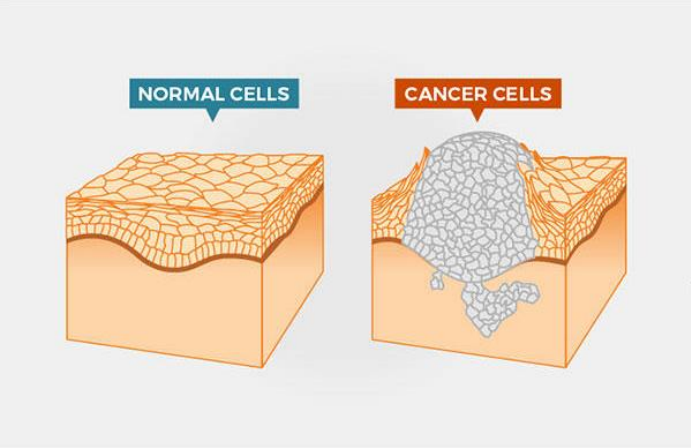
Analysis of Whole Slide Images of Skin Biopsies

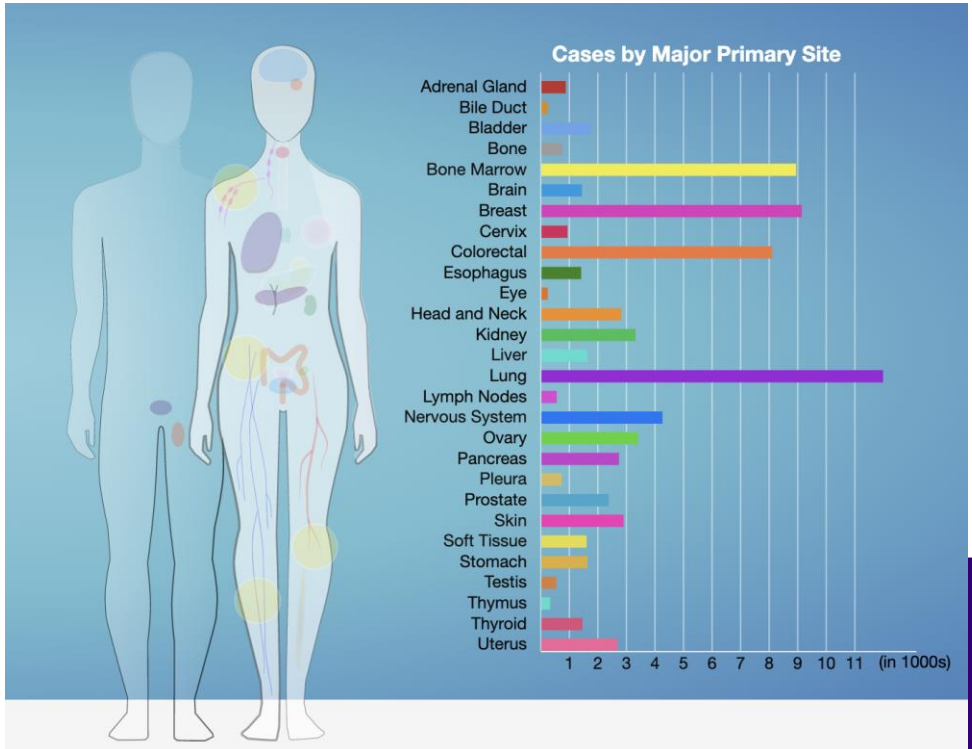
Linda G. Shapiro
Paul G. Allen School of Computer Science & Engineering
Department of Electrical & Computer Engineering
University of Washington

This work was sponsored by the National Cancer Institute under grants R01CA200690 andU01CA231782, by the Melanoma Research Alliance under grant 622600, and by the Department of Defense under grant W81XWH-20-1-0798.

What is Cancer?

Cancer is a disease in which some of the body's cells grow uncontrollably and spread to other parts of the body.

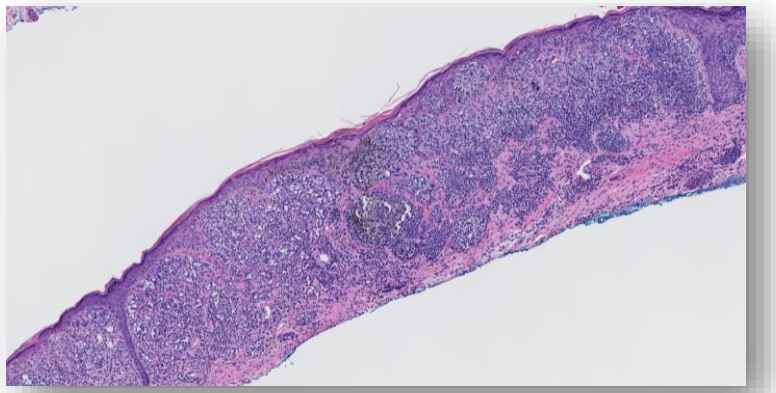




Figures are from National Institute of Cancer website.

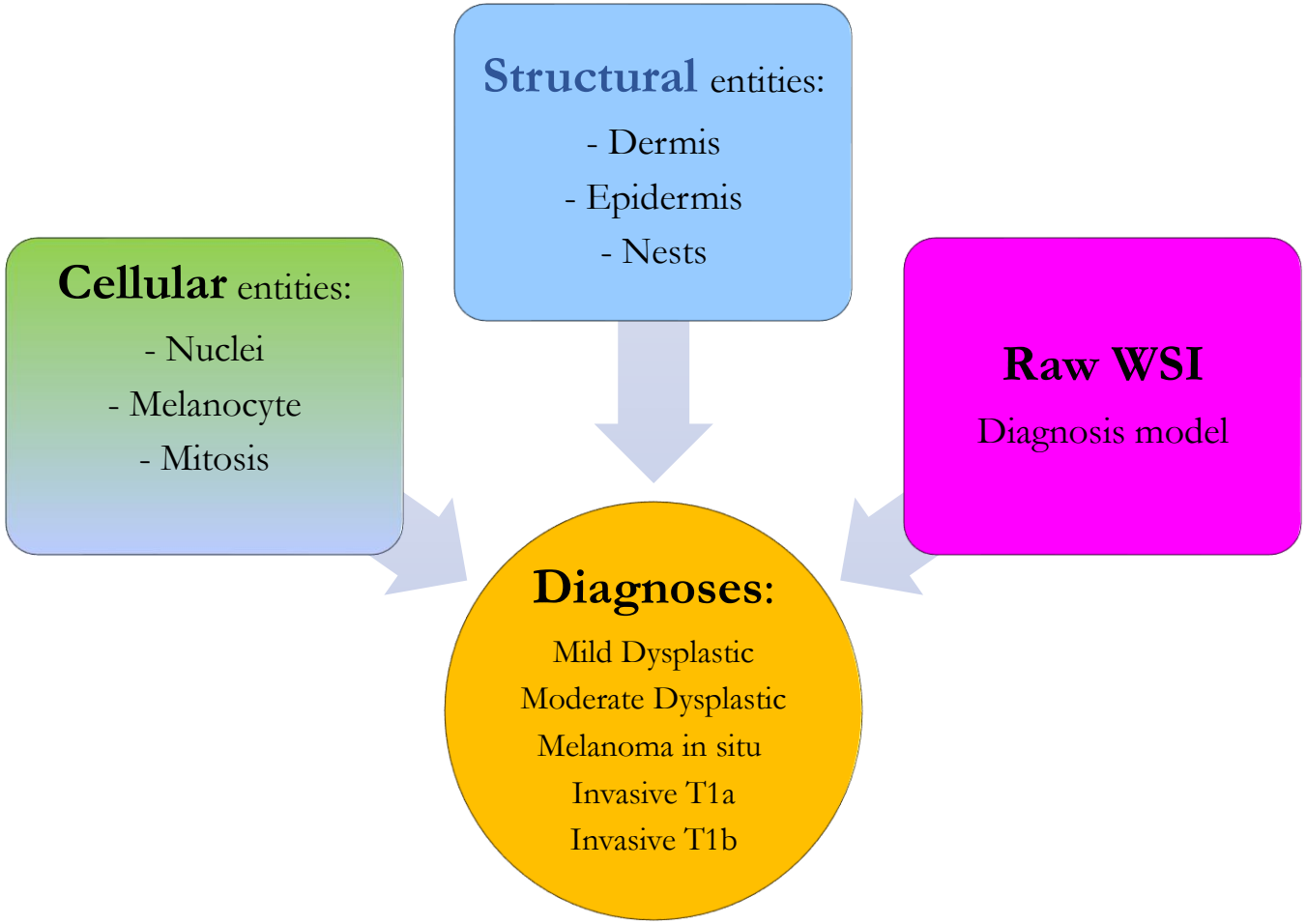
What is Melanoma?

- Melanoma is the most aggressive type of skin cancer.
- Pathologists look at a skin biopsy slide and determine if its overall structure is normal, abnormal, or malignant.
- Diagnostic errors are much more frequently than in other tissues and can lead to under- and over-diagnosis of cancer.
- Deep learning image analysis methods may improve and complement current diagnostic and prognostic capabilities.



An example of an Invasive Melanoma T1b in M-Path dataset.

Melanoma Diagnosis



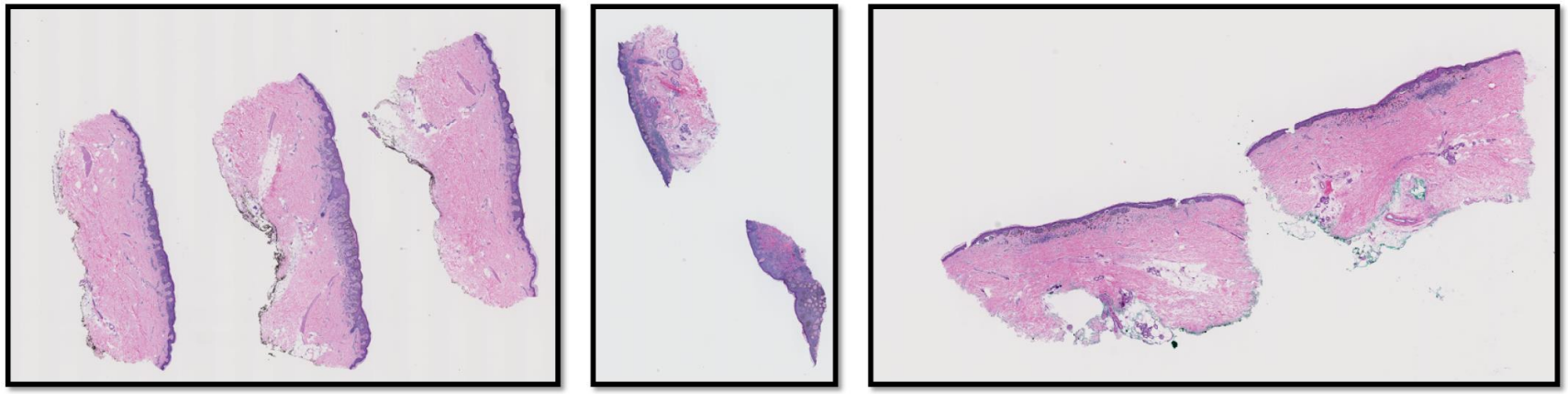
Our Work

- Cellular Level
  - Finding Mitotic Figures
  - Finding Melanocytes
- Structural Level
  - Segmenting into Epidermis/Dermis and Various Nest Structures
- Diagnosis Level
  - Diagnosing with a Multi-Resolution Transformer on Raw H&E (hematoxylin and eosin stained) images
  - Diagnosing with the Addition of Structures

An example of an Invasive Melanoma T1b in M-Path dataset.

Dataset

Our dataset comes from 240 H&E stained slides of skin biopsy images, acquired by the University of Washington School of Medicine in the MPATH study (R01 CA151306, PI Dr. Joann Elmore, now at UCLA).



