Deep Learning Predicts Breast Cancer Metastasis with Sentinel Lymph Nodes

Andy Nguyen, M.D., M.S.

Professor of Pathology and Laboratory Medicine
UTHealth, McGovern Medical School

8th DIGITAL PATHOLOGY & AI CONGRESS, USA
June 2022

UTHealth The University of Texas Health Science Center at Houston | McGovern Medical School

Outline of talk

- Deep learning has been shown to be useful to identification of breast cancer metastases from analyzing whole sections of slide images of sentinel lymph nodes.

- Our study focuses on breast cancer screening using Deep Learning with only a small set of image patches from any sentinel lymph node (positive or negative) to detect changes in tumor environment and not the tumor itself.

- Our approach is unique since it provides a very rapid screen rather than an exhaustive search for tumor in all fields of all lymph nodes.

Outline of talk (cont'd)

- This study involves breast pathologists in our department and uses our in-house breast cancer cases with scanners for whole slide imaging.

- We obtain excellent predictive results for cancer metastasis in this study, which provide a proof of concept for incorporating automated breast cancer metastatic screen into the digital pathology workflow to potentially augment the pathologists' productivity.

I have no relevant financial relationships with commercial interests to disclose

Deep learning and pathology image

- Machine learning: software algorithms that can learn from and make predictions on data - i.e. “gives software the ability to learn without being explicitly programmed”

- Numerous machine learning methods:
  Decision tree, Cluster analysis, Support vector machine, Random forest, Bayesian, Regression analysis, Neural network….

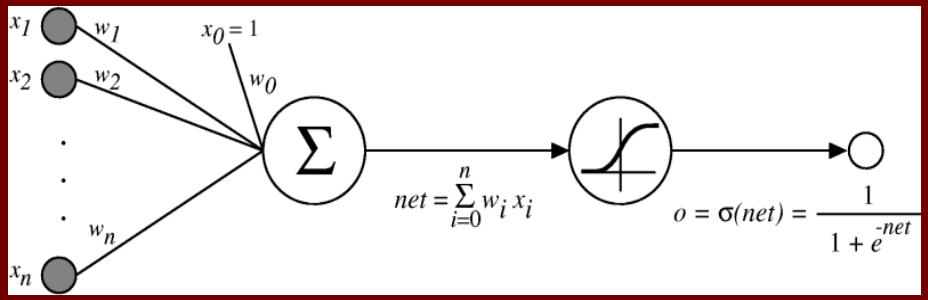
- Deep learning is the most recent and most disruptive method of machine learning; based on Neural network

- Big companies are analyzing large volumes of data for business analysis and decisions, using Deep Learning technology (Google’s search engine, Google Photo, automobile companies: self-driving cars, IBM’s Watson).

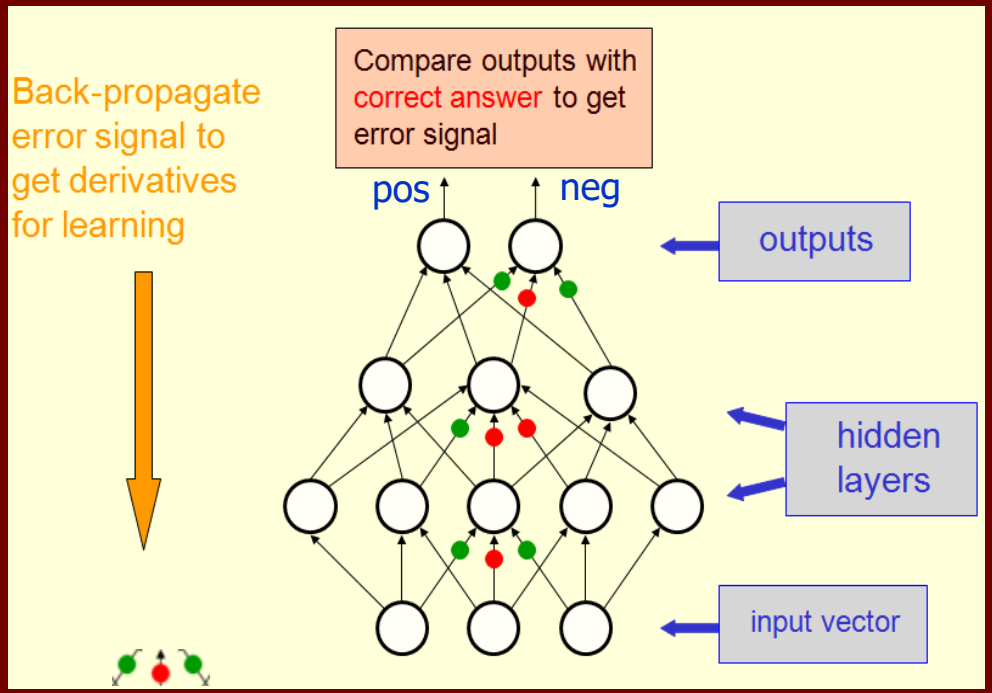
- The application of deep learning to digital pathology image has a promising start; it could impact personalized diagnostics, and treatment.

