Improving the Accuracy of Melanoma Diagnosis from Whole Slide Images

By: Shima Nofallah

Contact: shima@cs.washington.edu

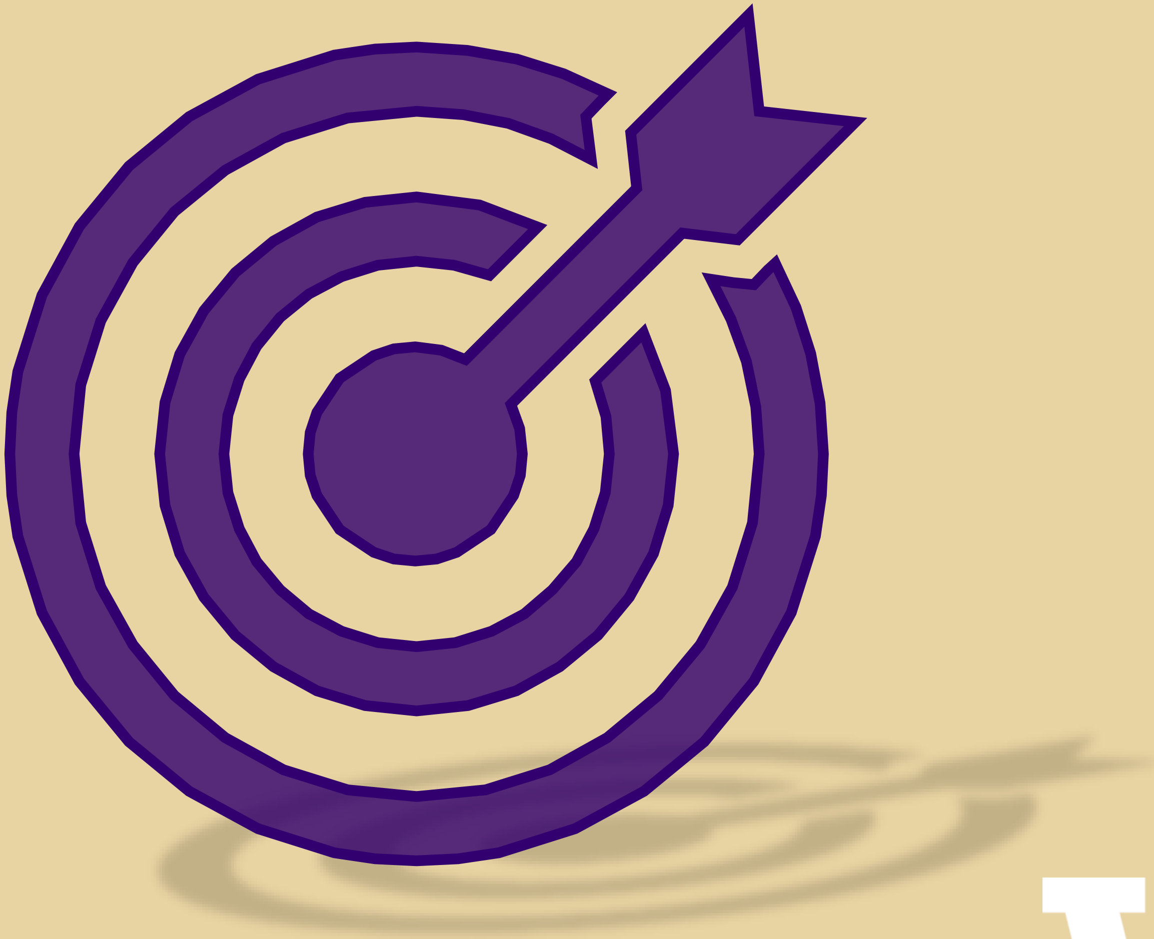
Autumn 2021

ELECTRICAL ENGINEERING
UNIVERSITY of WASHINGTON

Outline

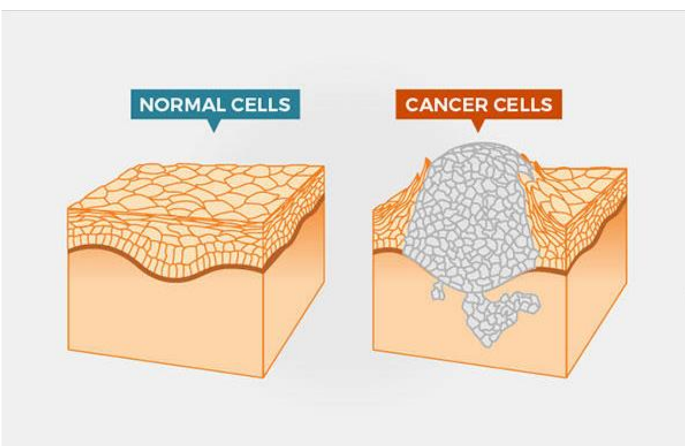
- Motivation and Goals
- Dataset
- Preliminary Work
  1. Mitosis Classification
  2. Segmentation using Coarse and Sparse Annotation
- Future work

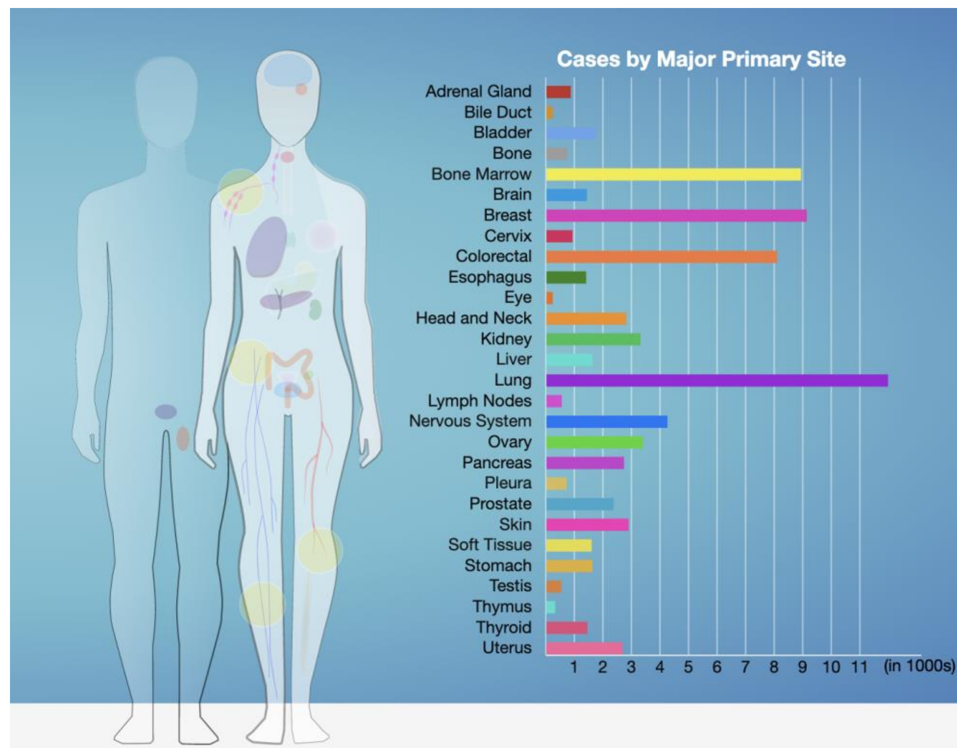
Motivation and Goals



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.

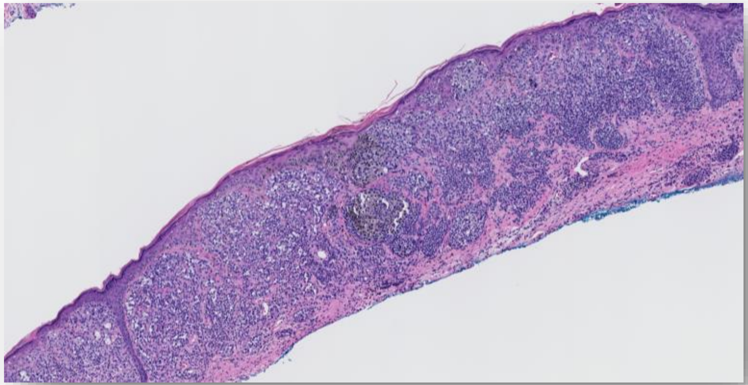




Figures are from National Institute of Cancer website.