Diagnostic Category	#Cases
Mildly Dysplastic Nevus	25
Moderately Dysplastic Nevus	36
Melanoma in Situ	60
Invasive Melanoma Stage T1a	58
Invasive Melanoma Stage $\geq$ T1b	61
Total	240

Machine learning techniques for mitoses classification

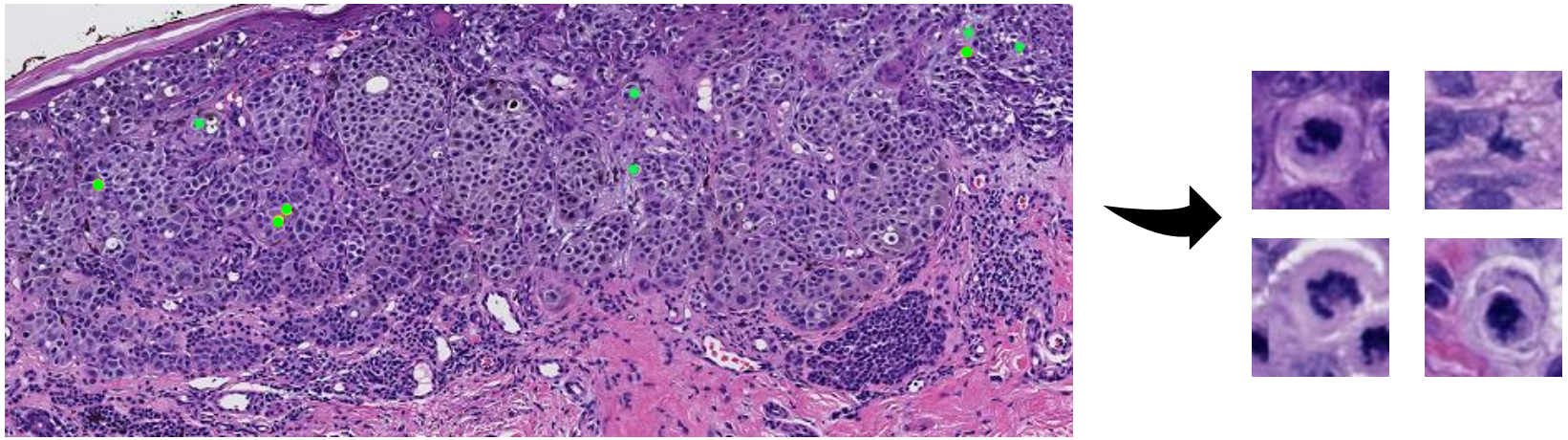
Shima Nofallah et al

Computerized Medical Imaging and Graphics
2021

Dataset for Mitosis Detection

Positive samples – class Mitosis

- About 600 mitoses marked by an expert pathologist (Dr. Stevan Knezevich).
- We cropped each mitosis in a 101*101 patch centered on the dot placing on the mitotic figure.

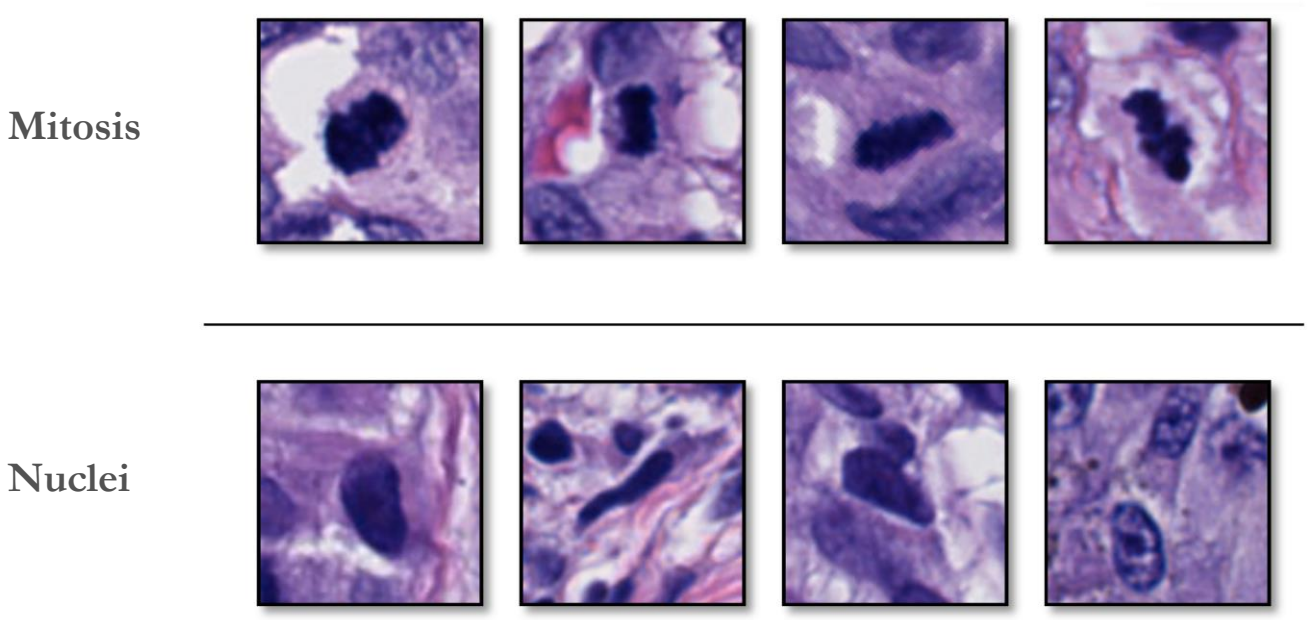


Example of expert pathologist markings of mitoses (Left) and sampled mitoses (Right)

Dataset

Negative samples – class NonMitosis

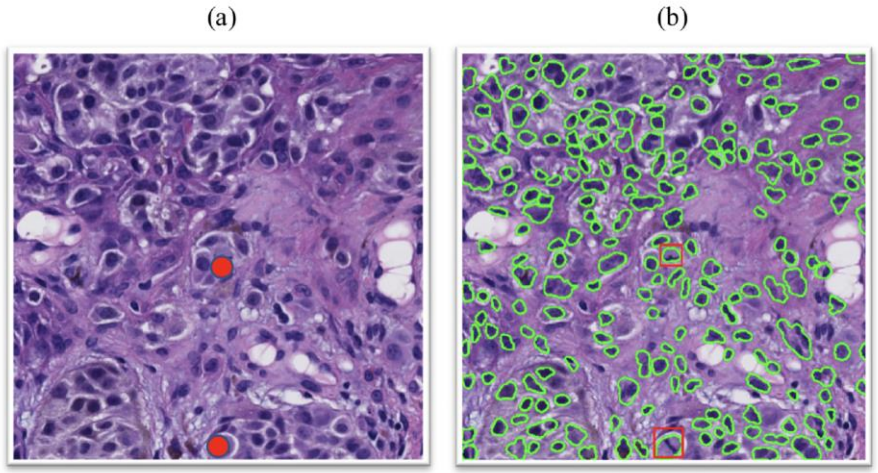
- Distinguishing mitoses from normal nuclei is a challenge.



Dataset

Negative samples – class NonMitosis

- We used a feature-based nuclei detector to find nuclei.
- We sampled them as negative cases for our dataset.



Examples of applying the nuclei segmentation on a crop of skin biopsy image (a) original crop (b) nuclei segmentation result. Two mitoses that are present in the original crop are marked with red dots for Visualization.

Method and Model

- We ran two separate experiments on two well-designed CNNs and compared their results:

1. Efficient Spatial Pyramid of Dilated Convolutions (ESPNet) [13]
   - A light model developed and published by a member of our group.

2. Densely Connected Convolutional Networks (DenseNet161) [14]
   - One of the well-known model in Deep Learning literature.

Results

Evaluation results of ESPNet and DenseNet161 on **Melanoma**

Metrics	ESPNet	DenseNet
Accuracy	0.984	0.988
Precision	0.961	0.984
Recall	0.976	0.968
F1 Score	0.968	0.976
Sensitivity	0.976	0.968
Specificity	0.987	0.995
FP, FN	5, 3	2, 4
TP, TN	122, 370	121 , 373
Training time	35m & 6s	106m & 32s

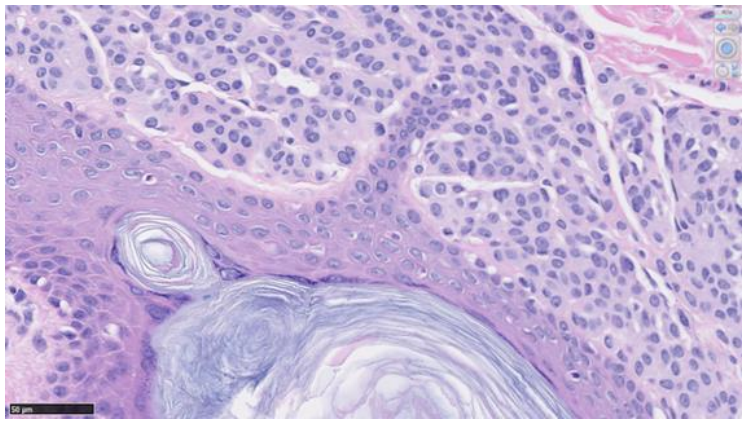
VSGD-Net: Virtual Staining Guided Melanocyte Detection on Histopathological Images

Kechun Liu et al

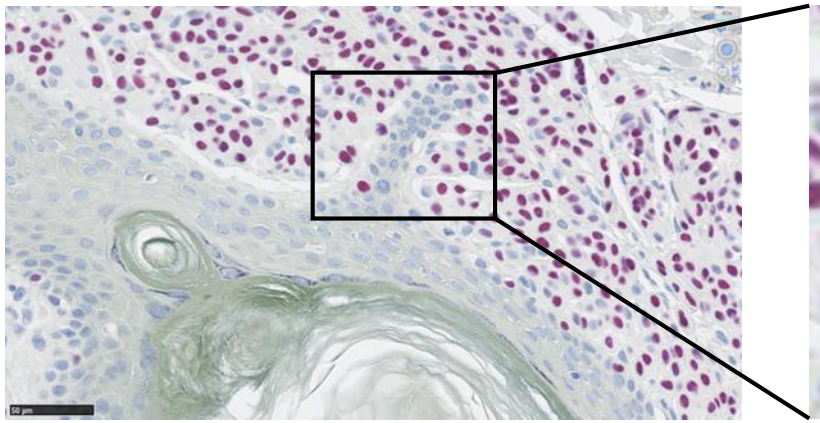
Submitted to WACV 2023

How to Detect Melanocytes

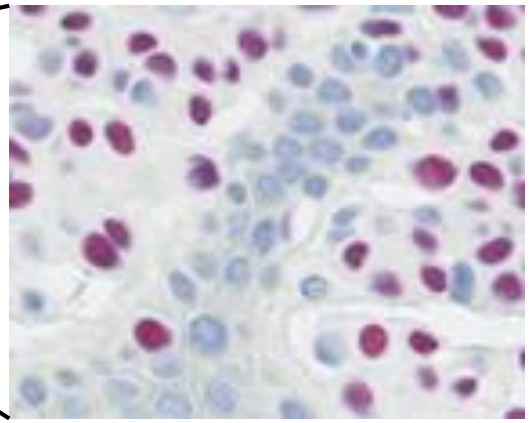
> Sox10 staining can highlight melanocytes, but it's not a routine procedure due to its **high cost**.



(a) H&E Staining



(b) Sox10 Staining – melanocytes are red



(c) Crop from Sox10

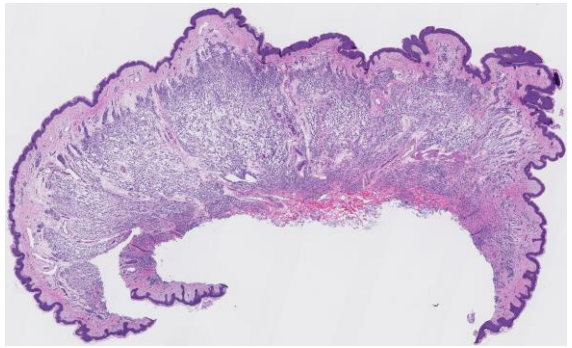
Goal

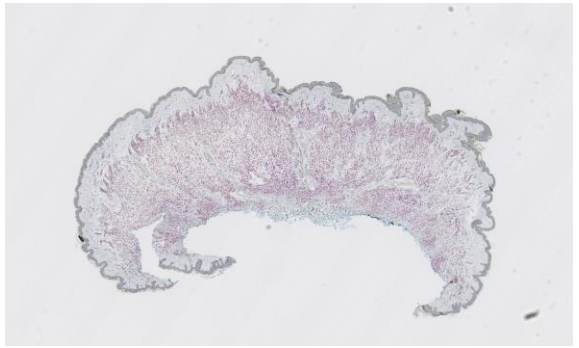
> Can we automatically **detect melanocytes** on **H&E** images?