Deep Learning: based on Neural Network

- Neural network (inspired by biological neural networks): artificial nodes ("neurons“) are connected together to form a network for prediction/classification



→ Fire/not Fire



Early Generation of Neural Networks with Supervised Training

Disadvantages

(1) Parameters often do not converge; not giving solution

(2) Model not scaling well

Deep Learning (3rd Gen Neural Network)

- Major breakthroughs in Deep Learning started in 2006 and helped it to outperform all other machine learning models

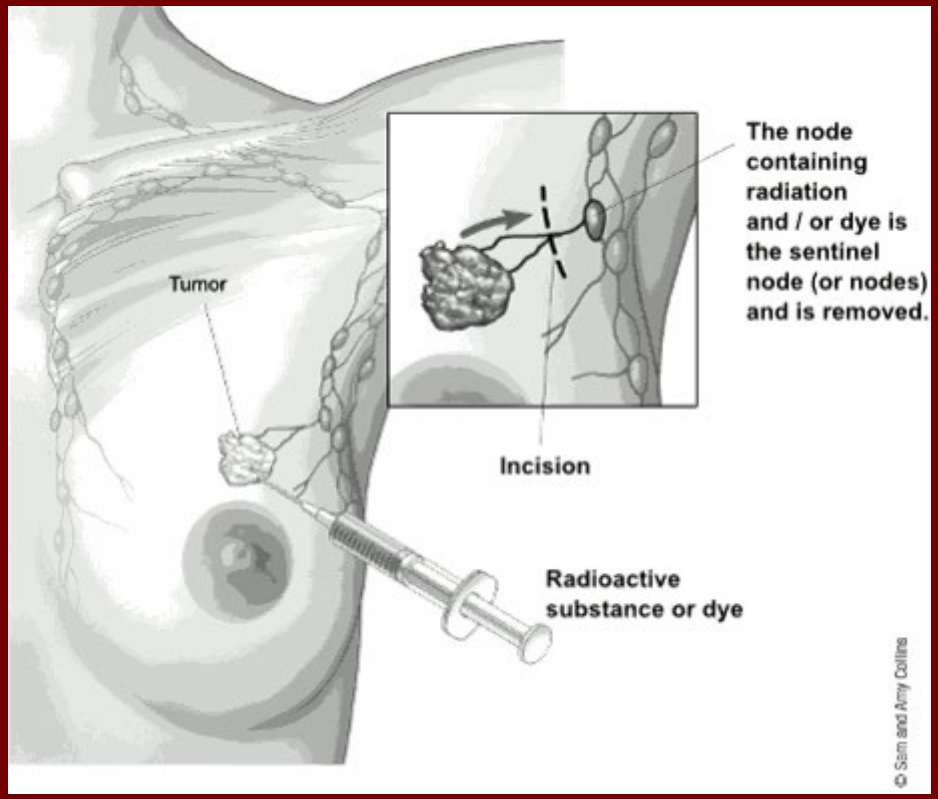
- Deep Learning algorithms:
  (1) Unsupervised learning ->allows a network to be fed with raw data (no known outcomes) and to automatically discover the representations needed for detection or classification
  (2) Extract high-level & complex data representations through multiple layers; avoid problems of last-gen networks (previous slide)

- Supporting hardware: multiple graphics processing units (GPU)

Sentinel lymph node biopsy

- In a sentinel lymph node biopsy, the surgeon finds and removes the first lymph node(s) to which a tumor is likely to spread (called the sentinel nodes).

- To do this, the surgeon injects a radioactive substance and/or a blue dye into the tumor, the area around it, or the area around the nipple. Lymphatic vessels will carry these substances along the same path that the cancer would likely take. The first lymph node(s) the dye or radioactive substance travels to will be the sentinel node(s).



Evaluation of breast sentinel lymph nodes

- In patients with breast cancer, the evaluation of breast sentinel lymph nodes is an important component of treatment. Patients with a sentinel lymph node positive for metastatic cancer will receive a more aggressive clinical management, including axillary lymph node dissection.

- The manual microscopic review of sentinel lymph nodes is time-consuming and laborious, particularly in cases in which the lymph nodes are negative for cancer or contain only small foci of metastatic cancer.

Diagnosis of Breast Cancer Metastasis

- Types of sentinel lymph nodes in terms of positive or negative for metastasis:

(a) Positive for metastasis:

- Macro metastasis: a tumor region of at least 2.0 mm, or
- Micro metastasis: a tumor region of at least 200 cells or size between 0.2 mm to 2.0 mm

(b) Negative:

- Negative, or
- Isolated tumor cells (ITC): a tumor region of up to 200 cells and/or smaller than 0.2 mm

Diagnosis of Breast Cancer Metastasis

- Potential morphologic features for metastasis:

  (a) Pleomorphic nuclei

  and

  (b) Tumor microenvironment:

  - Lymphocytic infiltrates in stroma (tumor-infiltrating lymphocytes, TILs)
  - Sinus
  - Follicular hyperplasia

Diagnosis of Breast Cancer Metastasis

Objectives of this study:

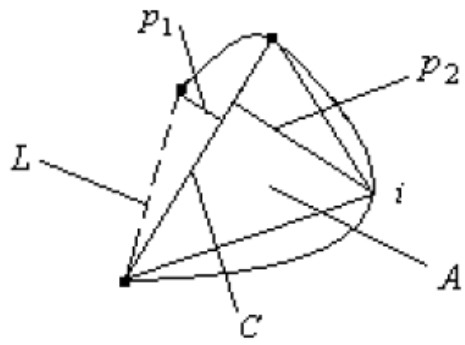
- How deep learning can be combined with microscopic examination by pathologists to achieve better accuracy in using reactive morphology to predict positive metastasis

- Feasibility of looking at negative slide or positive slide (uninvolved area of lymph node with metastasis) to predict metastasis

- Selection of interested areas for analysis: (1) interfollicular lymphocytes, (2) follicles, (3) sinus; and see if specific area is critical for prediction of metastasis.

Automated Diagnosis for Morphology

- Due to the large number of sentinel lymph nodes to screen for breast cancer metastasis, histopathologic screen often presents a challenge to the pathologists.
- An automated diagnosis for digital images would be helpful to assist the pathologists in daily work.
- Previous attempts to digitally classify histologic images were based on specific criteria (such as nuclear shape, nuclear size, texture, etc.). They were not very successful.
- Attention has turned to machine learning. In recent years, 'deep learning' techniques, especially 3rd generation neural network called convolutional neural network (CNN or ConvNet), has quickly become the state of the art in computer vision.



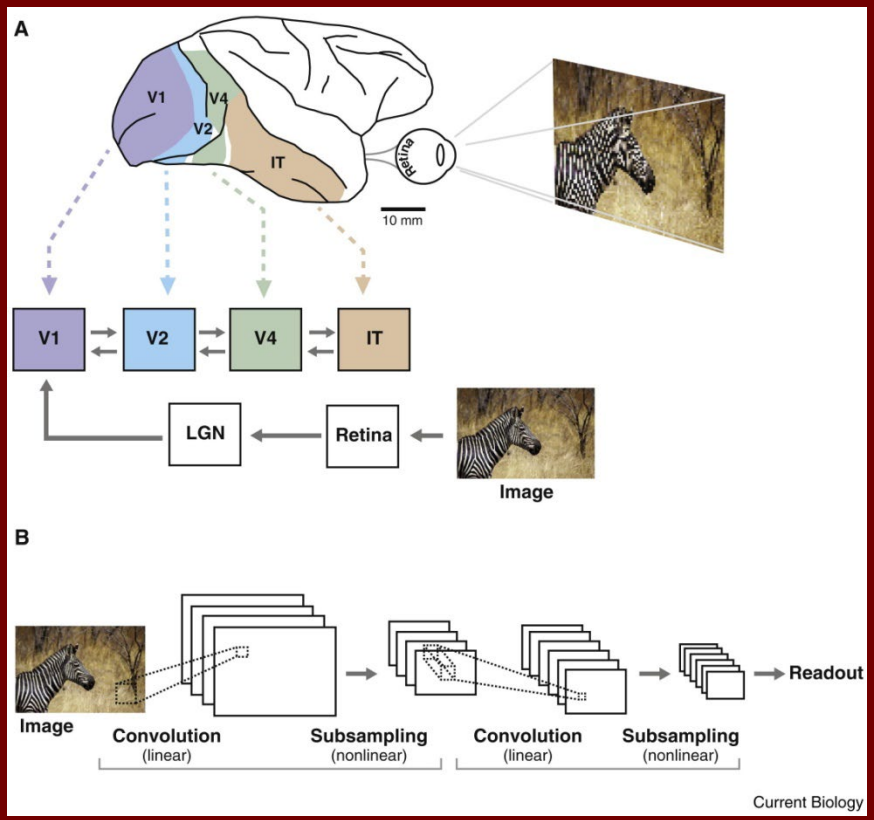
Shape and measures used to compute features.

1) Circularity $cir = \frac{4pA}{P^2}$
2) Aspect Ratio $ar = \frac{p_1+p_2}{C}$
3) Discontinuity Angle Irregularity $dar = \sqrt{\frac{(\sum|\theta_i-\theta_{i+1}|}{2\pi(n-2)}}$
   A normalized measure of the average absolute difference between the discontinuity angles of polygon segments made with its adjoining segments.
4) Length Irregularity $lir = \frac{\sum|L_i-L_{i+1}|}{K}$, where $K = 2P$ for $n > 3$ and $K = P$ for $n = 3$.
   A normalized measure of the average absolute difference between the length of a polygon segment and that of its preceding segment.
5) Complexity $com = 10^{\frac{3}{n}}$. A measure of the number of segments in a boundary group weighted such that small changes in the number of segments have more effect in low complexity shapes than in high complexity shapes.

