# Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis

Prepared by

Niloy Roy

# Histopathology

- Three Greek words:
  histos 'tissue',
  pathos 'suffering',
  logia 'study of‘
- Microscopic examination of tissue
- Diagnosis of cancers(lung, breast, prostate)
- Staining Techniques
  Hematoxylin and eosin (H&E)
  Immunohistochemical

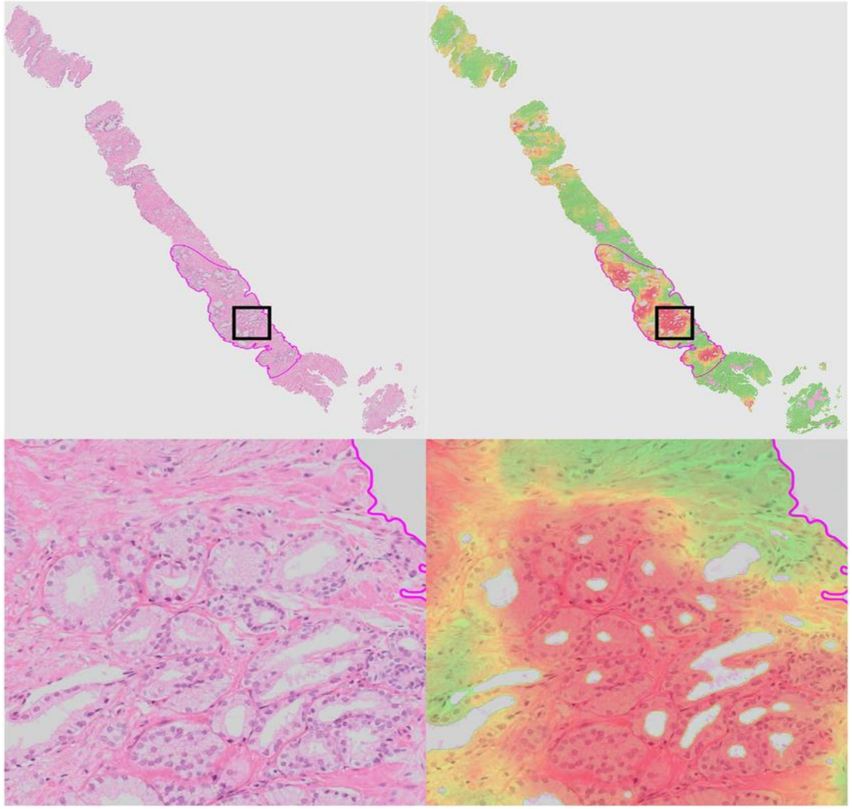


Figure: a whole slide prostate biopsy specimens with 30% cancer.

# Research Problem

**Challenges for pathologists**

- Increased workload
- Balance between efficiency and accuracy

**Reasons**

- Increase in cancer incidents
- Personalized medicines based on genetics, molecular profile, and disease stage
- Patient-specific treatment options (e.g., hormone therapy or HER2-targeted therapy)
- Increased parameters (e.g., surface area, mitotic counts) in quantification and standardization

# Context

Introduction of new technologies like digital whole slide images (WSI) paved the way of deploying deep learning image analysis techniques in diagnosis.

- Comprehensive overview of the entire slide on computer screen,
- Annotation, measurements by specialized software
- More objectivity and flexibility with less lab work
- Easier quantification and standardization

# Objective

Evaluation of performance for convoluted neural network (CNN) in two specific cases:

1. Prostate cancer identification in biopsy specimen

2. Breast cancer metastasis detection in sentinel lymph nodes

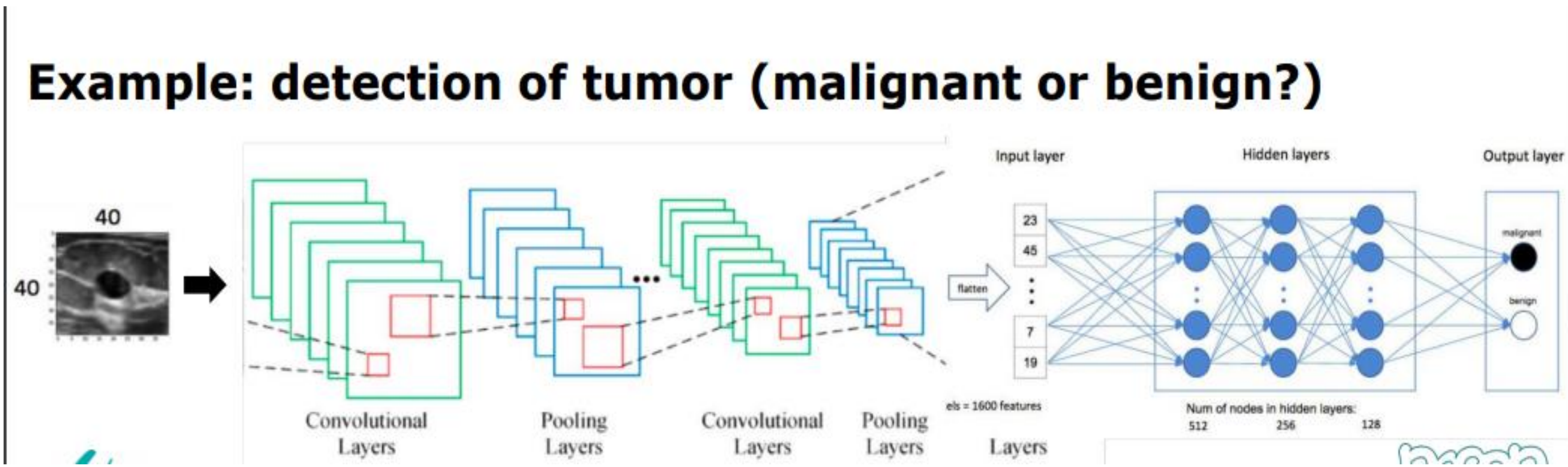
# Experiment

The experiment methodology can be discussed on the following points:

- CNN algorithm
- Data collection
- Materials
- Digitization and annotation
- Pre-processing Steps
- Model Training and application

# Convolution Neural Network

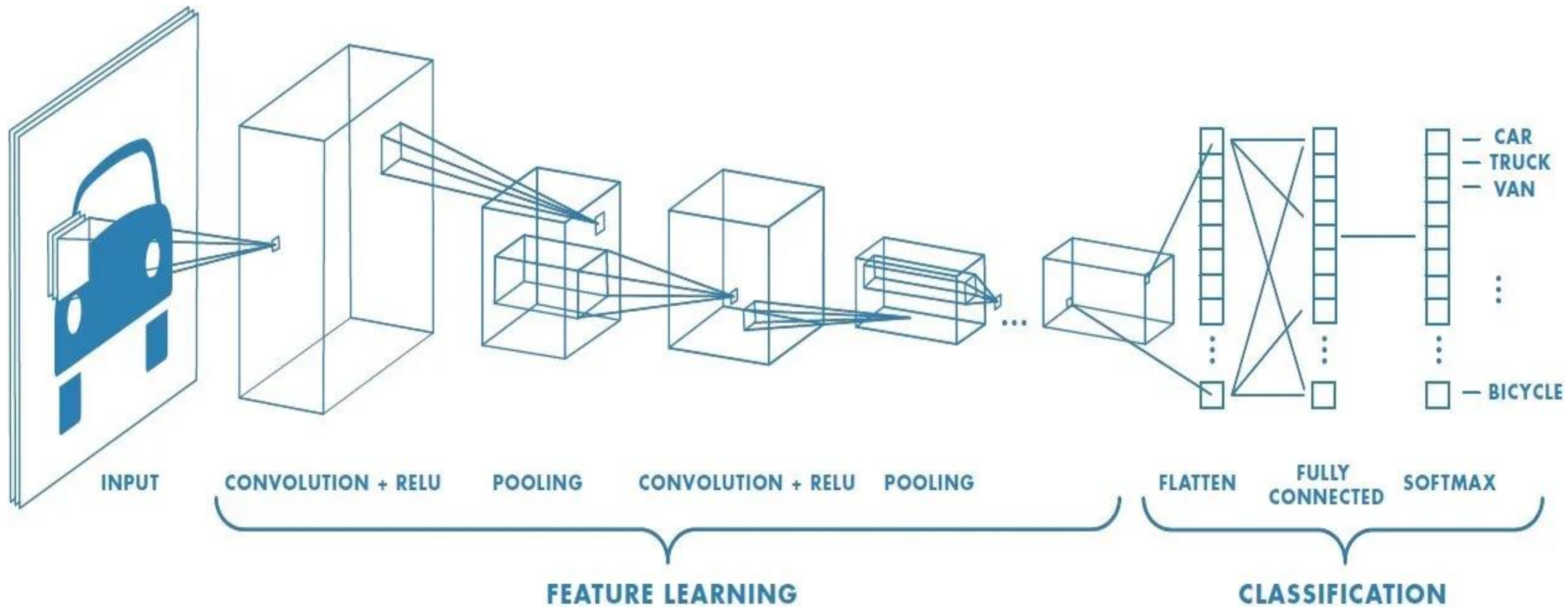
- State of the art in image recognition
- Learns relevant features from huge number of training images



# Convolution Neural Network

Pooling

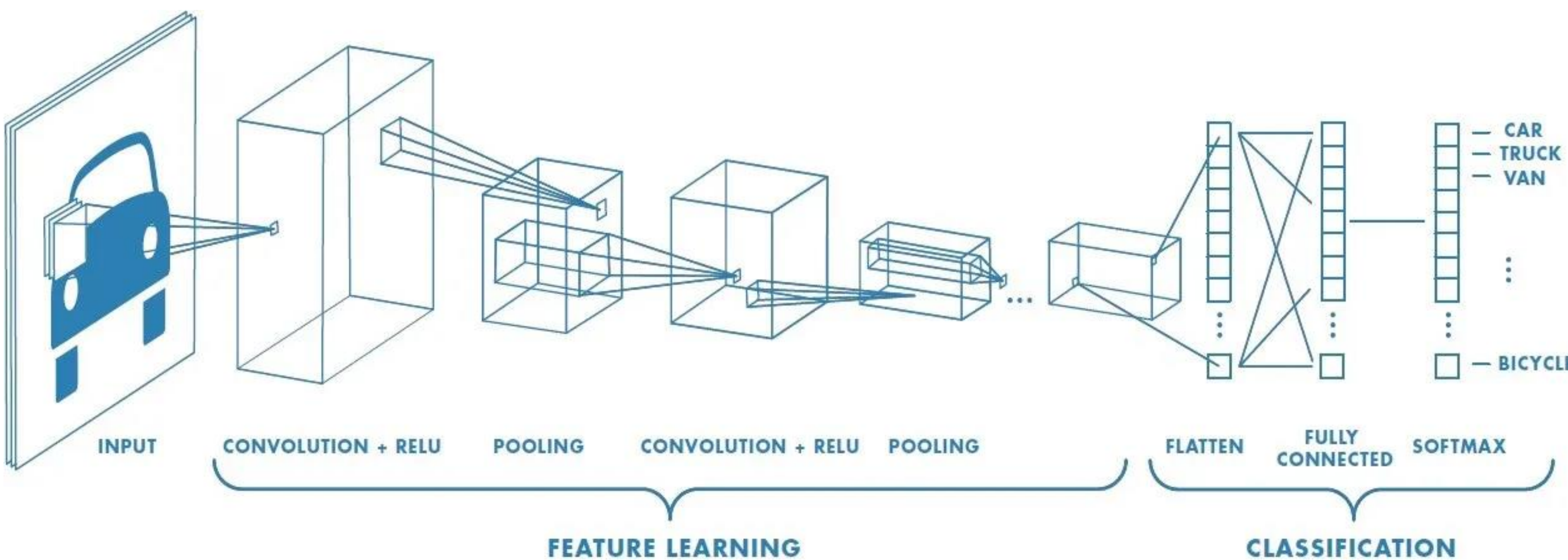
- Reduces the amount of data by dropping finer redundant details
- keeps most relevant structural properties in the feature maps
- Reduces computation load and training time