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.

Related Work

- There is various work related to the diagnosis of biopsy images of other types of cancers than melanoma, especially on breast histopathological Whole Slide Images (WSI) [1,2].
- Related work for melanoma diagnosis using skin biopsy WSI is very limited. There are works on staining other than H&E, such as Ki-67 stain [3].
- Most existing work on melanoma diagnosis are either binary classification system or on not very challenging categories to distinguish [4,5,6,7].

Melanoma Diagnosis

Structural entities:

- Dermis
- Epidermis
- Nests

Cellular entities:

- Nuclei
- Melanocyte
- Mitosis

Raw WSI

Diagnosis model

Diagnoses:

Mild Dysplastic

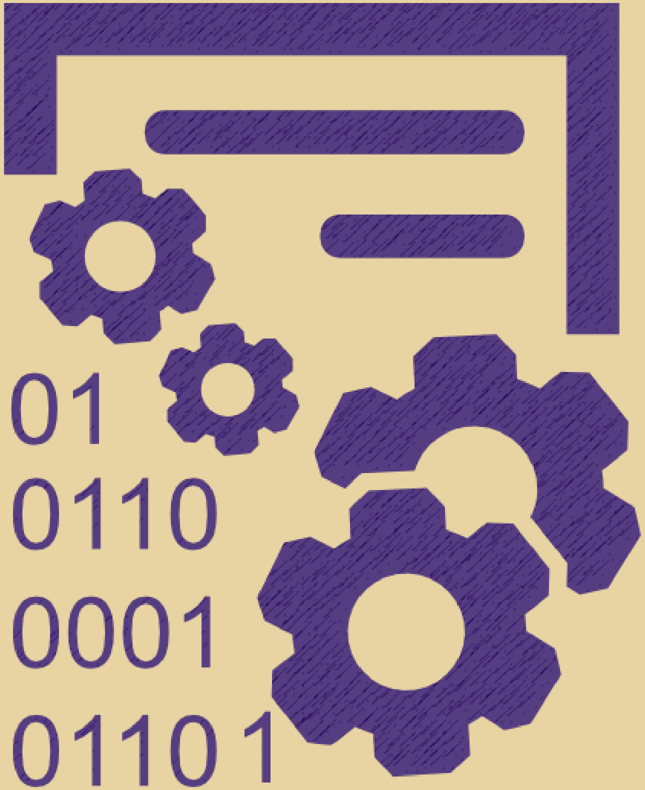
Moderate Dysplastic

Melanoma in situ

Invasive T1a

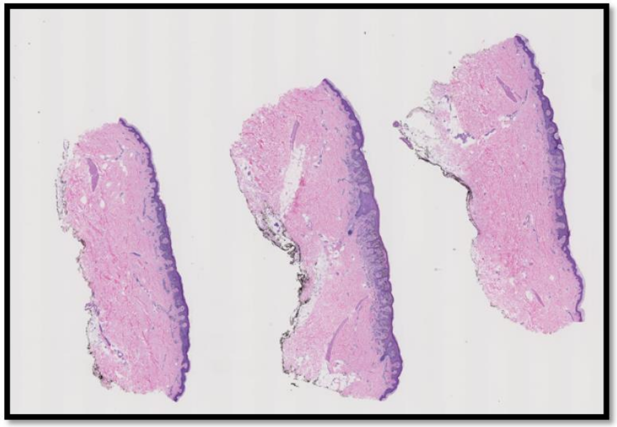
Invasive T1b

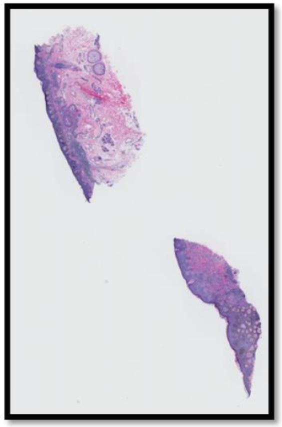
Dataset

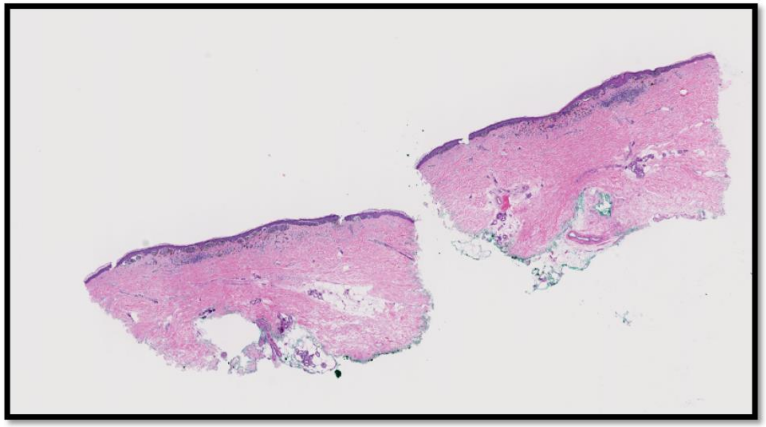


Dataset

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).







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

Dataset

- Each class varies in structure of tissues.
- Ground truth comes from 3 expert pathologists.
- For each WSI, we have one rectangular Region of Interest (ROI).

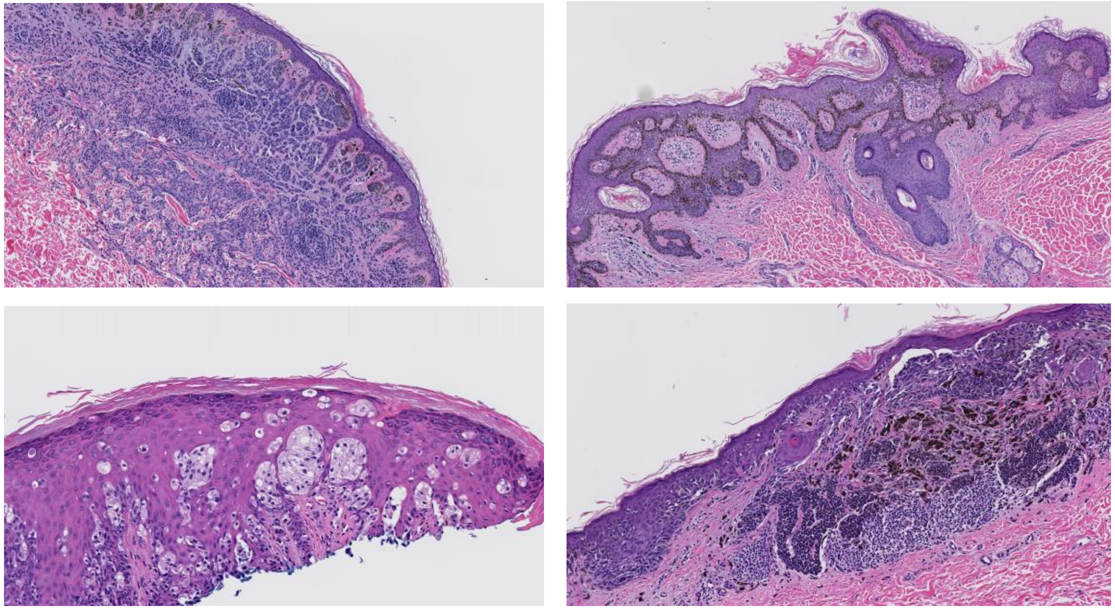
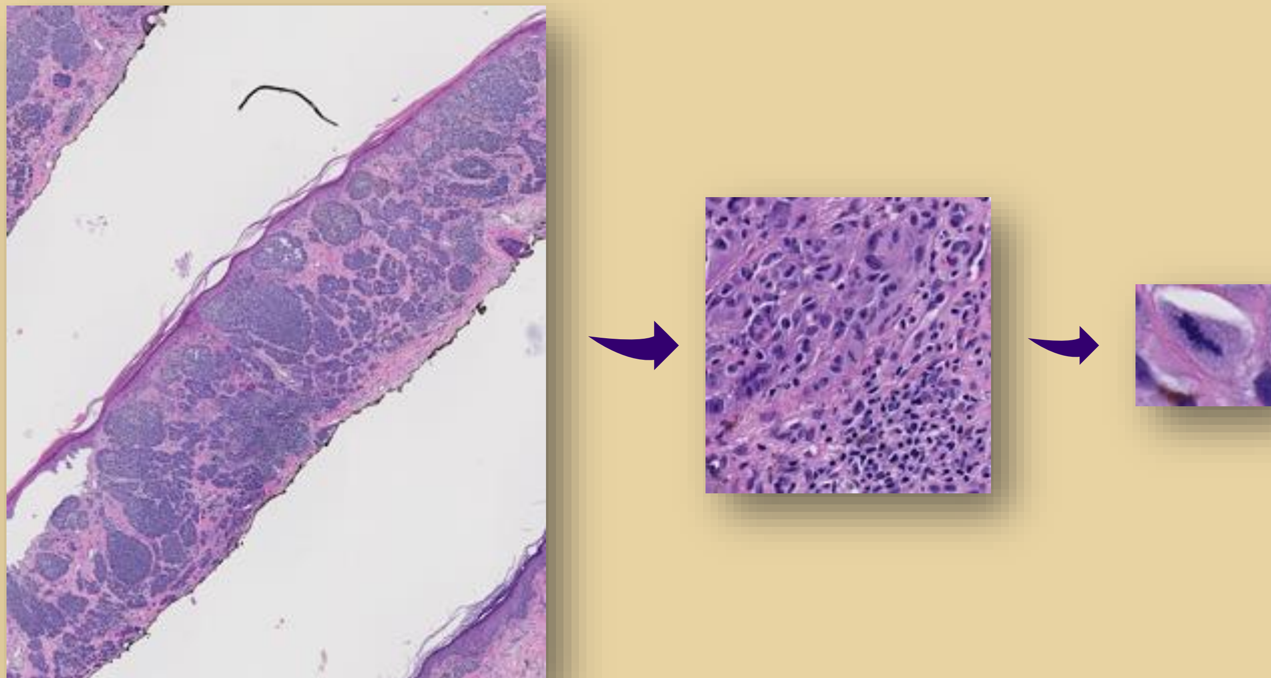


