On explainable neural networks in practice

Tomáš Brázdil

```
tensor([[[[-4.8129e-04, -5.6481e-04,  2.4511e-03],
          [ 1.2085e-03,  2.7492e-03,  1.7652e-03],
          [-3.9402e-03, -6.9425e-03, -7.2726e-03]],

         [[ 3.3416e-03, -9.7720e-03, -1.7814e-02],
          [ 4.2282e-03, -7.6528e-03, -1.6922e-02],
          [ 5.5957e-03, -5.9063e-03, -1.6483e-02]],

         [[-1.1220e-02, -1.1452e-02, -1.2850e-02],
          [-9.9574e-03, -8.5489e-03, -6.6620e-03],
          [-4.7562e-03, -3.6714e-03, -1.2898e-05]],

         ...,

         [[-1.5140e-02, -5.0111e-03,  4.2080e-03],
          [-1.6666e-02, -8.6684e-03, -5.1674e-05],
          [-1.8180e-02, -1.2616e-02, -5.7304e-03]],

         [[-8.0962e-03, -6.7141e-03, -2.9794e-03],
          [-1.1881e-02, -6.3858e-03,  2.7343e-03],
```

What is this good for?

```
tensor([[[[-4.8129e-04, -5.6481e-04,  2.4511e-03],
          [ 1.2085e-03,  2.7492e-03,  1.7652e-03],
          [-3.9402e-03, -6.9425e-03, -7.2726e-03]],

         [[ 3.3416e-03, -9.7720e-03, -1.7814e-02],
          [ 4.2282e-03, -7.6528e-03, -1.6922e-02],
          [ 5.5957e-03, -5.9063e-03, -1.6483e-02]],

         [[-1.1220e-02, -1.1452e-02, -1.2850e-02],
          [-9.9574e-03, -8.5489e-03, -6.6620e-03],
          [-4.7562e-03, -3.6714e-03, -1.2898e-05]],

         ...,

         [[-1.5140e-02, -5.0111e-03,  4.2080e-03],
          [-1.6666e-02, -8.6684e-03, -5.1674e-05],
          [-1.8180e-02, -1.2616e-02, -5.7304e-03]],

         [[-8.0962e-03, -6.7141e-03, -2.9794e-03],
          [-1.1881e-02, -6.3858e-03,  2.7343e-03],
```

What is this good for? Digital pathology …

Challenges in Pathology

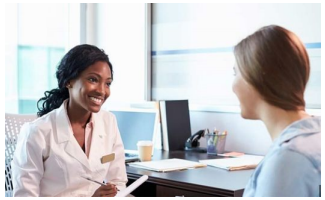
Laborious and time-consuming routine effort

Increasing workload due to cancer screening programs (cervix, breast, colorectal, recently prostate, lung)

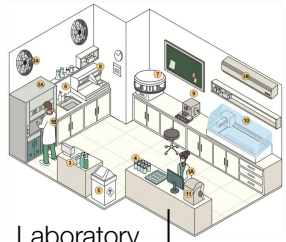
Few experienced pathologists

Human error prone: tired pathologist

Personal/spatial issues at smaller pathology departments, some pathologist working for part time for more laboratories



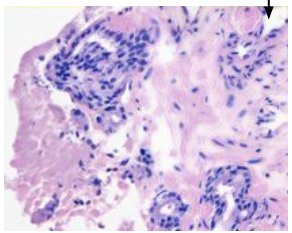
Tissue sample



Laboratory

Slide

Visual inspection



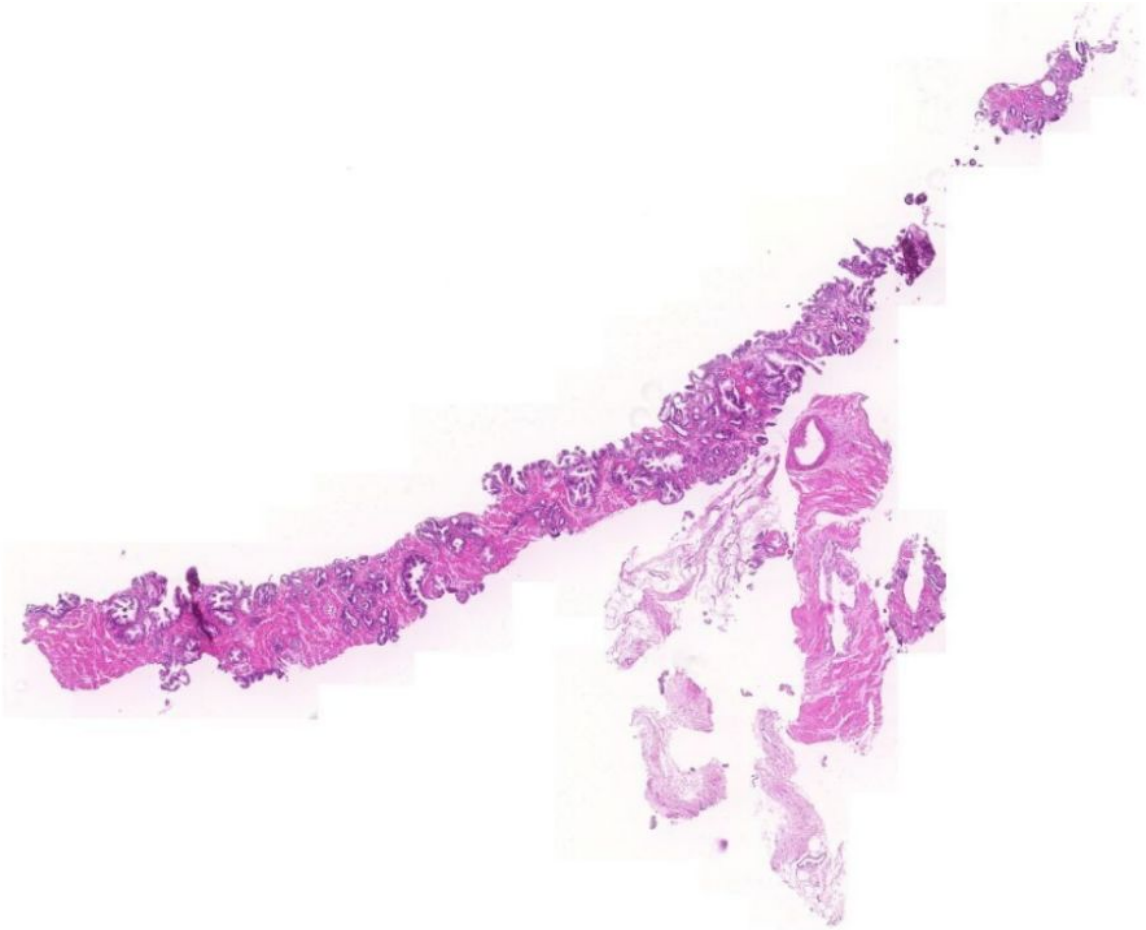
Cancer

Diagnosis

Cancer detection

Microscopic scan of tissue

- Magnification 20x
- 0.172 μm / pixel
- 105,185 px × 221,772 px
- Hematoxylin-eosin stained