CNN: Inspiration from the primate visual cortex

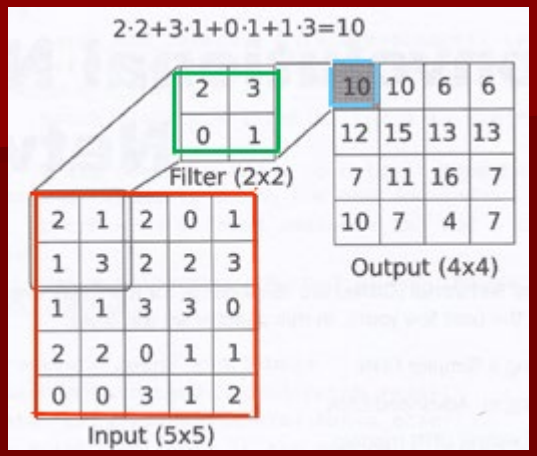
- A. The ventral visual pathway is organized as a hierarchical series of 4 interconnected visual areas called Brodmann areas. Neurons in early areas, such as area V1, respond to comparatively simple visual features of the retinal image, while later areas, such as area V4, respond to increasingly complex visual features

- B. The specialization of receptor cells are incorporated into the design of CNN as pairs of convolution operator followed by a pooling layer

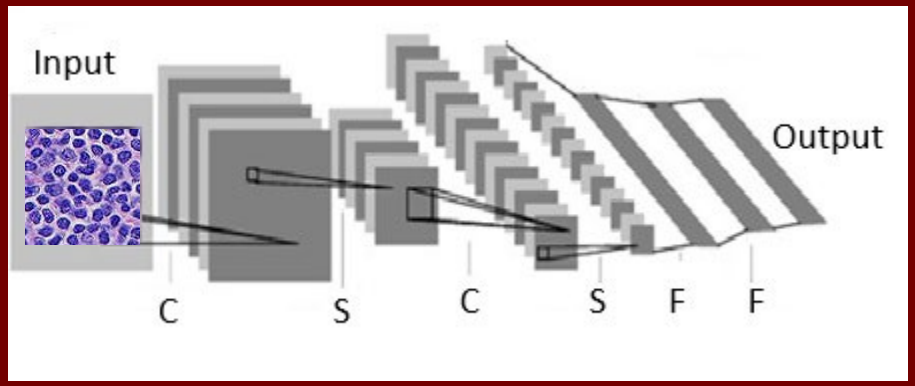


A rough correspondence between the areas associated with the primary visual cortex and the layers in a convolutional network.

Definition of Convolution



- Convolution: an operation in image processing using filters, to modify or detect certain characteristics of an image (Smooth, Sharpen, Intensify, Enhance). In CNN, it is used to extract features of images
- Mathematically, a convolution is done by multiplying the pixels' value in image patch by a filter (kernel) matrix [dot product]



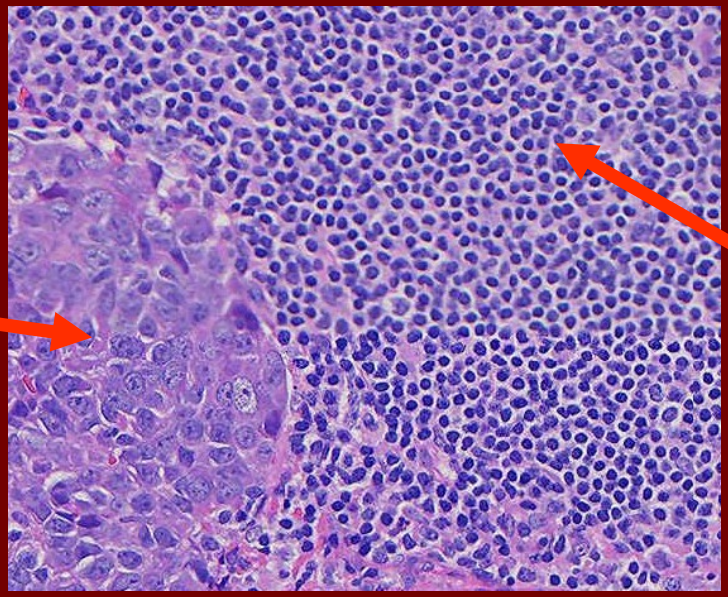
By moving the filter across input image, one obtains the final output as a modified filtered image

- The convolutional layers (C) perform 'feature extraction' consecutively from the image patch to higher level features.
- The max pooling layers (S) reduce image size by subsampling
- The last 'fully connected' layers (F): provide prediction