# Convolution Neural Network

Convolution layers

- Progressively identifies low-level geometric feature (like lines, edges, shapes, textures, etc.)
- The layers at the depth identify semantic features (like object classification, face recognition, etc.)



# Data

| Nr. of slides per category | Training | Validation | Test | Total |
|---|---|---|---|---|
| Cancer | 48 (62.94 ± 29.23) | 31 (62.32 ± 27.88) | 45 (64.90 ± 25.22) | 124 (64.02 ± 26.78) |
| 2 + 3 | 0 | 1 | 0 | 1 |
| 3 + 2 | 0 | 2 | 0 | 2 |
| 3 + 3 | 11 | 9 | 14 | 34 |
| 3 + 4 | 23 | 9 | 12 | 44 |
| 3 + 5 | 0 | 0 | 1 | 1 |
| 4 + 3 | 7 | 6 | 10 | 23 |
| 4 + 4 | 5 | 1 | 3 | 9 |
| 4 + 5 | 2 | 2 | 3 | 7 |
| 5 + 3 | 0 | 1 | 0 | 1 |
| 5 + 4 | 0 | 0 | 2 | 2 |
| Normal | 52 | 19 | 30 | 101 |
| Total | 100 | 50 | 75 | 225 |

**Table 1. Data details for the whole slide biopsy specimens used for the prostate cancer experiments.** The first column indicates the categories and the first row indicates the different data sets. For the cancer category, slide distribution is also indicated according to Gleason Score. The numbers between brackets for the 'Cancer'-row indicate the average volume percentage of cancer within the slides and the corresponding standard deviation.

# Data

| Nr. of slides per category | Training | Validation | Test | Consecutive | Total |
|---|---|---|---|---|---|
| *At least one macro-metastasis* | 18 | 5 | 7 | 16 | **46** |
| *No macro-metastasis, at least one micro-metastasis* | 29 | 8 | 8 | 4 | **49** |
| *No macro- or micro-metastases, at least one instance of ITC* | 1 | 0 | 1 | 22 | **24** |
| *No macro- or micro- metastases and no instances of ITC* | 50 | 20 | 26 | 56 | **152** |
| *Total* | 98 | 33 | 42 | 98 | 271 |

**Table 2. Data details for the whole slide sentinel lymph node specimens used for the breast cancer metastasis experiments.** The first column indicates the categories and the first row indicates the different data sets. (ITC = isolated tumor cells).

# Digitization and Annotation

Prostate Cancer

- Olympus VS120-S5 slide scanning system
- 40× objective (resultant pixel resolution of 0.16 microns)

Sentinel lymph node

- 3DHistech Pannoramic 250 Flash II slide scanner
- 20× objective (resultant pixel resolution of 0.24 microns).

For both cases, annotations made by free-hand drawing tool and checked by experienced pathologists.

# Pre-processing Steps

**Annotations were used to generate**:

1. Binary mask images (inside annotation label 1, outside label 0),
2. Binary tissue mask (To separate background from tissue)

$$OD_c = \log_{10}\frac{I}{I_{max}} \qquad (1)$$

Here $OD_c$ is the optical density of the channel c (Red, Green or Blue), $I$ is the intensity of the channel and $I_{max}$ is the maximum intensity, which is 255 due to 8-bit quantization. By thresholding the optical densities at 0.2, all background could be removed resulting in a binary mask where tissue is labeled 1 and background is labeled 0.

# CNN Training and application

- For training, open-source deep learning libraries were used,
- Tried patch sizes in pixels were : 64*64, 128*128 and 256*256 ( continued with 128*128)
- Small patches with enough information allows discrimination (with and without cancer).
- Too large patches makes discrimination difficult

