

# CAMELYON 2017 TEAM DTU

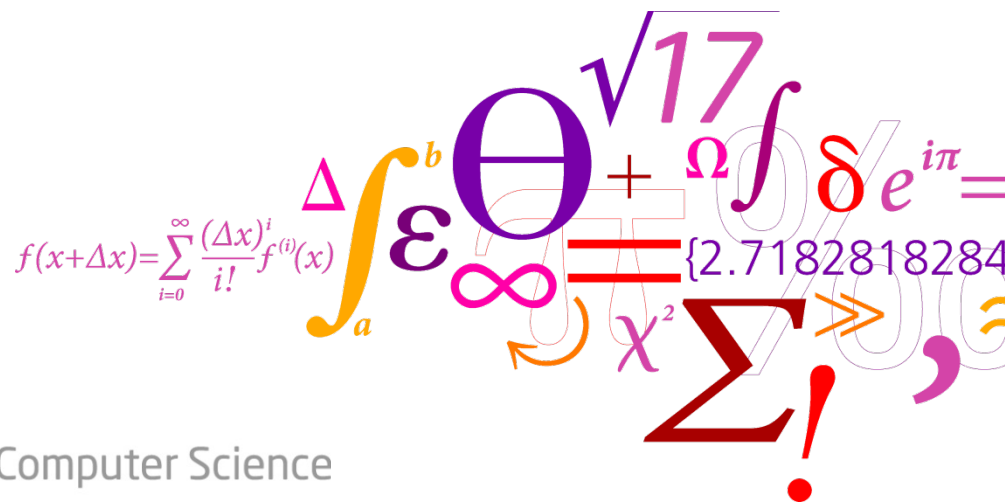
Jeppe Thagaard, MSc student

Biomedical Engineering, Technical University of Denmark

Thesis supervisors:

Anders B. Dahl PhD, Section for Image Analysis & Computer Graphics

Søren Hauberg PhD, Section for Cognitive Systems



DTU Compute

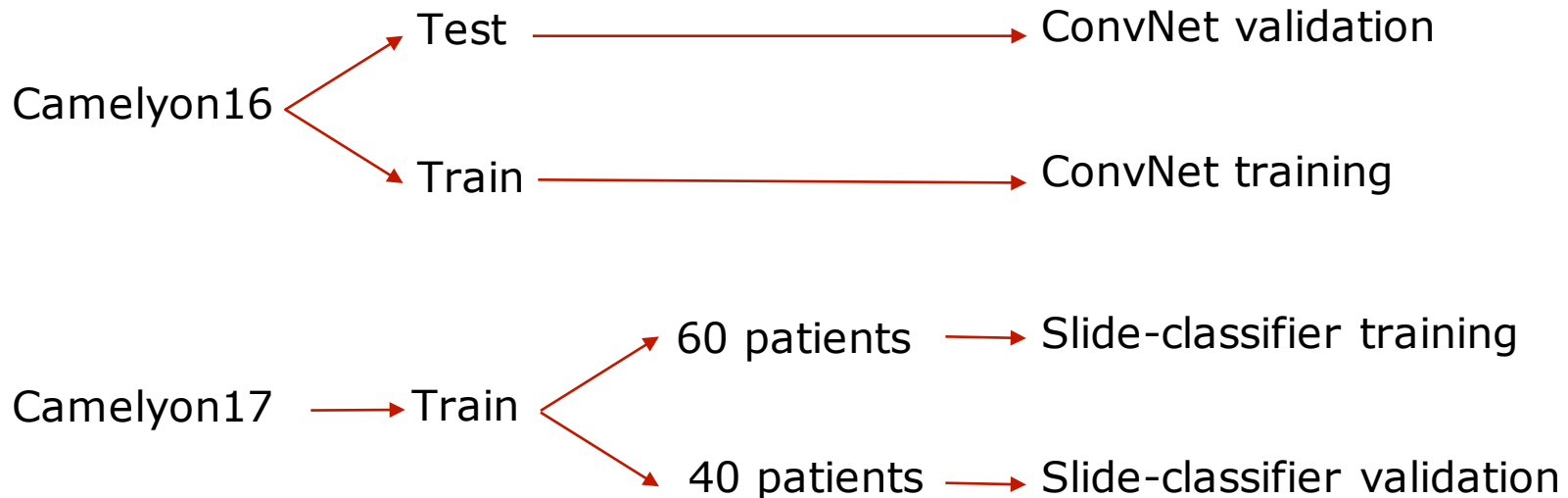
Department of Applied Mathematics and Computer Science