The Goal of our Study

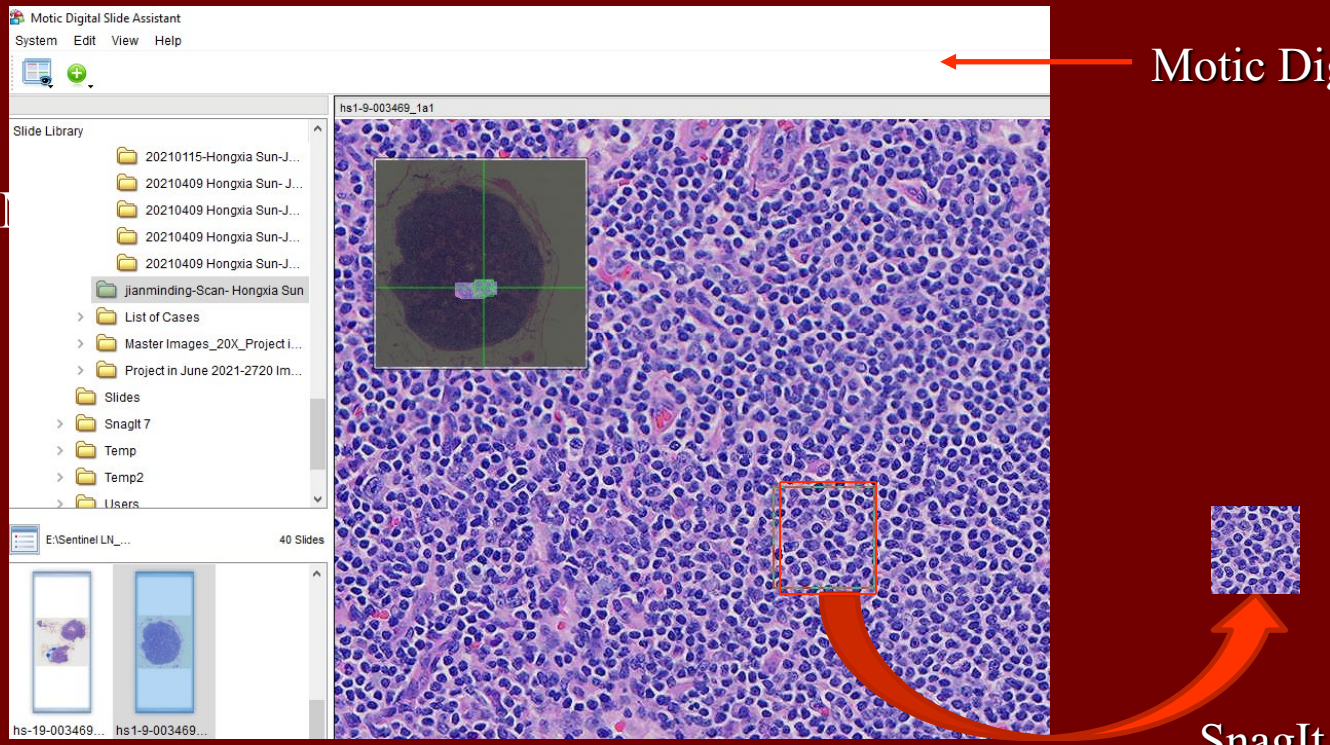
❑ Recent studies have shown promising results in using machine learning to detect breast cancer in whole slide imaging of sentinel lymph nodes (examples: Camelyon16, ICIAR 2018). However, they require extensive scanning and analysis of all the lymph node slides for each case.

❑ We explore how Deep Learning could be used for breast cancer screening with only a small set of image patches (5) from any sentinel lymph node. Our goal is to detect changes in tumor environment and not the tumor itself. Our approach is unique since it provides a very rapid screen rather than an exhaustive search for tumor in all fields of all lymph nodes



Digital Images of Sentinel Lymph Nodes

- Data source: Whole Slide Image (WSI) of sentinel lymph nodes using Motic scanners in the Pathology Department of University of Texas-Houston Medical School.
- Motic Digital Slide Assistant software: to view WSIs
- SnagIt (TechSmith Corp, Okemos, Michigan, USA) to capture and automatically save 100x100 image patch in files with format xxx.jpg



Motic Digital Assistant

SnagIt

Whole Slide Imaging (WSI) of Sentinel Lymph Nodes

- WSI's were obtained with Motic Easy Scan (Motic Instrument, Ricmond, BC, Canada) whole slide imaging systems. Our study includes 34 cases with approximate equal distribution in 4 diagnostic categories:
  1. Macro metastasis,
  2. Micrometastasis,
  3. Isolated tumor cells,
  4. Negative

A positive WSI and a negative WSI were selected for each case to obtain a total of 68 WSIs

- 40 image patches (100x100) were obtained from each WSI -> 68 x 40= 2720 image patches