# Results

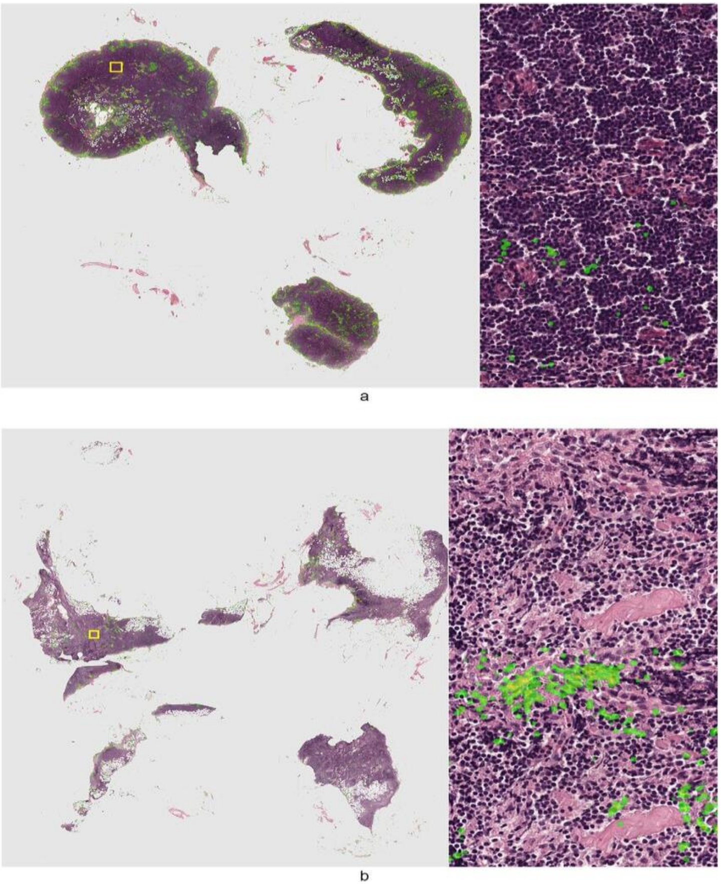



**Figure**: Representative examples of normal lymph nodes from the consecutive set. Metastases likelihood maps are overlaid on the original H&E image. Transparent/green means a low likelihood, whereas red indicates
a high likelihood of metastasis. On the right side of the whole slide images the areas indicated by the yellow squares are shown at full-resolution.

# Results

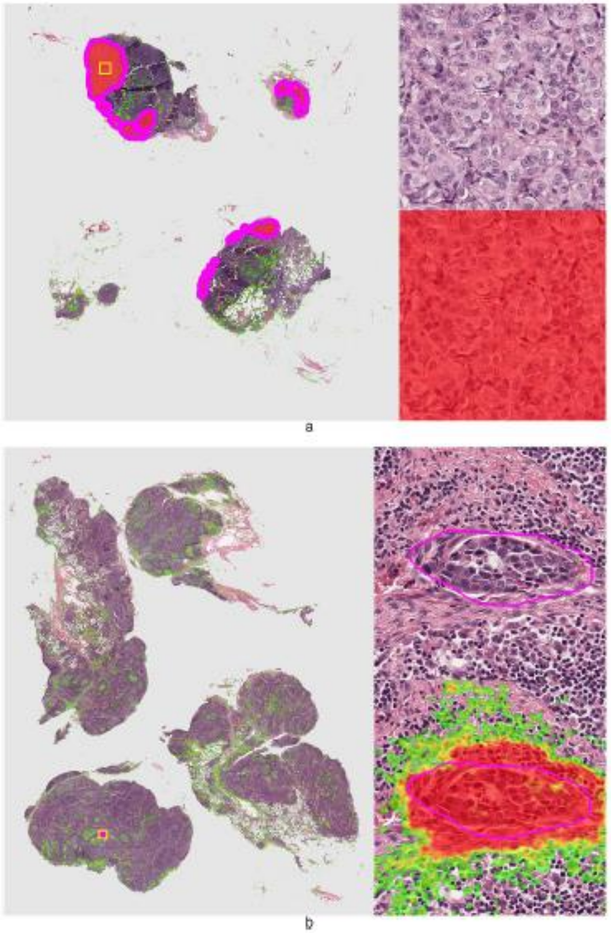



**Figure**: Representative examples of lymph nodes with macro-metastases (top image) and a single micro-metastasis (bottom image) from the test set. Metastases likelihood maps are overlaid on the original
H&E image. Transparent/green means a low likelihood, whereas red indicates a high likelihood of metastasis.
Magenta contours indicate the ground truth annotation. On the right side of the whole slide images the areas indicated by the yellow squares are shown at full-resolution.