Figure3. Examples of Skin Biopsies with Different Diagnosis. Top Left: Benign.
Top Right: Atypia. Bottom Left: Melanoma in Situ. Bottom Right: Invasive Melanoma (T1a).

1. Mitosis Classification



Paper: Machine Learning Techniques for Mitoses Classification*

*Published in the journal of Computerized Medical Imaging and Graphics

https://doi.org/10.1016/j.compmedimag.2020.101832

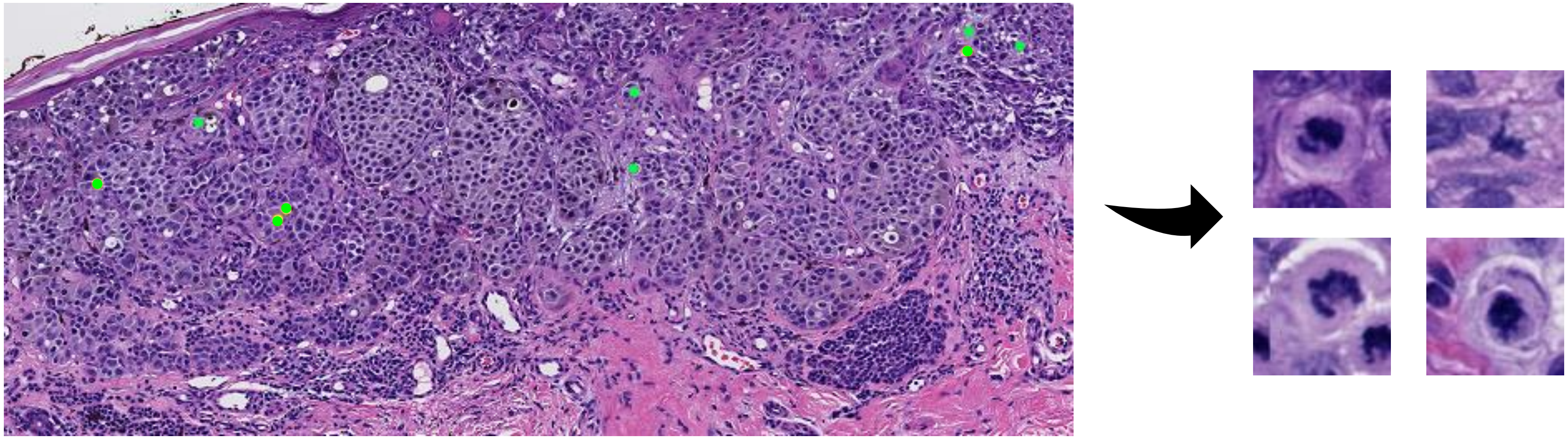
Related Work

- Among the independent predictors of melanoma-specific survival, mitotic rate is the strongest prognostic factor after tumor thickness [8].
- Various approaches have been applied to detect mitotic figures. Probability based methods [9], graph-based multi-resolution approach [10], used morphological features [11], and CNN-based method [12] are some of the approaches in detection of mitosis in biopsy images.
- Most of these methods have been applied on breast biopsy images.

Dataset

Positive samples – class Mitosis

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

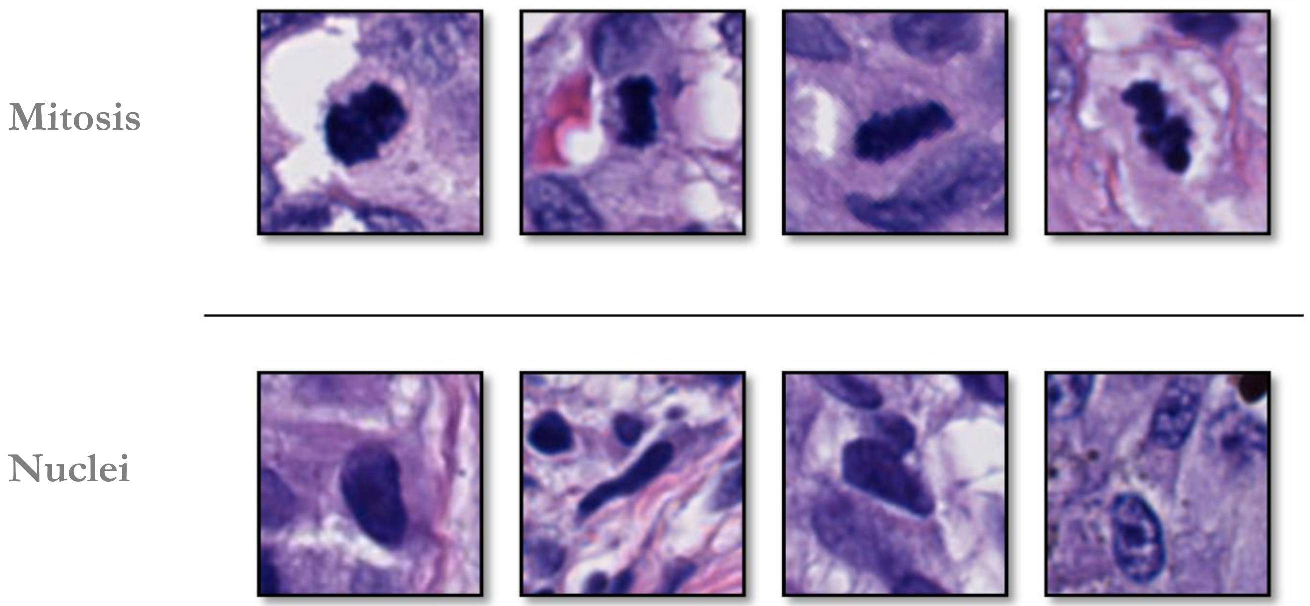


Example of expert pathologist markings of mitoses (Left) and sampled mitoses (Right)