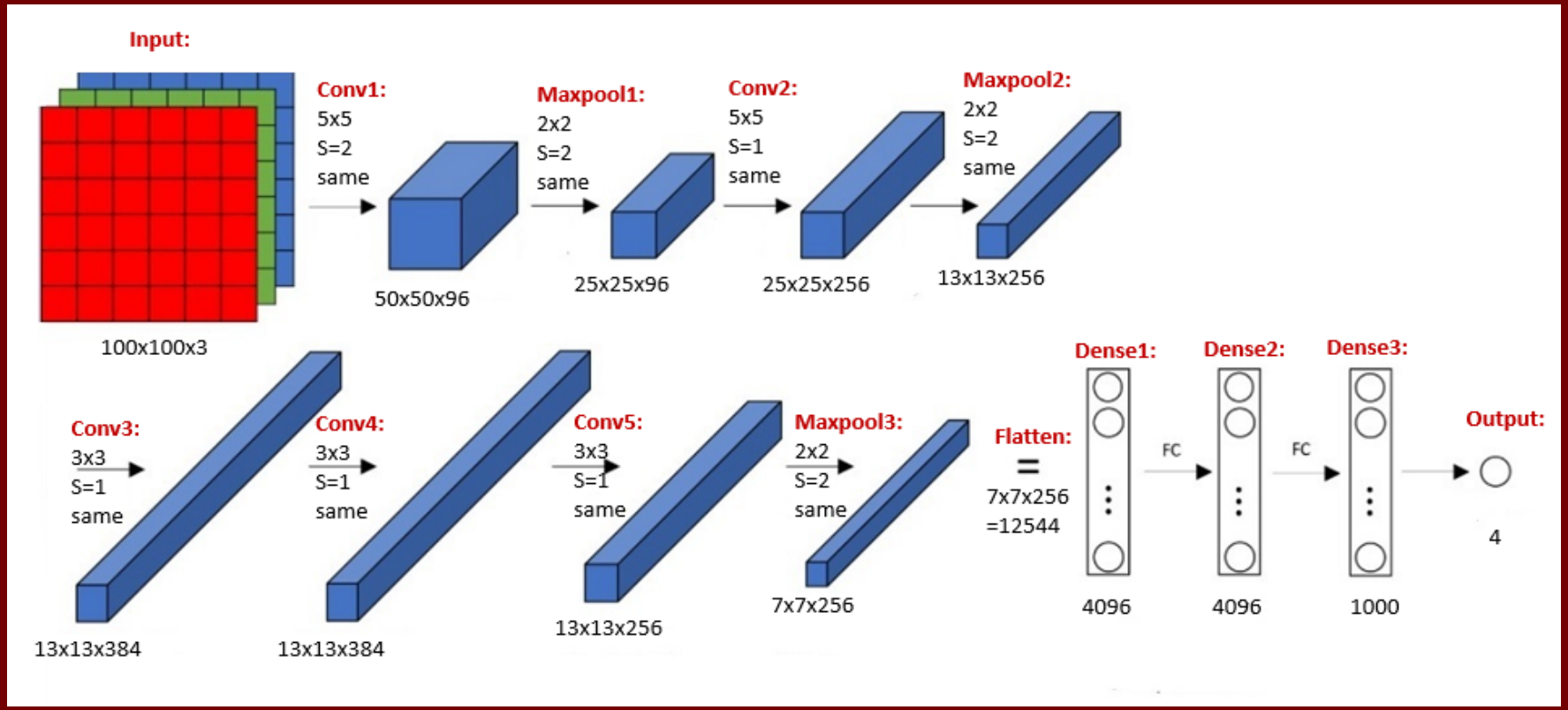
Our Programming Platform

- We design a CNN model in Python language, commonly used in deep learning (together with TensorFlow and Keras libraries);
  Windows 10 Professional 64bit

- Keras allows for parallel computing using graphic processing unit (GPU) with Compute Unified Device Architecture (CUDA)

- Hardware:
  9th Gen Intel® Core™ i7 9700 (8-Core, 12MB Cache, 4.7GHz)
  RAM: 32GB DDR4 at 2666MHz
  GPU: NVIDIA® GeForce RTX™ 2070, 8GB GDDR6
  (2304 CUDA cores)

Schematics of our Deep Learning Model (14 Layers)



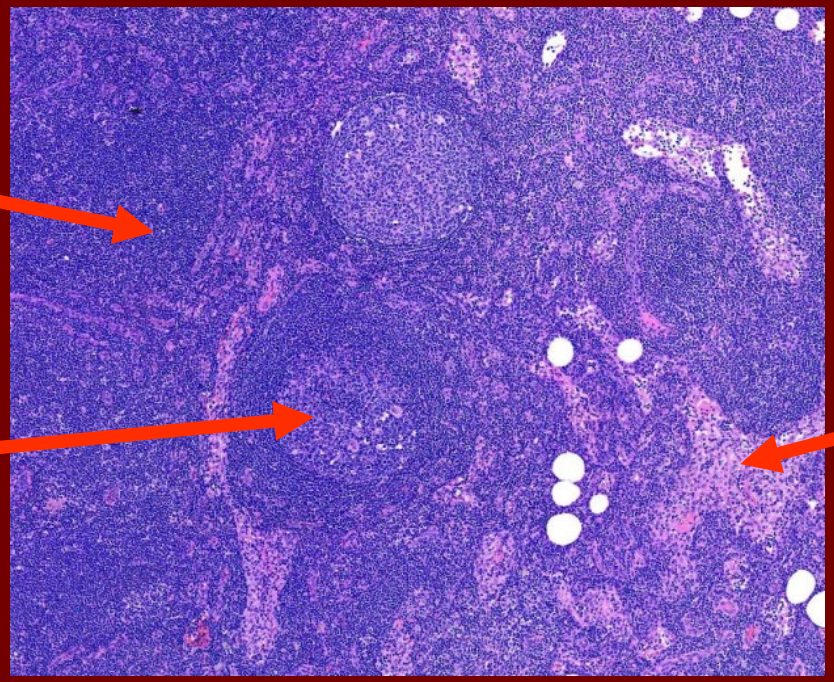
Schematics of our Deep Learning Model (14 Layers)

```
Model: "sequential_5"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_13 (Conv2D)           (None, 50, 50, 96)        7296
_________________________________________________________________
max_pooling2d_7 (MaxPooling2 (None, 25, 25, 96)        0
_________________________________________________________________
conv2d_14 (Conv2D)           (None, 25, 25, 256)       614656
_________________________________________________________________
max_pooling2d_8 (MaxPooling2 (None, 13, 13, 256)       0
_________________________________________________________________
conv2d_15 (Conv2D)           (None, 13, 13, 384)       885120
_________________________________________________________________
conv2d_16 (Conv2D)           (None, 13, 13, 384)       1327488
_________________________________________________________________
conv2d_17 (Conv2D)           (None, 13, 13, 256)       884992
_________________________________________________________________
max_pooling2d_9 (MaxPooling2 (None, 7, 7, 256)         0
_________________________________________________________________
flatten_3 (Flatten)          (None, 12544)             0
_________________________________________________________________
dense_9 (Dense)              (None, 4096)              51384320
_________________________________________________________________
dense_10 (Dense)             (None, 4096)              16781312
_________________________________________________________________
dense_11 (Dense)             (None, 1000)              4097000
_________________________________________________________________
dense_12 (Dense)             (None, 4)                 4004
=================================================================
Total params: 75,986,188
Trainable params: 75,986,188
Non-trainable params: 0
_________________________________________________________________
```