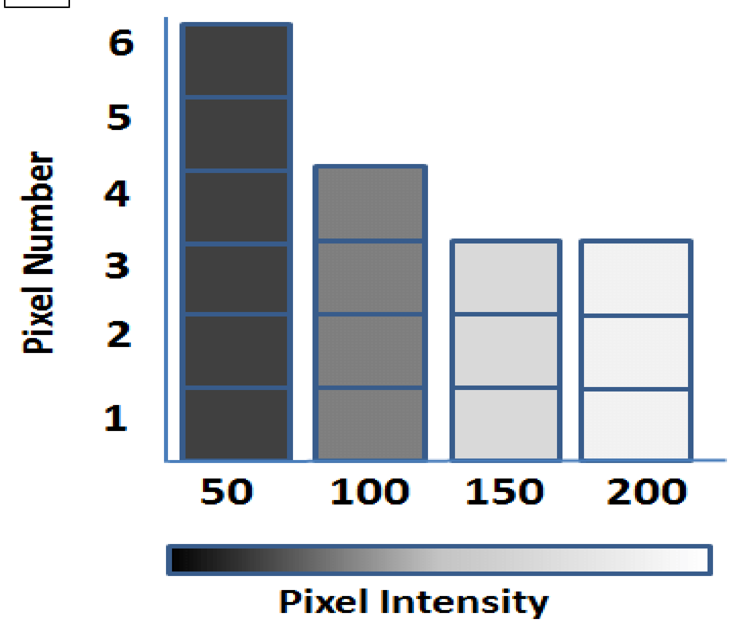
# Results

## Cumulative histograms of pixel intensity

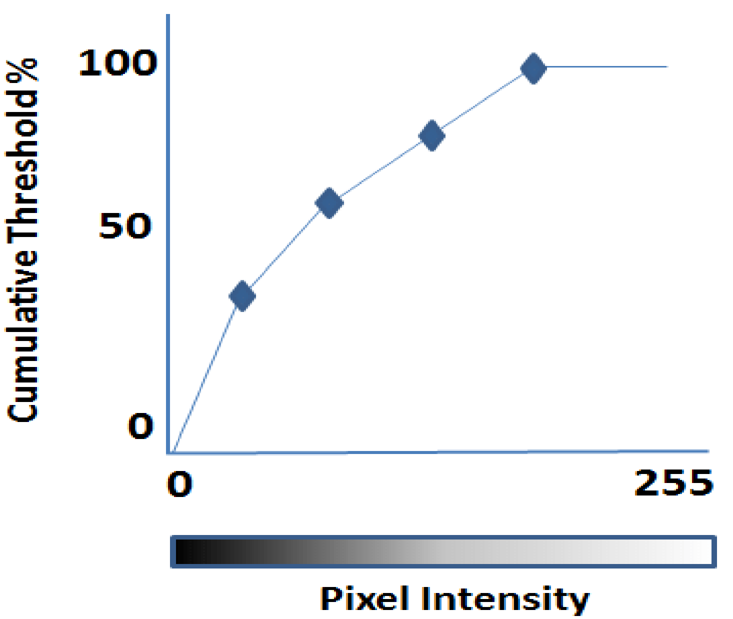
A

| Pixel Intensity | Number of Pixels | Cumulative Threshold % |
|---|---|---|
| 50 | 6 | (6/16) x 100 = 37.5 % |
| 100 | 4 | ((6 + 4)/16) x 100 = 62.5 % |
| 150 | 3 | ((10 + 3)/16 x 100 = 81.25 % |
| 200 | 3 | ((13 + 3)/16 x 100 = 100% |
| Total Pixels | 16 | |

B



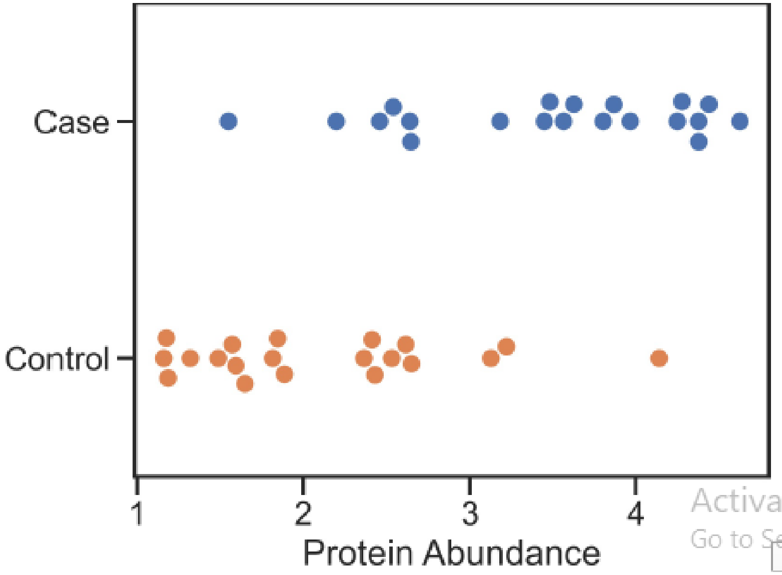

C



# Results

## Sensitivity and specificity are useful metrics

- Sensitivity is the proportion of positives (cases) that are correctly identified.
- Specificity is the proportion of negatives (controls) that are correctly identified.



# Results

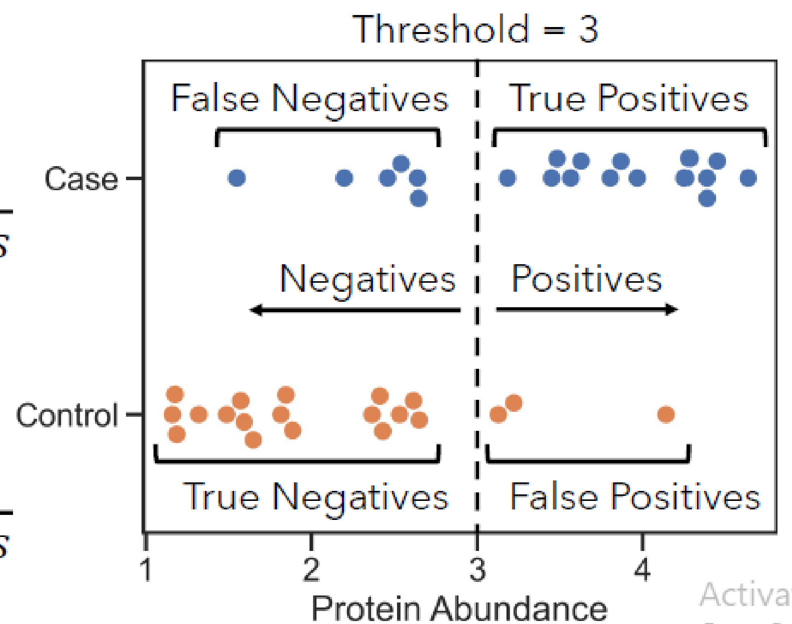
## How do we calculate sensitivity and specificity?