Dataset

Negative samples – class NonMitosis

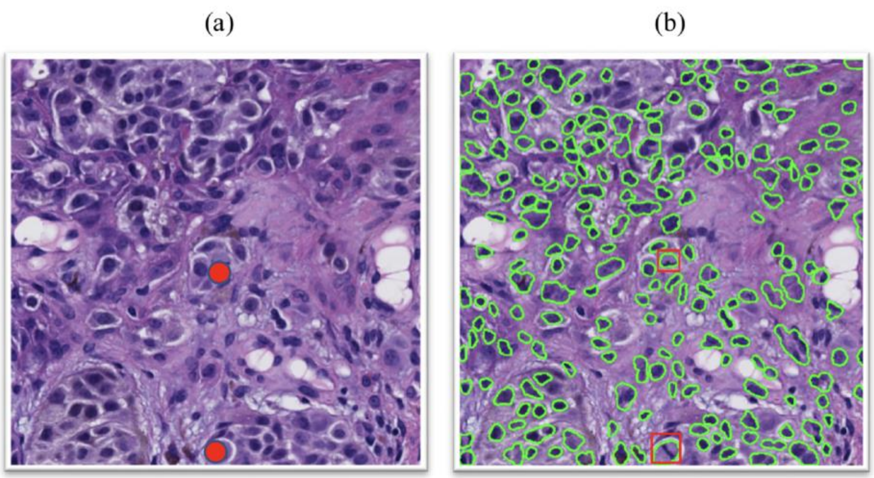
- Distinguishing mitoses from normal nuclei is a challenge.



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.

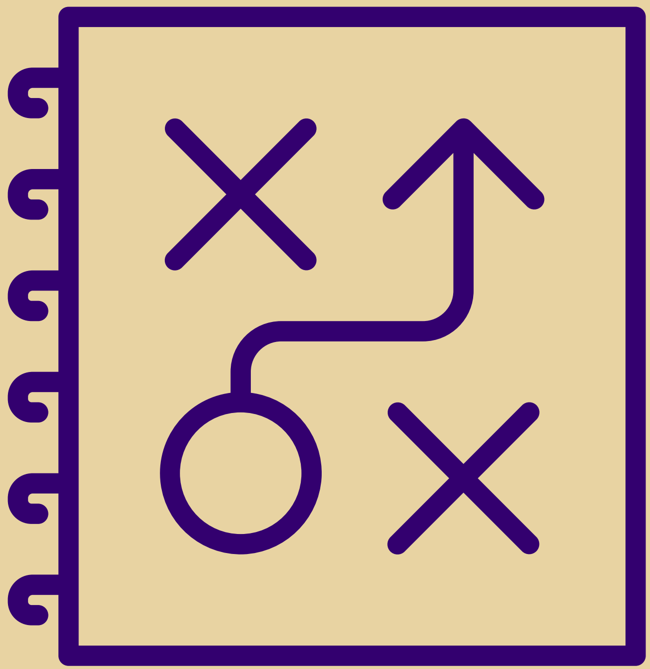
Preprocessing

> **Data augmentation:**
  - Rotations of 45, 90, 135 or 225 degrees.
  - Mirroring horizontal and vertical.

> **The final dataset:**
  - 4364 mitosis samples.
  - 12640 non-mitosis samples.

> **Dataset randomly split:**
  - Training: 60%
  - Validation: 20%
  - Testing: 20%

Method and Model

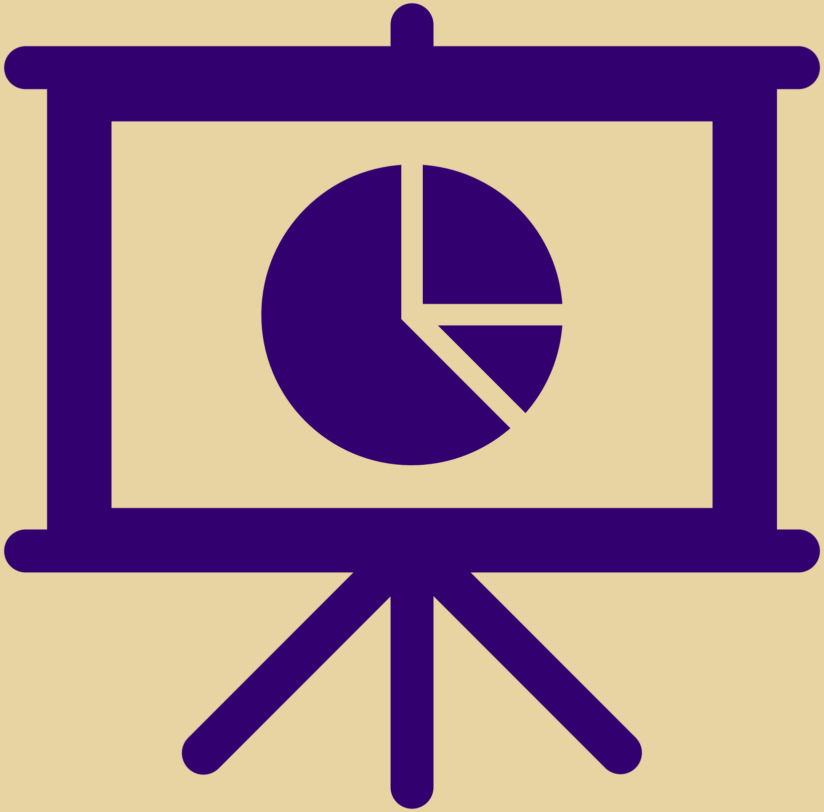


Method and Model

- In recent years, with the development of fast and accessible GPUs, Convolutional Neural Networks (CNNs) have dominated computer vision research due to their impressive performance, and mitosis detection is not an exception.

- 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



Method and Model

- Hyperparameters
  - Adam optimizers.
  - learning rate decay schedule with step size = 5 and $\boldsymbol{\gamma} = 0.1$.
  - 20 epochs.
  - cross-entropy loss function.

- Evaluation Metrics
  - Accuracy = (TP+TN)/(TP+FP+FN+TN)
  - Precision = TP / (TP + FP)
  - Recall = TP / (TP + FN)
  - $\text{F1 score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}}$
  - Sensitivity = TP / (TP + FN)
  - Specificity = TN / (TN + FP)

Results

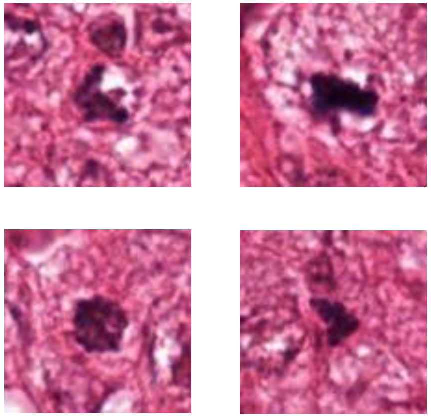
Evaluation results of ESPNet and DenseNet161 on **Melanoma**

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

Results

> ## Supplementary Dataset

- There is not a public marked dataset of mitoses in skin biopsies.
- Therefore, we used the MITOS dataset which is a public mitosis dataset of breast biopsies.
  - The MITOS dataset contains 50 images stained with Hematoxylin & Eosin.
  - A total of 800 mitoses are visible in MITOS



Examples of Mitoses in MITOS public dataset of breast biopsy images.

Results

- To generalize on a public dataset, we used MITOS dataset from ICPR12 challenge, which is a breast biopsy dataset for mitosis.