---

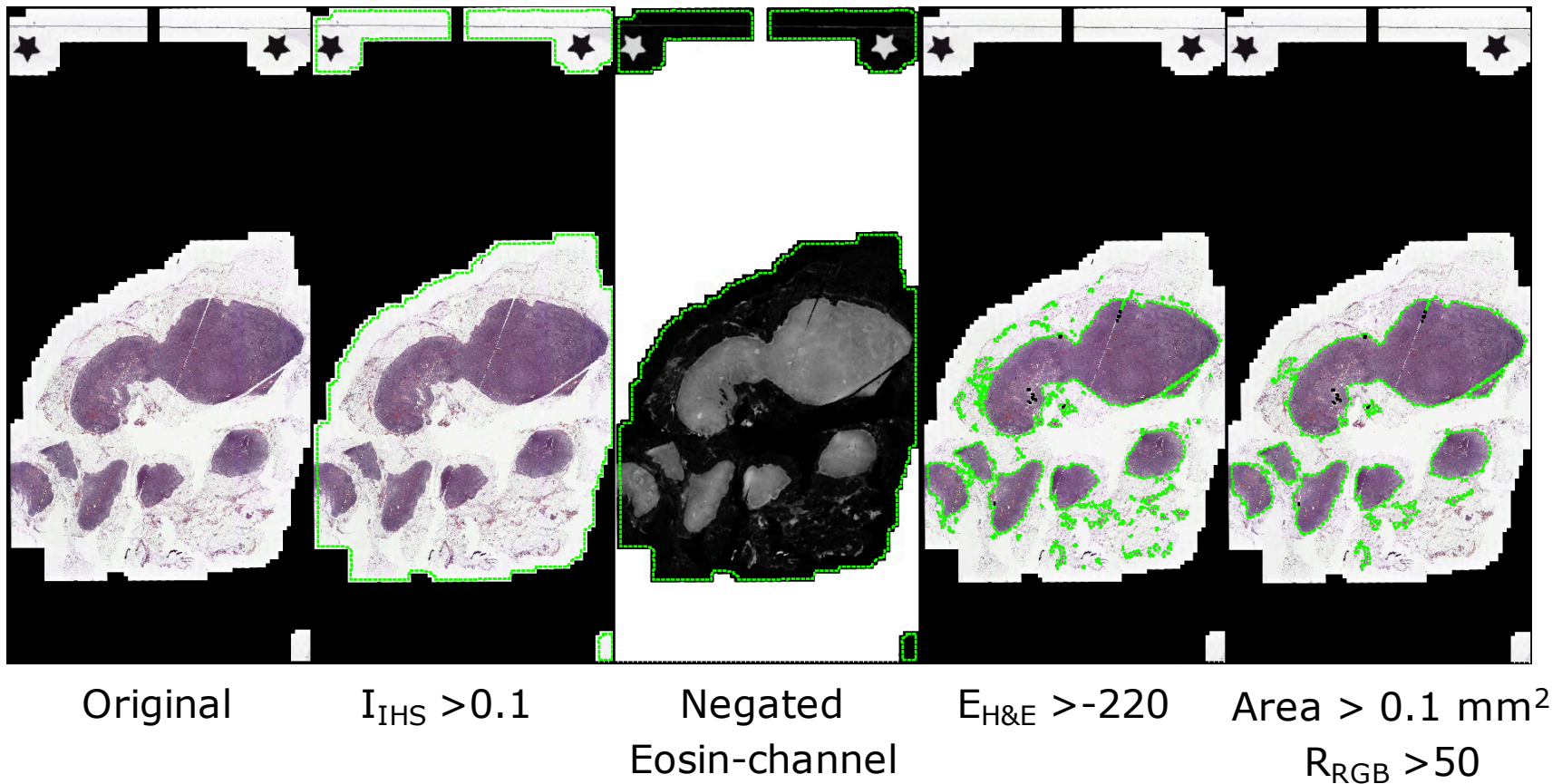
# Overview

- Tissue detection using Visiopharm Integrator System (VIS)
- Patch-classification with convolutional neural networks
- Post-processing and high-level features with VIS
- Slide-classification with Random Forest classifier

- Data split:



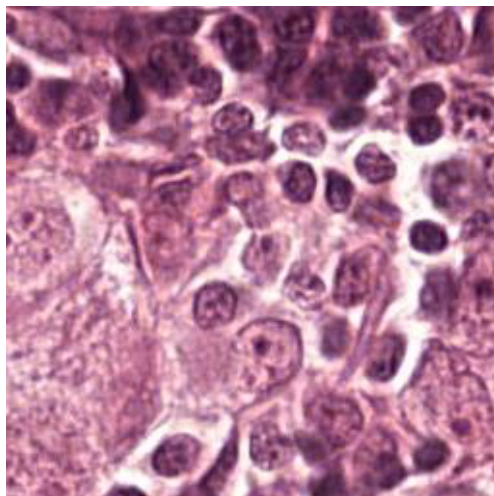
