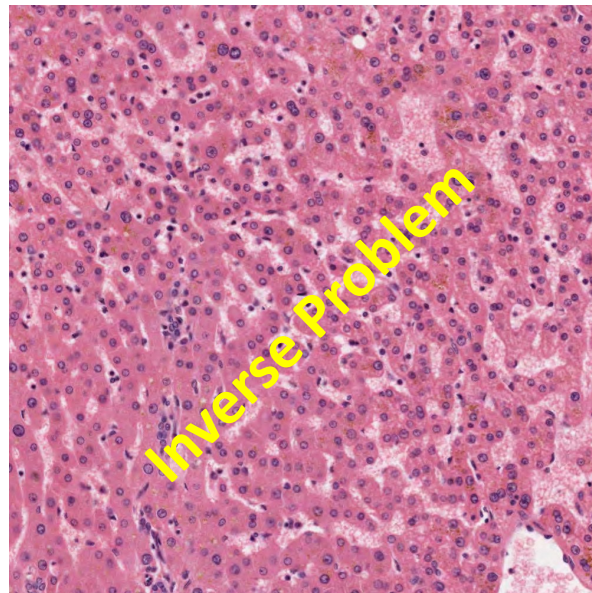
Stain Separation/Deconvolution

Problem:

Given an image containing both H & E contents, separate the H & E channels

Why?

Measure stain intensity for downstream analysis



H&E image

The Ruifrok-Johston (RJ) Method

- The RGB values of I cannot directly be used for stain measurement due to a **nonlinear** relationship between them and the stain concentration
- So we use the local **optical density (OD)** which is defined as,

$$\begin{aligned} d(i,j) &= \alpha c(i,j) x \\ &= \log_{10} (I_0 / I(i,j)) \end{aligned} \quad (5)$$

- And define the **stain matrix** as,

$$M = \begin{bmatrix} \textcolor{red}{R} & \textcolor{green}{G} & \textcolor{blue}{B} \\ p_{1r} & p_{1g} & p_{1b} \\ p_{2r} & p_{2g} & p_{2b} \\ p_{3r} & p_{3g} & p_{3b} \end{bmatrix} \begin{matrix} \mathbf{H} \\ \mathbf{E} \\ \mathbf{BG} \end{matrix}$$

RJ Stain Deconvolution

- The observed optical density d is a linear function of the stain concentrations c present in the sample:

$$\vec{d} = \vec{c}\hat{M}$$

- This gives a simple relationship to estimate c from d as follows,

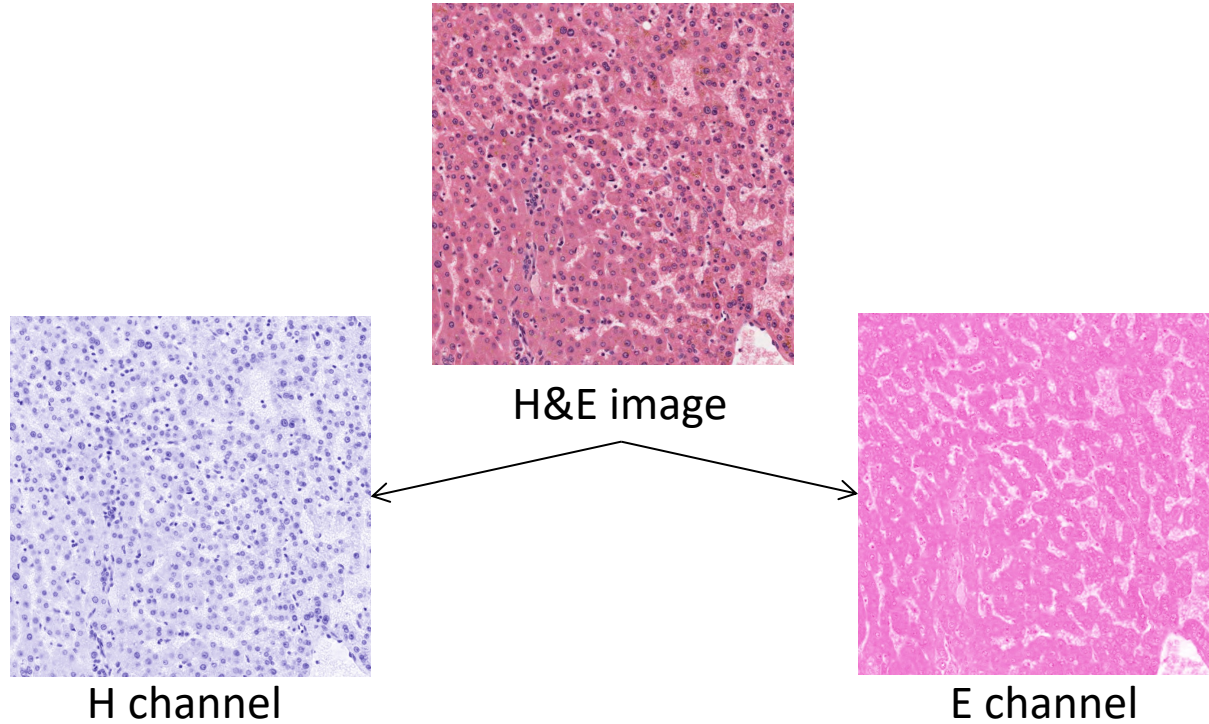
$$\vec{c} = \vec{d}(\hat{M})^{-1}$$

- Advantages:** Simple and fast

$$\hat{M} = \begin{bmatrix} \overset{\text{R}}{\hat{p}_{1r}} & \overset{\text{G}}{\hat{p}_{1g}} & \overset{\text{B}}{\hat{p}_{1b}} \\ \hat{p}_{2r} & \hat{p}_{2g} & \hat{p}_{2b} \\ \hat{p}_{3r} & \hat{p}_{3g} & \hat{p}_{3b} \end{bmatrix} \begin{matrix} \text{H} \\ \text{E} \\ \text{BG} \end{matrix}$$

$$\hat{p}_{ir} = p_{ir} / \sqrt{p_{ir}^2 + p_{ig}^2 + p_{ib}^2}$$

RJ Stain Deconvolution – Results



RJ Stain Deconvolution – Limitations

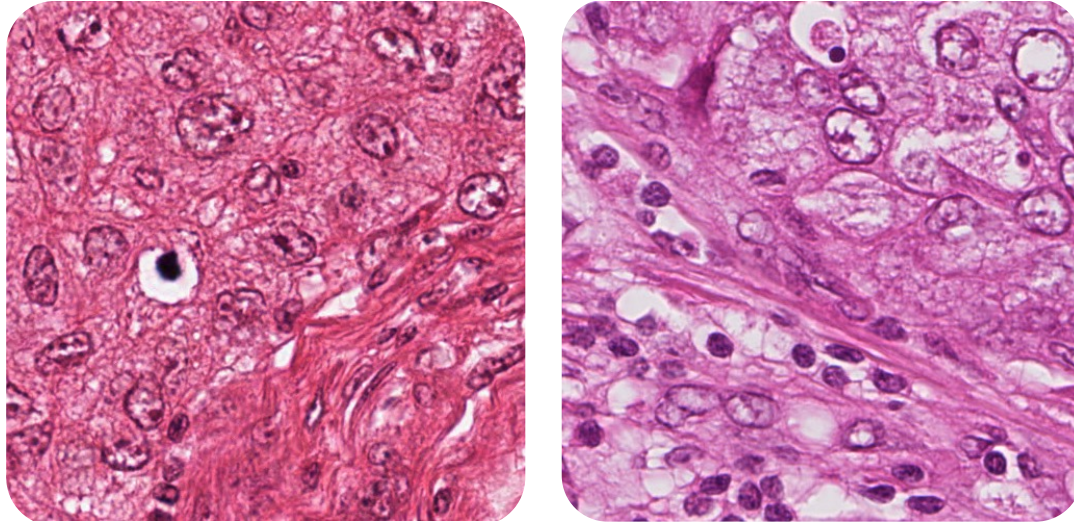
- Requires the estimation of stain matrix

$$\hat{M} = \begin{bmatrix} \overset{\text{R}}{\hat{p}_{1r}} & \overset{\text{G}}{\hat{p}_{1g}} & \overset{\text{B}}{\hat{p}_{1b}} \\ \hat{p}_{2r} & \hat{p}_{2g} & \hat{p}_{2b} \\ \hat{p}_{3r} & \hat{p}_{3g} & \hat{p}_{3b} \end{bmatrix} \begin{matrix} \mathbf{H} \\ \mathbf{E} \\ \mathbf{BG} \end{matrix}$$

- The **DAB** absorption does not follow the Beer-Lambert law → the OD does not vary linearly with the stain concentration (due to scattering characteristics of DAB)