- Sensitivity $= 14/(14+6) = 0.70$

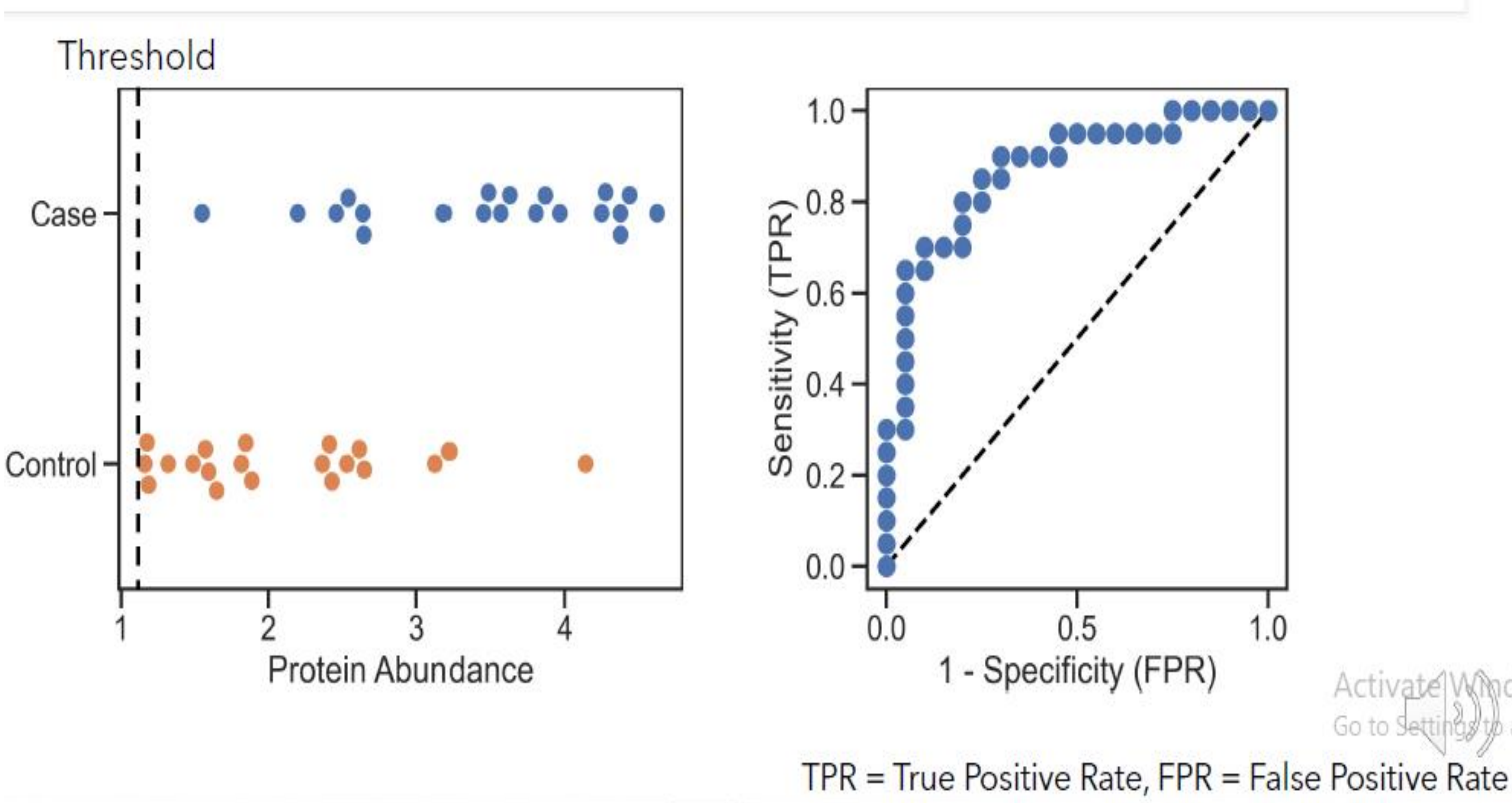
$$= \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

- Specificity $= 17/(17+3) = 0.85$

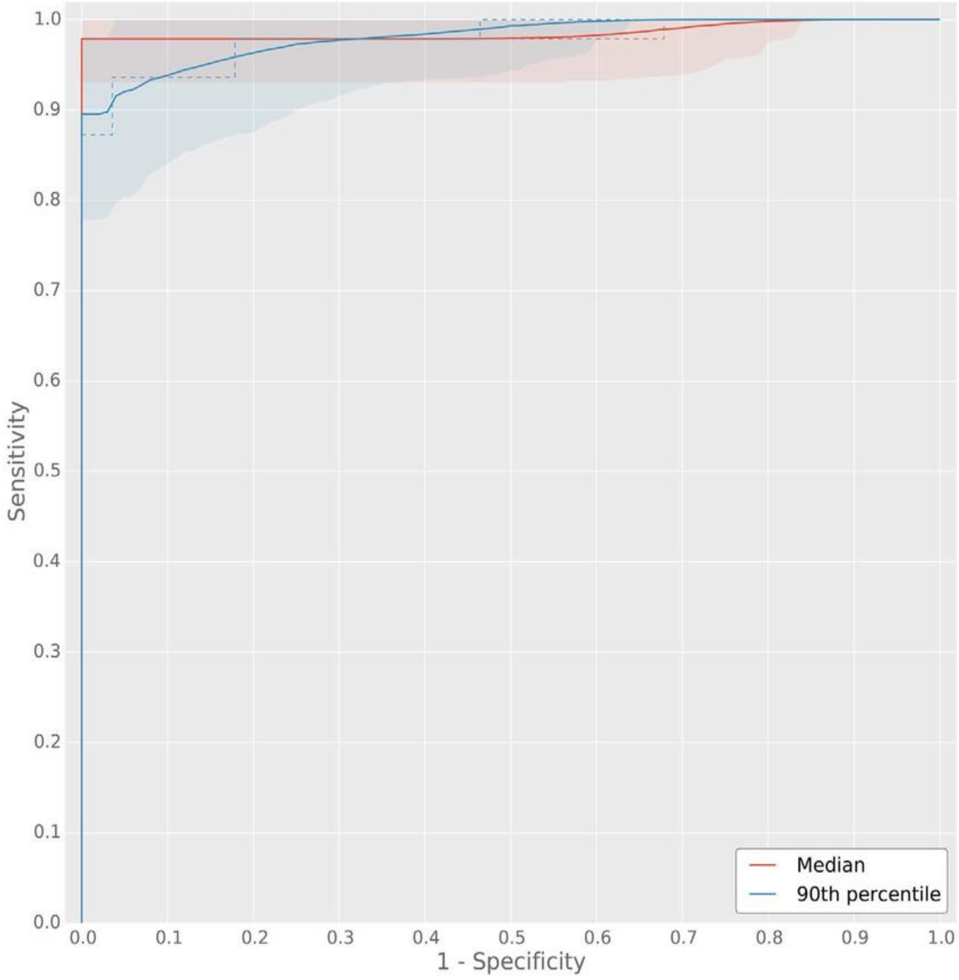
$$= \frac{True\ Negatives}{True\ Negatives + False\ Positives}$$