# Tissue detection



## Patch size: How much detail?

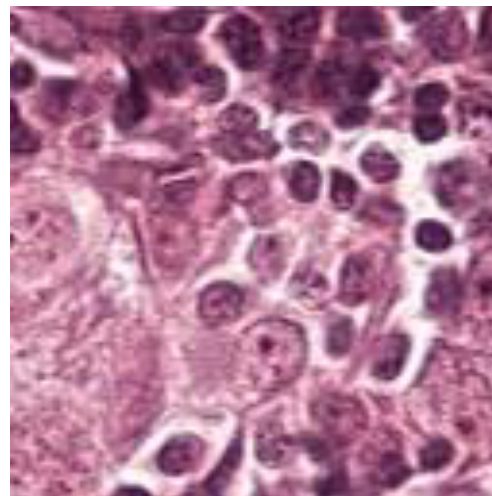
- Same receptive field in tissue, same performance
- 128 x 128 pixels at 20X magnification
- Smaller input size – less computational

40X



256 x 256

20X



128 x 128

# Sampling & model iteration

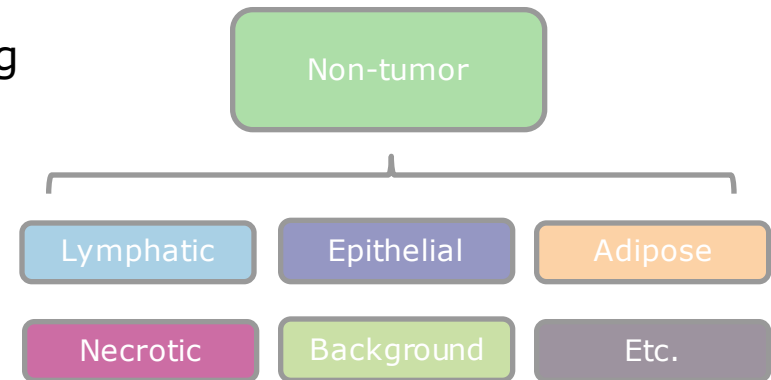
- Sampling in 2 iterations on Cam16 training

1<sup>st</sup>:

- Uniform random sampling (URS)
- 2000 tumor per slide
- 1000 normal per slide

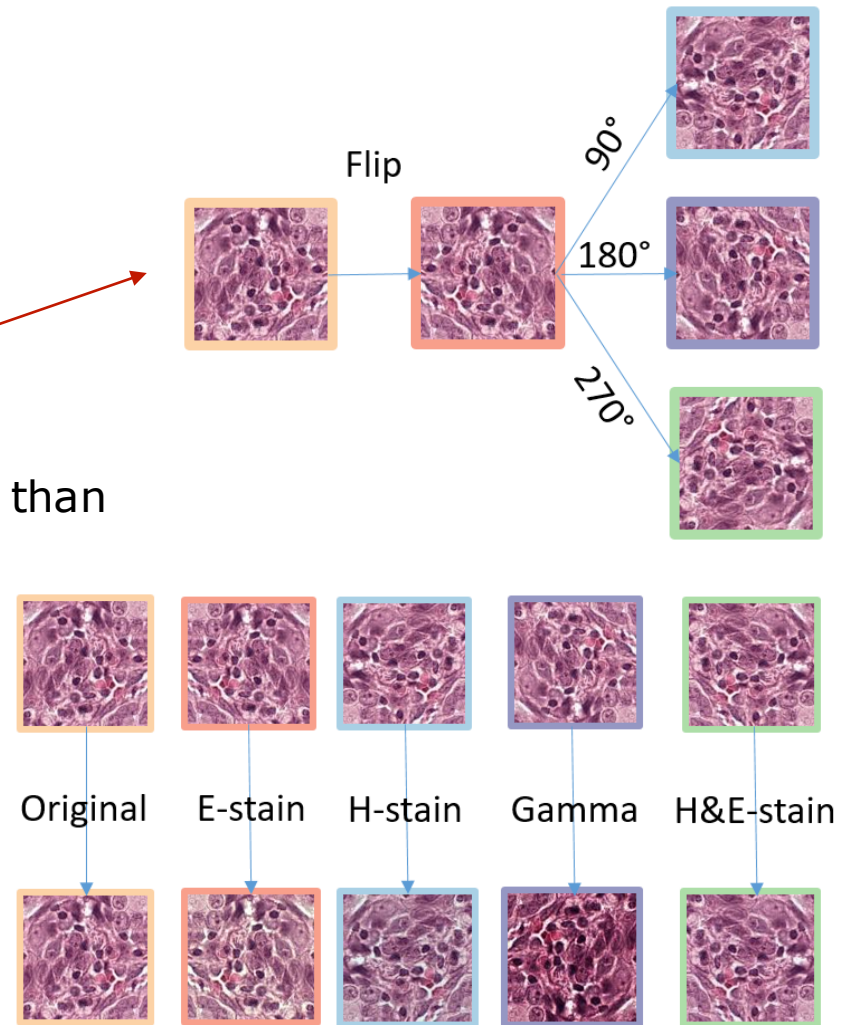
2<sup>nd</sup>:

- Run inference and obtain errors ( $N = 347000$ )
- Resample  $N$  tumor patches on tumor slides with URS
- Resample  $\frac{N}{2}$  normal patches on normal slides and  $\frac{N}{2}$  normal patches on tumor slides with URS
- More difficult dataset: 350K tumor, 700K normal