Stain Normalization (SN)

Stain Variability



Variation in the colour appearance of stains
due to various reasons (see next slide)

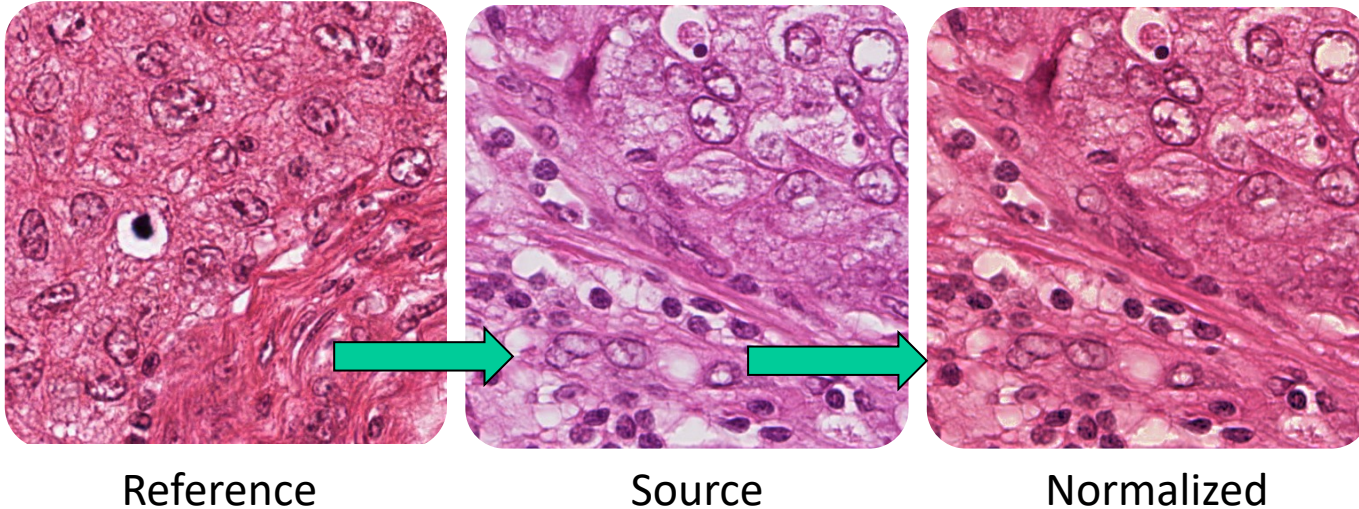
Stain Variability

- Stain variability results from methods and protocols used to prepare the specimen:
 - Type and duration of fixation
 - Consistency and thickness of the sections
 - Room or specimen/dye temperatures
 - Stain reactivity (pH, enzyme concentration) for different manufacturers/batches
 - Incubation times
 - Scanner equipment/settings



Stain Normalization (SN)

The **problem** of ‘normalizing’ the stain colour distribution according to some ‘reference’ image



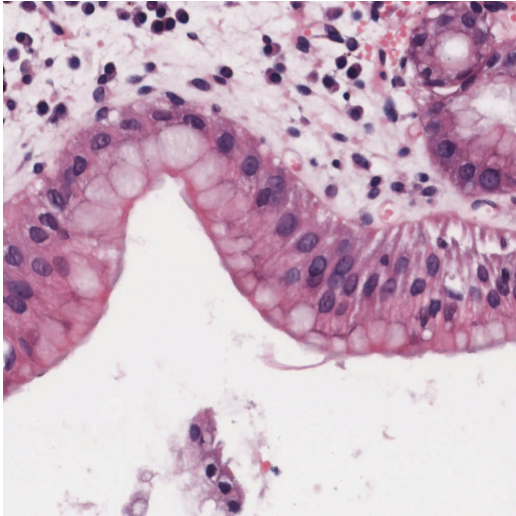
SN via Stain Separation

1. Perform stain deconvolution to get the stain channels.
2. Intensity normalization (e.g., via histogram mapping) in each stain channel
3. RGB image reconstruction from the normalized stain channels.

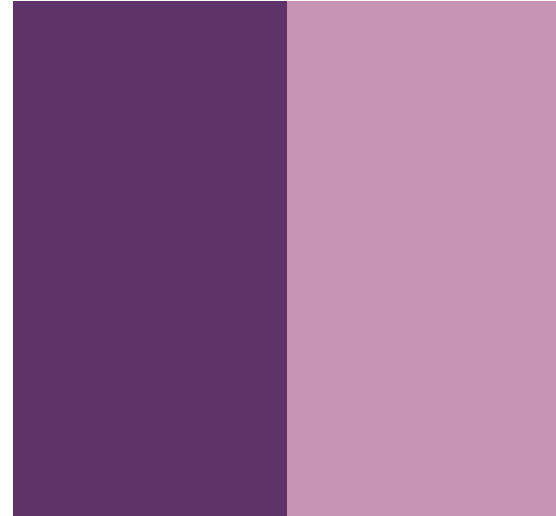


[SN1]. PCA in the OD Space

Presents a framework for estimating color deconvolution matrix.



**Estimated
Stains**



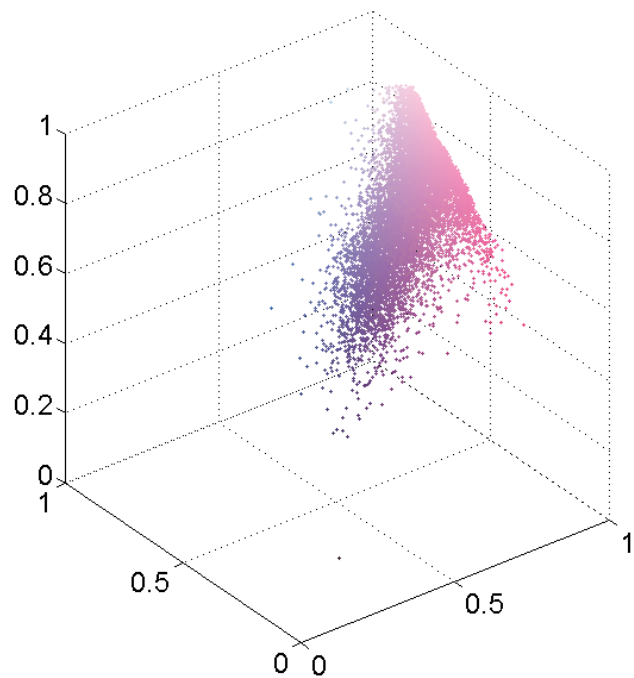
Macenko et al., ISBI 2009

[SN1]. The Algorithm

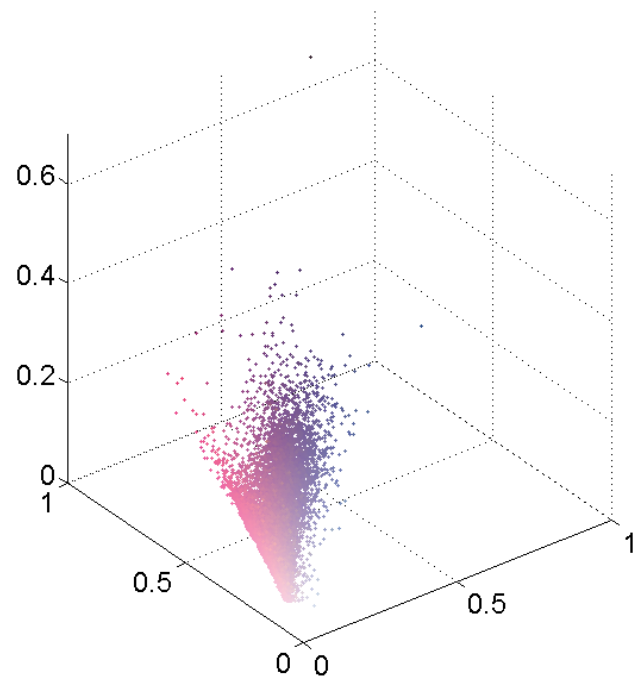
1. RGB to linear space (Optical density space, OD)
2. PCA on the OD channels
3. Project data points on the top 2 principal components and unit normalize
4. Calculate angle of each point in the projection with the 1st principal component
5. Find the robust extremes (1st and 99th percentile) of angles as they represent the stain colors
6. Convert robust extremes back to OD space