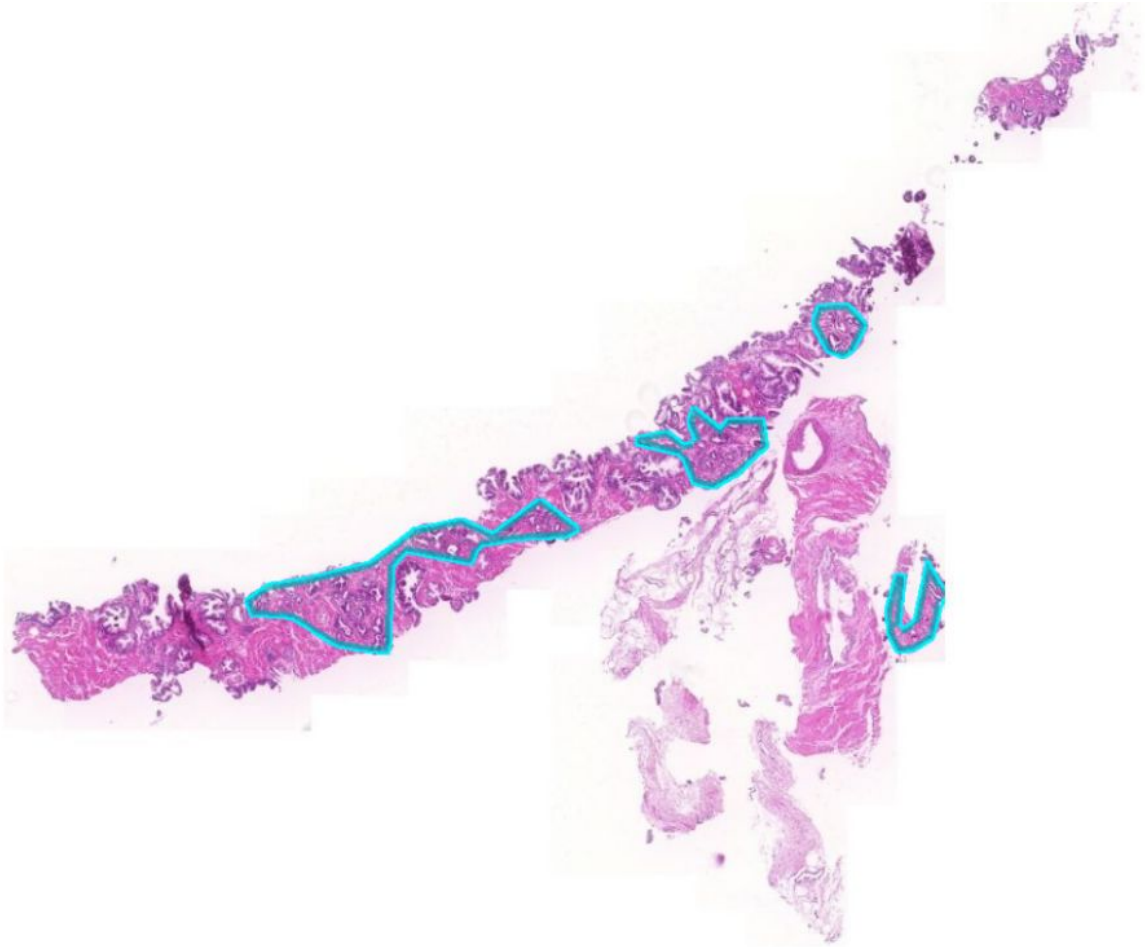


Cancer detection

Microscopic scan of tissue

- Magnification 20x
- 0.172 μm / pixel
- 105,185 px × 221,772 px
- Hematoxylin-eosin stained

Tumor annotation



Cancer detection

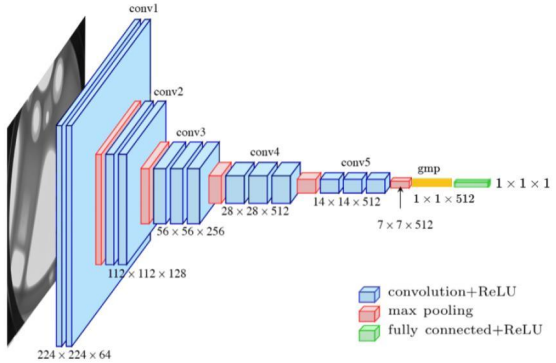
Microscopic scan of tissue

- Magnification 20x
- 0.172 μm / pixel
- 105,185 px × 221,772 px
- Hematoxylin-eosin stained

Tumor annotation

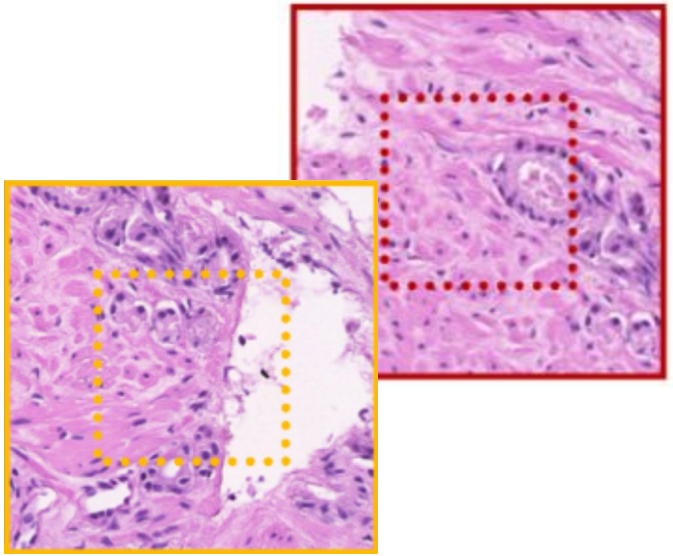
Tumor prediction using ML

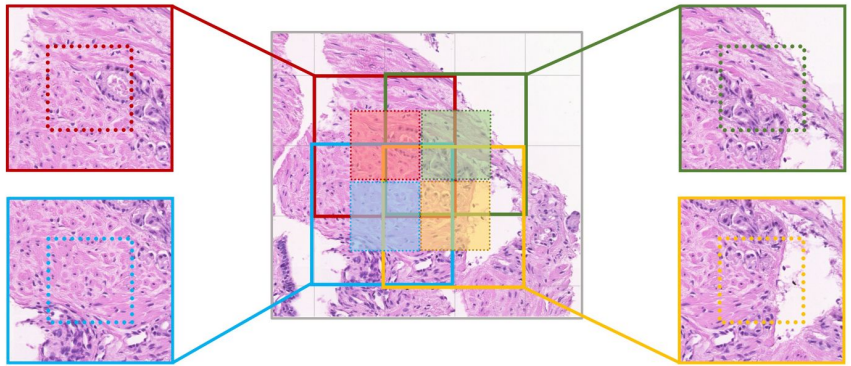
(VGG-16 with an alternative head)