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

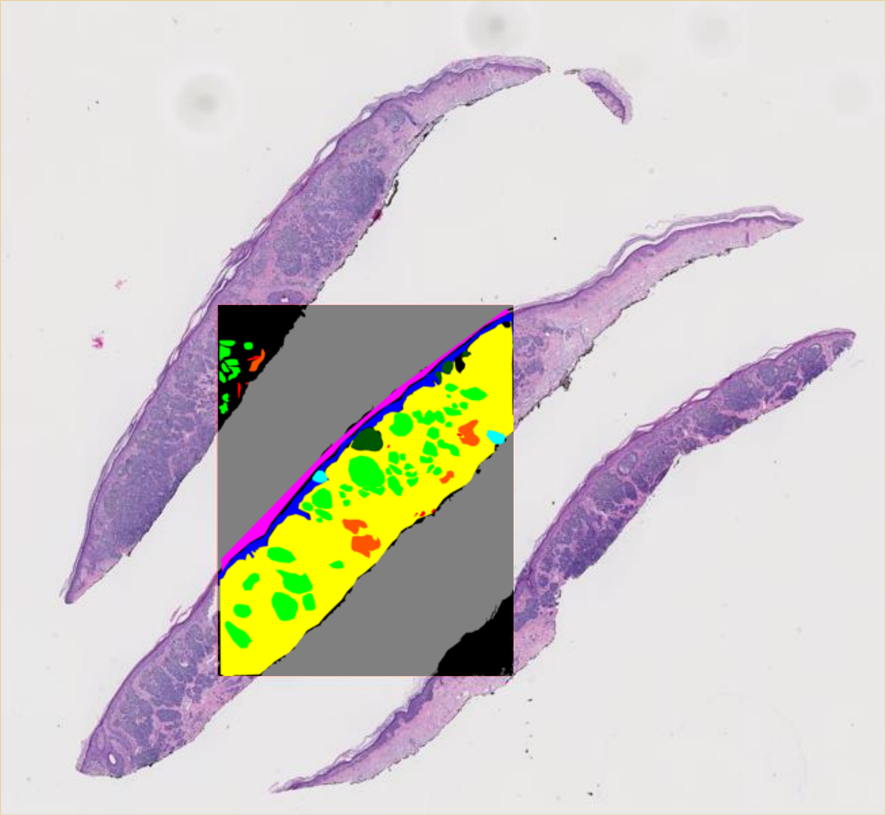
Method	**ESPNet** (Our trained model)	**DenseNet** (Our trained model)	[Saha et al., 2018]	[Dodballapur et al., 2019]	[Li et al., 2018]	[López-Tapia et al., 2019]	[Cireşan et al., 2013]
Precision	0.916	**0.939**	0.92	0.93	0.854	N/A	0.866
Recall	0.866	**0.916**	0.88	0.80	0.812	N/A	0.70
F1 Score	0.890	**0.927**	0.90	0.87	0.832	0.826	0.782

Discussion

Discussion

- We achieved very high scores in both of our experiments.
- **Melanoma:**
  - DenseNet161 performed slightly better than ESPNet.
  - Training time of ESPNet is significantly less than Densenet161.
- **MITOS:**
  - Both of ESPNet and DenseNet performed significantly better than the classifier of ICPR12 winner with significance level of 0.01.
  - DenseNet161 is significantly better than the classifier of the out-performer at significance level of 0.05.
  - ESPNet is not significantly better than the out-performer.

2. Segmentation using Coarse and Sparse Annotation



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

*Submitted to the Journal of Digital Imaging

Related Work

Various approaches have been developed to overcome imperfect and limited data annotation and vary with the specific challenges posed by the specific dataset on which they were developed.

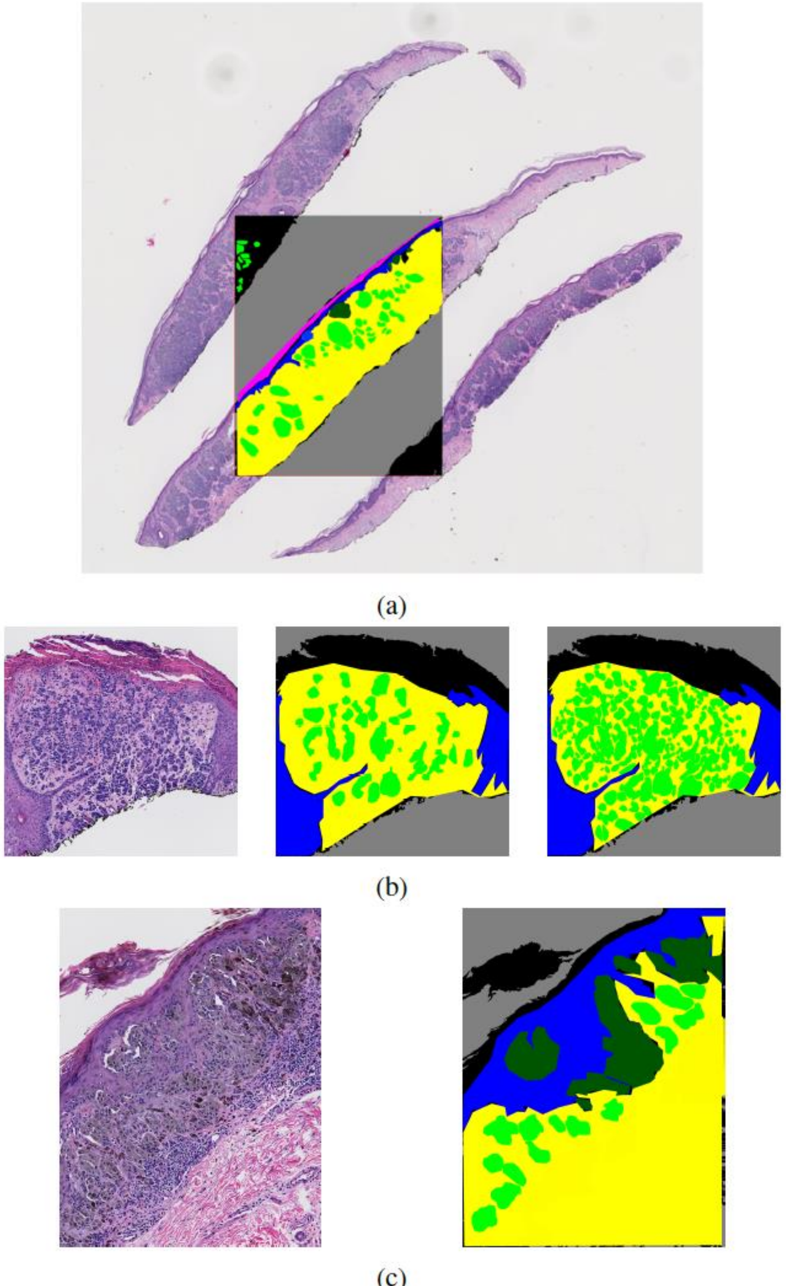
- When a small portion of an image is fully annotated, different methods of augmentation have proven to be helpful [15,16].
- Active learning is another popular method in the case of limited annotation [17,18,19].
- Changing loss function, or using external data also generated promising segmentation result.

Dataset

- We obtained coarse and sparse annotations only on the ROI images by an expert pathologist (Dr. 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.

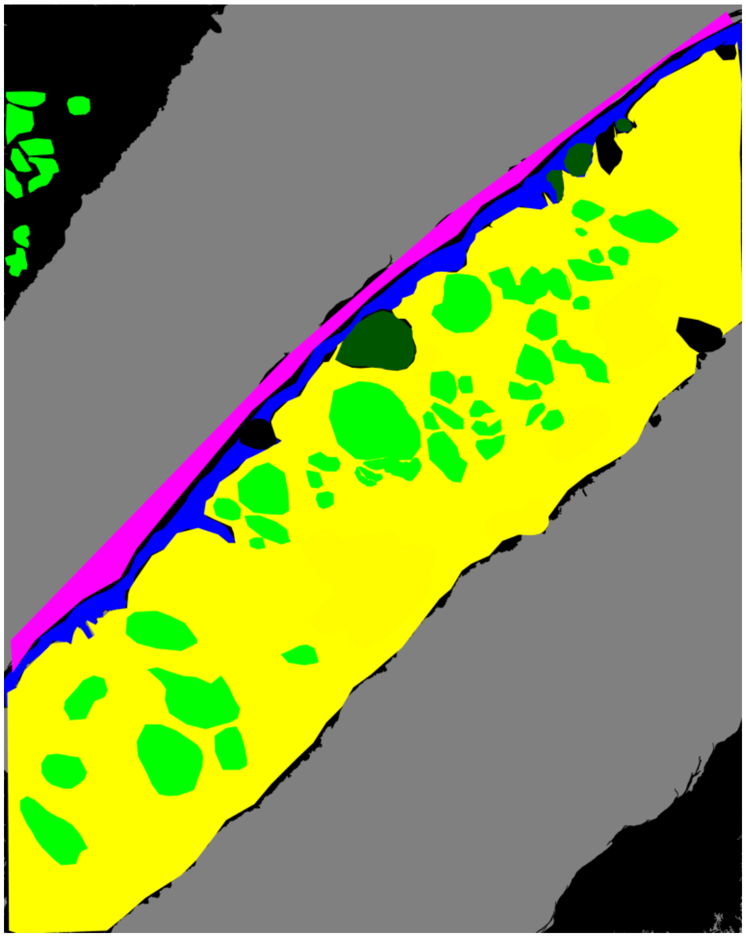


(a)