Preliminary Results

- Preliminary study to look at different areas of interests in WSIs to see which one would be of most predictive value (positive vs negative metastasis):
  (1) interfollicular lymphocytes,
  (2) follicles,
  (3) sinus

- Preliminary results indicated that areas containing interfollicular lymphocytes are of most predictive value. Subsequent our study has been focusing on this parameter alone.



Materials / Methods

- 2160 images (79%) were used for training the model,

  240 (9%) for validation, and

  320 (12%) for testing

Display of Analysis Results

```
Test loss: 2.1046812295913697
Test accuracy: 0.503125
Actual (rows) Predicted (columns)
```

col_0	0	1	2	3
row_0				
0	42	36	2	0
1	20	54	3	3
2	12	2	65	1
3	22	1	57	0

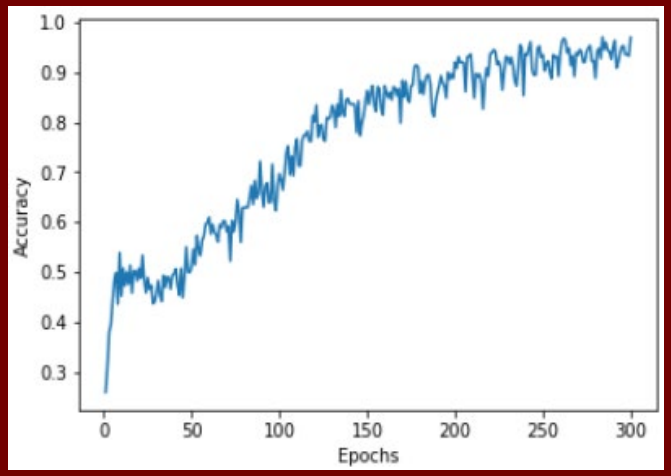
```
RESULT FOR EACH IMAGE PATCH:

Observed DX=  3 Predicted DX=  1
Observed DX=  3 Predicted DX=  0
Observed DX=  3 Predicted DX=  2
Observed DX=  3 Predicted DX=  0
Observed DX=  3 Predicted DX=  2

Observed DX=  2 Predicted DX=  2
Observed DX=  2 Predicted DX=  2
Observed DX=  2 Predicted DX=  0
Observed DX=  2 Predicted DX=  0
Observed DX=  2 Predicted DX=  2

Observed DX=  1 Predicted DX=  1
Observed DX=  1 Predicted DX=  1
Observed DX=  1 Predicted DX=  1
Observed DX=  1 Predicted DX=  0
Observed DX=  1 Predicted DX=  0
```

...

RESULTS: ACCURACY (User 1)

Observed DX \ Predicted DX	Neg (0)	ITC (1)	Micro Met (2)	Macro Met (3)
Neg (0)	42	36	2	0
ITC (1)	20	54	3	3
Micro Met (2)	12	2	65	1
Macro Met (3)	22	1	57	0

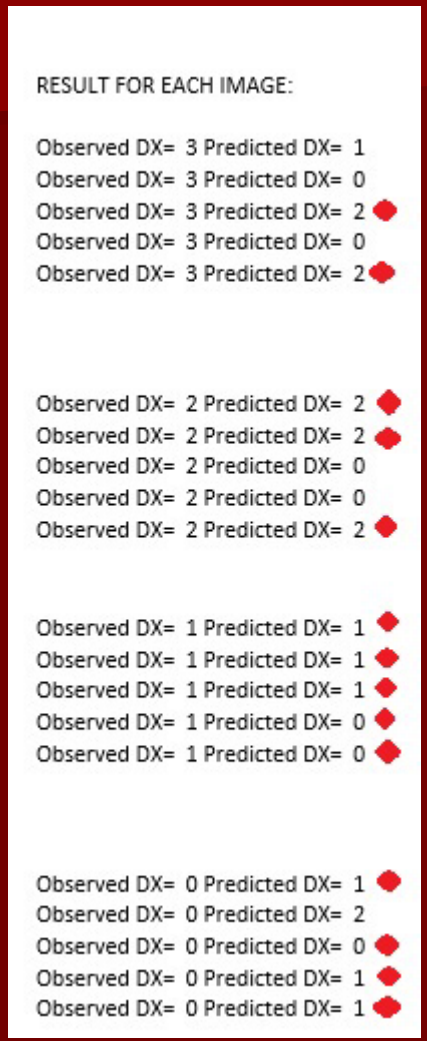
Step 1: Image-by-image:
Accuracy: 161/320= 50.31%