AI models training

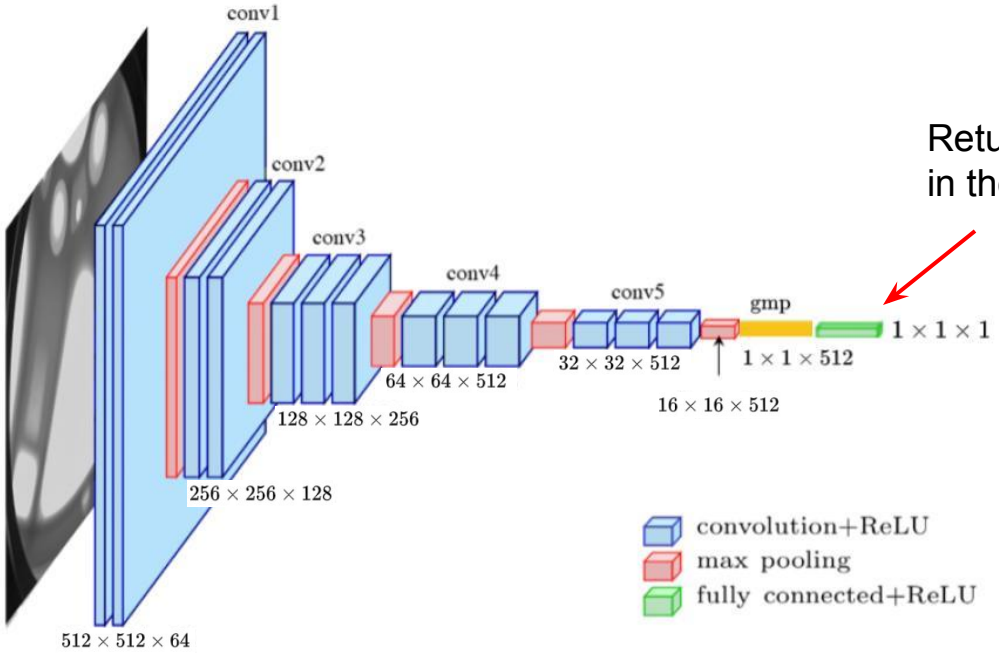
- Training data
  - Provided by MMCI
  - 785 scans, 166 patients
- Model trained on patches 512 x 512 px
  - Patches cover the tissue and overlap (stride 128)
  - 7,878,675 patches for training
- Binary classification problem (cancer positive/negative)
  - A patch labeled positive iff its center square intersects the tumor annotation





The neural network - modified VGG-16

Input patch
(imagine it's purple!)



Returns probability of cancer
in the center of the input patch

AI model for testing

- Testing
  - 87 scans
  - 98 % AUC in patch-level tumor detection

100 % prediction accuracy in slide-level tumor detection with a threshold close to 1
(slide level tumor probability = maximum of patch level probabilities)

Prediction with threshold 0.5

		PREDICTED	
		Pos	Neg
TRUE	Pos	24796	4340
	Neg	5345	158754

AI model for testing

- Testing
  - 87 scans
  - 98 % AUC in patch-level tumor detection

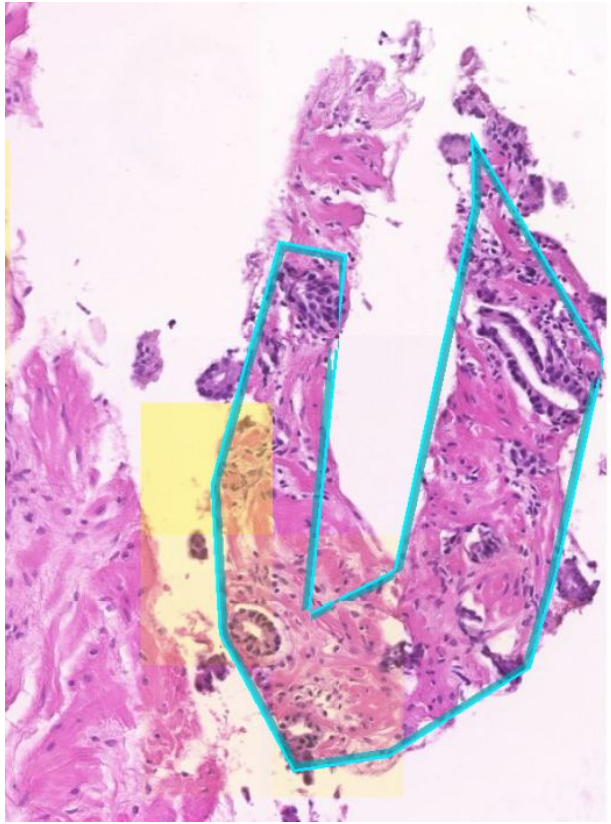
100 % prediction accuracy in slide-level tumor detection with a threshold close to 1
(slide level tumor probability = maximum of patch level probabilities)

Prediction with threshold 0.5

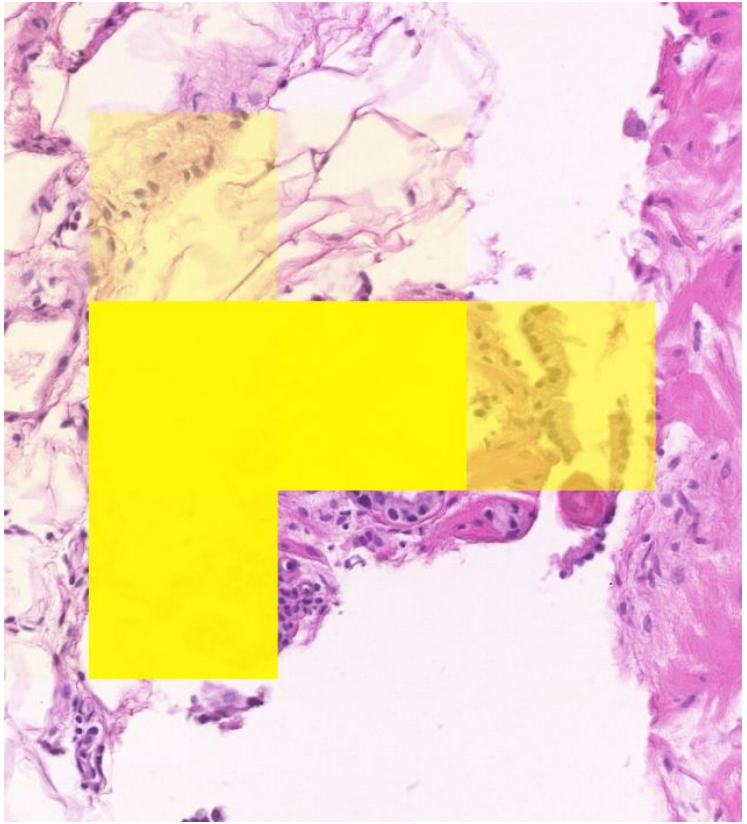
		PREDICTED Pos	PREDICTED Neg
TRUE	Pos	24796	4340
	Neg	5345	158754

… does it work in practice??