(b)

(c)

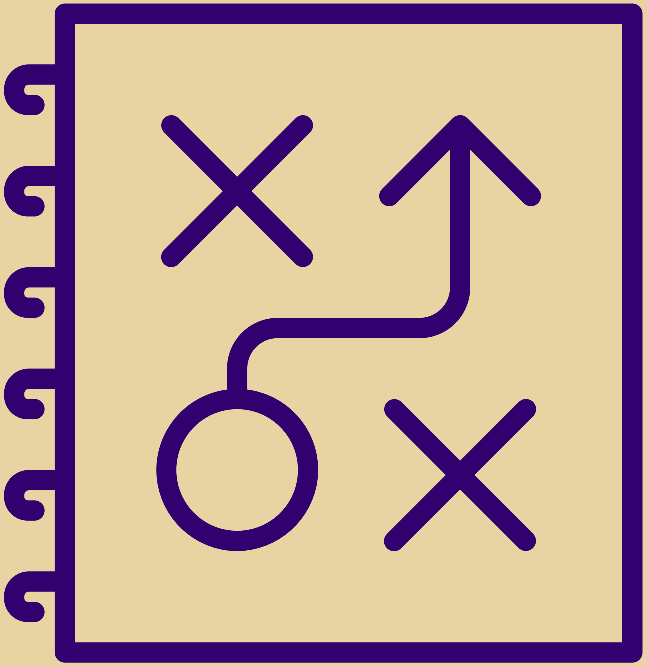
Dataset



Labels and colors:

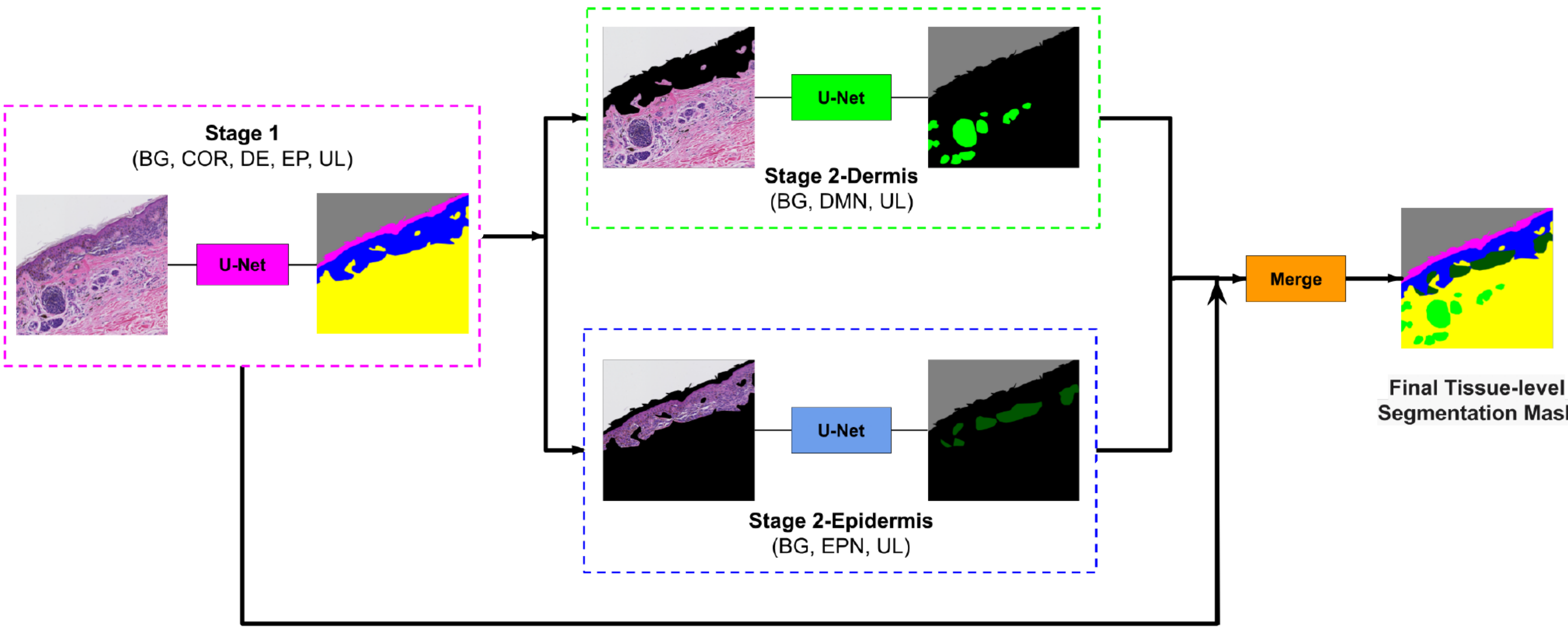
- Corneum (COR)
- Epidermis (EP)
- Epidermal Nests (EPN)
- Dermis (DE)
- Dermal Nests (DMN)
- Background (BG)
- Unlabeled (UL)