1. Address the limitations:
   -- Visual similarity of melanocytes to other cells in H&E images
   -- High cost of Sox10 staining
2. Reduce the burden on pathologists
3. Aid in melanoma diagnosis in the future

Dataset for Melanocyte Detection, including both H&E and Sox10 Staining

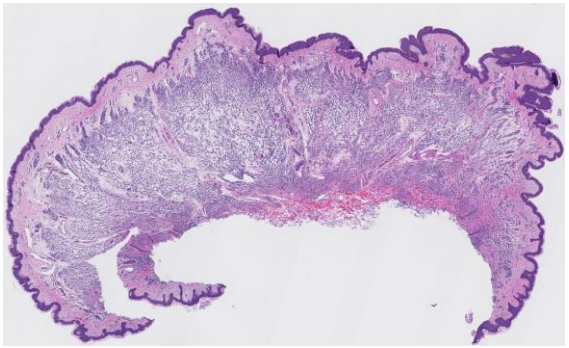
- Our dataset consists of skin tissues of 15 cases from a private dermatopathology lab, including 3 cases for each MPATH diagnostic category[1] [8].
- We stain each glass slide with H&E first, then de-stain and re-stain the same glass slide in Sox10.
- Each skin tissue is cut into multiple (4-6) thin slices for microscopic examination, resulting in **75 slices** at 20x magnification.



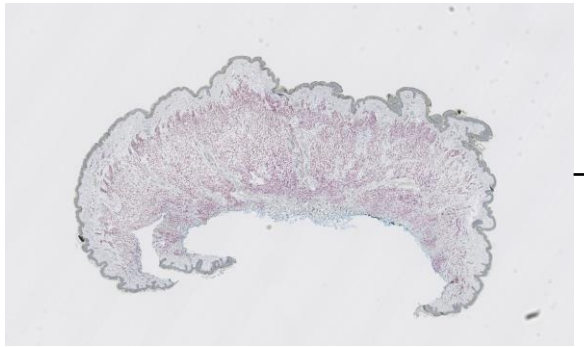


[1] Class 1-5: Benign mildly atypical nevi, Moderate dysplastic nevi, Melanoma in situ, Invasive melanoma T1a, and Invasive melanoma T1b.

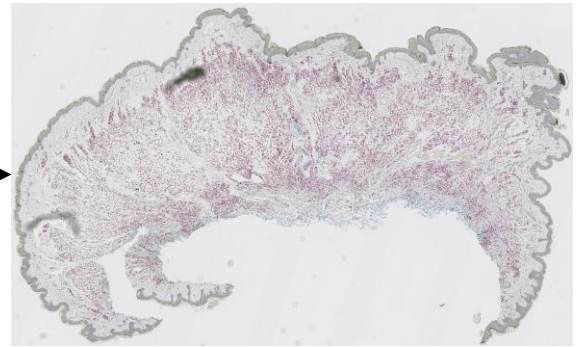
Dataset - Preprocessing



(a) Raw H&E stained image



(b) Raw Sox10 stained image



(c) Registered Sox10 stained image

- First, we register raw Sox10 images (b) into aligned Sox10 images (c) using template H&E images (a) with the Histokat software [1].
- Then, we apply a Random Forest classifier to classify pixels into melanocyte or non-melanocyte.
- At last, the pretrained NuSeT [2] separates touching nuclei and refine the masks.

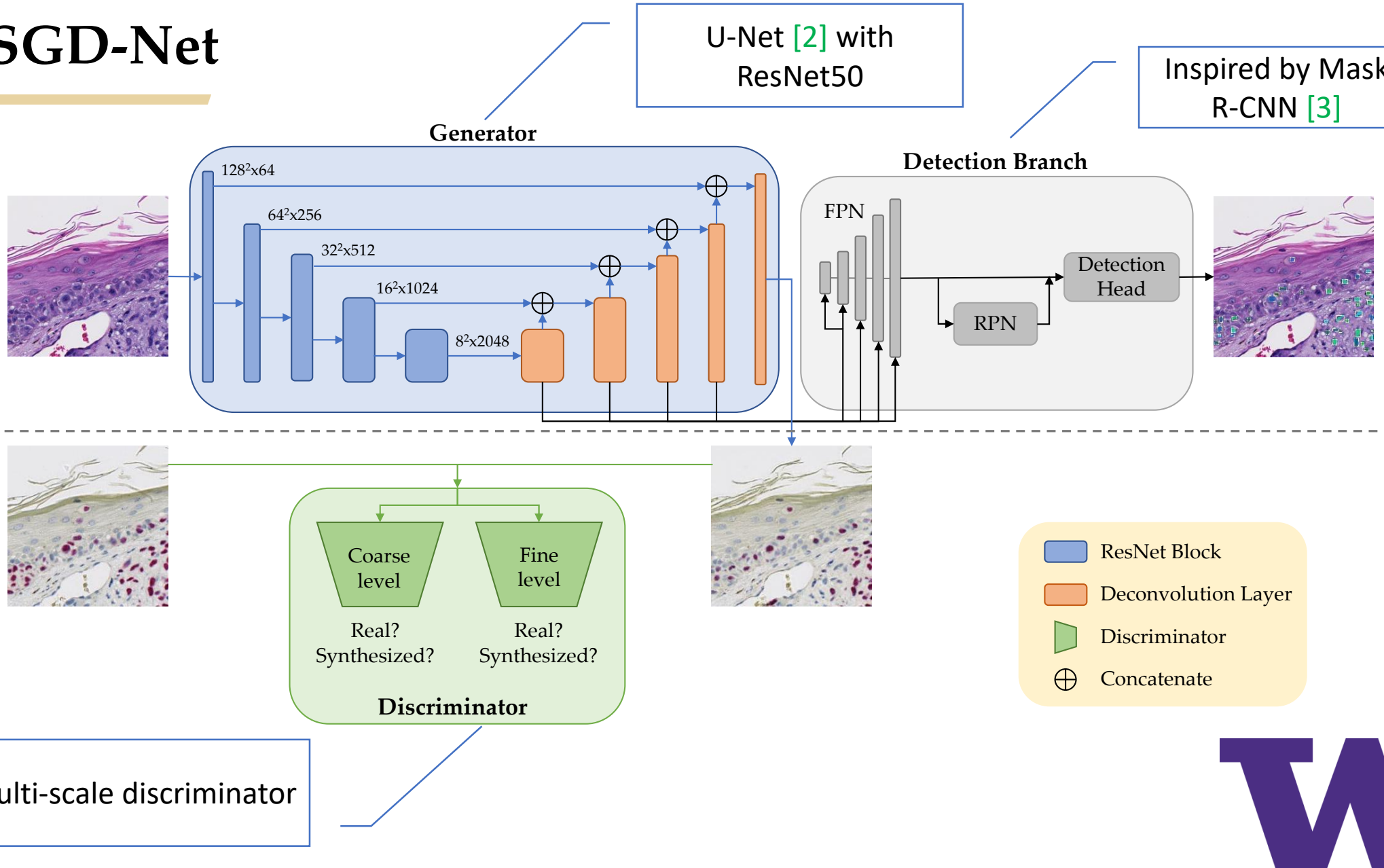
1. Lotz, J., Weiss, N., van der Laak, J., StefanHeldmann: High-resolution Image Registration of Consecutive and Re-stained Sections in Histopathology. arXiv:2106.13150 [cs, eess] (Jun 2021)
2. Yang, L., Ghosh, R.P., Franklin, J.M., Chen, S., You, C., Narayan, R.R., Melcher, M.L., Liphardt, J.T.: NuSeT: A deep learning tool for reliably separating and analyzing crowded cells. PLOS Computational Biology 16(9), e1008193 (Sep 2020), publisher: Public Library of Science

Dataset

> To fit images into memory as well as keep adequate information, we crop the registered paired images into 256x256 patches with 10x magnification and exclude the background patches, leaving 25,314 patches to use.

Sub-dataset	From ? cases	# Patches
Train	10 cases	14,630
Validation		1,032
Test	5 cases	9,652

VSGD-Net


U-Net [2] with ResNet50
Inspired by Mask R-CNN [3]
Generator
Detection Branch
128^2x64
64^2x256
32^2x512
16^2x1024
8^2x2048
FPN
RPN
Detection Head
Coarse level
Fine level
Real? Synthesized?
Real? Synthesized?
Discriminator
ResNet Block
Deconvolution Layer
Discriminator
Concatenate
Multi-scale discriminator

Main Results

Table 1: Results of different methods.

Method	**Precision**	**Recall**	F_1	**Jaccard**
RLS [16]	0.443	0.570	0.499	0.332
Nuclei Classification	0.693	0.506	0.585	0.413
Mask R-CNN [7] (w/o image resize)	0.698	0.500	0.583	0.411
Mask R-CNN [7] (w image resize)	0.735	0.514	0.605	0.434
U-Net [21]	0.630	0.639	0.635	0.465
StarDist[22]	**0.745**	0.426	0.542	0.372
HoverNet[6]	0.729	0.499	0.592	0.421
CHR-Net [3]	0.607	0.688	0.645	0.476
Ours (w Pix2PixHD [27] generator)	0.663	0.645	0.654	0.486
Ours	0.623	**0.733**	**0.674**	**0.508**

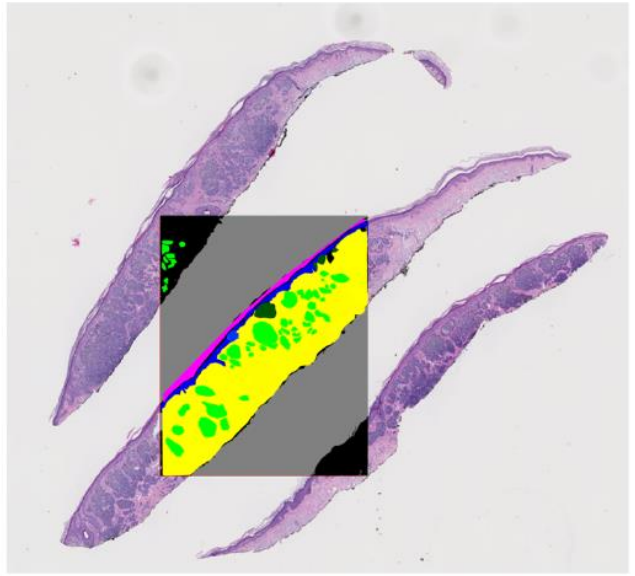
Segmenting Skin Biopsy Images with Coarse and Sparse Annotations using U-Net

Shima Nofallah et al.

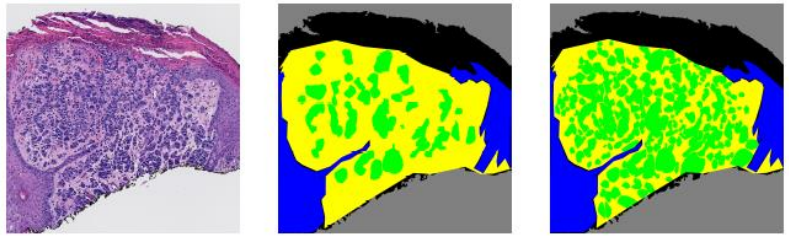
Journal of Digital Imaging 2022

Dataset for Structure Detection (Rough Semantic Segmentation)

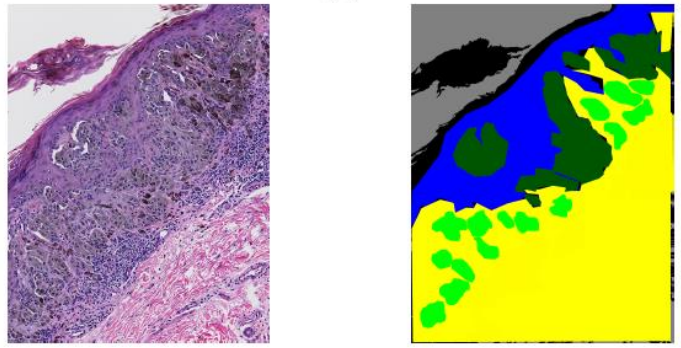
- We obtained coarse and sparse annotations only on the ROI (region of interest) images by an expert pathologist (Dr. Mojgan Mokhtari).
- Not only are the annotations not on the full WSI, but they are also sparse within the annotated ROI.
- Moreover, the annotations are coarse, i.e., they are not pixel-level accurate.