# Data augmentation

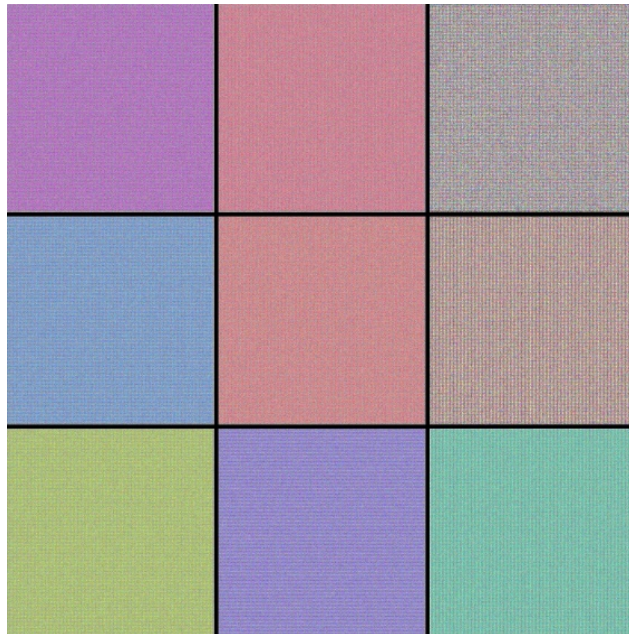
- Motivation
  - Rotational invariant
  - Stain/color invariant
- 4 spatial transformation
- Augment more stain variation digitally than available in the training data
  - H&E de-convolution
  - Randomly vary stains with  $\pm 10\%$
  - Convolve back to RGB
  - Stain variation
- Gamma-correction
  - Random  $\gamma$ -value [0.4-1.6]
  - Scanner setting variations
- 5x amount of data: 5.2M patches



