Pathologists' viewing behaviors contribute to diagnostic accuracy of melanocytic skin lesions

Fatemeh Ghezloo

February 28, 2022
Outline

- Introduction
- Material and Methods
- Results
- Application
- Conclusions
Skin cancer is the most commonly diagnosed cancer in the United States!

1 in 5 Americans will develop skin cancer by the age of 70.¹ More than 2 people die of skin cancer in the U.S. every hour.¹

¹. https://www.skincancer.org/skin-cancer-information/skin-cancer-facts/
Where do skin cancers start?

2. https://smart.servier.com
Melanoma diagnosis: Traditional Microscopy to Whole Slide Imaging
Pathologists’ Viewing behavior: Why is it important?

➢ More cases are being interpreted in digital format
➢ Computer-based technologies are being developed for diagnosis

However,
There is discordance among pathologists' diagnoses even when they observe similar features on a biopsy sample slide. Errors made in cancer diagnosis are one of the main reasons that cause death.

Diagnostic accuracy varies due to the interaction among
➢ Case characteristics
➢ Pathologist characteristics
➢ Visual search process

Pathologists’ Viewing behavior:
Study workflow

- Tracking Devices
- Interpretation session
- Data Analysis

➢ Eye fixation
➢ Saccade
➢ Timestamp
➢ Zoom level
➢ Mouse click
➢ Screen viewport
Study Goals

- We outline the types of data points that can be gathered to describe pathologists’ viewing behavior using viewport data.
- How do specific viewing behaviors contribute to diagnostic accuracy?
- How are pathologists’ characteristics associated with specific viewing patterns?
Outline

➢ Introduction
➢ Material and Methods
➢ Results
➢ Application
➢ Conclusions
M-PATH data

<table>
<thead>
<tr>
<th>Class</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class 1</td>
<td>Nevus</td>
</tr>
<tr>
<td>Class 2</td>
<td>Moderate atypia</td>
</tr>
<tr>
<td>class 3</td>
<td>Melanoma in situ</td>
</tr>
<tr>
<td>Class 4</td>
<td>Stage pT1a invasive melanoma</td>
</tr>
<tr>
<td>Class 5</td>
<td>Stage pT1b invasive melanoma</td>
</tr>
</tbody>
</table>
Data collection

Digital skin biopsy images
3 Experts in dermatopathology
Consensus ROI and diagnosis

32 Pathologists
180 Cases
Divided into 5 groups of 36
Survey
Interpretation
Viewport log
Diagnosis
## Pathologists’ characteristics

<table>
<thead>
<tr>
<th>Pathologist Characteristics</th>
<th>Number (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>13 (40.6%)</td>
</tr>
<tr>
<td>Female</td>
<td>19 (59.4%)</td>
</tr>
<tr>
<td><strong>Age (years)</strong></td>
<td></td>
</tr>
<tr>
<td>20 – 49</td>
<td>12 (37.5%)</td>
</tr>
<tr>
<td>50 – 64</td>
<td>20 (62.5%)</td>
</tr>
<tr>
<td><strong>Board certification/ Dermatopathology Fellowship training</strong></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>10 (31.3%)</td>
</tr>
<tr>
<td>No</td>
<td>22 (68.8%)</td>
</tr>
<tr>
<td><strong>Experience with interpreting melanocytic skin lesions (years)</strong></td>
<td></td>
</tr>
<tr>
<td>&lt; 5</td>
<td>3 (9.4%)</td>
</tr>
<tr>
<td>5 – 9</td>
<td>10 (31.3%)</td>
</tr>
<tr>
<td>10 – 19</td>
<td>9 (28.1%)</td>
</tr>
<tr>
<td>&gt; 20</td>
<td>10 (31.3%)</td>
</tr>
<tr>
<td><strong>Caseload of melanocytic skin lesions (%)</strong></td>
<td></td>
</tr>
<tr>
<td>&lt; 10</td>
<td>14 (43.8%)</td>
</tr>
<tr>
<td>10 – 24</td>
<td>13 (40.6%)</td>
</tr>
<tr>
<td>25 – 49</td>
<td>5 (15.6%)</td>
</tr>
</tbody>
</table>
Viewport data