Observed DX \ Predicted DX	Neg (0)	ITC (1)	Micro Met (2)	Macro Met (3)
Neg (0)	42	36 ↓	2	0
ITC (1)	20 →	54	3	3
Micro Met (2)	12	2	65	1 ↓
Macro Met (3)	22	1	57 →	0

Step 2: Group Ranking (0~1, 2~3):
Significantly better accuracy:
275/320= 85.93%

Step 3: Evaluation for Case-by-case (majority voting ≥3/5), with group ranking (0~1, 2~3), User 1



Case-by-case (set of 5):

2/5->incorrect

3/5->correct

5/5->correct

4/5->correct

- For each test case, the predicted diagnosis is combined from the prediction for 5 images (at least 3 or more have to agree), a process known as "majority voting"

-> Optimal accuracy:
59 sets /64 sets =0.9218
92.18%

All measurements, User 1

Evaluation for Case-by-case (majority voting ≥3/5), with group ranking (0~1, 2~3)

- Accuracy=Correct results/All test Cases

- Sensitivity=TP/(TP+FN)
- Specificity=TN/(TN+FP)

- Positive Predictive Value=TP/(TP+FP)
- Negative Predictive Value=TN/(TN+FN)

Where:
TP: true positive=27
TN: true negative=32
FN: false negative=5
FP: false positive=0

- Accuracy= 59/64= 92.18%

- Sensitivity=27/(27+5)= 84.4%
- Specificity=32/(32+0)= 100%

- Positive Predictive Value=27/(27+0)= 100%
- Negative Predictive Value=32/(32+5)= 86.5%

All measurements, Means of 3 users in inter-user validation

Evaluation for Case-by-case (majority voting ≥3/5), with group ranking (0~1, 2~3)

- Accuracy=Correct results/All test Cases

- Sensitivity=TP/(TP+FN)
- Specificity=TN/(TN+FP)

- Positive Predictive Value=TP/(TP+FP)
- Negative Predictive Value=TN/(TN+FN)

Where:
TP: true positive
TN: true negative
FN: false negative
FP: false positive

User	Accuracy	Sensitivity	Specificity	PPV	NPV
User 1	92.19	76.88	95	93.89	80.42
User 2	87.5	78.75	89.38	88.11	80.79
User 3	93.75	78.12	91.88	90.58	80.77
Means	**91.15**	**77.92**	**92.09**	**90.86**	**80.66**

- Accuracy= 91.15%

- Sensitivity= 77.92%
- Specificity=32/(32+0)= 92.09%

- Positive Predictive Value= 90.86%
- Negative Predictive Value= 80.66%

SUMMARY

- Our study focuses on breast cancer screening using Deep Learning with only a small set of image patches from any sentinel lymph node (positive or negative) to detect changes in tumor environment and not the tumor itself.

- Our approach is unique since it provides a very rapid screen rather than an exhaustive search for tumor in all fields of all lymph nodes.

- We obtain excellent predictive results for cancer metastasis from this study, 91% accuracy, 78% sensitivity, 92% specificity for a set of 5 (100x100) image patches from each test case.

- Current limitations include:

(a) the model was only validated on one hardware platform (Motic scanner),
(b) representative images require preselection of lymphocyte-rich areas
(c) lack of explicit diagnostic criteria [inherent to DL]

- Our preliminary study provided a proof of concept for incorporating automated breast cancer screen using digital microscopic images into the pathology workflow to augment the pathologists' QA. This could have significant impact on health economics.

- Future studies will need to:

(a) include more hardware platforms and many more cases for training, and validation,
(b) automated segmentation of WSIs for lymphocyte-rich areas.

SUMMARY (Cont'd)

- Potential role for this model in clinical work as a QA tool:
  1.Positive cases by histology-> final diagnosis of metastasis
  2.Negative cases by histology-> screening with this model for metastasis
  A.Negative screen-> final diagnosis of negative metastasis
  B.Positive screen-> re-examination of histology to rule out false-negative (original) diagnosis

Acknowledgement

Breast cancer domain expertise:
Breast subspecialty pathologists, Department of Pathology,
University of Texas-Houston, Medical School:

- Sun H, MD
- Zhang S, MD
- Wang XI, MD
- Saluja K, MD

Whole slide scan technical expertise:

- Ding J, MD

Pathology resident

- Kareem Allam, MD

Medical student

- Kevin Chiu


Delicate Arch, Moab National Park, UT
May 2021