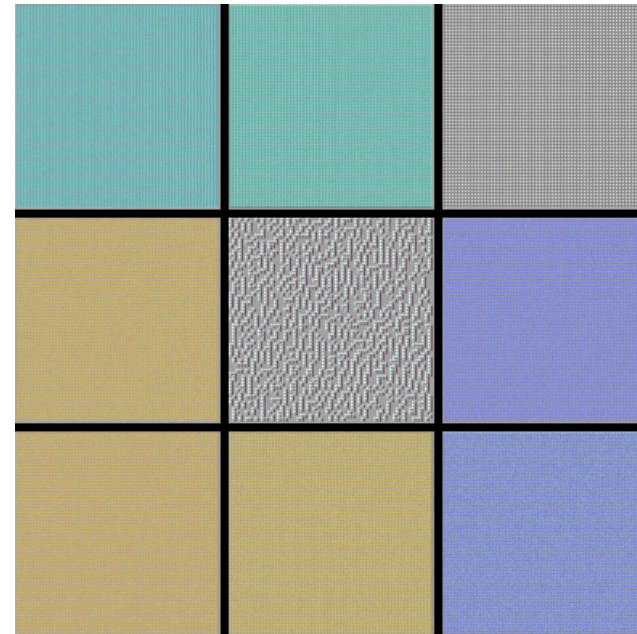


# Filters of first convolutional layer

Without augmentation



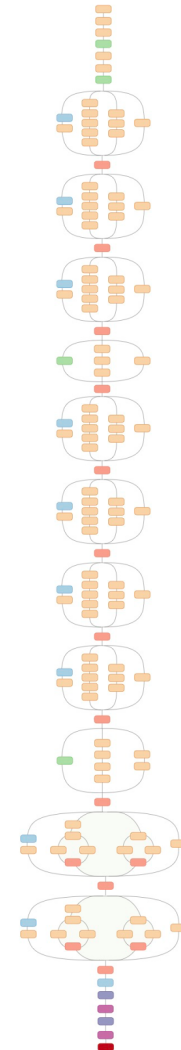
With augmentation



Visualization: Weight activation maximization using backpropagation

# ConvNet details

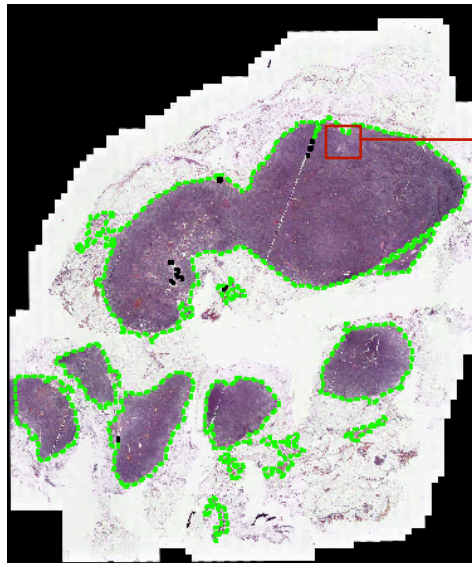
- Slightly changed Inception-V3 (C. Szegedy 2015)
- Extra fully-connected layer
- Dropout ( $p = 0.2$ ) to prevent overfitting
- Optimizer
  - SGD with Nesterov's momentum (0.9)
- Mini-batch size: 32
- Learning rate schedule:
  - Initial: 0.1
  - Dropped 50% every 125K gradient update
- Single NVIDIA GTX 1080 GPU
- Patch-based performance (full sampling)
  - Camelyon16 test: 96.6 % accuracy



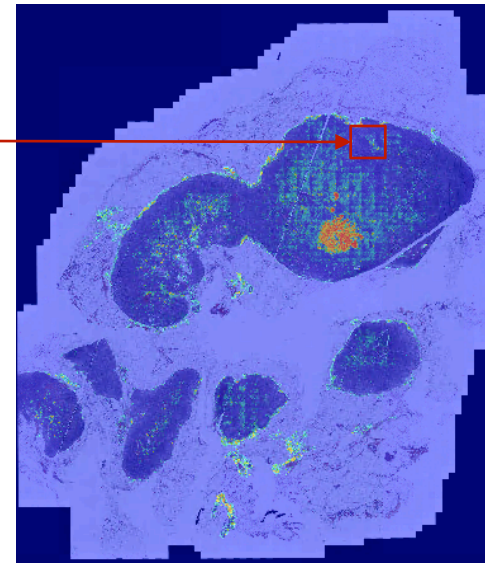


# WSI inference

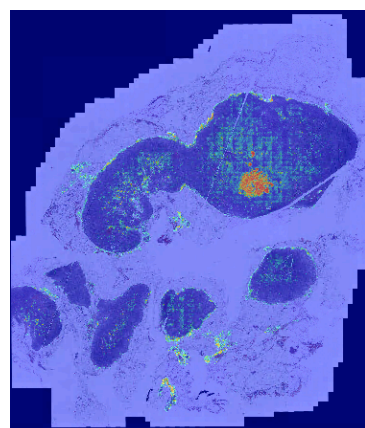
- Sliding field-of-view (FOV)
- Step-length: 128 pixels
- Each FOV is assigned a probability  $p \in [0,1]$
- No ensemble average due to limited time



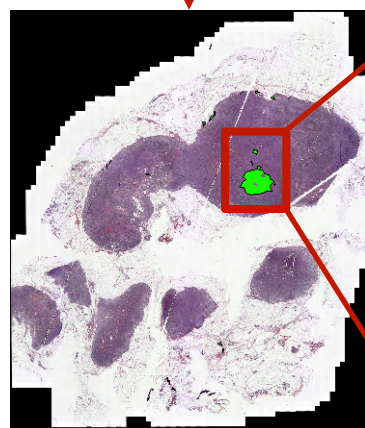
ConvNet



# Slide-based features



$p > 0.5$   $I_{IHS} < 200$



5 largest objects

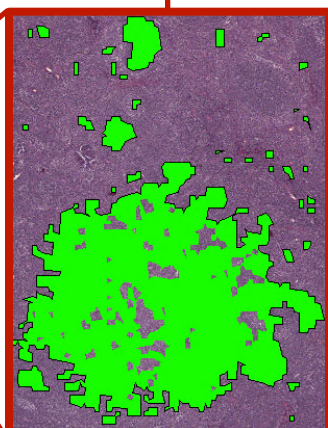
Area  
Max probability  
Mean probability  
Mean filter  
response of  
polynomial blob  
filter  
...