<table>
<thead>
<tr>
<th>ID</th>
<th>X</th>
<th>Y</th>
<th>Width</th>
<th>Height</th>
<th>Zoom</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
<td>6</td>
<td>50</td>
<td>100</td>
<td>2</td>
<td>1:20:30</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>6</td>
<td>50</td>
<td>100</td>
<td>2</td>
<td>1:20:33</td>
</tr>
<tr>
<td>3</td>
<td>20</td>
<td>13</td>
<td>50</td>
<td>100</td>
<td>2</td>
<td>1:20:34</td>
</tr>
<tr>
<td>4</td>
<td>23</td>
<td>15</td>
<td>30</td>
<td>80</td>
<td>4</td>
<td>1:20:38</td>
</tr>
<tr>
<td>5</td>
<td>23</td>
<td>4</td>
<td>25</td>
<td>70</td>
<td>5</td>
<td>1:21:10</td>
</tr>
<tr>
<td>6</td>
<td>40</td>
<td>4</td>
<td>25</td>
<td>70</td>
<td>5</td>
<td>1:21:50</td>
</tr>
<tr>
<td>7</td>
<td>44</td>
<td>4</td>
<td>25</td>
<td>70</td>
<td>5</td>
<td>1:22:30</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>10</td>
<td>60</td>
<td>110</td>
<td>2</td>
<td>1:23:40</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>11</td>
<td>60</td>
<td>110</td>
<td>2</td>
<td>1:24:10</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>12</td>
<td>60</td>
<td>110</td>
<td>2</td>
<td>1:24:56</td>
</tr>
</tbody>
</table>
Viewing behaviors

- Total interpretation time: Duration of Interpretation
- Average zoom level: Average of zoom levels used during an interpretation
- Maximum zoom level: Maximum of zoom levels used during an interpretation
- Zoom level variance: Variance of zoom levels used during an interpretation
- Magnification: Percentage of viewports associated with consecutive zooming in
- ROI time Percentage: Percentage of time spent viewing consensus ROI to the total time
- Scanning Percentage: Percentage of viewports associated with fixed zoom level and panning
Statistical analysis

- Crossed-level structure of cases and pathologists
- Both case and pathologist contribute to the variation of the outcome

Cross-classified multilevel model

\[ y_i = \beta_0 + \beta_1 x_i + u_{\text{pathologist}(i)} + u_{\text{case}(i)} + e_i \]

Outline

➢ Introduction
➢ Material and Methods
➢ Results
➢ Application
➢ Conclusions
Viewing behaviors visualization
How do specific viewing behaviors contribute to diagnostic accuracy?

- Diagnostic accuracy is the agreement of a pathologist's diagnosis with the consensus diagnosis
- Separate models with each viewing behavior as the predictor variable
- Diagnostic accuracy is the outcome
- All models are adjusted for pathologists’ experience level and fellowship training

<table>
<thead>
<tr>
<th>Predictor Variable</th>
<th>OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total time</td>
<td>1.33 (1.09, 1.62)</td>
<td>0.005</td>
</tr>
<tr>
<td>Average zoom</td>
<td>1.26 (1.03, 1.54)</td>
<td>0.023</td>
</tr>
<tr>
<td>Maximum zoom</td>
<td>1.24 (1.03, 1.50)</td>
<td>0.026</td>
</tr>
<tr>
<td>Zoom variance</td>
<td>1.37 (1.11, 1.68)</td>
<td>0.003</td>
</tr>
<tr>
<td>Magnification</td>
<td>0.76 (0.63, 0.92)</td>
<td>0.006</td>
</tr>
<tr>
<td>ROI time percentage</td>
<td>1.35 (1.07, 1.69)</td>
<td>0.011</td>
</tr>
<tr>
<td>Scanning percentage</td>
<td>1.21 (1.00, 1.47)</td>
<td>0.054</td>
</tr>
</tbody>
</table>
Statistical analysis:
Diagnostic accuracy - Multivariate model

Viewing Behaviors

<table>
<thead>
<tr>
<th>Predictor Variables</th>
<th>OR (95% CI)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total time</td>
<td>1.25 (1.01, 1.54)</td>
<td>0.0360</td>
</tr>
<tr>
<td>Zoom variance</td>
<td>1.22 (0.98, 1.53)</td>
<td>0.0786</td>
</tr>
<tr>
<td>ROI time percentage</td>
<td>1.38 (1.10, 1.73)</td>
<td>0.0058</td>
</tr>
<tr>
<td>Scanning percentage</td>
<td>1.20 (0.98, 1.47)</td>
<td>0.0716</td>
</tr>
</tbody>
</table>

✓ Diagnostic accuracy is the outcome
✓ Model is adjusted for pathologists’ experience level and fellowship training
How are pathologists’ characteristics associated with specific viewing patterns?