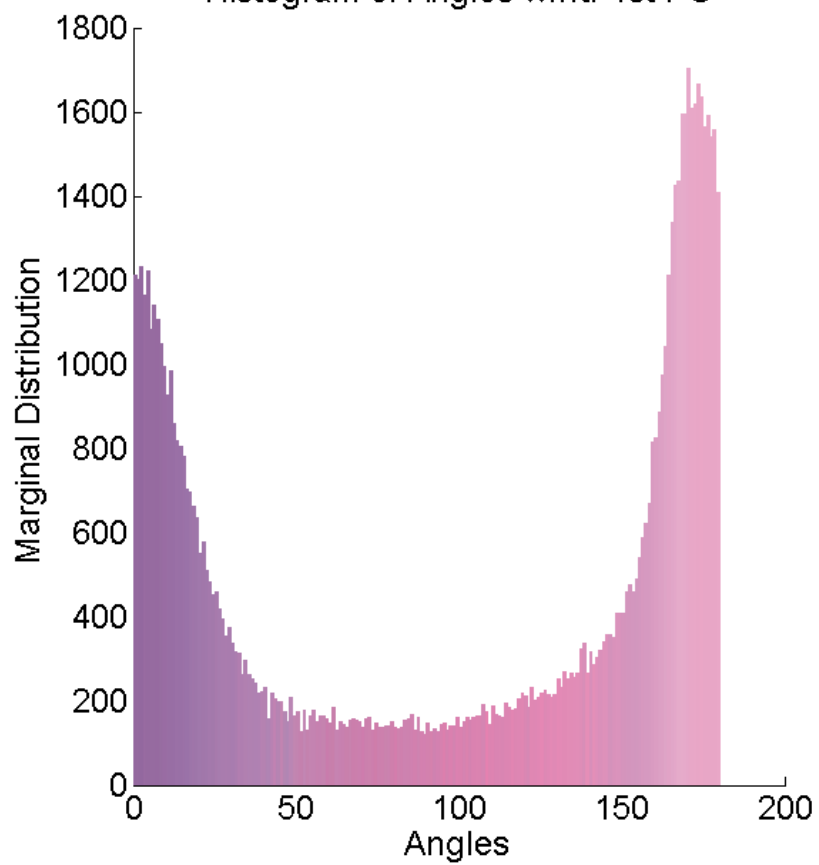
Pixels in RGB Space



Pixels in Optical Density Space



Histogram of Angles w.r.t. 1st PC

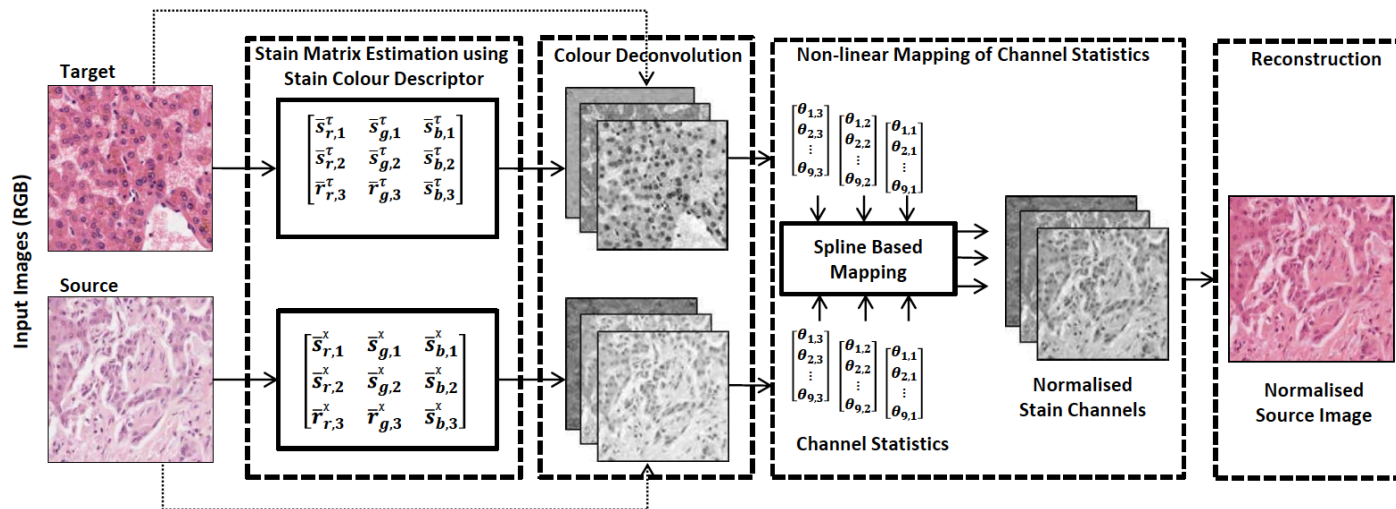


[SN2]. Nonlinear Mapping of Channel Stats

1. Estimate the stain matrix and deconvolve the image into H&E channels
2. Divide the OD intensities in each of the deconvolved channels into 3 classes: BG, STAINED, and OTHER
3. For each class, calculate 3 stats: mean, 5th percentile, and 95th percentile
4. Compute the B-spline mapping function using the 9 input-output pairs and also using some identity pairs

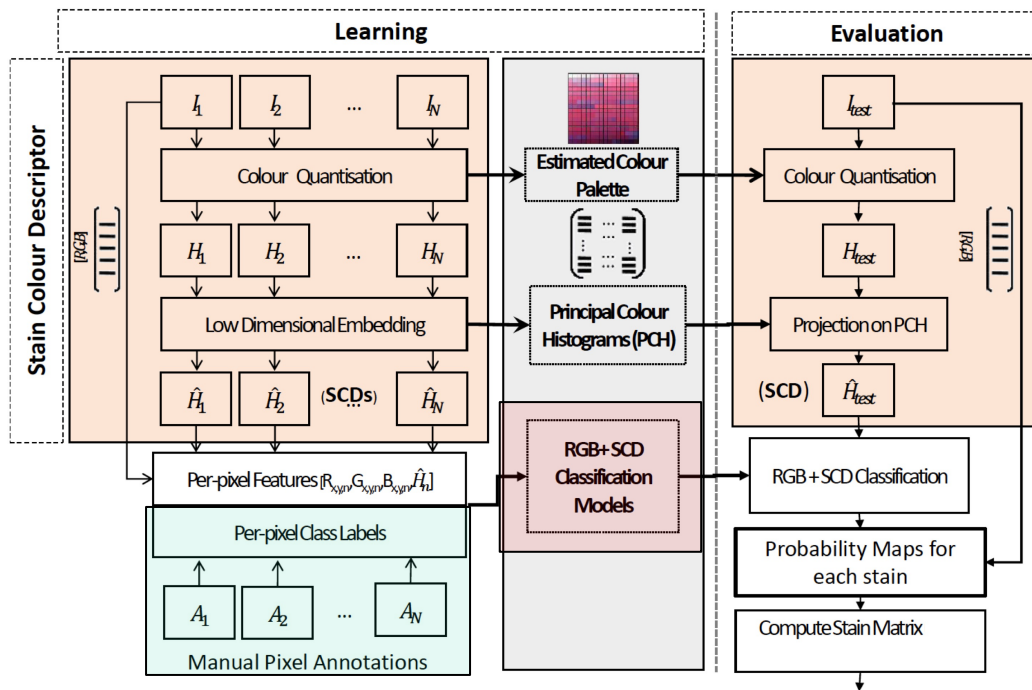


[SN2]. Warwick-Leeds SN – Overview



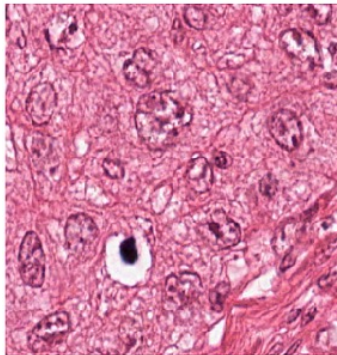
Source: Adnan Khan, PhD Thesis (University of Warwick, 2014)

[SN2]. Stain Matrix Estimation

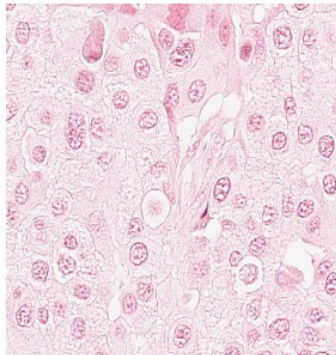


Source: Adnan Khan, PhD Thesis (University of Warwick, 2014)

