Outline

> Background and Goal
> Dataset
> Related Work
> Our Work
  – HATNet
  – ScAtNet
> Next Step
Background
What is Melanoma?

- Melanoma is the most aggressive type of skin cancer.
- Melanoma occurs when UV radiation triggers DNA damages in the melanocytes.
- The “gold standard” for diagnosis of invasive melanoma relies on the visual assessments of skin biopsy images by pathologists.

An example of an Invasive Melanoma T1b in M-Path dataset.
Why melanoma diagnosis?

> Unfortunately, diagnostic errors are common
> Computer-aided diagnostic system can be a second reader and help reduce uncertainties
Goal

Model

Diagnosis

MMD
MIS
pT1a
pT1b
Dataset
## Melanoma Dataset

<table>
<thead>
<tr>
<th>Diagnostic</th>
<th>Training</th>
<th>Validation</th>
<th>Test</th>
<th>Total</th>
<th>Average WSI size (in pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMD</td>
<td>26</td>
<td>6</td>
<td>29</td>
<td>61</td>
<td>11843 × 10315</td>
</tr>
<tr>
<td>MIS</td>
<td>25</td>
<td>5</td>
<td>30</td>
<td>60</td>
<td>9133 × 8501</td>
</tr>
<tr>
<td>pT1a</td>
<td>33</td>
<td>6</td>
<td>34</td>
<td>73</td>
<td>9490 × 7984</td>
</tr>
<tr>
<td>pT1b</td>
<td>18</td>
<td>6</td>
<td>22</td>
<td>46</td>
<td>14858 × 12154</td>
</tr>
<tr>
<td>Total</td>
<td>102</td>
<td>23</td>
<td>115</td>
<td>240</td>
<td>11130 × 9603</td>
</tr>
</tbody>
</table>
Difficulties in diagnosis

Size of whole slide images

An example image from ImageNet [500 x 375]

An example WSI at 10x [15264 x 19824]
Difficulties in diagnosis

Size of whole slide images

Dataset size

TABLE 1: Statistics of skin biopsy whole slide image (WSI) dataset. The average WSI size is computed at a magnification factor of x10. Diagnostic terms for the dataset used in this study are as follows: mild and moderate dysplastic nevi (MMD), melanoma in situ (MIS), invasive melanoma stage pT1a (pT1a), invasive melanoma stage ≥ pT1b (pT1b).
Difficulties in diagnosis

Size of whole slide images

Dataset size

cancerous structure vs. normal structure
Related Work
Related Work

> Multiple Instance Learning

Negative Bag  Positive Bag
Related Work

> Multiple Instance Learning
Related Work

> **Multiple Instance Learning**
  + reduce high computational cost
  + effective in learning instance/bag-wise representation
  - Does not allow long-range/global feature interaction
  - Prone to label ambiguity/noise
Related Work

> Segmentation-based methods

Related Work

> Segmentation-based methods
  + Learns global representation
  + More effective (better performance) on small dataset
  - Require fine tissue-level segmentation masks
  - Diagnostic performance highly dependent on segmentation quality
Related Work

> Visual Transformers
Difficulties in diagnosis

Size of whole slide images

Dataset size

cancerous structure vs. normal structure

Multiple Instance Learning

Visual Transformer

Segmentation-based Methods
HATNet

Image

Bags

Words

CNN

\( \mathbf{I} \in \mathbb{R}^d \)

Bag-to-image attention

Sec. 4.4

Classifier

Benign

Atypia

DCIS

Invasive

\( \mathbf{B}_{i,j} \)

\( n \)

\( d \)

\( \mathbf{B}_{i,j} \in \mathbb{R}^d \)

Word-to-bag attention

Sec. 4.2

\( \mathbf{B}_{w2b} \)

\( m \)

\( n \)

\( d \)

\( \mathbf{B}_{w2w} \)

\( \mathbf{B}_{i,\text{con}} \)

\( \mathbf{B}_{i} \)

\( \mathbf{B}_{w2b} \)

\( \mathbf{B}_{w2w} \)

\( \mathbf{B}_{b2b} \)

Multi-head attention

Feed forward network (FFN)

\( \mathbf{B}^i_{\text{con}} \)

\( \mathbf{B}^i_{w2w} \)

\( \mathbf{B}^i_{w2b} \)

Multi-head attention

Function \( \psi \)

Linear

Softmax

Dot-product

Word-to-bag attention

Word-to-word attention
HATNet (on a breast dataset)

- Outperforms CNN-based methods by a large margin
- Significant overlap between top bags, words and annotations of clinical biomarkers
- Learned representations from clinically relevant tissue structures without any supervision
ScAtNet
ScAtNet: Soft Label

Invasive T1a Skin Biopsy Image (or Class 3)

<table>
<thead>
<tr>
<th>TS 1 (w/o ROI)</th>
<th>TS 2 (w/ ROI)</th>
<th>TS 3 (w/o ROI)</th>
</tr>
</thead>
</table>

| 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 |
ScAtNet

- Outperforms MIL and CNN based methods
- Achieves comparable performance to 187 practicing U.S. pathologists
- Saliency analysis shows that ScAtNet learns to weigh features from different scales

<table>
<thead>
<tr>
<th>Input scales</th>
<th>Accuracy</th>
<th>F1</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5x</td>
<td>0.55</td>
<td>0.55</td>
<td>0.55</td>
<td>0.85</td>
<td>0.75</td>
</tr>
<tr>
<td>10x</td>
<td>0.60</td>
<td>0.60</td>
<td>0.60</td>
<td>0.87</td>
<td>0.77</td>
</tr>
<tr>
<td>12.5x</td>
<td>0.61</td>
<td>0.61</td>
<td>0.61</td>
<td>0.87</td>
<td>0.78</td>
</tr>
<tr>
<td>7.5x 10x</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.88</td>
<td>0.79</td>
</tr>
<tr>
<td>7.5x 12.5x</td>
<td>0.64</td>
<td>0.64</td>
<td>0.64</td>
<td>0.88</td>
<td>0.80</td>
</tr>
<tr>
<td>10x 12.5x</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.88</td>
<td>0.79</td>
</tr>
<tr>
<td>7.5x 10x 12.5x</td>
<td>0.63</td>
<td>0.63</td>
<td>0.63</td>
<td>0.88</td>
<td>0.79</td>
</tr>
</tbody>
</table>

(a) Overall performance of ScAtNet
Semantic Segmentation-based Method
How do we combine everything?

Diagnoses:
- Mild Dysplastic
- Moderate Dysplastic
- Melanoma in situ
- Invasive T1a
- Invasive T1b
Acknowledgement

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
Dr. Donald Weaver

Collaborators:
Shima Nofallah
Ximing Lu
Dr. Sachin Mehta