(a)

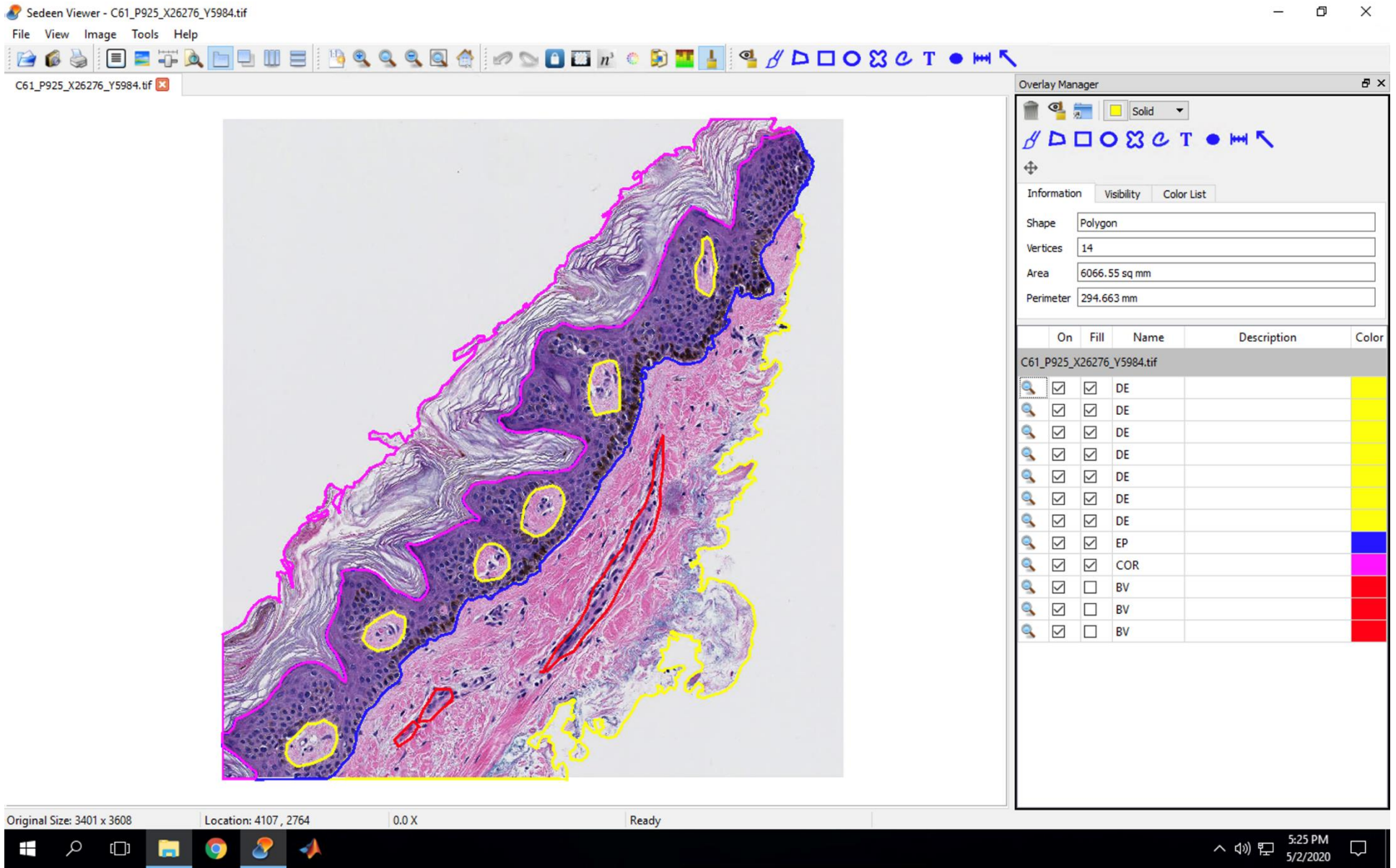


(b)

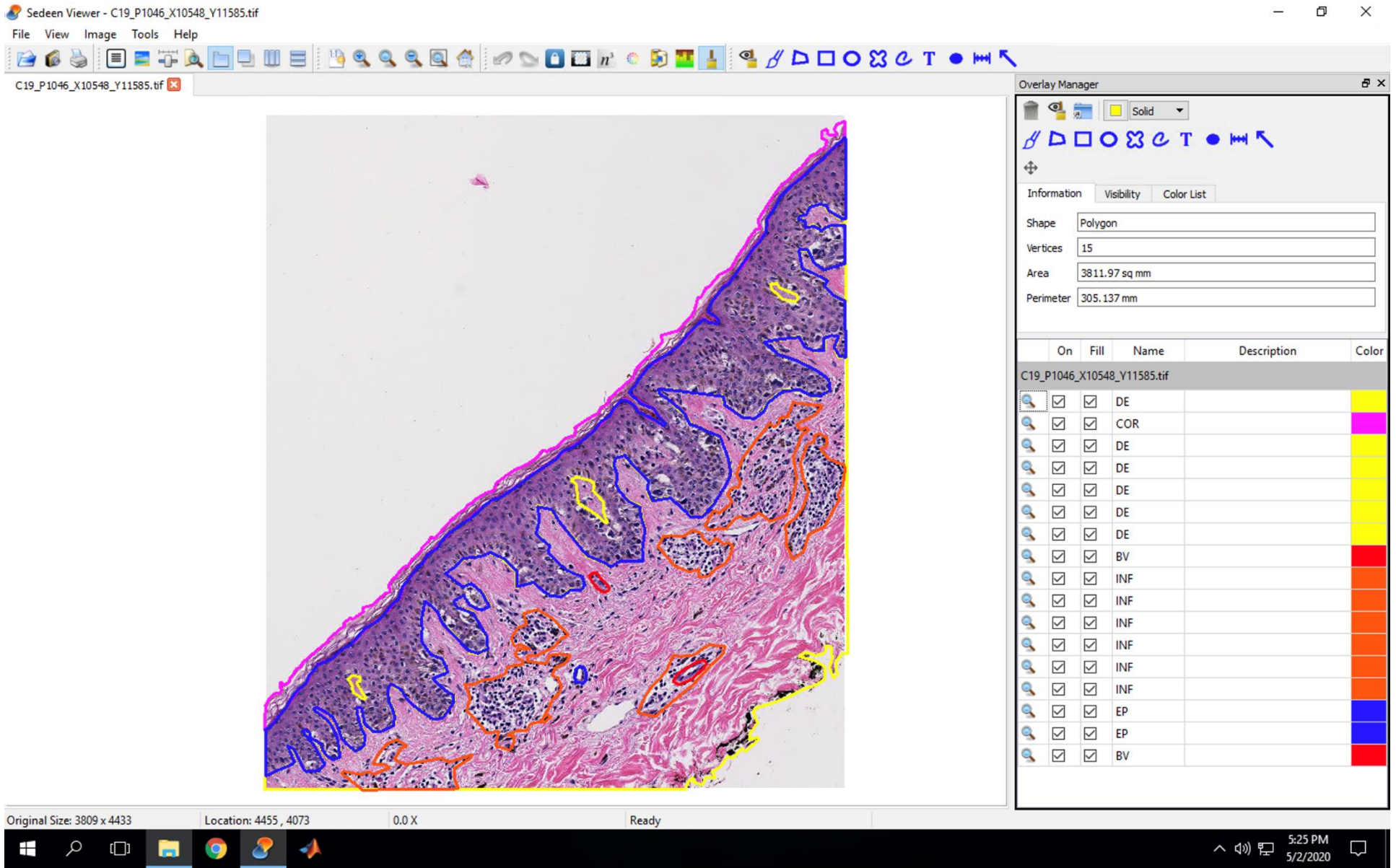


(c)

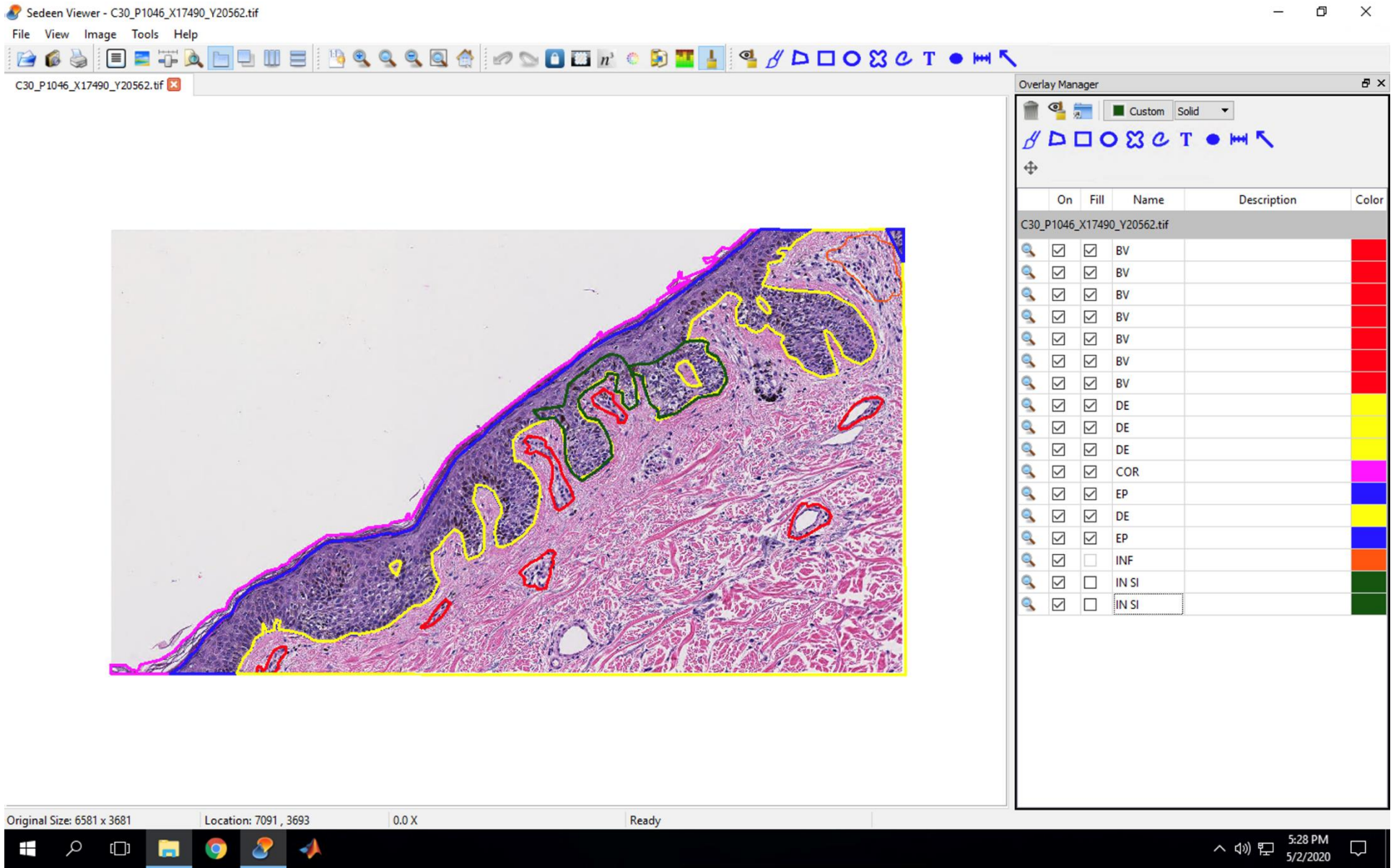
Benign


Sedeen Viewer - C61_P925_X26276_Y5984.tif
File View Image Tools Help
C61_P925_X26276_Y5984.tif
Overlay Manager
Solid
Information Visibility Color List
Shape Polygon
Vertices 14
Area 6066.55 sq mm
Perimeter 294.663 mm
On Fill Name Description Color
C61_P925_X26276_Y5984.tif
DE
DE
DE
DE
DE
DE
DE
EP
COR
BV
BV
BV
Original Size: 3401 x 3608
Location: 4107 , 2764
0.0 X
Ready
5:25 PM
5/2/2020