- Separate models with each characteristic as the predictor variable
- Pathologists’ viewing behaviors as the outcome

<table>
<thead>
<tr>
<th>Pathologist’s characteristic, clinical experience and ratings of difficulty and confidence on melanocytic skin lesions</th>
<th>Average zoom</th>
<th>Maximum zoom</th>
<th>Zoom variance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Contrast</td>
<td>P-value</td>
<td>Contrast</td>
</tr>
<tr>
<td>Clinical characteristics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>0.03</td>
<td>0.878</td>
<td>0.05</td>
</tr>
<tr>
<td>Age</td>
<td>0.29</td>
<td>0.192</td>
<td>0.39</td>
</tr>
<tr>
<td>Clinical Experience Level</td>
<td>-0.62</td>
<td>0.003</td>
<td>-0.45</td>
</tr>
<tr>
<td>Board certification/ Fellowship training</td>
<td>-0.35</td>
<td>0.015</td>
<td>-0.29</td>
</tr>
<tr>
<td>Experience with melanocytic skin lesions</td>
<td>0.12</td>
<td>0.286</td>
<td>0.12</td>
</tr>
<tr>
<td>Caseload of melanocytic skin lesions</td>
<td>-0.05</td>
<td>0.765</td>
<td>-0.03</td>
</tr>
<tr>
<td>Ratings on melanocytic skin lesions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulty level</td>
<td>-0.05</td>
<td>0.765</td>
<td>-0.03</td>
</tr>
<tr>
<td>Confidence level</td>
<td>0.21</td>
<td>0.044</td>
<td>0.19</td>
</tr>
</tbody>
</table>
Discussion

- How do specific viewing behaviors and search patterns contribute to diagnostic accuracy?
  - More total time, zoom variables, and ROI time, and scanning percentage
  - Less magnification
  - Higher accuracy

- How are pathologists’ characteristics associated with specific viewing patterns?
  - Higher age category
  - Lower confidence level
  - Higher zoom variance
  - Higher average zoom, maximum zoom, and zoom variance

- How do viewing patterns change as pathologists gain more expertise in diagnosing melanocytic lesions?
  - Having Board certification and/or fellowship training
  - Higher caseload of melanocytic skin lesions
  - Lower average zoom, maximum zoom, and zoom variance
Outline

- Introduction
- Material and Methods
- Results
- Application
- Conclusions
Problem: Predicting and Localizing diagnostically relevant Regions of Interest (ROIs)

1. Extracting viewing ROIs
2. Bag of Words Visual model
3. Binary classification

WSI To Image patches
Feature Extraction from each patch
Clustering patches Using K-means

ROI localization
Bag of Words

Whole slide image with a red Sliding window (3600 x 3600)

Image Patches (120 * 120)

Clusters of Patches

L*a*b color and LBP texture histograms

Extracting viewing ROIs

Bag of Words Visual model

Binary classification

WSI To Image patches

Feature Extraction from each patch

Clustering patches Using K-means
ROI localization
Training classifier

- Binary classification (relevant vs non-relevant)
  - Logistic regression
  - SVM

Whole slide image with ROI selected by expert pathologists marked in red
ROI localization
Extracting Viewing ROIs

➢ Use the viewing data from pathologists to extract regions
➢ Use these regions as ground truth for testing the classifier

1. **Zoom peaks** are the log entries where the zoom level is higher than the previous and the next entries. A zoom peak identifies a region where the pathologist intentionally zoomed to look at a higher magnification.

2. **Slow pannings** are the log entries where the zoom level is the same as the previous entry, and the displacement is small. Slow pannings are intended for investigating a slightly larger and closer area without completely moving the viewport.

3. **Fixations** are the log entries where the duration is longer than 2 seconds. Fixations identify the areas to which a pathologist focused extra attention by looking at them longer.
ROI localization - Results

Ground Truth Viewing ROIs

Logistic Regression

SVM
Outline

- Introduction
- Material and Methods
- Results
- Application
- Conclusions
Limitations

High quality digital slide preparation is costly.

Ethical issues must be resolved when sharing patient information across larger platforms.

Experimental sample sizes in pathology studies are often small.

A thorough understanding of a variety of statistical principles is required.

Access to a multidisciplinary team of professionals, including statisticians and pathologists.
Conclusions

- Diagnosis of pathology slides is a complex task and requires years of training
- It is essential to study pathologists’ viewing behaviors
- Digital pathology has made it possible to record and study these viewing behaviors
- We showed various ways of quantifying these behaviors
- We investigated the association of these viewing behaviors with accuracy
- The results of such behavioral studies can be beneficial in various areas
  - Improving the training and education of younger pathologists
  - Determining the reasons for diagnostic errors
  - Assisting with the development of computer-aided tools for diagnosis purposes