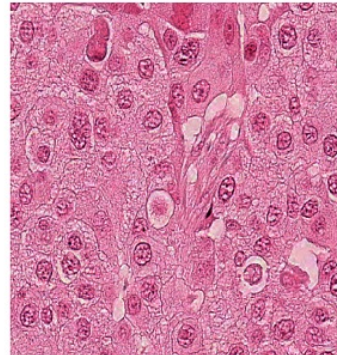
Reference



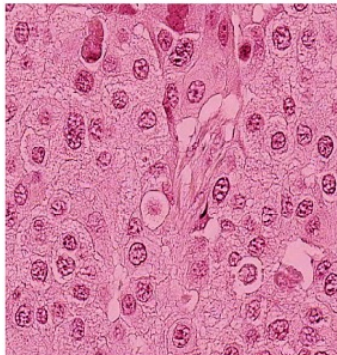
Source



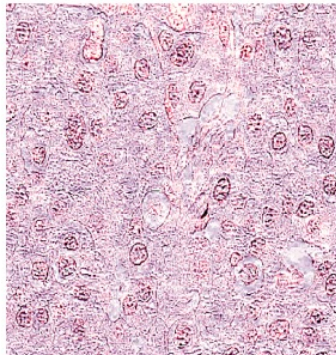
Histogram Specification



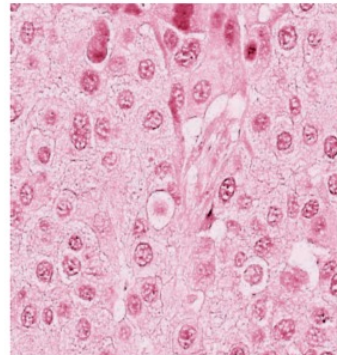
Reinhard et al. [1]



Macenko et al. [2]