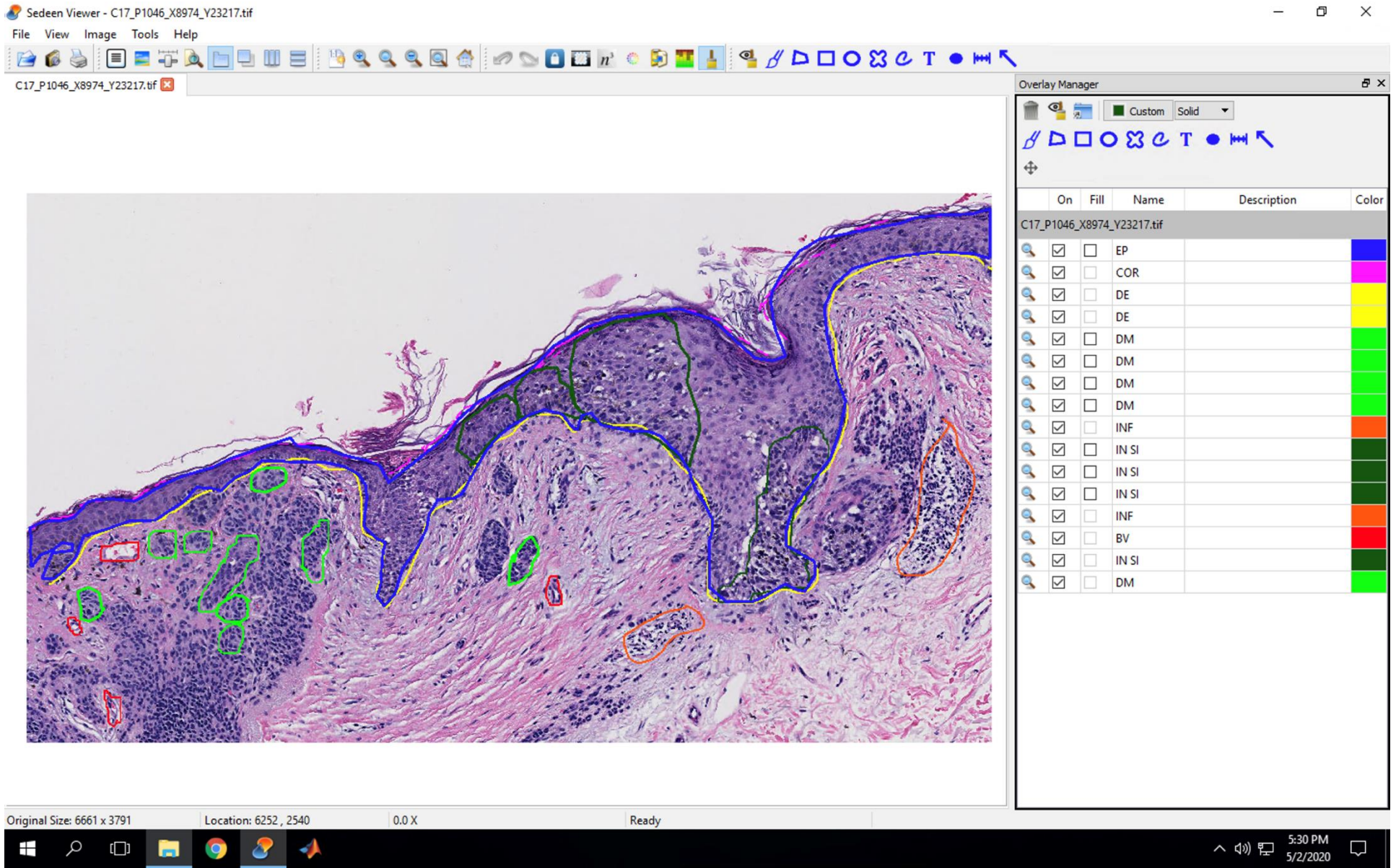
Atypia


Sedeen Viewer - C19_P1046_X10548_Y11585.tif
File View Image Tools Help
C19_P1046_X10548_Y11585.tif
Overlay Manager
Solid
Information Visibility Color List
Shape Polygon
Vertices 15
Area 3811.97 sq mm
Perimeter 305.137 mm
On Fill Name Description Color
C19_P1046_X10548_Y11585.tif
DE
COR
DE
DE
DE
DE
DE
BV
INF
INF
INF
INF
INF
INF
EP
EP
BV
Original Size: 3809 x 4433
Location: 4455 , 4073
0.0 X
Ready
5:25 PM
5/2/2020

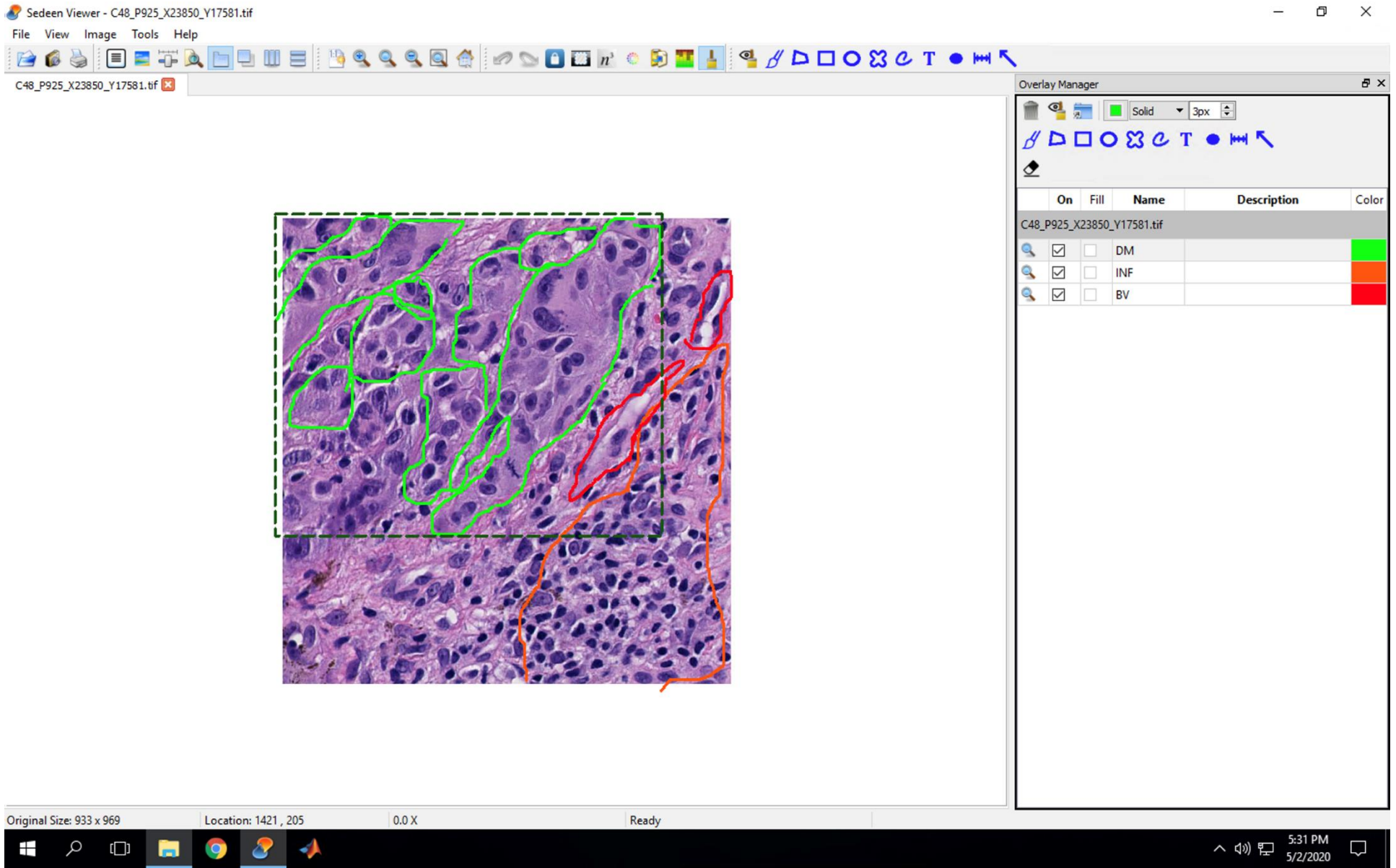
Melanoma In Situ



Invasive Cancer Class IA



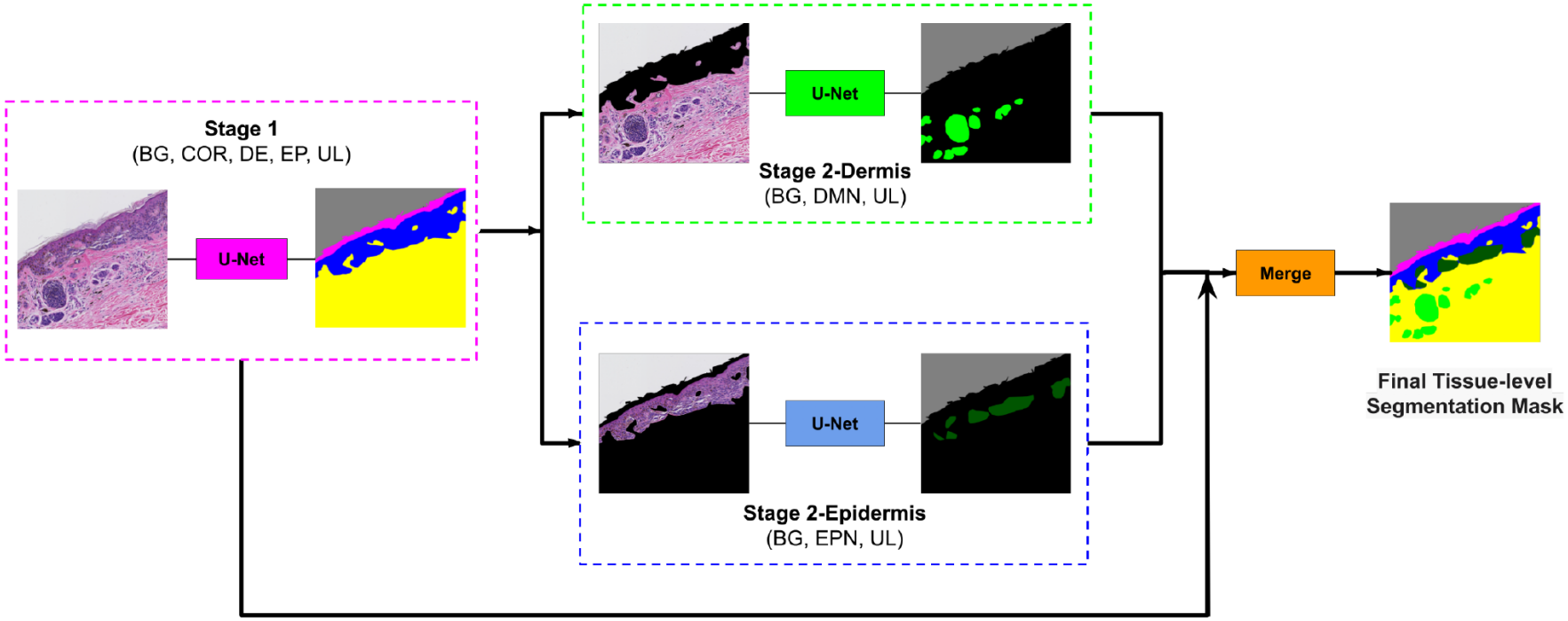
Invasive Cancer Class IB



Method and Model

Two stage method: first big tissue structures, then smaller tissue structures.

BG=Background, COR=Stratum Corneum, DE=Dermis, EP=Epidermis, UL=Unlabeled, DMN=Dermal Nests, EPN= Epidermal Nests.



Results

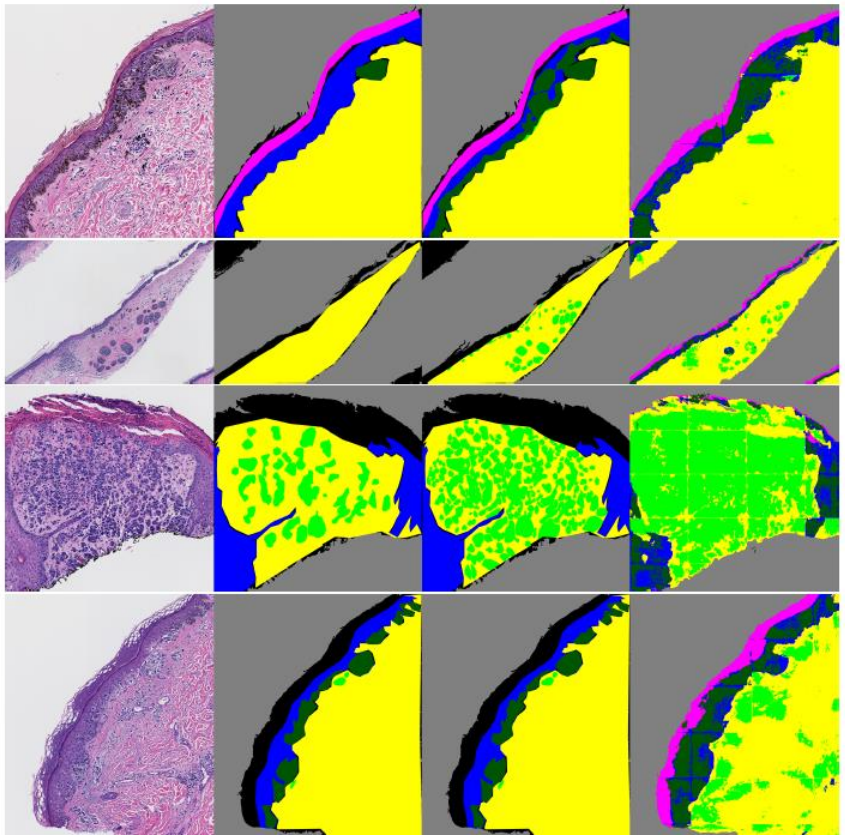
- Evaluation of the segmentation model on **ROI** testing set.