# Results


Threshold
Case
Control
1
2
3
4
Protein Abundance
Sensitivity (TPR)
1.0
0.8
0.6
0.4
0.2
0.0
0.0
0.5
1.0
1 - Specificity (FPR)
TPR = True Positive Rate, FPR = False Positive Rate

# Results




Receiver operating characteristic (ROC) curves for the cumulative histogram analysis in the whole-slide prostate biopsy experiment. Two cumulative histogram parameters were used to obtain ROC curves, the median and 90th-percentile of the cumulative histogram of the whole slide images. The median ROC curve has a higher area under the curve (AUC), however, the 90th-percentile ROC curve shows higher specificity at high sensitivity.

# Results

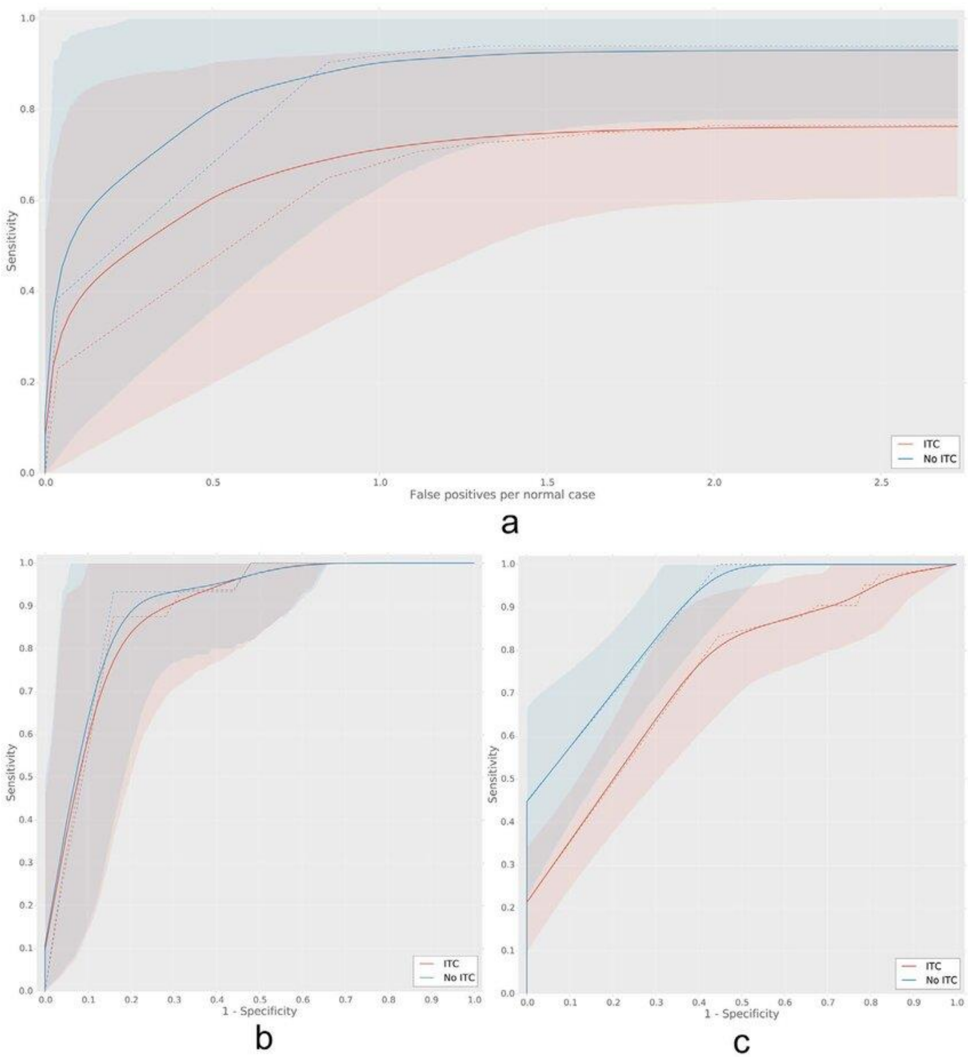



**Figure**. Bootstrapped FROC and ROC curves for the lymph node experiments. Subfigure (**a**) contains the FROC curve on the test set, (**b**) contains the ROC curve on the test set and (**c**) contains the ROC curve on the consecutive data. Curves for both including (red) and excluding isolated tumor cells (ITCS (blue) from the analysis are shown. Solid lines indicate the mean bootstrapped ROC curve, the shaded areas indicate the 95thpercentile confidence intervals and the dashed line indicates the raw ROC curve.