… how to persuade pathologists that it works??



Should be yellow completely!



Should not be yellow at all!

Interpretation of the behavior

- What exactly is the network searching for?

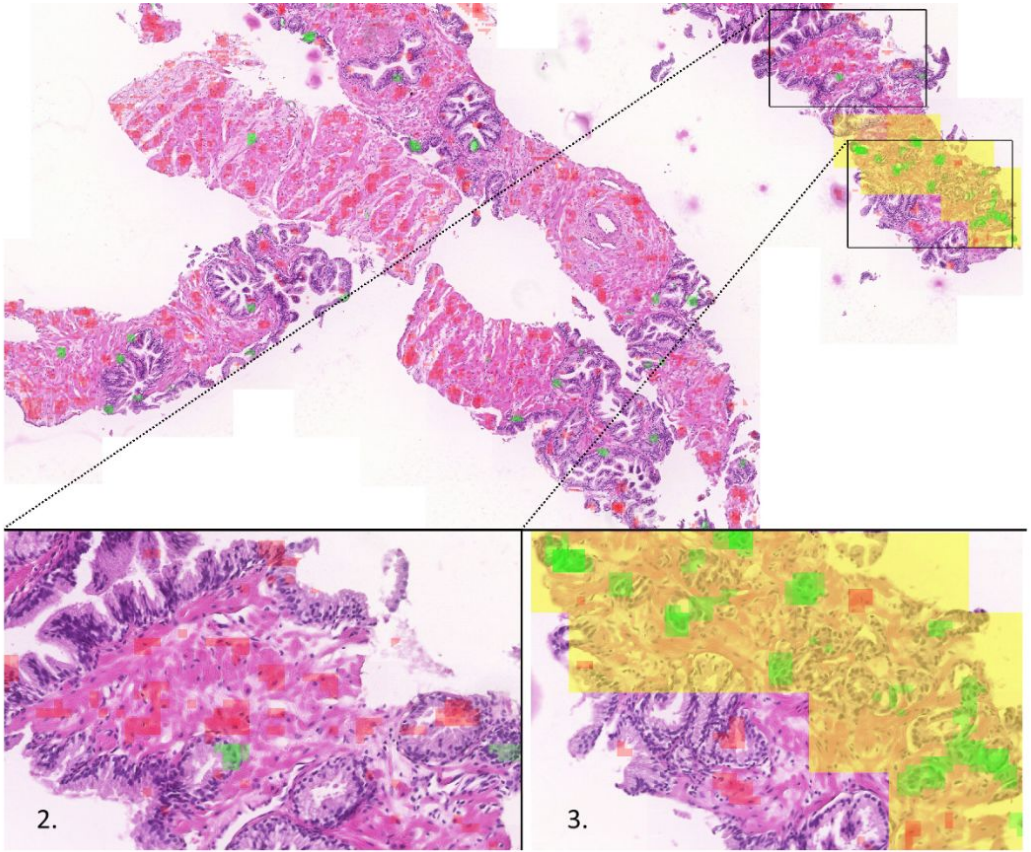
Interpretation of the behavior

- What exactly is the network searching for?
- Does it understand cancer?

Interpretation of the behavior

- What exactly is the network searching for?
- Does it understand cancer?
- Does it (at least) look for sensible patterns?
  - How to find out what patterns it looks for?
  - How to explain that the patterns make sense?
  - How to make sure that we have understood all patterns?

Interpretation



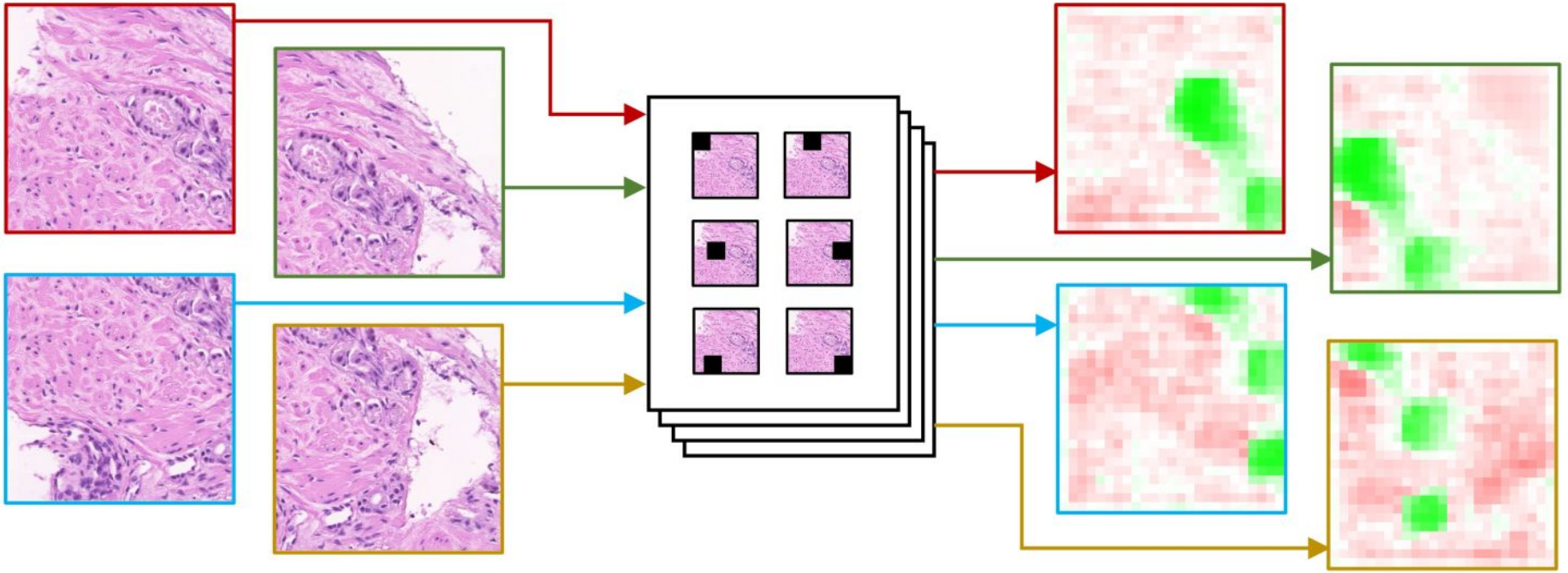
Tumor prediction by AI

Areas with positive impact on the prediction

Areas with negative impact on the prediction

Using simple occlusion sensitivity analysis

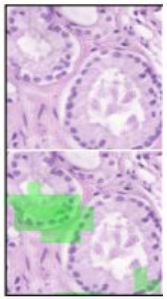
Occlusion sensitivity analysis



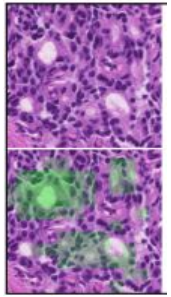
Catalog of typical patterns

Pro cancer:

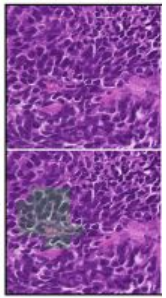
Single chain of nuclei



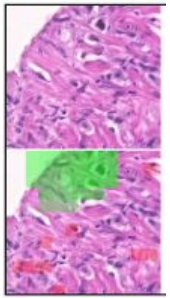
Small round hole



High nuclear density

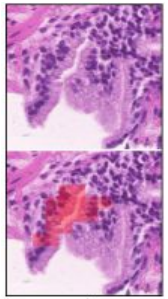


Large nuclei with halo

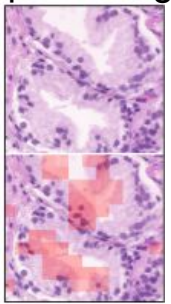


Con cancer:

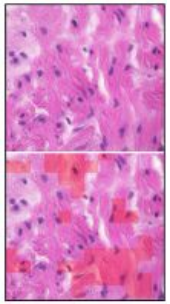
Double chains of nuclei



Chain of nuclei with eosinophilic neighborhood



Low cellular density



Interpretable patterns

- Randomly selected >600 points (xPOI) with “high” occlusion sensitivity
  - Square region 15 x 15 px around the point
  - Either green or red color prevails in the square
- Tissue surrounding xPOIs classified by the catalog of typical tissue patterns

90 % of identified patterns have a known pattern!

Morphological pattern under attribution	WSIs w/ carcinoma						WSIs w/o carc.	Tot. %
	Gleason							
	3+3 (N=14)	3+4 (N=3)	4+3 (N=11)	4+4 (N=5)	4+5 (N=4)	**Total (N=37)**	N=49	
Single chain of nuclei (TP1)	*52*	*7*	*65*	*6*	*2*	**132**	-	**132 (20.4%)**
Small round hole (TP2)	*12*	*2*	*24*	*11*	*8*	**57**	-	**57 (8.8%)**
High nuclear density (TP3)	*2*	*1*	*17*	*17*	*11*	**48**	-	**48 (7.4%)**
Larger nucleus with perinuclear halo (TP4)	*1*	*0*	*6*	*1*	*8*	**16**	-	**16 (2.5%)**
Undefined	*1*	*0*	*0*	*0*	*0*	**1**	-	**1 (0.2%)**
Single chain of nuclei (FP1)	5	0	0	1	0	**6**	29	**35 (5.4%)**
Small round hole (FP2)	4	0	0	0	0	**4**	35	**39 (6.0%)**
High nuclear density (FP3)	5	3	0	2	0	**10**	8	**18 (2.8%)**
Larger nucleus with perinuclear halo (FP4)	0	1	2	0	0	**3**	12	**15 (2.3%)**
Undefined	1	0	0	0	0	**1**	1	**2 (0.3%)**
Two-layered chain of nuclei (TN1)	13	2	11	11	6	**43**	29	**72 (11.1%)**
Areas of low nuclear density with eosinophilic background (TN2)	23	6	4	2	2	**37**	125	**162 (25.1%)**
Chain of nuclei with abundant slightly eosinophilc neighborhood (TN3)	5	0	1	1	2	**9**	30	**39 (6.0%)**
Undefined	4	1	0	2	0	**7**	3	**10 (1.5%)**

Interpretable patterns

- Randomly selected >600 points (xPOI) with “high” occlusion sensitivity
  - Square region 15 x 15 px around the point
  - Either green or red color prevails in the square
- Tissue surrounding xPOIs classified by the catalog of typical tissue patterns

90 % of identified patterns have a known pattern!

But does it make sense in pathology??

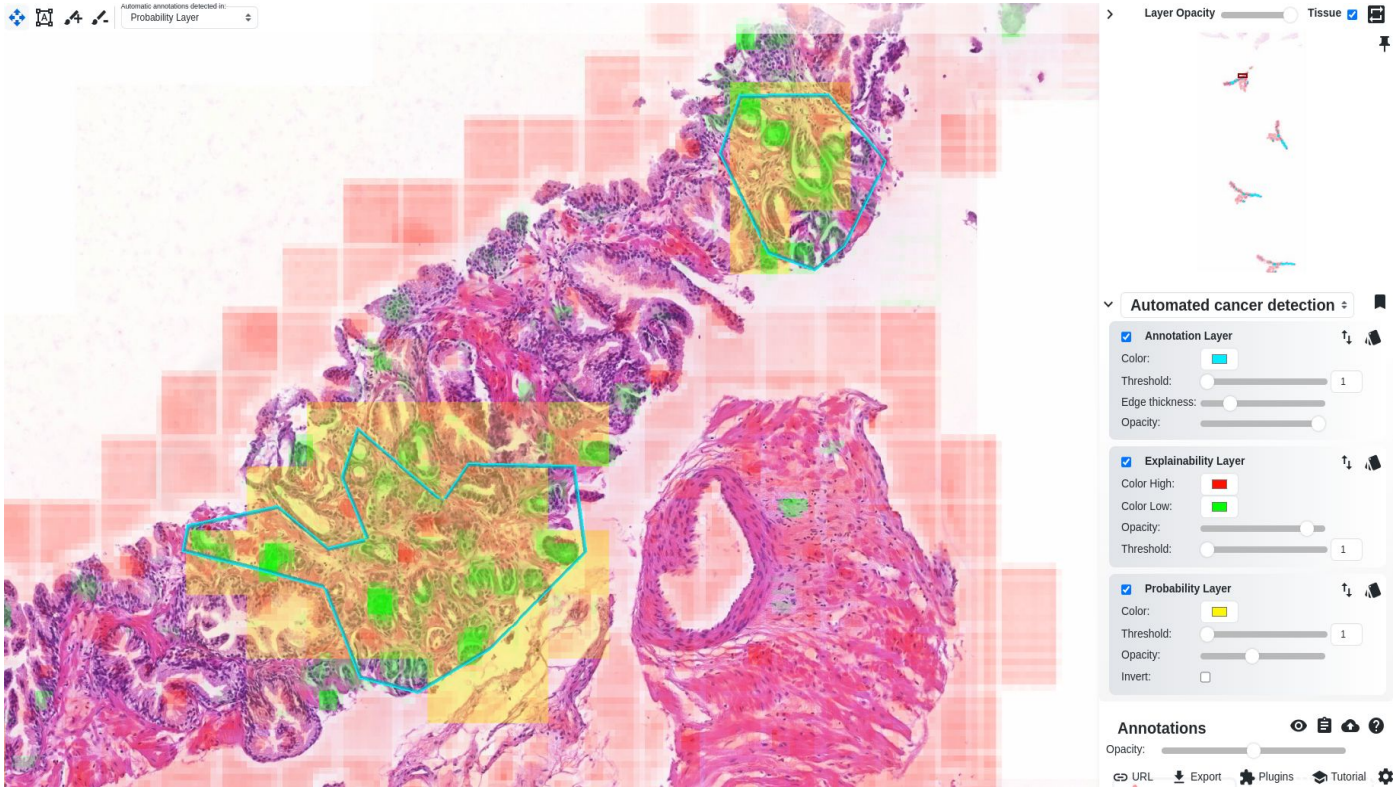
Morphological pattern under attribution	WSIs w/ carcinoma						WSIs w/o carc.	Tot. %
	Gleason							
	3+3 (N=14)	3+4 (N=3)	4+3 (N=11)	4+4 (N=5)	4+5 (N=4)	**Total (N=37)**	N=49	
Single chain of nuclei (TP1)	*52*	*7*	*65*	*6*	*2*	**132**	-	**132 (20.4%)**
Small round hole (TP2)	*12*	*2*	*24*	*11*	*8*	**57**	-	**57 (8.8%)**
High nuclear density (TP3)	*2*	*1*	*17*	*17*	*11*	**48**	-	**48 (7.4%)**
Larger nucleus with perinuclear halo (TP4)	*1*	*0*	*6*	*1*	*8*	**16**	-	**16 (2.5%)**
Undefined	*1*	*0*	*0*	*0*	*0*	**1**	-	**1 (0.2%)**
Single chain of nuclei (FP1)	5	0	0	1	0	**6**	29	**35 (5.4%)**
Small round hole (FP2)	4	0	0	0	0	**4**	35	**39 (6.0%)**
High nuclear density (FP3)	5	3	0	2	0	**10**	8	**18 (2.8%)**
Larger nucleus with perinuclear halo (FP4)	0	1	2	0	0	**3**	12	**15 (2.3%)**
Undefined	1	0	0	0	0	**1**	1	**2 (0.3%)**
Two-layered chain of nuclei (TN1)	13	2	11	11	6	**43**	29	**72 (11.1%)**
Areas of low nuclear density with eosinophilic background (TN2)	23	6	4	2	2	**37**	125	**162 (25.1%)**
Chain of nuclei with abundant slightly eosinophilc neighborhood (TN3)	5	0	1	1	2	**9**	30	**39 (6.0%)**
Undefined	4	1	0	2	0	**7**	3	**10 (1.5%)**