Segmentation stage	Dice score	IoU
Stage 1 (all tissues)	0.942	0.906
Stage 2-Dermis (DMN)	0.558	0.638
Stage 2-Epidermis (EPN)	0.332	0.558

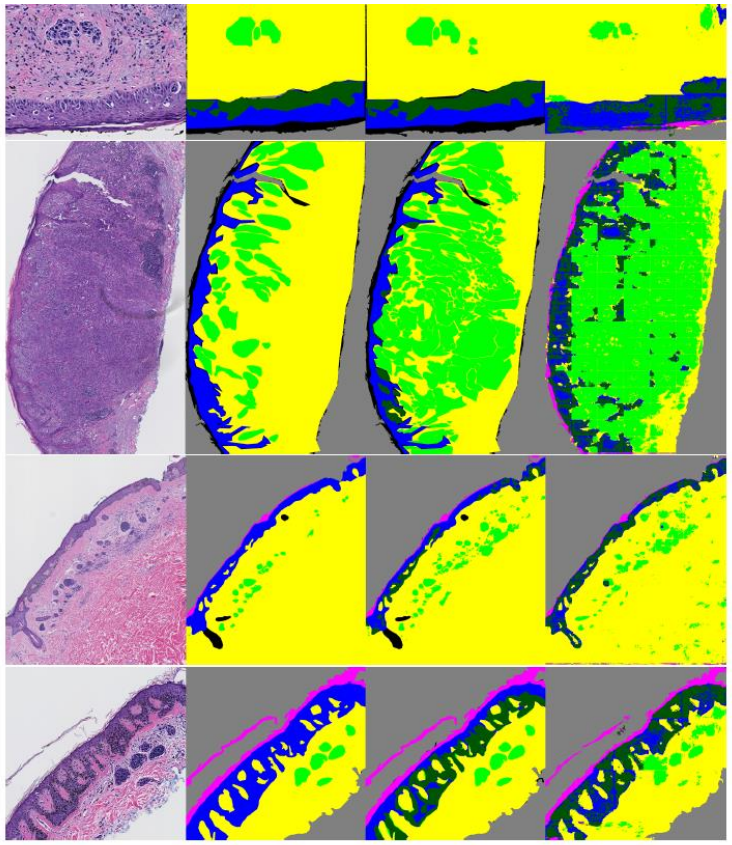
$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)} \qquad IoU = \frac{TP}{TP + FN + FP}$$

Some Results - ROI testing set

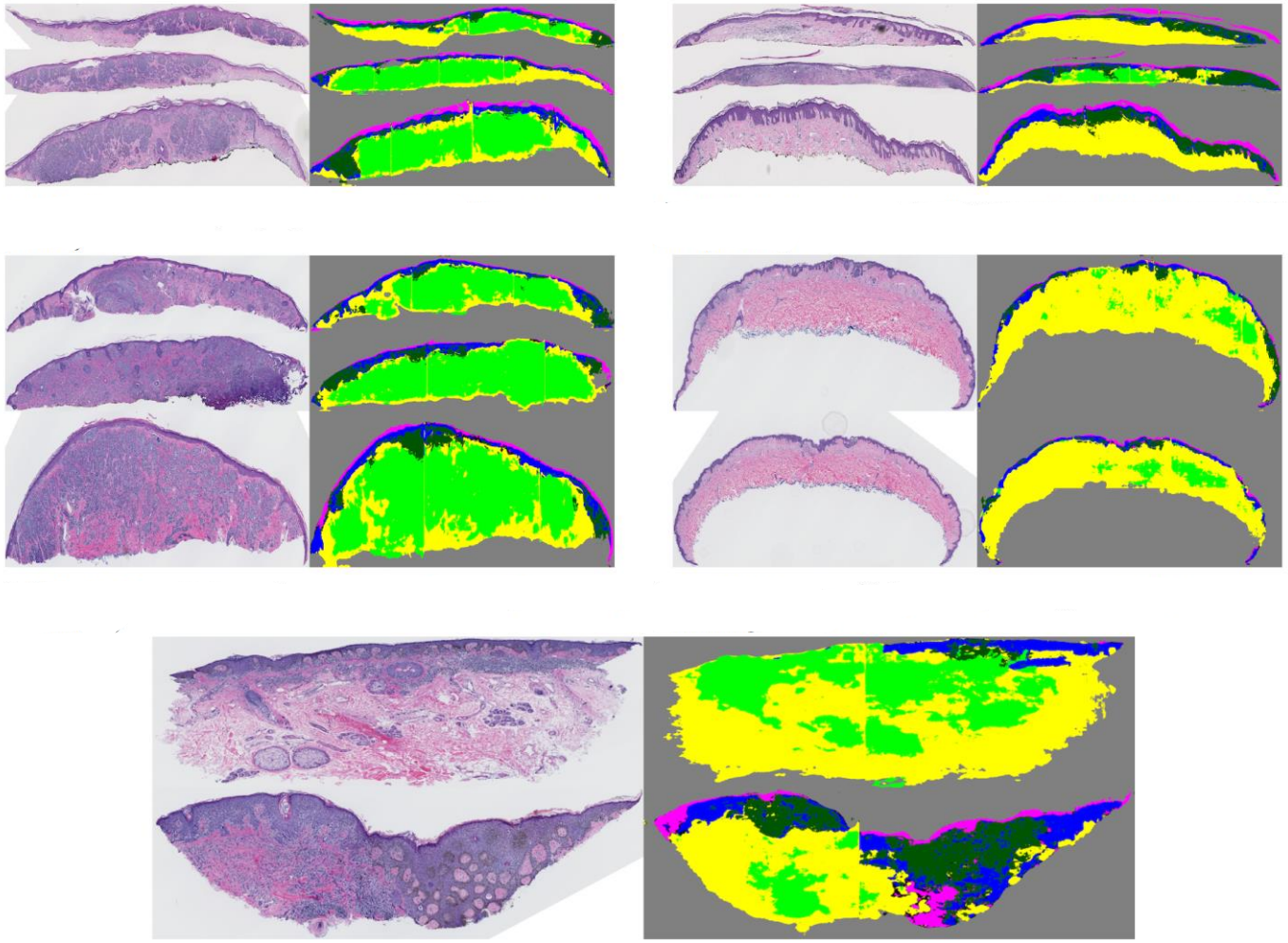
(a) ROI image (b) Sparse annot. (c) Fine annot. (d) Predicted mask



(a) ROI image (b) Sparse annot. (c) Fine annot. (d) Predicted mask



Results - Generating WSI Segmentation Masks

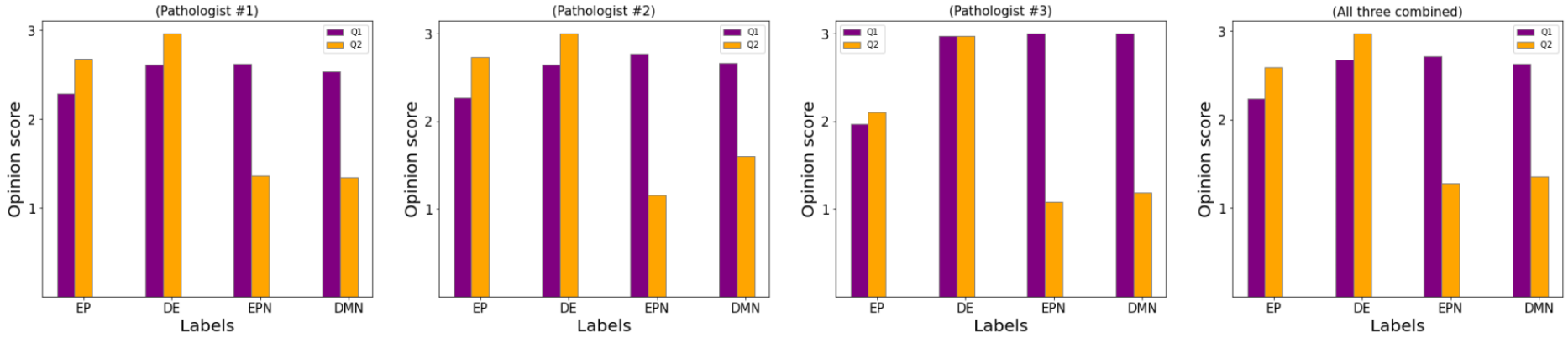


Subjective Assessment with Pathologists

Qualitatively evaluation the WSI segmentation, with these questions:

- **Q1:** How much of the tissue/area that is present in the corresponding WSI has been correctly identified by the model? Rate Low, Medium, or High.

- **Q2:** How much of the label identified by the model is the correct tissue/area? Rate Low, Medium, or High.

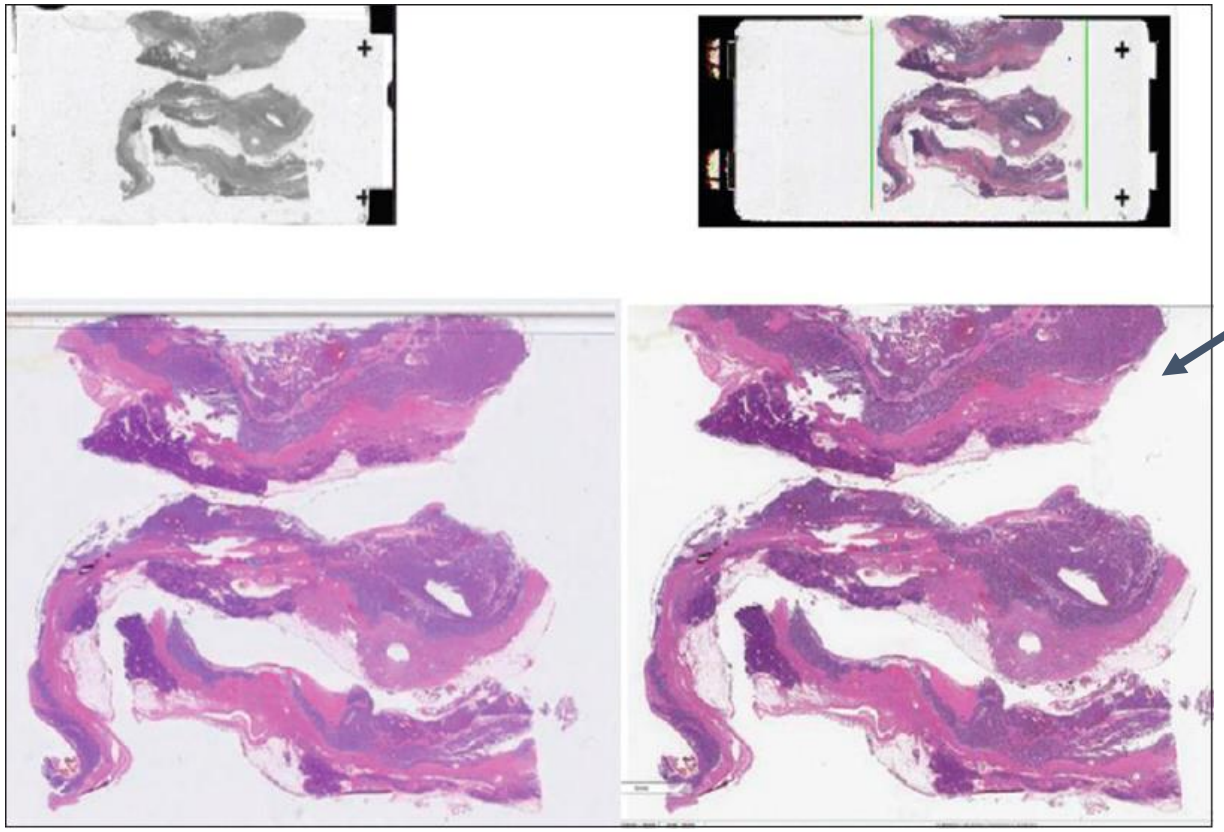


(a) Individual survey #1 (b) Individual survey #2 (c) Individual survey #3 (d) Combination of all three