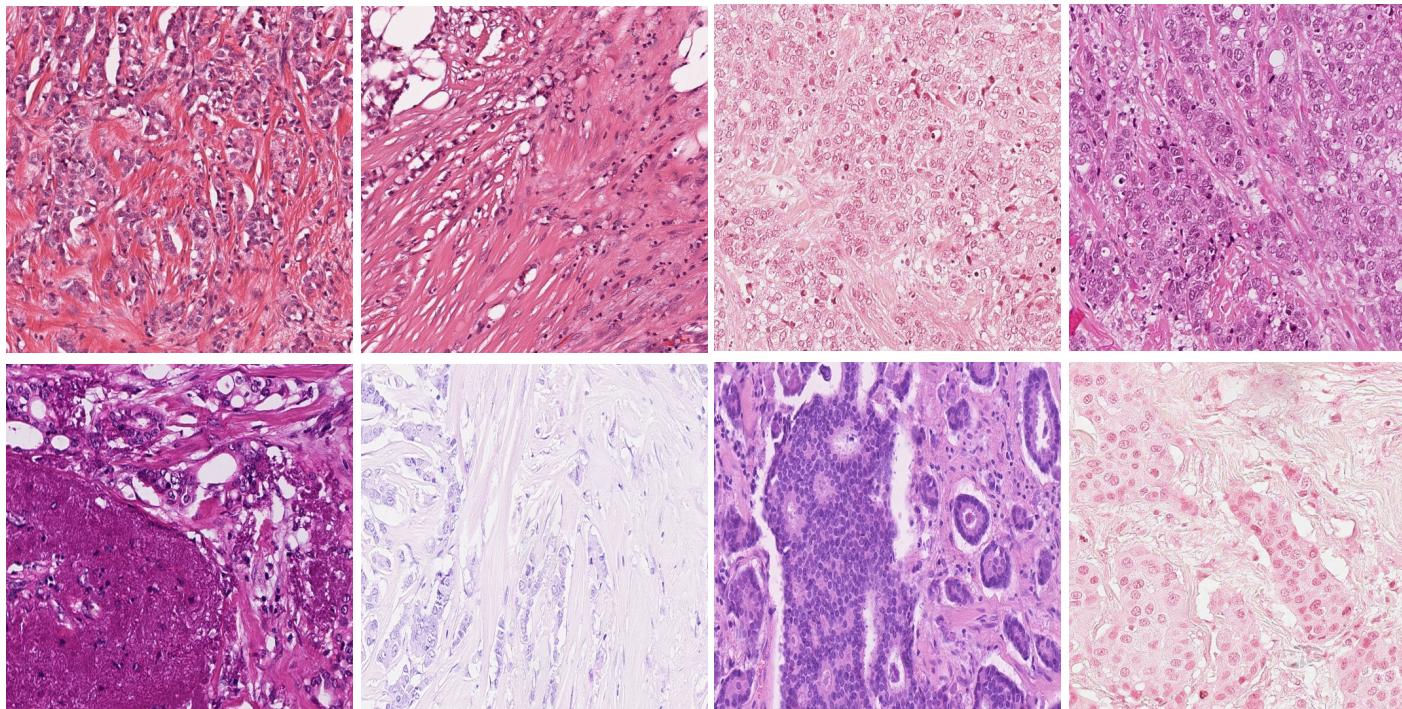


Proposed Method





BEFORE

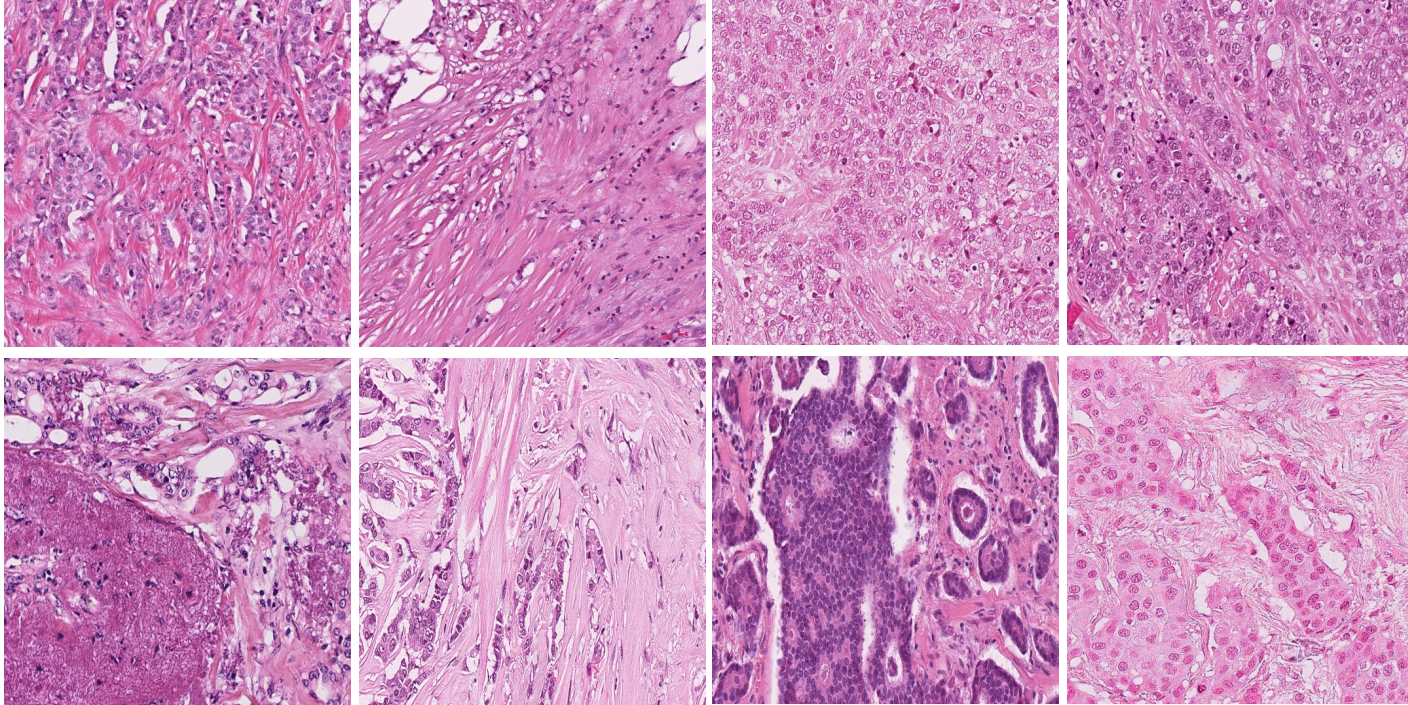


Adnan Khan

Khan et al., IEEE Transactions on Biomedical Engineering, June 2014



AFTER

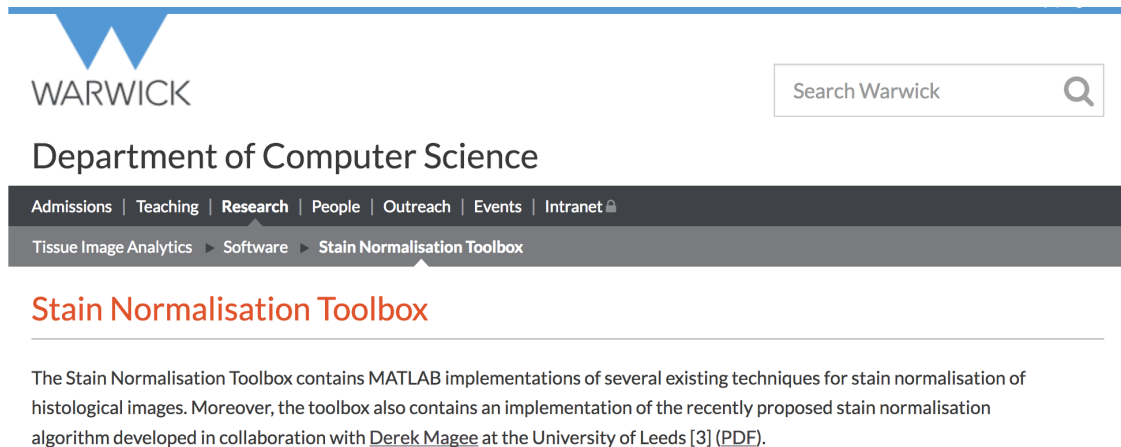


Adnan Khan

Khan et al., IEEE Transactions on Biomedical Engineering, June 2014

SN Toolbox

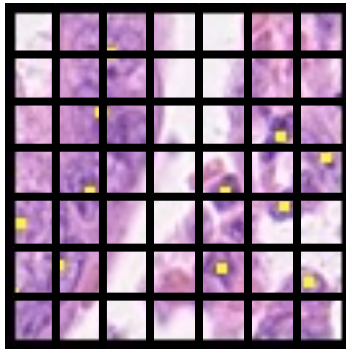
Publicly available toolbox consisting of some of the leading algorithms (including our own) for normalization of stain colors in histology images



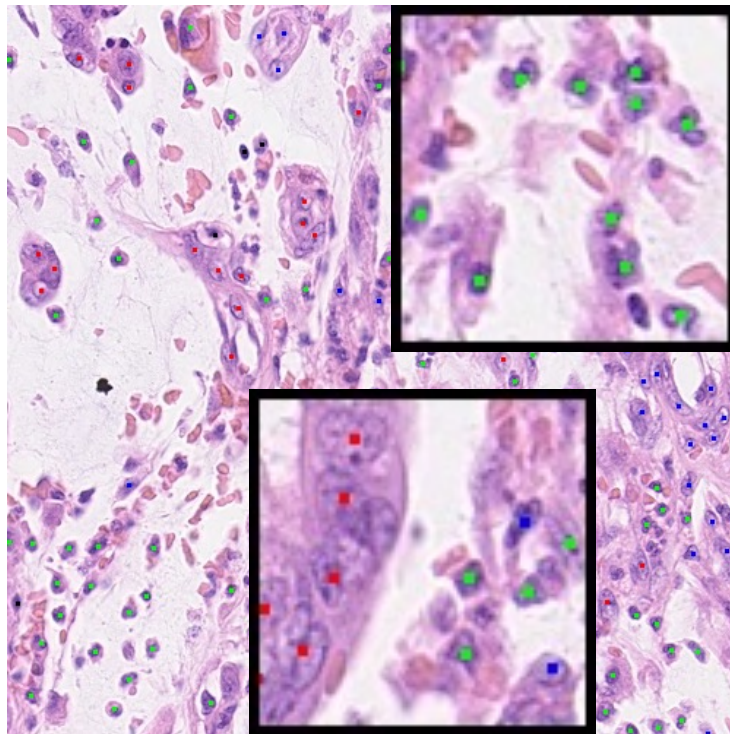
<http://www.warwick.ac.uk/TIA/software/sntoolbox>
replaced by **TIAtoolbox** (in Python)

Cell Detection as Patch Classification

- Divide the image into small patches
- For each patch, find if there is a nucleus within the patch or not
- Patches can be overlapping – ie, sliding windows



Cell Classification



Tumor

Immune

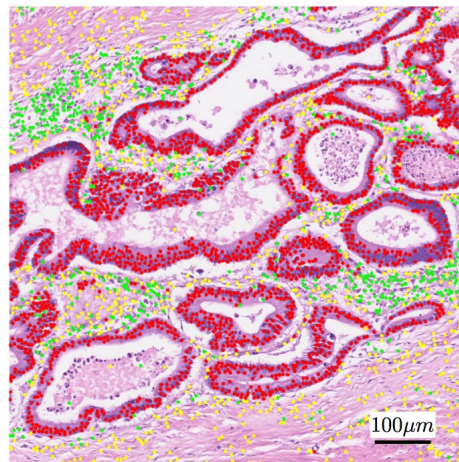
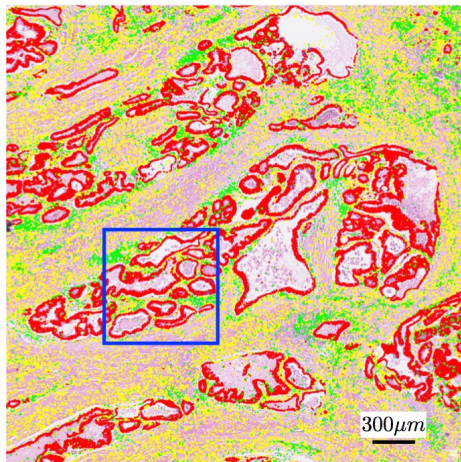
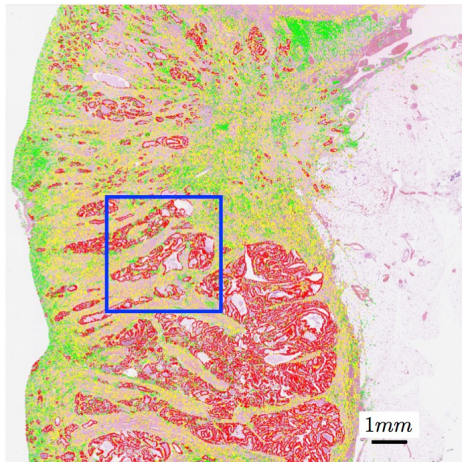
Fibroblast

Other

Detection and Segmentation of Cell Nuclei & Glands



Detecting & Classifying Cell Nuclei in H&E Images

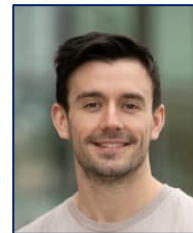
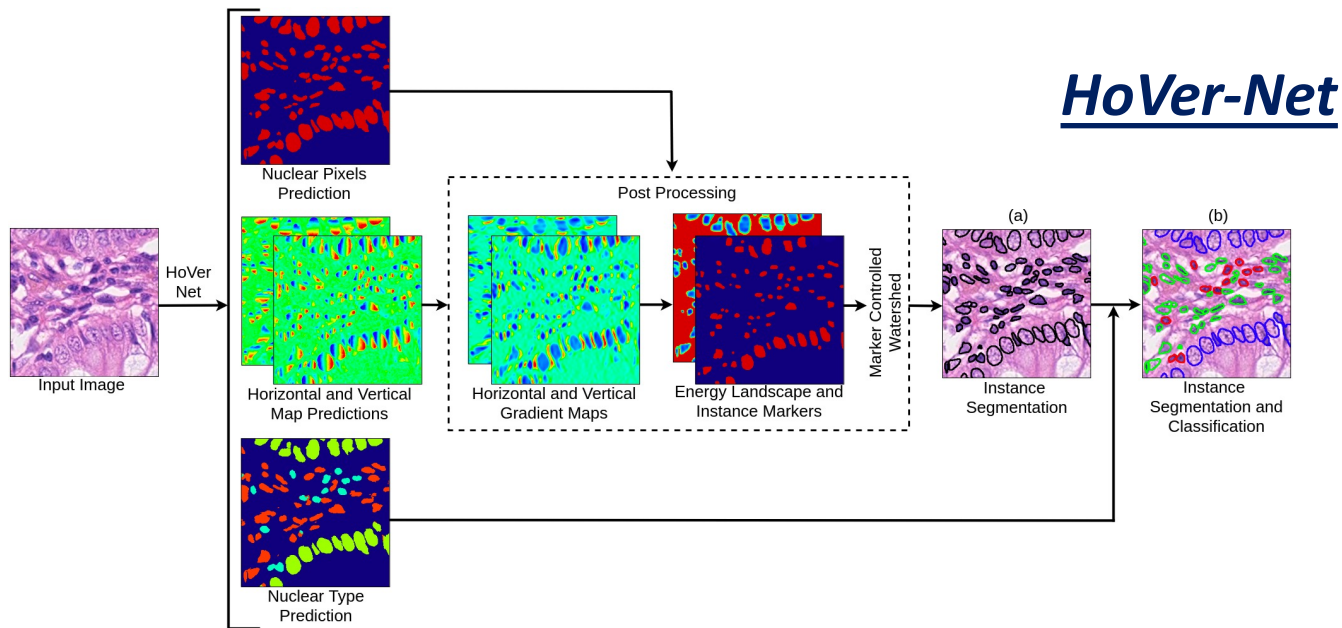