# Results

| FROC analysis | 1 FP | 2 FP |
|---|---|---|
| Sensitivity (incl. ITC) | 0.71 (0.39–0.93) | 0.74 (0.59–0.94) |
| Sensitivity (excl. ITC) | 0.90 (0.63–0.99) | 0.93 (0.78–1.0) |
| **ROC analysis** | **Area under the curve** | **Specificity at 99.9% sensitivity** |
| Test (incl. ITC) | 0.88 (0.77–0.97) | 0.39 (0.33–0.90) |
| Test (excl. ITC) | 0.90 (0.79–0.98) | 0.39 (0.32–0.94) |
| consecutive (incl. ITC) | 0.74 (0.65–0.82) | 0.02 (0.01–0.30) |
| consecutive (excl. ITC) | 0.88 (0.81–0.93) | 0.44 (0.43–0.69) |

**Table:** Free-response receiver operating characteristic (FROC) and receiver operating characteristic (ROC) analysis in the sentinel lymph node experiment. Mean bootstrap values are given for sensitivity (FROC analysis), area under the curve (ROC analysis) and specificity at 99.9% sensitivity (ROC analysis).
95th-percentile confidence intervals obtained through bootstrapping are shown between brackets. (FP = False positive detections per tumor-negative image).

# Advantages

- Substantial gains in excluding tumor-negative slides (For prostate cancer slides, up to 32% and For the sentinel lymph nodes, specificity was even higher at 44% )
- High sensitivity achieved in both cases (was 0.99 for the prostate cancer experiment (median analysis) and 0.88 for the sentinel lymph node experiment (consecutive set)
- For metastases, localization accuracy was high(90% at 1 false positive per normal image)

# Disadvantages

**For prostate cancer detection**

- Cutting and processing often cause deforms, tears and abnormal appearance

- False positive region inside the biopsy specimen

# Disadvantages

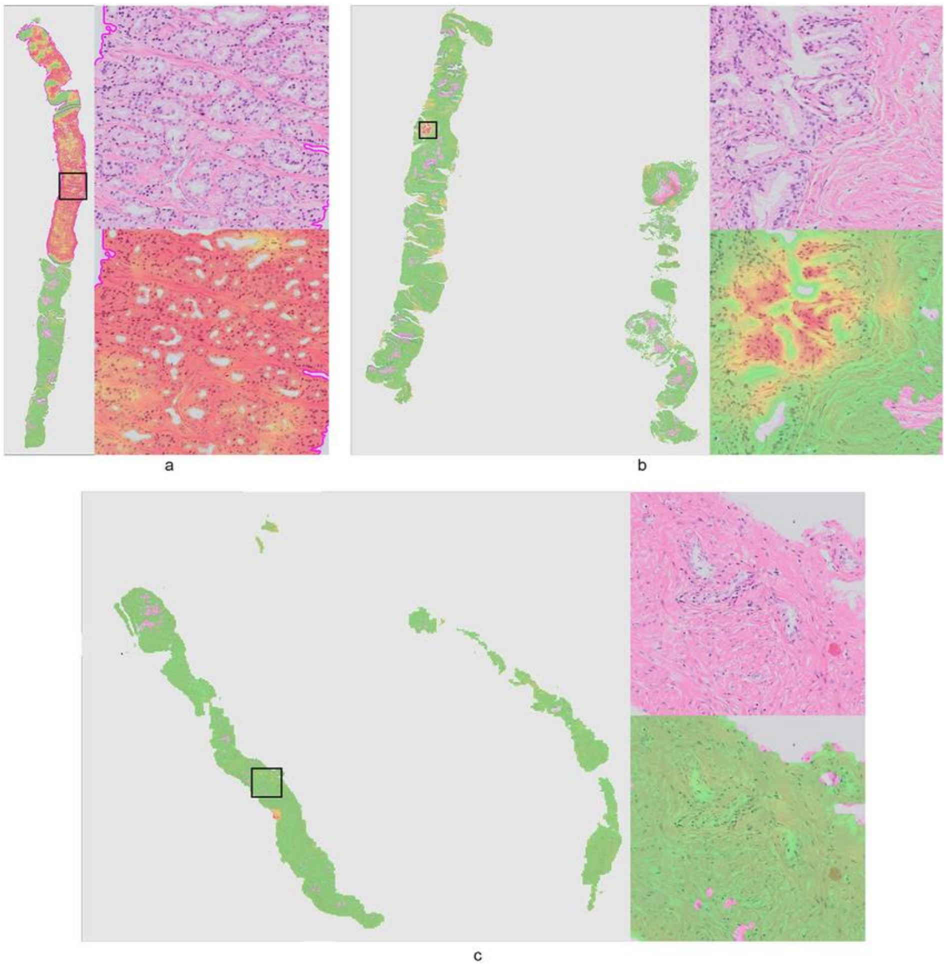



**Figure.** Three representative examples of a whole slide prostate biopsy specimen. Each example (a–c) shows the complete field of view with the cancer likelihood map as an overlay. Red indicates a high likelihood of cancer, whereas transparent/green indicates a low likelihood. Example (a) contains around 40% cancer (indicated by the magenta outline), examples (b,c) do not contain cancer. Close-up sub-images are shown for the areas indicated by black square. For example (b) we choose to highlight a small false positive area caused by tissue deformation at the edges of the biopsy.

# Disadvantages

**For sentinel lymph nodes**

- Identification of isolated tumor cells (ITC) lowers accuracy performance (0.74 area under the curve )
- Importance of ITC is debated
- CNNs for macro and micro metastases detection
- Immunohistochemistry for ITCs

# Conclusion

- First paper discussing applicability of ‘deep learning’ in WSI diagnosis of the two examples
- Potential future research topics
    - Quick extraction of relevant cases from huge clinical databases,
    - Automatic annotation of disease areas for fast quantification,
    - Application in immunohistochemistry (Efficacy of drugs, expression of genes)