Scale-Aware Transformers for Diagnosing Melanocytic Lesions

Wenjun Wu et al.
IEEE Access 2021

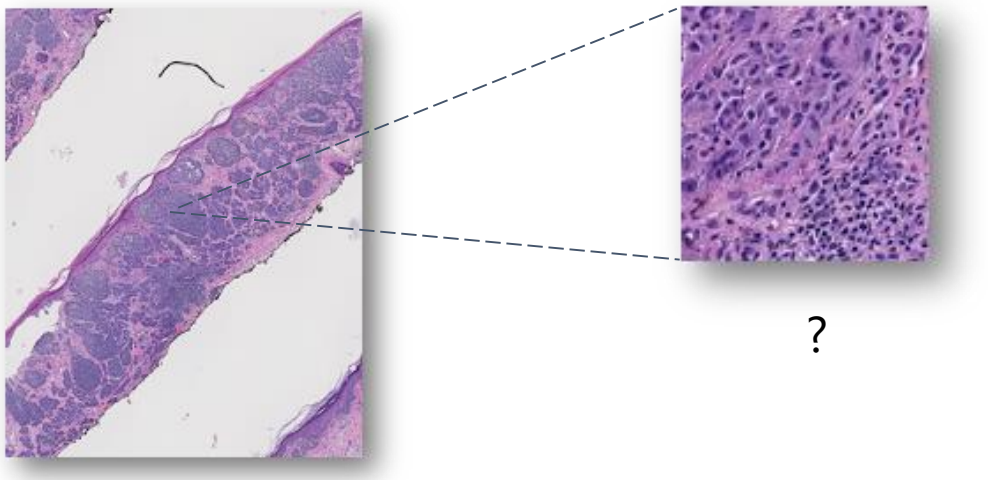
Diagnosing Whole Slide Images with Transformers



Multiple tissues

Difficulties in diagnosis

Mixed normal and cancerous tissue

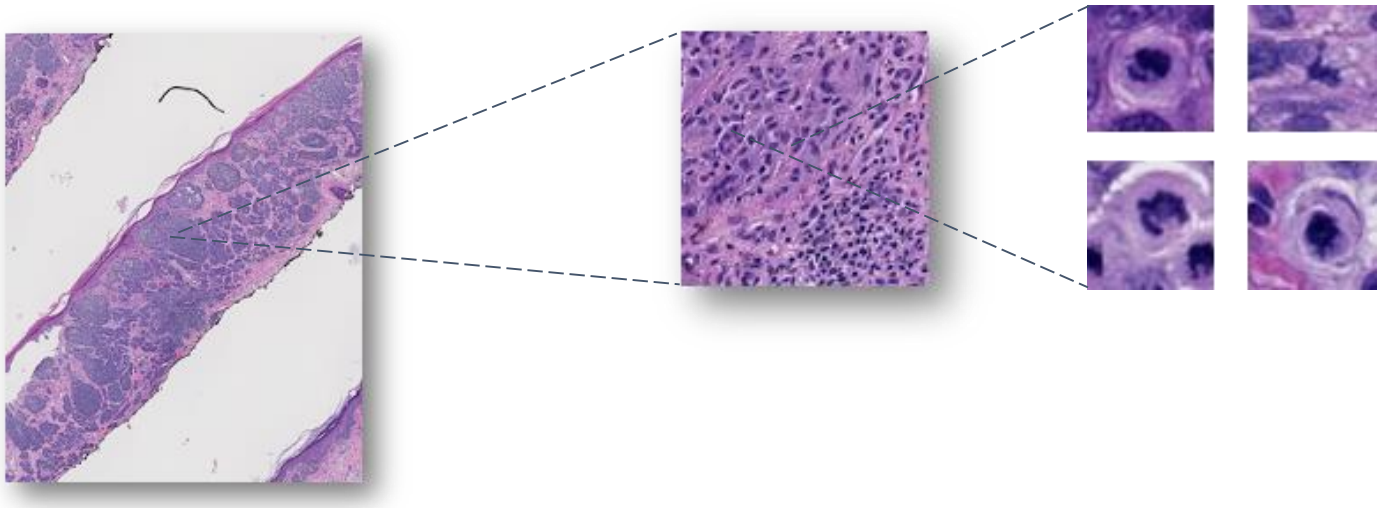


?

Difficulties in learning to diagnose

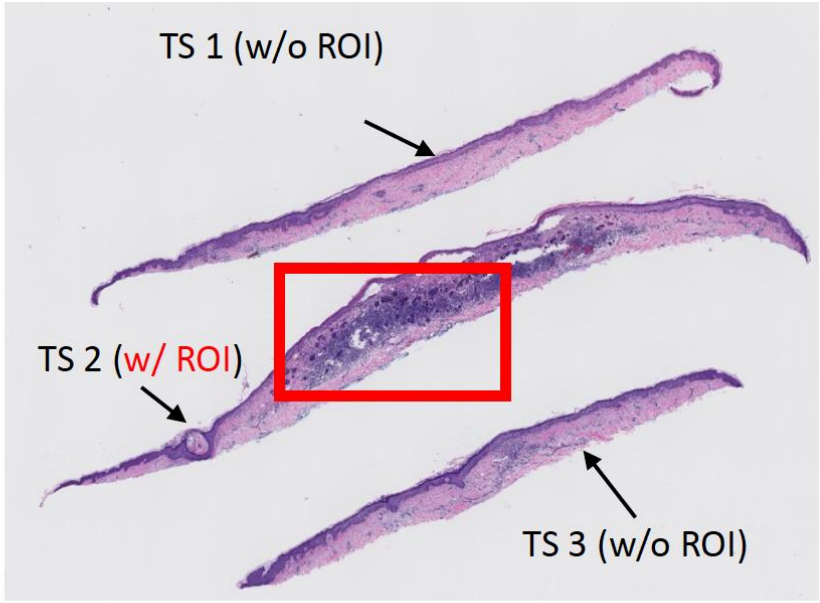
Mixed normal and cancerous tissue

Feature is dependent on resolution



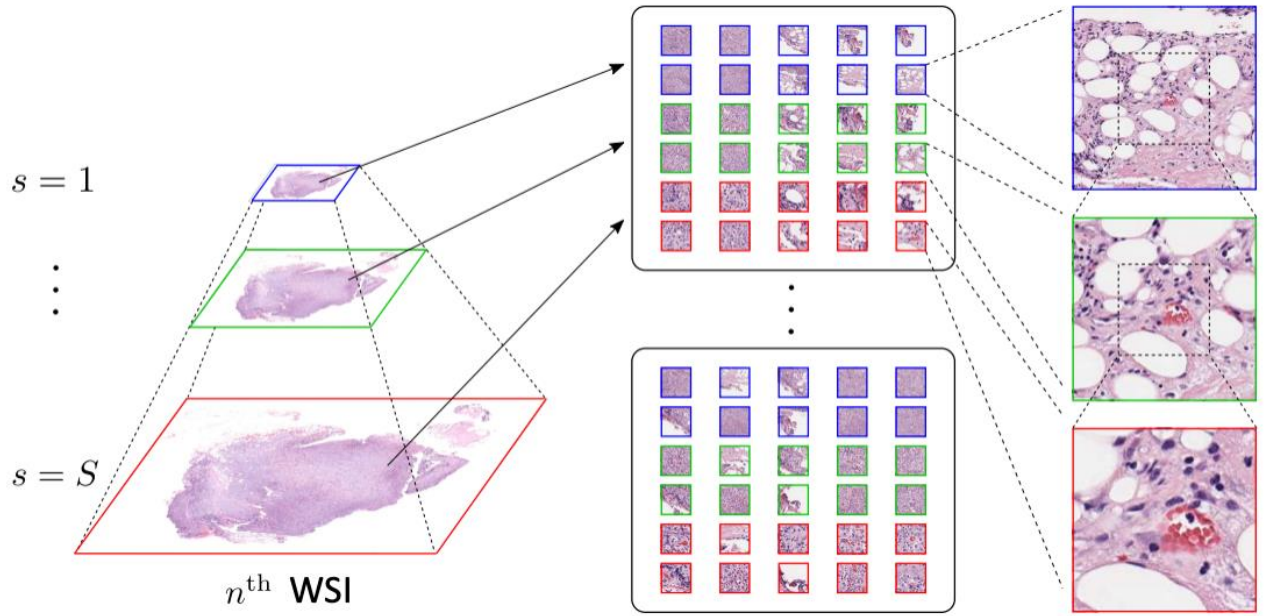
Dataset Problems

Invasive T1a Skin Biopsy Image
(or Class 3)

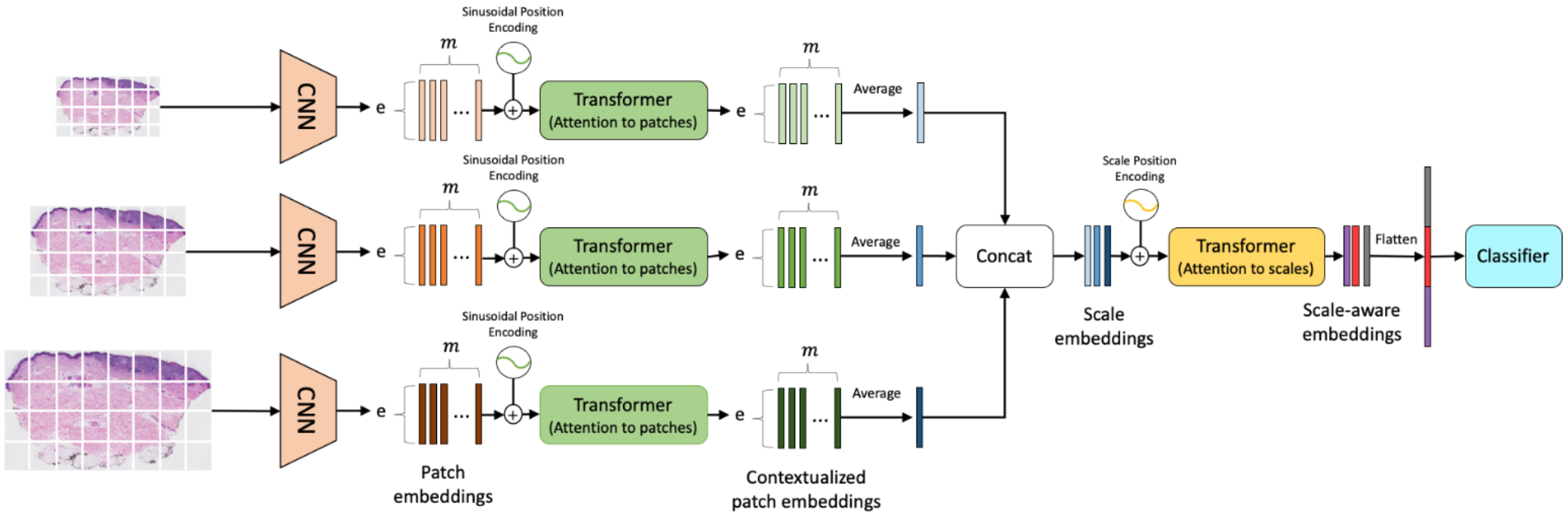


Key Ideas

- Self-attention-based framework for classifying WSIs at multiple input scales
- A soft label assignment method to reduce ambiguities