Korsuk
Sirinukunwattana

Sirinukunwattana et al., *IEEE Transactions Medical Imaging*
special issue on **Deep Learning in Medical Imaging** (May 2016)



Nucleus Segmentation & Classification



Simon Graham

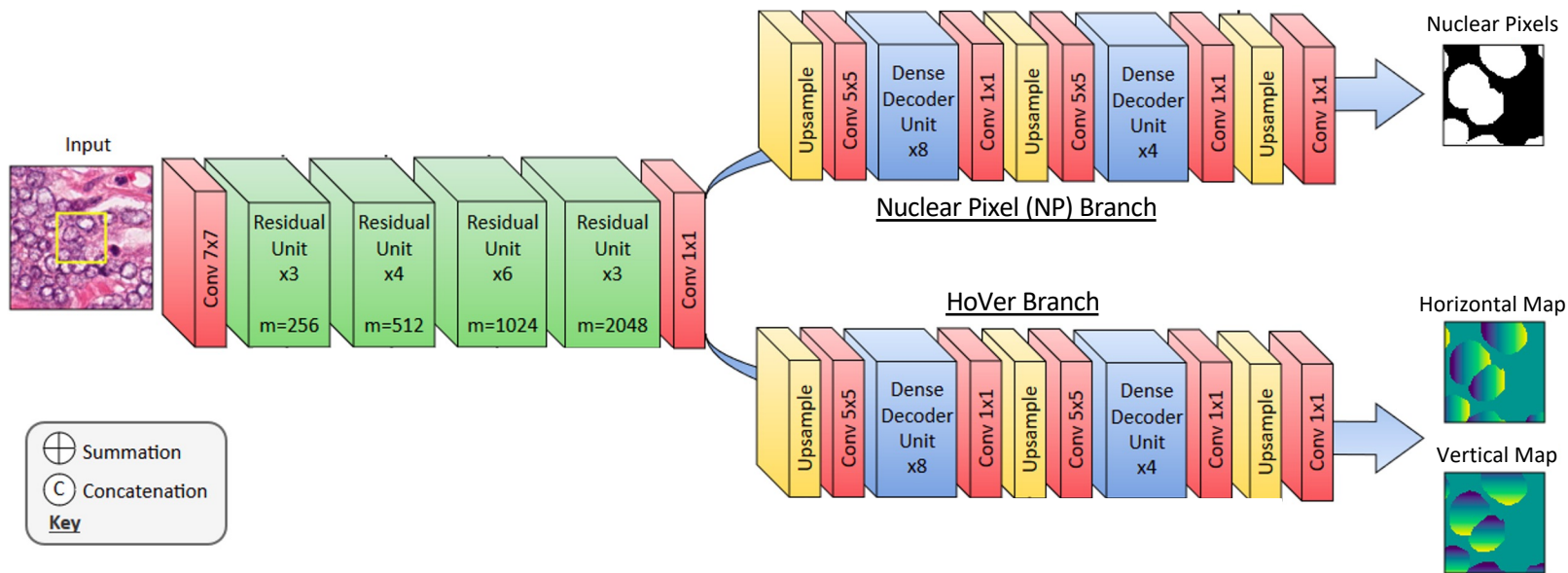


Dang Vu

Graham *et al.*, **Medical Image Analysis** (Dec 2019)



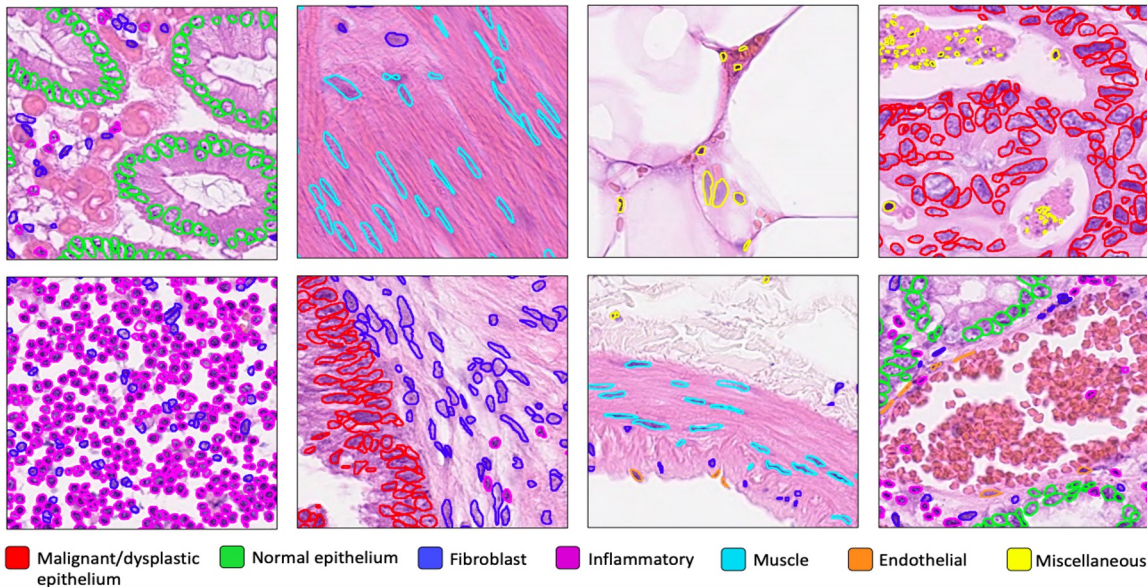
HoVer-Net: Nucleus Segmentation & Classification



Graham *et al.*, Medical Image Analysis (Sep 2019)



HoVer-Net: Nuclear Segmentation & Classification



Simon Graham



Dang Vu

Graham *et al.*, Medical Image Analysis (Dec 2019)



HoVer-Net: Experimental Results

Methods	Kumar					CoNSeP					CPM-17				
	DICE	AJI	DQ	SQ	PQ	DICE	AJI	DQ	SQ	PQ	DICE	AJI	DQ	SQ	PQ
Cell Profiler (Carpenter et al., 2006)	0.623	0.366	0.423	0.704	0.300	0.84	-	-	-	-	0.570	0.338	0.368	0.702	0.261
QuPath (Bankhead et al., 2017)	0.698	0.432	0.511	0.679	0.5	-	-	-	-	-	0.398	0.320	0.717	0.230	-
FCN8 (Long et al., 2015)	0.797	0.281	0.434	0.714	-	-	-	-	-	-	0.397	0.575	0.750	0.435	-
FCN8 + WS (Long et al., 2015)	0.797	0.429	0.590	0.719	-	-	-	-	-	-	0.397	0.575	0.750	0.435	-
SegNet (Badrinarayanan et al., 2017)	0.811	0.377	0.545	0.719	0.5	-	-	-	-	-	0.491	0.679	0.778	0.531	-
SegNet + WS (Badrinarayanan et al., 2017)	0.811	0.508	0.677	0.719	0.5	-	-	-	-	-	0.594	0.779	0.784	0.614	-
U-Net (Ronneberger et al., 2015)	0.758	0.556	0.691	0.6	0.5	-	-	-	-	-	0.643	0.778	0.734	0.578	-
Mask-RCNN (He et al., 2017)	0.760	0.546	0.704	0.7	0.5	-	-	-	-	-	0.684	0.848	0.792	0.674	-
DCAN (Chen et al., 2016)	0.792	0.525	0.677	0.7	0.5	-	-	-	-	-	0.828	0.561	0.732	0.740	0.545
Micro-Net (Raza et al., 2018)	0.797	0.560	0.692	0.747	0.5	-	-	-	-	-	0.857	0.668	0.836	0.788	0.661
DIST (Naylor et al., 2018)	0.789	0.559	0.601	0.732	0.44	-	-	-	0.28	0.398	0.826	0.616	0.663	0.754	0.504
CNN3 (Kumar et al., 2017)	0.762	0.508	-	-	-	-	-	-	-	-	-	-	-	-	-
CIA-Net (Zhou et al., 2019)	0.818	0.620	0.754	0.762	0.577	-	-	-	-	-	-	-	-	-	-
DRAN (Vu et al., 2018)	-	-	-	-	-	-	-	-	-	-	0.862	0.683	0.811	0.804	0.657
HoVer-Net	0.826	0.618	0.770	0.773	0.597	0.853	0.571	0.702	0.778	0.547	0.869	0.705	0.854	0.814	0.697