Feature type	Scale	Feature	Features (≈50 µm ⌀)	Found	xPatterns
Architectural	L	distorted gland architecture		No*	
		small uniform glands infiltrate in between normal glands	OoS	No*	
		poorly formed fused, cribriform or glomeruloid glands, high nuclear density in Gleason pattern 4	SL, HCD	Yes	small round hole, high nuclear density
		solid sheets, cords, medium or large nests with rosettes, comedo type necrosis	SL, HCD	Yes	small round hole, high nuclear density
	M	small caliber glands	SLE, SL	Yes	single chain of nuclei, small round hole
		crowded or compact gland clusters	HCD	Yes	high nuclear density
		blue mucin	AA	No	
Intraluminal		eosinophilic amorphous secretions	AA	No	
		crystalloids		No	
		rigid or sharp gland lumina, may have periglandular clefts	SL, SLE	Yes	small round hole, single chain of nuclei
		glands lack basal cells (single-layered epithelium in Gleason pattern 3)	SLE	Yes	single chain of nuclei
		infiltrative single cells in Gleason pattern 5	HNH	Yes	larger nucleus with perinuclear halo
Cytoplasmic		cuboidal to low cylindrical cells with modest cytoplasm	SLE, SL	Yes	single chain of nuclei, small round hole
Nuclear		enlarged hyperchromatic nuclei	HNH	Yes	larger nucleus with perinuclear halo
	S	prominent enlarged nucleoli often eosinophilic	OoS	N/A	
		multiple nucleoli located in periphery	OoS	N/A	

Mapping identified patterns to “textbook” features used in cancer diagnostics

- OoS – out of scale
- SL – small lumina
- HCD – high cellular density
- SLE – single-layered epithelium
- AA – acellular areas
- HNH - hyperchromatic nuclei with Halo

xOpat toolbox



https://rationai-vis.ics.muni.cz/visualization-demo/client/index.php

AI **diagnosis and explanation**

Fast zooming, ergonomic

Flexible, fast adaptation to different tasks

Interactive, allows annotation etc.

Web interface, uploading of WSI, automated analysis

Deployment at MMCI

- Our system:
  - Pathologists upload patient’s scans via web form
  - Our GPU servers execute the NN inference and produce the probability overlay map
  - The online viewer overlays the map over the scan

Deployment at MMCI

- Our system:
  - Pathologists upload patient’s scans via web form
  - Our GPU servers execute the NN inference and produce the probability overlay map
  - The online viewer overlays the map over the scan
- We need
  - Fast network communication of large images (compression is the way to go)
  - Fast inference on large images (millions of patches, need to utilize sufficient hw)
  - Very fast operation of the viewer
    - pathologists are **extremely** fast and efficient when zooming with their microscopes
    - … most of them are very impatient

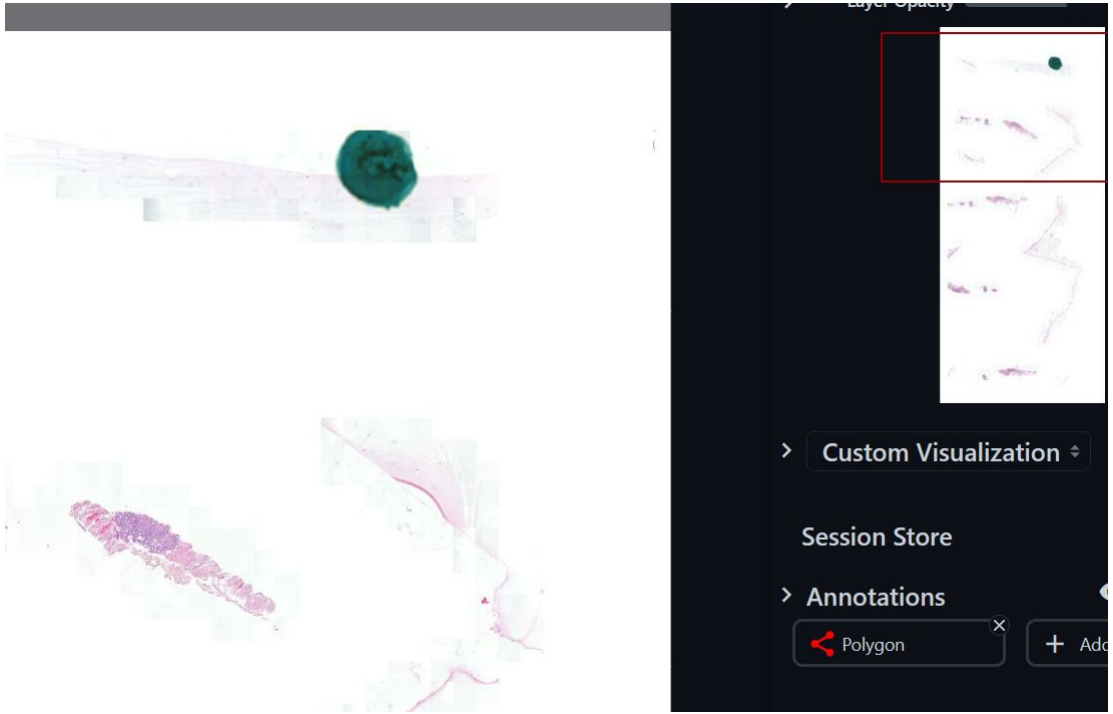
Deployment at MMCI

- Our system:
  - Pathologists upload patient's scans via web form
  - Our GPU servers execute the NN inference and produce the probability overlay map
  - The online viewer overlays the map over the scan
- We need
  - Fast network communication of large images (compression is the way to go)
  - Fast inference on large images (millions of patches, need to utilize sufficient hw)
  - Very fast operation of the viewer
    - pathologists are **extremely** fast and efficient when zooming with their microscopes
    - … most of them are very impatient
- Our system has examined approx. 50 patients;
  our pathologist uses it as an assistant system