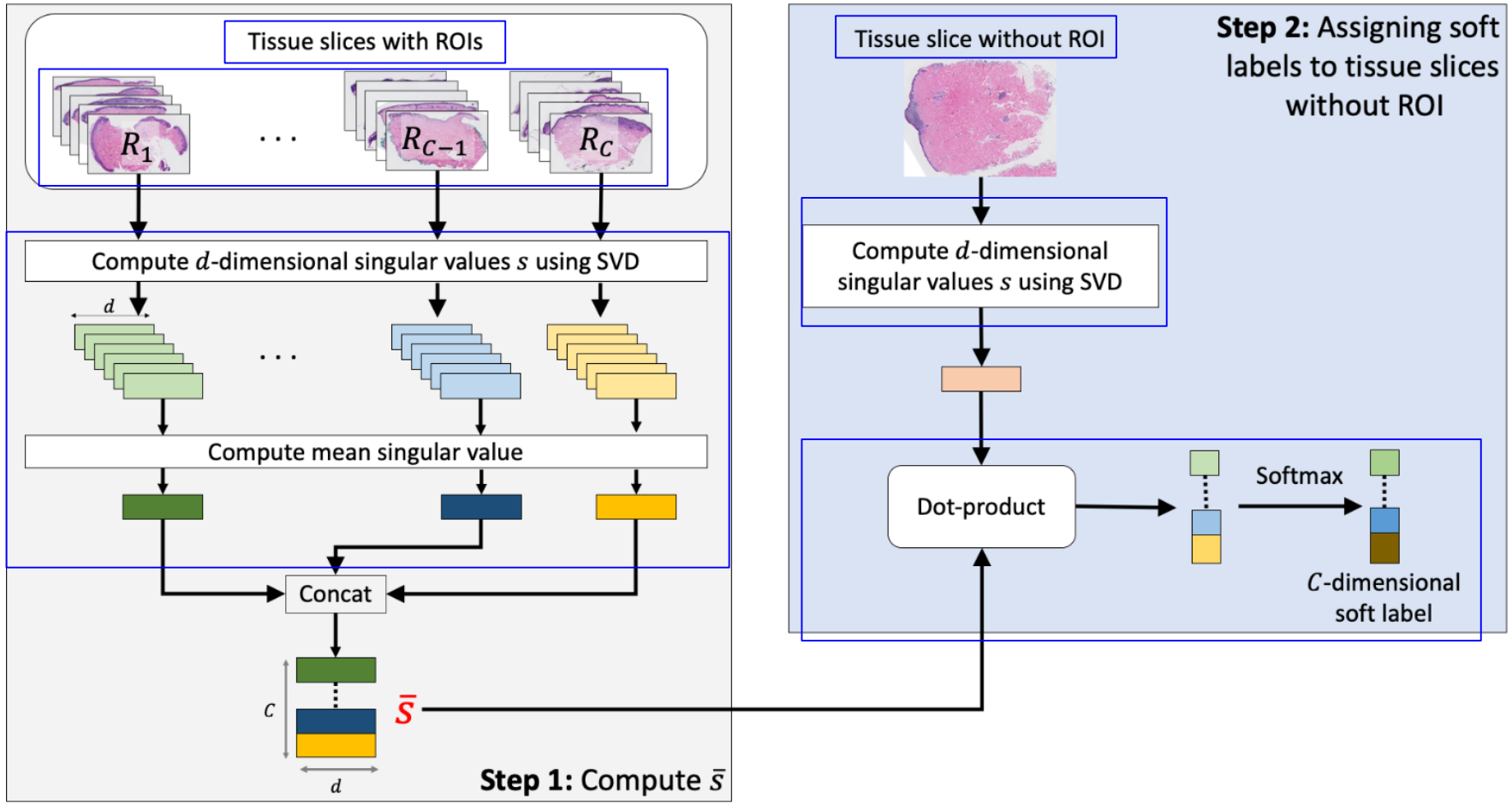
Scale-Aware Transformers for Diagnosing Melanocytic Lesions



Soft labels

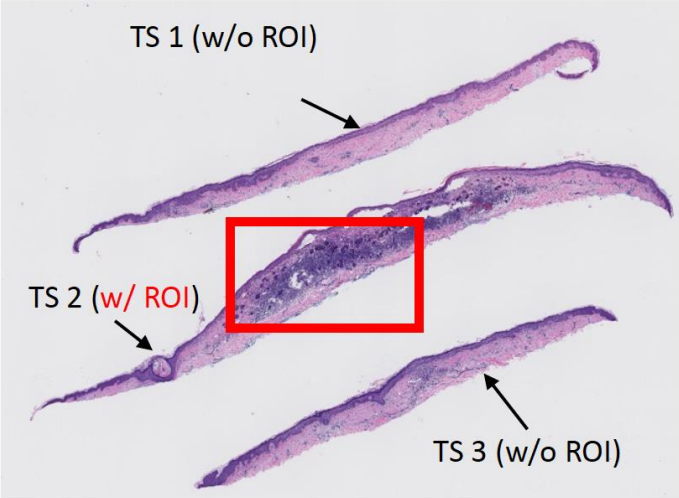
Tissues in each class C are represented by a mean singular value vector for that class obtained from tissue slices WITH ROIs.

For slices with no ROI, the dot product of its SVD vector is taken with the mean vector from each class.



Soft labels

Invasive T1a Skin Biopsy Image (or Class 3)



Hard Label (one-hot encoding)

TS 1	0	0	1	0
TS 2	0	0	1	0
TS 3	0	0	1	0

Label smoothing (smoothing=0.1)

TS 1	0.033	0.033	0.9	0.033
TS 2	0.033	0.033	0.9	0.033
TS 3	0.033	0.033	0.9	0.033

Constrained label smoothing

TS 1	0.5	0.5	0	0
TS 2	0	0	1	0
TS 3	0.5	0.5	0	0

Soft labels (ours)

TS 1	0.54	0.46	0	0
TS 2	0	0	1	0
TS 3	0.28	0.72	0	0

Experimental Result: baseline methods

Row #	Method	Accuracy	F1	Sensitivity	Specificity	AUC
R1	Patch-based (SSC)	0.35	0.35	0.35	0.79	0.67
R2	Patch-based (MSC)	0.40	0.40	0.40	0.80	0.68
R3	Penultimate-weighted (SSC)	0.44	0.44	0.44	0.81	0.67
R4	Hypercolumn-weighted (SSC)	0.43	0.43	0.43	0.43	0.67
R5	Streaming CNN (SSC)	0.32	0.32	0.32	0.77	0.58
R6	ChikonMIL (SSC)	0.56	0.56	0.56	0.85	0.74
R7	MS-DA-MIL (SSC)	0.49	0.49	0.49	0.83	0.68
R8	MS-DA-MIL (MSC*)	0.58	0.58	0.58	0.86	0.75
R9	`ScAtNet` (SSC)	0.60	0.60	0.60	0.87	0.77
R10	`ScAtNet` (MSC)	**0.64**	**0.64**	**0.64**	**0.88**	**0.79**

TABLE 2: Comparison of overall performance with state-of-the-art WSI classification methods across different metrics on the test set. Here, SSC denotes single input scale (10×). MSC denotes multiple input scales (7.5×, 10×, 12.5×). MSC* denotes multiple input scales (10×, 20×)

Experimental Result: soft label

Method	Accuracy	Specificity	AUC
Hard labels	0.50	0.83	0.73
Label smoothing	0.50	0.83	0.71
Constrained label smoothing	0.56	0.85	0.77
Soft labels (Ours; Section III-C)	**0.60**	**0.87**	**0.77**

Comparison of the performance of different labeling methods.

Experimental Result: pathologists performance

Comparisonto 187 practicing pathologists in a study

Diagnostic	Accuracy		F1		Sensitivity		Specificity	
Category	PG	Ours	PG	Ours	PG	Ours	PG	Ours
MMD	0.92	0.79	0.71	0.75	0.92	0.79	0.76	0.89
MIS	0.46	0.40	0.49	0.44	0.46	0.40	0.85	0.84
pT1a	0.51	0.65	0.62	0.63	0.51	0.65	0.95	0.84
pT1b	0.72	0.77	0.72	0.74	0.78	0.77	0.97	0.92
Overall	0.65	0.64	0.65	0.64	0.65	0.64	0.88	0.88

Comparison of ScAtNet with pathologists' (PG) performance.

Current Work

- Using the segmentation labels to boost the performance of the transformer classifiers to perform diagnosis
- Using the size and shape pattern of the melanocytes to improve diagnosis
- Teaching transformers which patches are most important due to their content

References

[1] Caner Mercan, Selim Aksoy, Ezgi Mercan, Linda G Shapiro, Donald L Weaver, and Joann G Elmore. Multi-instance multi-label learning for multi-class classification of whole slide breast histopathology images. IEEE transactions on medical imaging, 37(1):316–325, 2017.

[2] Haomiao Ni, Hong Liu, Kuansong Wang, Xiangdong Wang, Xunjian Zhou, and Yueliang Qian. Wsi-net: Branch-based and hierarchy-aware network for segmentation and classification of breast histopathological whole-slide images. In International Workshop on Machine Learning in Medical Imaging, pages 36–44. Springer, 2019.

[3] Salah Alheejawi, Richard Berendt, Naresh Jha, Santi P Maity, and Mrinal Mandal. Automated proliferation index calculation for skin melanoma biopsy images using machine learning. Computerized Medical Imaging and Graphics, 89:101893, 2021.

[4] Cheng Lu and Mrinal Mandal. Automated analysis and diagnosis of skin melanoma on whole slide histopathological images. Pattern Recognition, 48(8):2738–2750, 2015.

[5] Hongming Xu, Cheng Lu, Richard Berendt, Naresh Jha, and Mrinal Mandal. Automated analysis and classification of melanocytic tumor on skin whole slide images. Computerized medical imaging and graphics, 66:124–134, 2018.

[6] Mike Van Zon, Nikolas Stathonikos, Willeke AM Blokx, Selim Komina, Sybren LN Maas, Josien PW Pluim, Paul J Van Diest, and Mitko Veta. Segmentation and classification of melanoma and nevus in whole slide images. In 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), pages 263–266. IEEE, 2020.

[7] Achim Hekler, Jochen Sven Utikal, Alexander H Enk, Carola Berking, Joachim Klode, Dirk Schadendorf, Philipp Jansen, Cindy Franklin, Tim Holland-Letz, Dieter Krahl, et al. Pathologist-level classification of histopathological melanoma images with deep neural networks. European Journal of Cancer, 115:79–83, 2019.

References

[8] John F Thompson, Seng-Jaw Soong, Charles M Balch, Jeffrey E Gershenwald, Shouluan Ding, Daniel G Coit, Keith T Flaherty, Phyllis A Gimotty, Timothy Johnson, Marcella M Johnson, et al. Prognostic significance of mitotic rate in localized primary cutaneous melanoma: an analysis of patients in the multi-institutional american joint committee on cancer melanoma staging database. Journal of Clinical Oncology, 29(16):2199, 2011.

[9] Olcay Sertel, Umit V Catalyurek, Hiroyuki Shimada, and Metin N Gurcan. Computer-aided prognosis of neuroblastoma: Detection of mitosis and karyorrhexis cells in digitized histological images. In 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pages 1433–1436. IEEE, 2009.

[10] Ludovic Roux, Daniel Racoceanu, Nicolas Loménie, Maria Kulikova, Humayun Irshad, Jacques Klossa, Frédérique Capron, Catherine Genestie, Gilles Le Naour, and Metin N Gurcan. Mitosis detection in breast cancer histological images an icpr 2012 contest. Journal of pathology informatics, 4, 2013.

[11] Humayun Irshad, Ludovic Roux, and Daniel Racoceanu. Multi-channels statistical and morphological features based mitosis detection in breast cancer histopathology. In 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pages 6091–6094. IEEE, 2013.

[12] Dan C Cire¸san, Alessandro Giusti, Luca M Gambardella, and Jürgen Schmidhuber. Mitosis detection in breast cancer histology images with
deep neural networks. In International Conference on Medical Image Computing and Computer-assisted Intervention, pages 411–418. Springer, 2013.

[13] Sachin Mehta, Mohammad Rastegari, Anat Caspi, Linda Shapiro, and Hannaneh Hajishirzi. Espnet: Efficient spatial pyramid of dilated convolutions for semantic segmentation. In ECCV, 2018b.

References

[14] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017.

[15] Hao Zheng, Lin Yang, Jianxu Chen, Jun Han, Yizhe Zhang, Peixian Liang, Zhuo Zhao, Chaoli Wang, and Danny Z Chen. Biomedical image segmentation via representative annotation. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 33, pages 5901–5908, 2019.

[16] Zeju Li, Konstantinos Kamnitsas, and Ben Glocker. Overfitting of neural nets under class imbalance: Analysis and improvements for segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 402–410. Springer, 2019.

[17] Jamshid Sourati, Ali Gholipour, Jennifer G Dy, Sila Kurugol, and Simon K Warfield. Active deep learning with fisher information for patch-wise semantic segmentation. In Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, pages 83–91. Springer, 2018.

[18] Dwarikanath Mahapatra, Behzad Bozorgtabar, Jean-Philippe Thiran, and Mauricio Reyes. Efficient active learning for image classification and segmentation using a sample selection and conditional generative adversarial network. In International Conference on Medical Image Computing and Computer-Assisted Intervention, pages 580–588. Springer, 2018.

[19] Yishuo Zhang and Albert CS Chung. Deep supervision with additional labels for retinal vessel segmentation task. In International conference on medical image computing and computer-assisted intervention, pages 83–91. Springer, 2018.

Acknowledgment

Research reported in this study was supported by grants R01CA200690 and U01CA231782 from the National Cancer Institute of the National Institutes of Health, 622600 from Melanoma Research Alliance, and W81XWH-20-1-0798 from the United States Department of Defense.

PI:
Dr. Linda Shapiro

PI:
Dr. Joann Elmore

Pathologists:
Dr. Stevan Knezevich
Dr. Caitlin May
Dr. Oliver Chang
Dr. Mojgan Mokhtari

Collaborators:
Wenjun Wu
Dr. Sachin Mehta
Dr. Annie Lee
Dr. Daniella Witten
Dr. Ezgi Mercan

Lab mates:
Nicholas Nuechterlein
Dr. Beibin Li
Kechun Liu
Dr. Shima Nofallah
Fatemeh Ghezloo