42 features

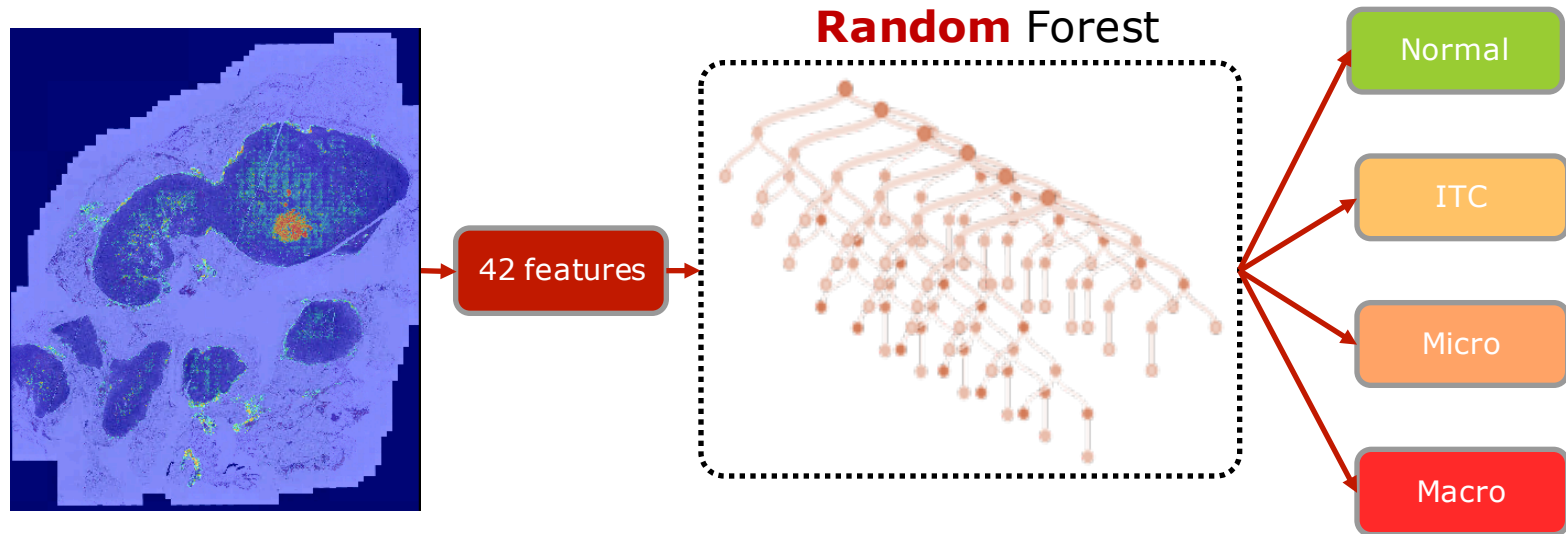
Max and mean  
probability in all  
objects

Area

<0.3mm



# Slide classification



- Hyperparameters tuned on Camelyon17 training split using 3-fold cross-validation
- Tested on Camelyon17 validation split
- Features with highest feature importance:
  - Maximum probability and area of largest object

# Patient classification

- Rule-based scheme only
- Official pN-staging system in Camelyon17

pN-stage	Rule
pN0	No micro-metastases or macro-metastases or ITCs found.
pN0(i+)	Only ITCs found.
pN1mi	Micro-metastases found, but no macro-metastases found.
pN1	Metastases found in 1–3 lymph nodes, of which at least one is a macro-metastasis.
pN2	Metastases found in 4–9 lymph nodes, of which at least one is a macro-metastasis.

# THANK YOU

## Contact

E-mail: [jeppe\\_thagaard@icloud.com](mailto:jeppe_thagaard@icloud.com)

Twitter: [@Jeppethagaard](https://twitter.com/Jeppethagaard)

LinkedIn: [www.linkedin.com/in/jeppethagaard](https://www.linkedin.com/in/jeppethagaard)

## DTU Compute

Department of Applied Mathematics and Computer Science

---