Method and Model



Method and Model

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



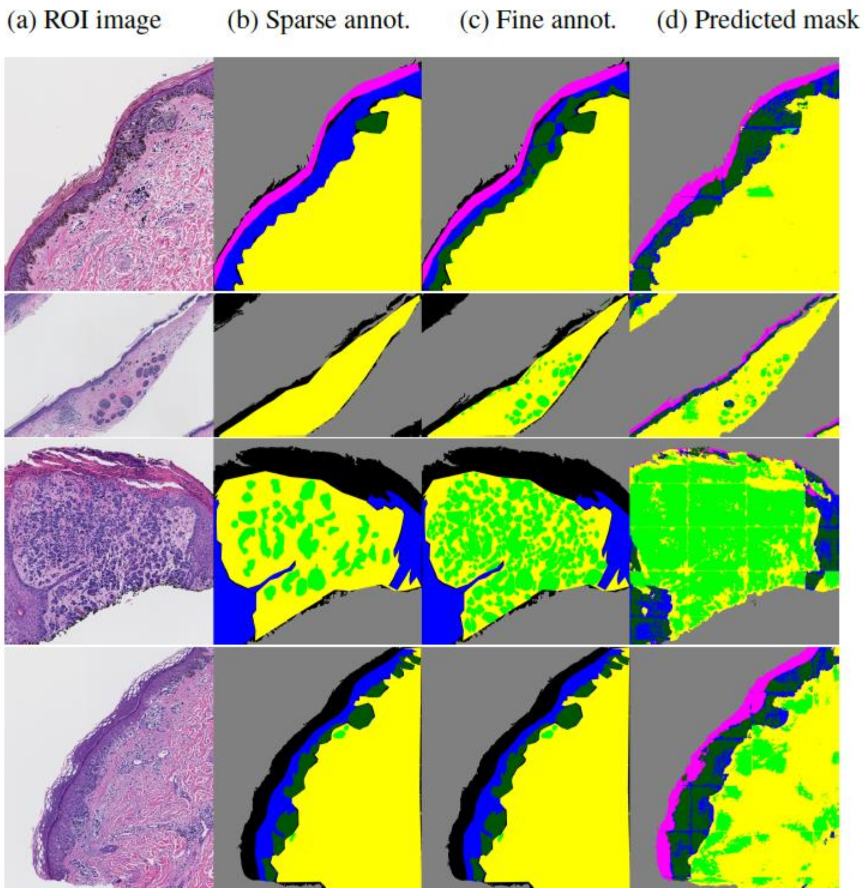
Results

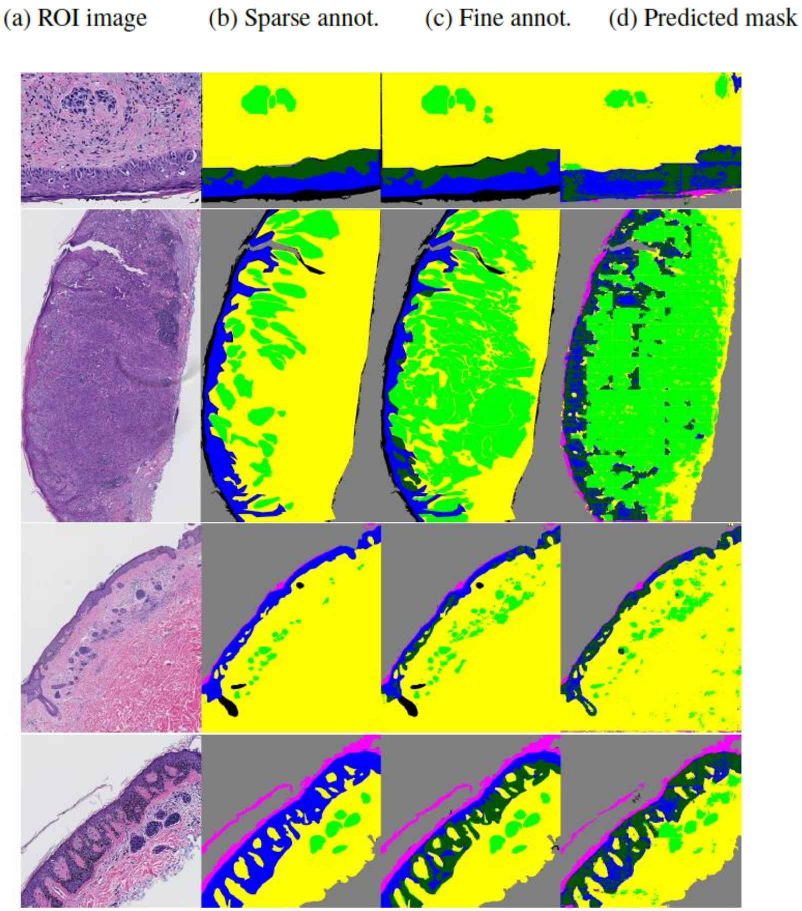
Results

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

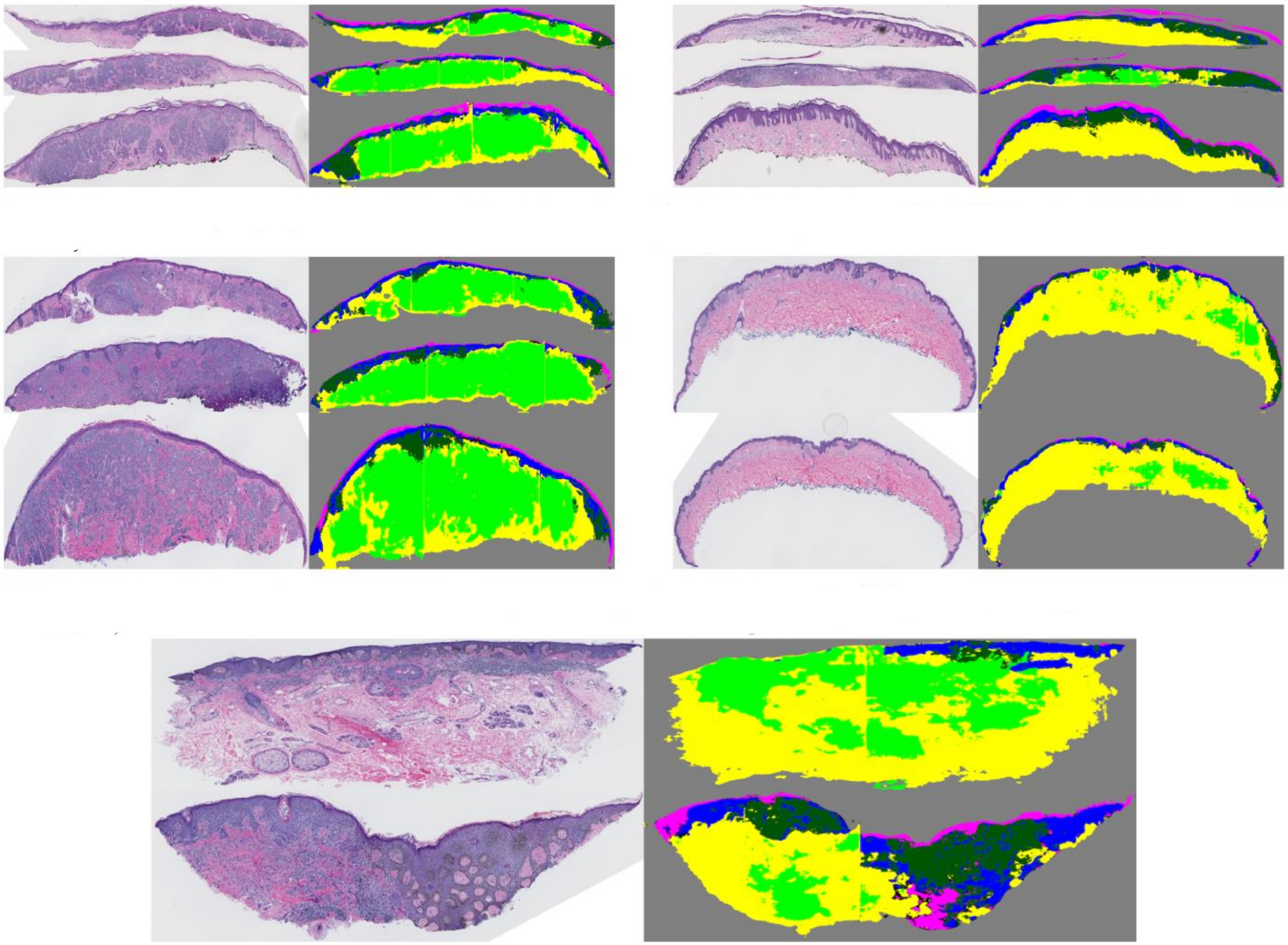
Segmentation stage	Dice score	IoU
Stage 1 (all tissues)	0.942	0.906
Stage 2-Dermis (DMN)	0.558	0.638
Stage 2-Epidermis (EPN)	0.332	0.558

Results - ROI testing set





Results - Generating WSI Segmentation Masks

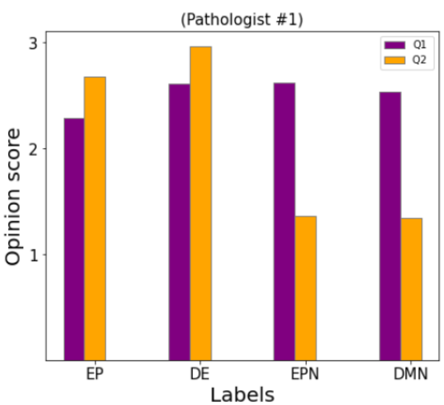


Subjective Assessment with Pathologists

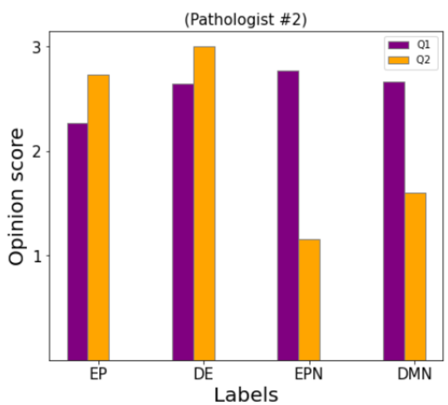
Qualitatively evaluation the WSI segmentation, with these questions:

- **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.

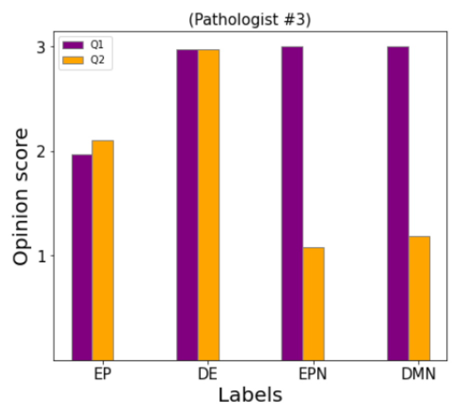
- **Q2:** How much of the label identified by the model is the correct tissue/area? Rate Low, Medium, or High.