So far the operation at MMCI was almost flawless - the system simply looks for the pro/con cancer patterns mentioned previously

So far the operation at MMCI was almost flawless - the system simply looks for the pro/con cancer patterns mentioned previously

Just on Friday last week … no cancer detected (huge tumor present) since the green dot completely confused the part detecting the tissue in the scan



Conclusions

- How to help pathologists in their routine work using ML methods?

Conclusions

- How to help pathologists in their routine work using ML methods?
- Easy, construct and train a model that will be
  - Reliable
  - Understandable

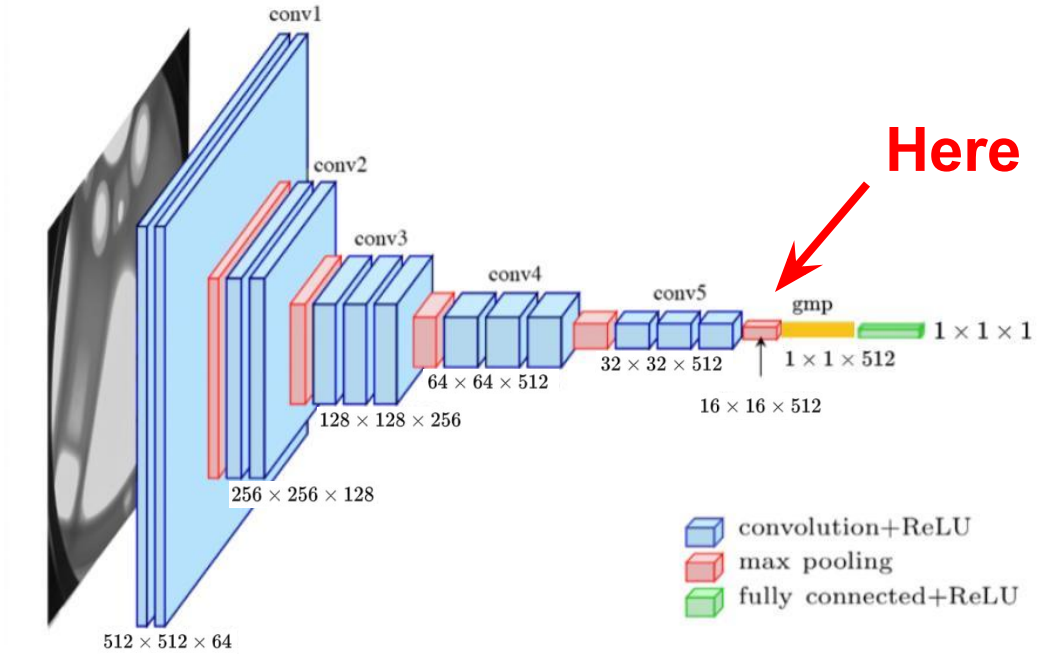
Conclusions

- How to help pathologists in their routine work using ML methods?
- Easy, construct and train a model that will be
  - Reliable
  - Understandable
- Both reliability and understandability are subjective
  - Pathologists cannot afford “obvious” mistakes
  - Pathologists understand pathology, not computer science!
  - Computer scientists do not understand pathology
  - Computer scientists do not see “obvious” mistakes