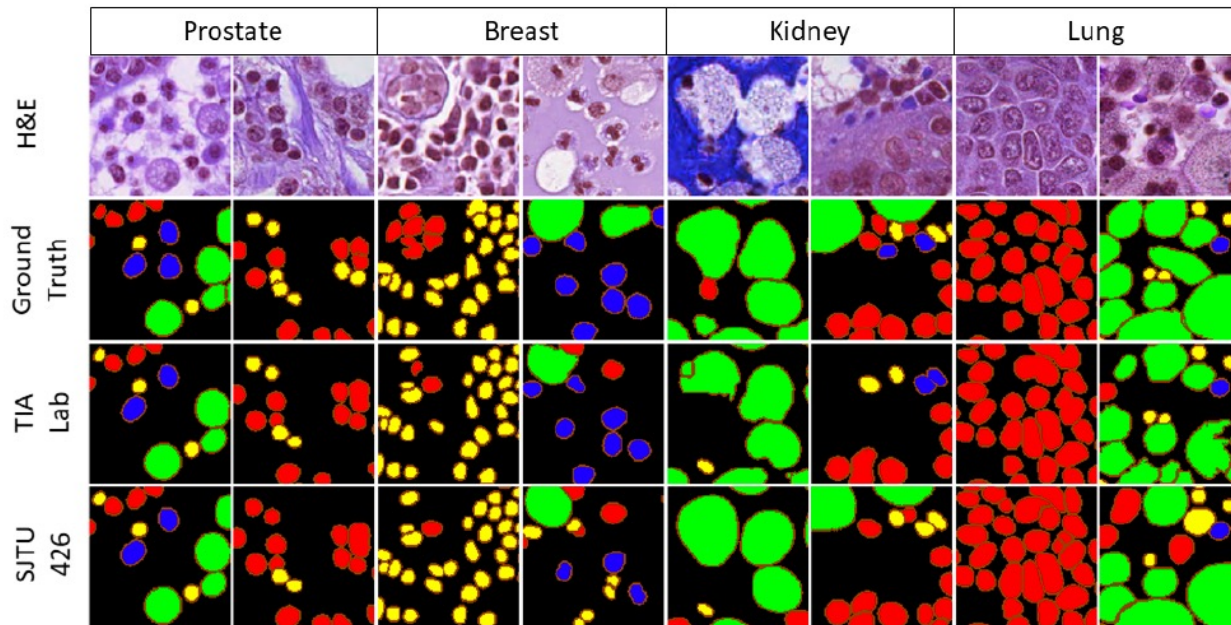
CoNSeP Dataset

24K nuclei:
segmentation
masks & class
labels

Graham et al., Medical Image Analysis (Dec 2019)

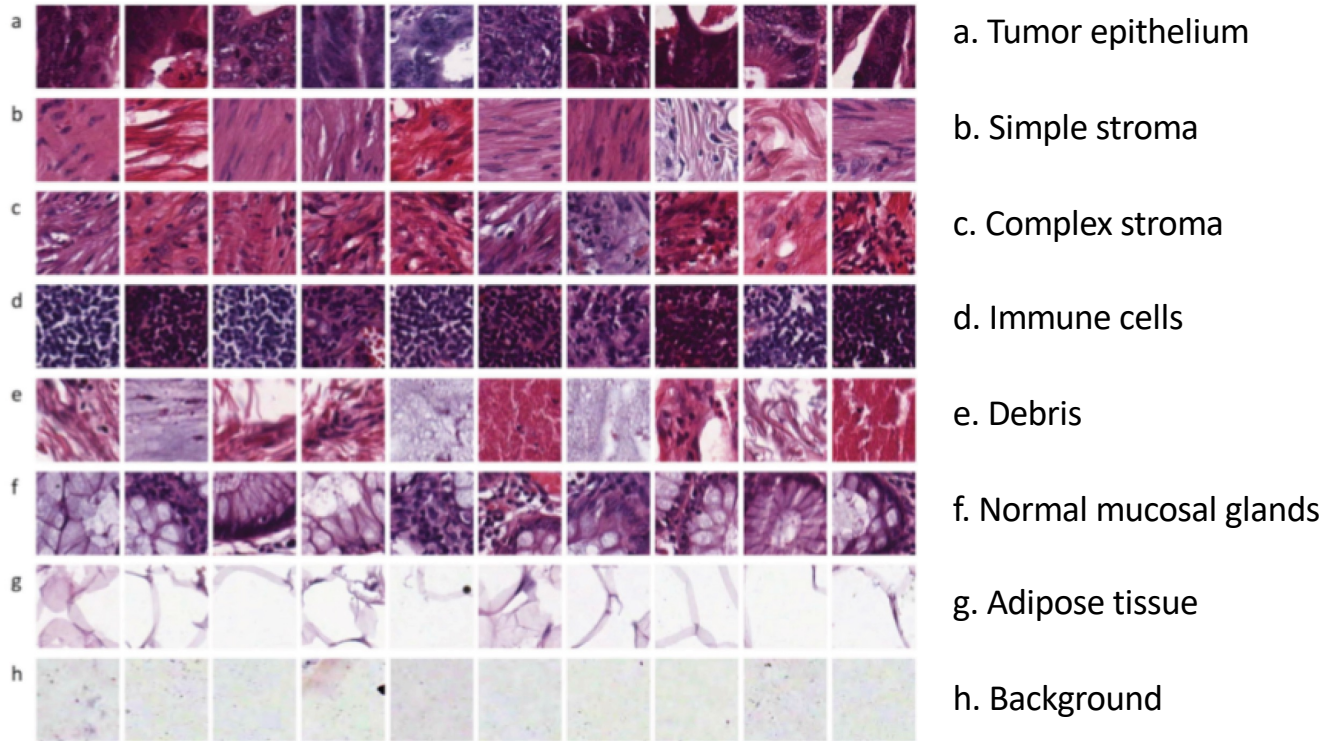


The MONuSaC Challenge Contest



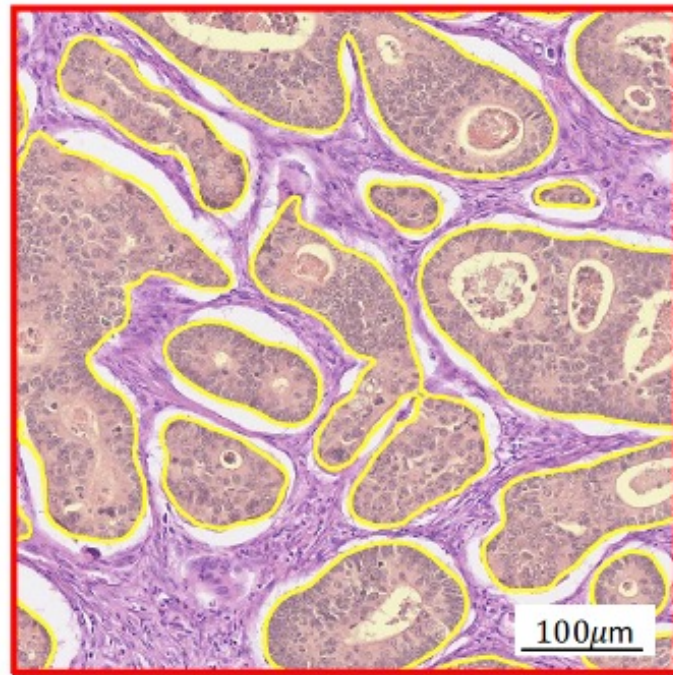
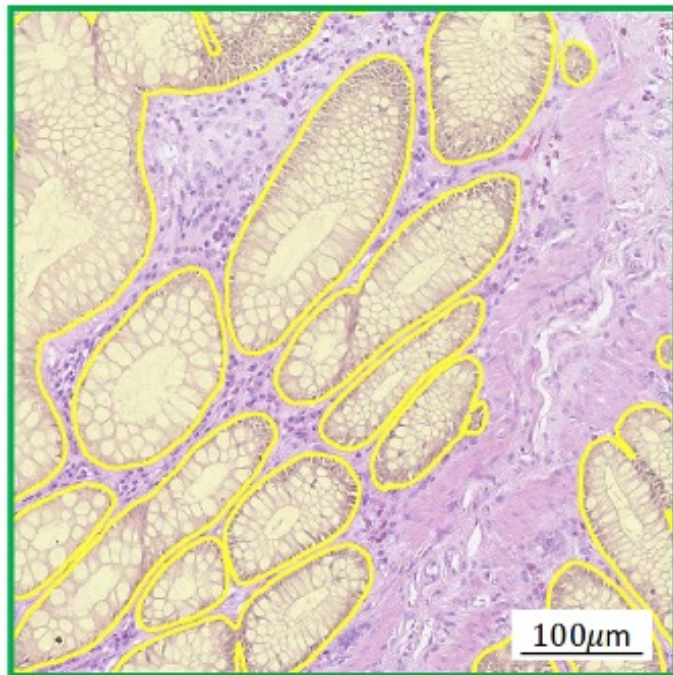
Verma *et al.*, *IEEE Transactions Medical Imaging* (Dec 2021)