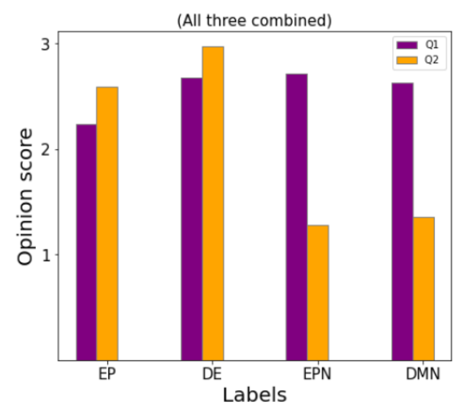
(a) Individual survey #1



(b) Individual survey #2



(c) Individual survey #3



(d) Combination of all three

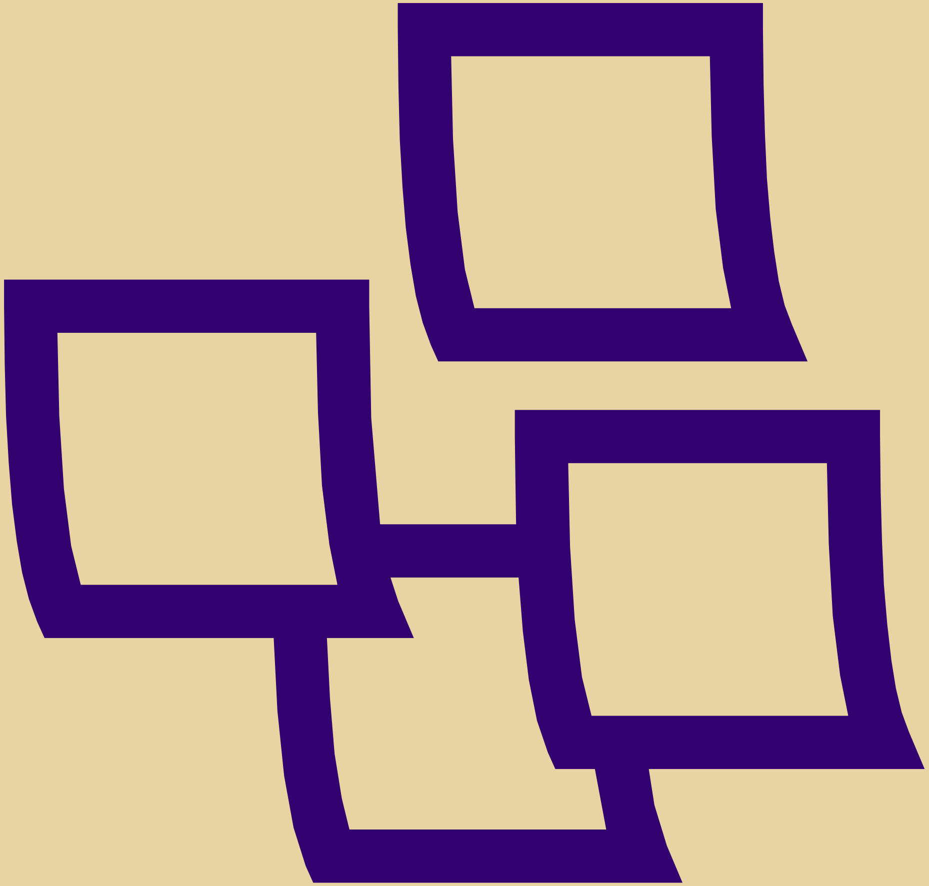
Discussion

ELECTRICAL ENGINEERING
UNIVERSITY *of* WASHINGTON

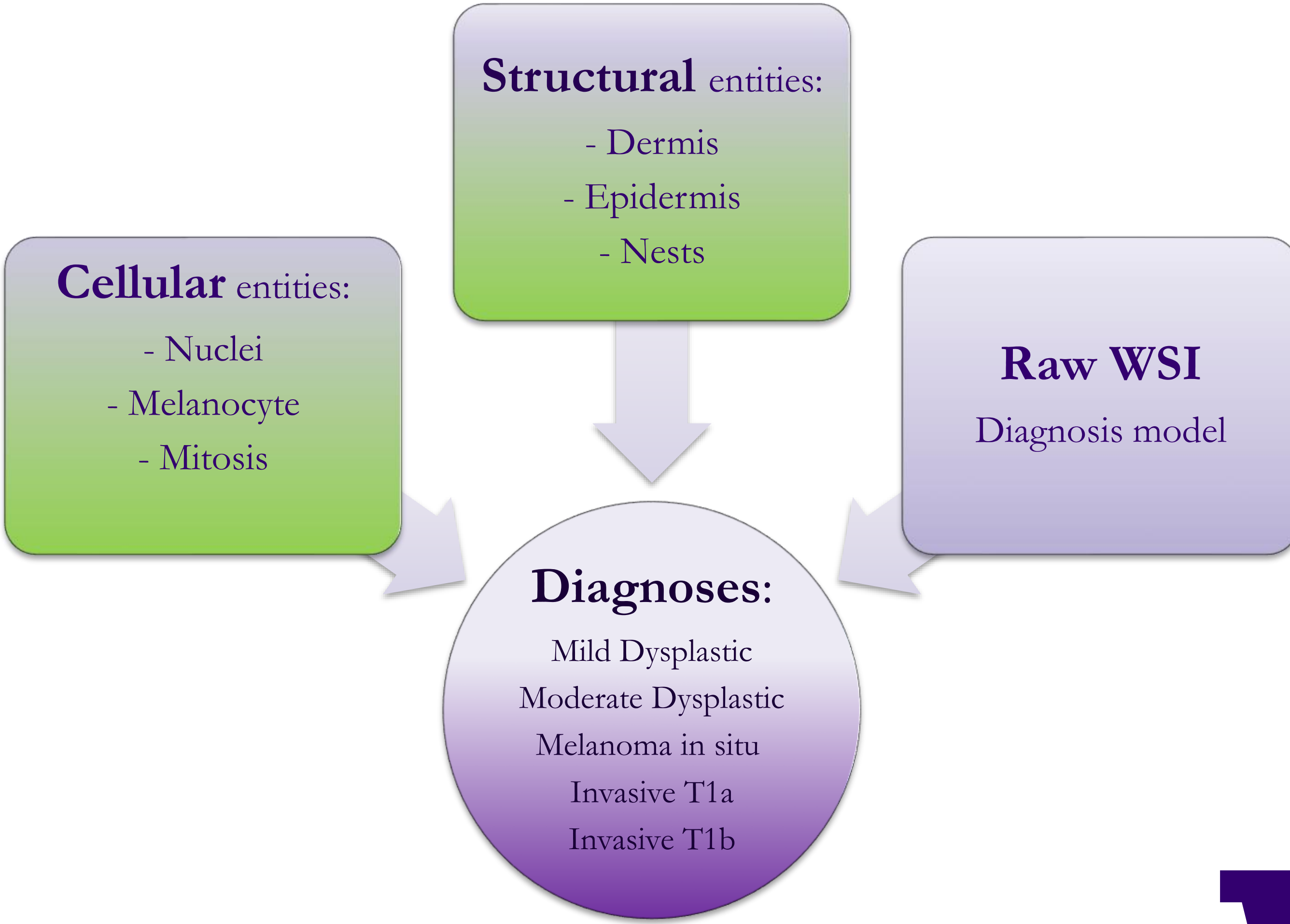
Discussion

- Training a segmentation model generally requires a large, high-quality annotated ground-truth. However, medical datasets require expert-level annotation as ground-truth.
- Our system was able to generate segmentation masks for both epidermis/dermis and nests with high-quality performance, indicating that having **sparse** annotation on important tissues has the potential for producing a useful segmentation model.
- Our results suggest that both the DMN and EPN can be over-labeled by the model, highlighting the problems that **coarse** annotation can cause for the system, especially on a small dataset in which the ground-truth did not clearly distinguish.
- **Sparse, but fine**, annotation on a small region of the WSI may be enough for training a better segmentation model.

Future Work



Melanoma Diagnosis



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.

Advisor:
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
Beibin Li
Kechun Liu

Thank you for your attention.

ELECTRICAL ENGINEERING
UNIVERSITY *of* WASHINGTON