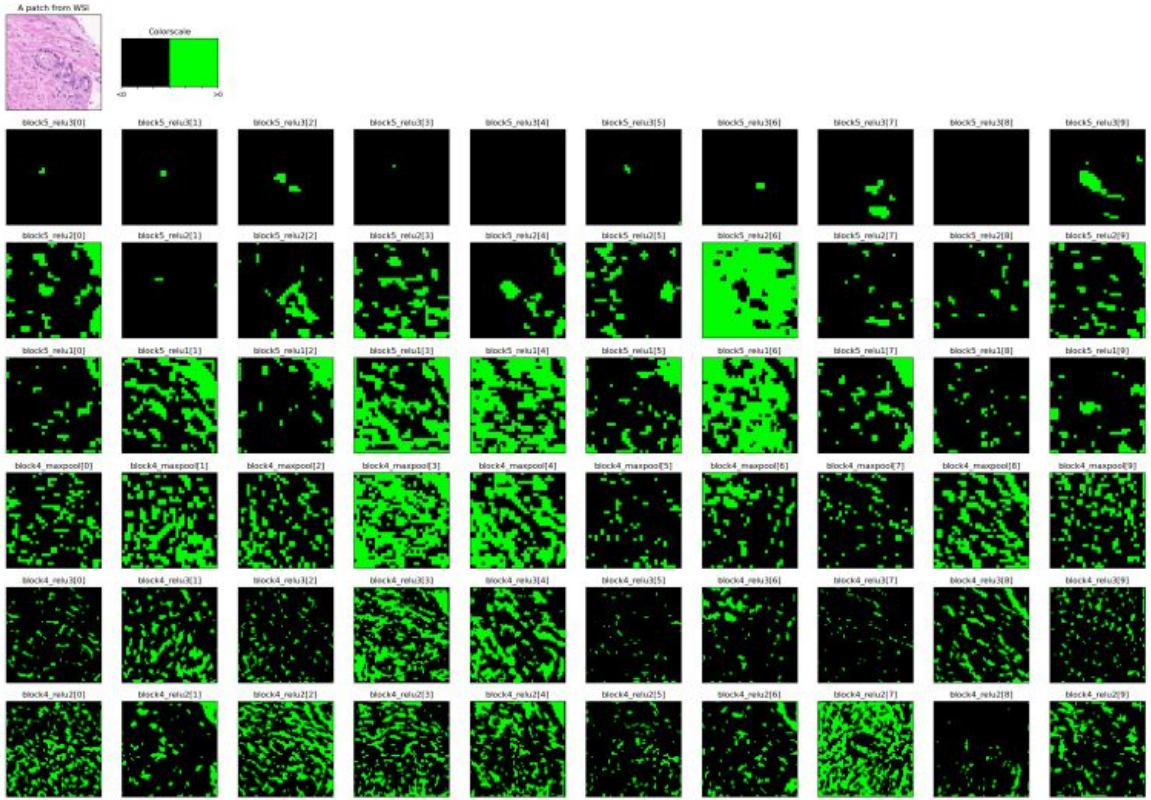
Conclusions

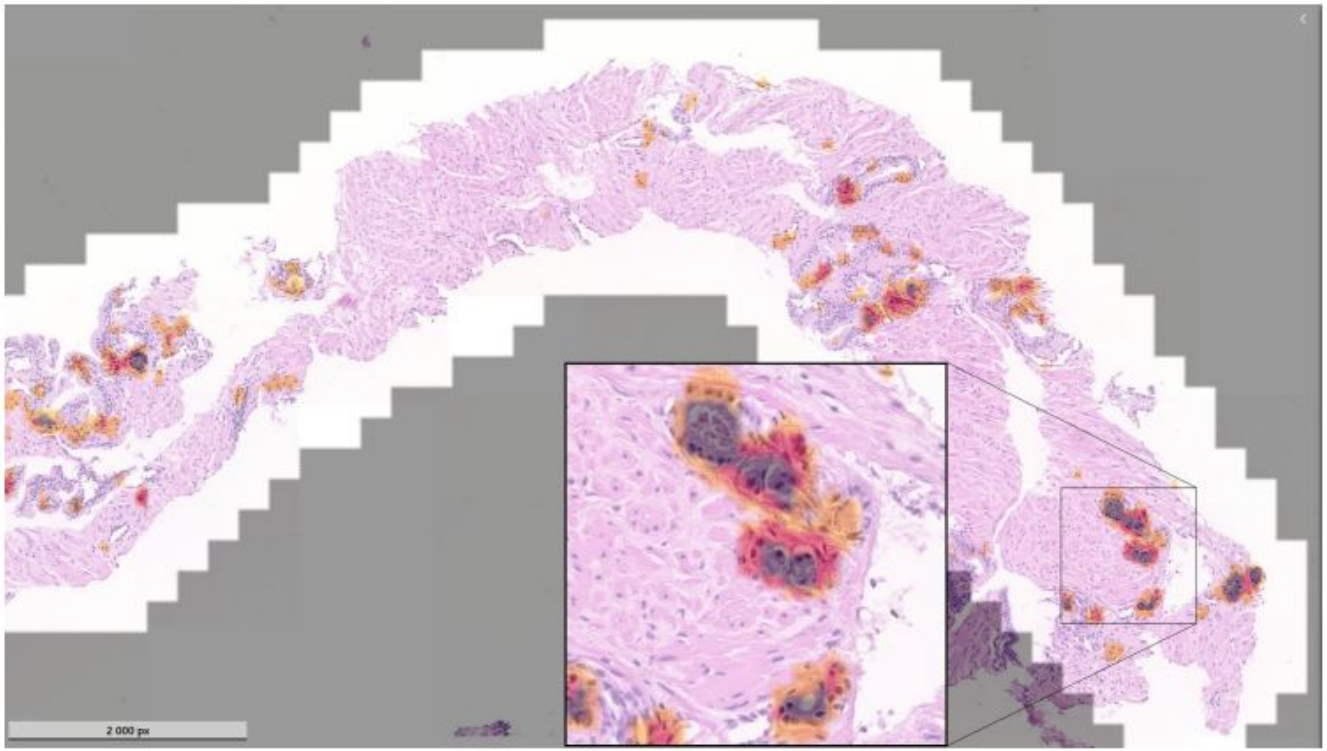
- How to help pathologists in their routine work using ML methods?
- Easy, construct and train a model that will be
  - Reliable
  - Understandable
- Both reliability and understandability are subjective
  - Pathologists cannot afford “obvious” mistakes
  - Pathologists understand pathology, not computer science!
  - Computer scientists do not understand pathology
  - Computer scientists do not see “obvious” mistakes
- The technology must be ready and reliable - we are working with highly efficient professionals who cannot afford to play with weird occasionally non-functioning toys!

What does the network really think?



What does the network really think? (selected maps)



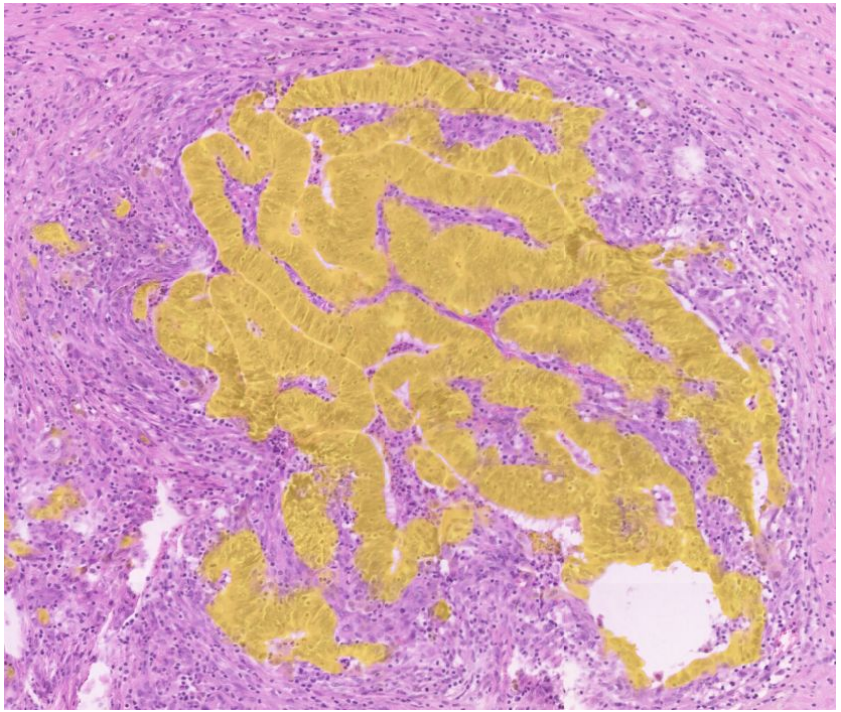


- For each input pixel consider “corresponding” pixels from 512 feature maps
- I.e. for each input pixel we get a “fingerprint” vector dim. 512 measuring “stimulation” of feature maps at the spatial position of the pixel
- Cluster input pixels according to these 512 dimensional fingerprint vectors

Virtual staining

- AI based epithelium segmentation
- **New virtual staining method**

- Model predicts immunohistochemical staining based on H&E inputs
- Trained on scans with dual staining
  - H&E first
  - Re-stained using an immunohistochemical staining
- Trained on scans of the breast and colon cancer, successfully transfered to scans of the prostate cancer



People

- Memorial Masaryk Cancer Institute
  - **Department of pathology**

    MUDr. Rudolf Nenutil, CSc., MUDr. Michal Tichý, Ph.D.
- Masaryk University
  - **Faculty of informatics**

    Doc. RNDr. Tomáš Brázdil, Ph.D., Mgr. Matej Gallo,

    Bc. Jiří Horák, Bc. Adam Bajger, …
  - **Institute for Computer Science**

    Doc. RNDr. Petr Holub, Ph.D.