Tissue Classification

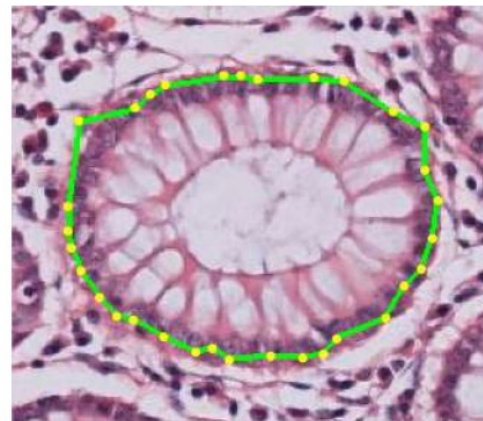
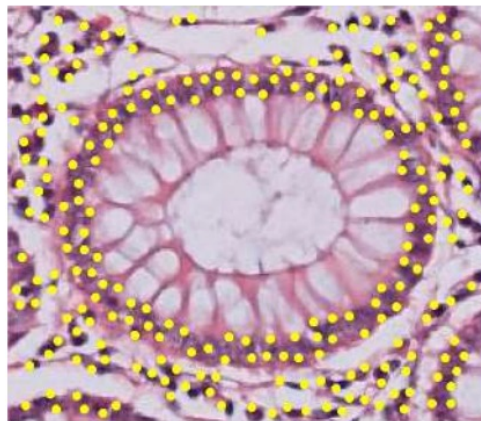


Source: Kather *et al.* (Sci Rep, 2016)

Gland Segmentation



Gland Segmentation with Bayesian Inference



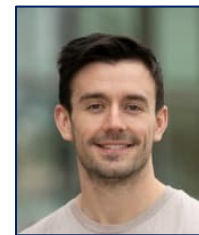
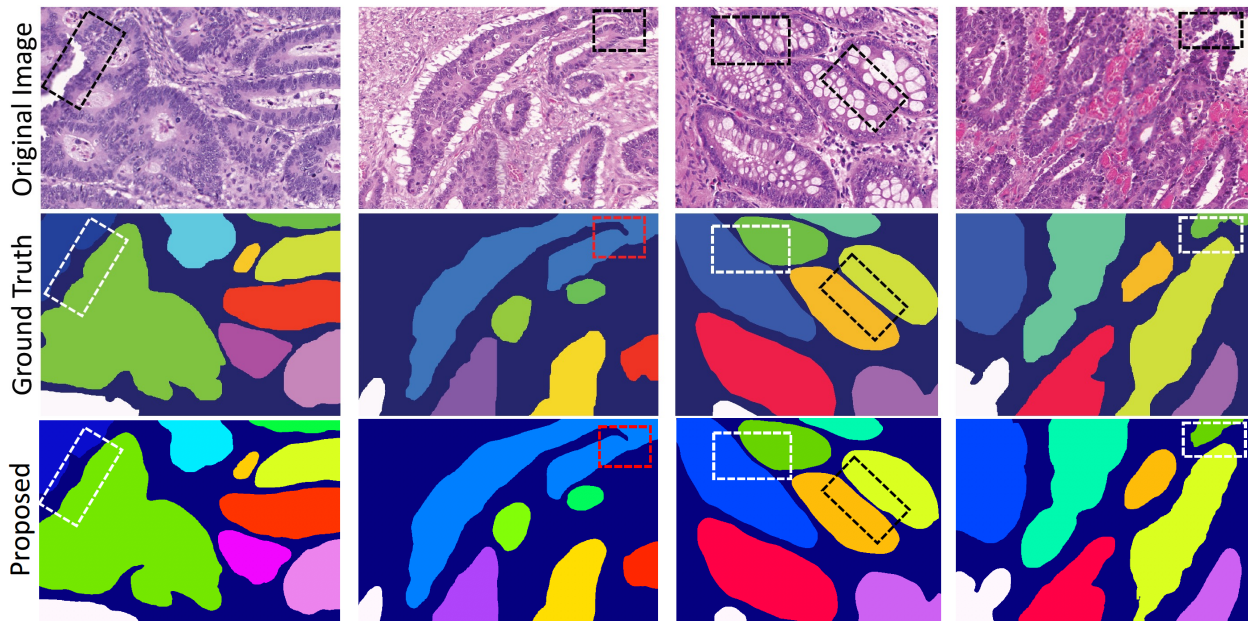
A Stochastic Polygons Model for Glandular Structures in Colon Histology Images

Korsuk Sirinukunwattana, David R. J. Snead, and Nasir M. Rajpoot[†], *Senior Member, IEEE*

Sirinukunwattana *et al.* (IEEE TMI, 2015)



MILD-Net for Gland Segmentation

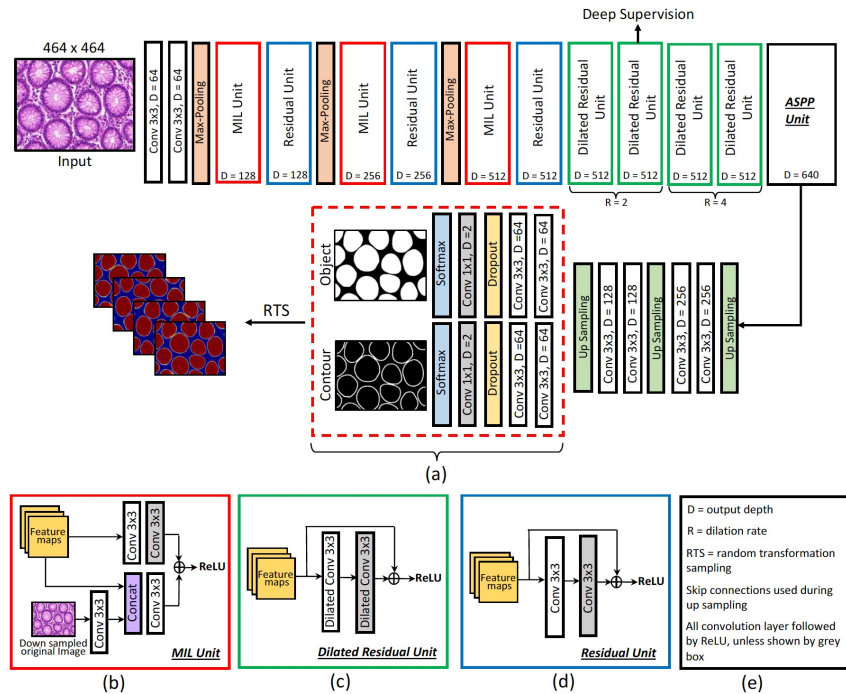


Simon Graham

Graham *et al.*, **Medical Image Analysis** (Feb 2019)

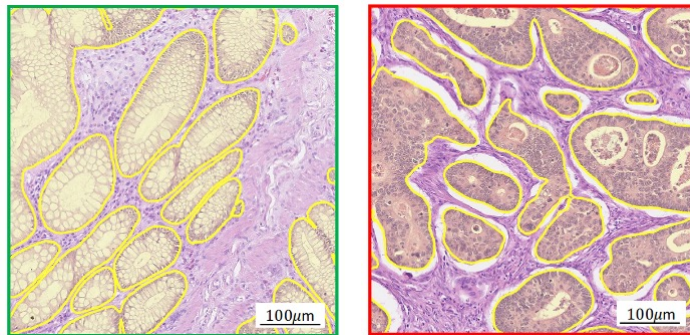


MILD-Net for Gland Segmentation



Graham *et al.*, Medical Image Analysis (Feb 2019)

Gland Segmentation Challenge (GlaS)



GlaS@MICCAI'2015: Gland Segmentation Challenge Contest



[Home](#) | [Details](#) | [Rules](#) | [Important Dates](#) | [Registration](#) | [Download](#) | [Evaluation](#) | [Submission](#) | [Results](#) | [Contact](#) | [Organisers](#)



GlaS@MICCAI 2015 (Munich, 5 Oct 2015)

And that's all, folks! (Part 1)



Source: I Heart Histo

Fundamentals of Computational Pathology: Part I

Nasir Rajpoot

Professor of Computational Pathology

Wolfson Fellow, The Royal Society

Director, Tissue Image Analytics (TIA) Centre

University of Warwick, UK

Croucher ASI AI for Biomedicine (Fri 28 Nov 2025)

