CellMap: Interactive Tissue Analysis

Yidi Ed Liu
u1267187@utah.edu
u1267187

Sukrut Shishupal
sukrut.shishupal@utah.edu
u1401898

Repository link: https://github.com/sukrut-shishupal/Vis_for_data_science

1. Background and Motivation

The intersection of data visualization and medical imaging presents unique challenges and opportunities that our team is eager to explore. With advancing digital pathology, there is a need for tools that can enhance the interpretability of complex biomedical datasets for clinical and research purposes. We are drawn towards learning the power of technology to solve real-world problems, particularly data visualization to transform how medical data is analyzed and understood.

Histopathological analysis used to rely heavily on manual inspection, which is time-consuming and subject to variability based on multiple observations. We aim to create a visualization tool that assists in the objective analysis of tissue samples and enhance diagnostic accuracy. Our project aims to improve the visualization tools that can lead to better clinical outcomes by enabling precise and faster diagnosis.

Through this project, we leverage our background in computer science and improve our skills in D3.js to contribute to the field of digital pathology by creating a tool that bridges the gap between technical data analysis and practical medical application.

2. Primary Objectives

Our project's primary objective is to improve histopathological image analysis through an interactive visualization tool designed to enhance diagnostic precision. By incorporating predictions from machine learning models, our tool will enable users to visualize their model outputs with greater control, examining the data from multiple perspectives.

Pathologists and researchers will be able to interact with the data, adjusting parameters, zooming, and filtering to examine specific features. The tool will provide immediate visual feedback, fostering a learning environment. We will apply advanced image processing techniques to improve image quality and feature extraction, revealing patterns and details not easily visible in raw images.

3. Data

We will be using the Kaggle Breast cancer histopathological image dataset (https://www.kaggle.com/datasets/paultimothymooney/breast-histopathology-images/data). Since the data is on Kaggle, it is publicly available.

4. Data Processing

The raw histopathological images will undergo a cleaning process, which includes removing incomplete files and standardizing image dimensions and color scales to ensure uniformity across This design displays the original image as a background with the model's prediction layered on top. A control panel allows users to select different parameters to display various predictions or switch between images. Extract these features, which will be used as overlays to highlight areas of interest in tissue samples. As the processed images will feed into our neural network model, image processing quality will directly impact prediction accuracy, making it a crucial step in our workflow.

5. Visualization Design

a) Alternative design 1: Prediction over the original imaging

This will be the most basic design. By taking predictions (with different model parameters), our visualization will appear as a dashboard, where the main window will display the original image as a 'background', and the prediction on top of the image. There will also be a control panel area on the dashboard, where the users can select different parameters to display different predictions or select between different images.

b) Alternative design 2: Prediction over the adjusted imaging

How we process the image will affect our model performance. On the basis of alternative design 1, we will enable another control panel for users to select different images and process them, and then the predictions and the comparative image should change accordingly.

c) Alternative design 3: Statistics

Browsing through the image list and randomly selecting from it is pointless. To avoid this, we will enable one or more sub-windows in the dashboard to display certain statistics graph(s) (e.g. ROC curve, feature weight, etc) to assist users.

d) General Ideas:

i. Interactive heatmaps over tissue images.
ii. Dynamic filtering of image layers to reveal underlying structures.

e) Prototype Designs:

i. Basic heatmap overlay.
ii. Interactive slider for threshold adjustments.
iii. Multi-layer toggling with feature highlighting.

f) Final Design:

i. Combines the multi-layer toggling from design 3 with the interactive sliders of design 2 for a comprehensive visualization tool.
ii. Justification: This design enables detailed analysis while being intuitive for users of varying expertise.

6. Must-have Features

a) Accurate heatmap generation.
b) Interactive elements for user-controlled visualization.
c) Responsive design for various devices.

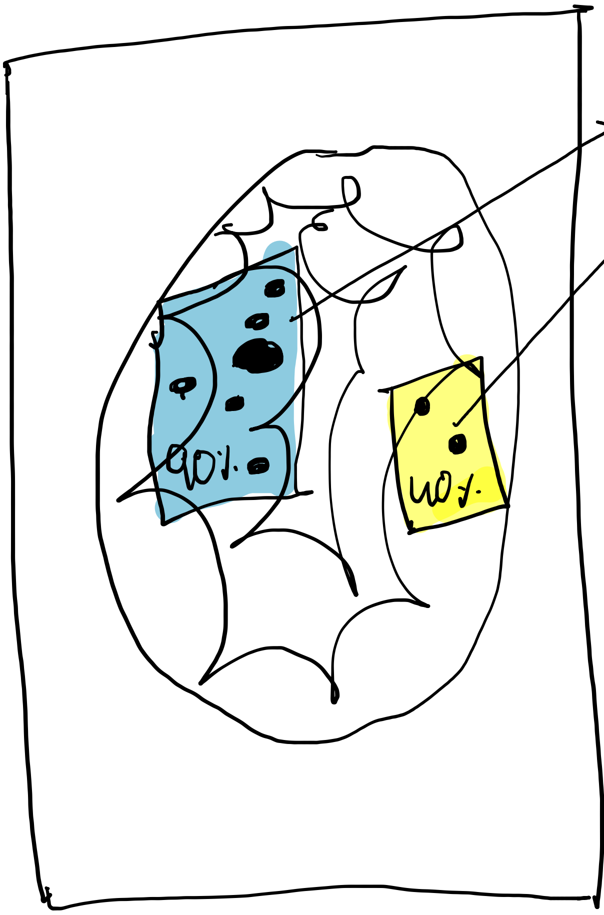
7. Optional Features

a) Upload a different dataset to view in the visualization tool
b) Api to enable users with different models

8. **Project Schedule**

<table>
<tr><th></th><th>Project Plan</th><th>Project Event</th></tr>
<tr><td>Week 4 (Sep 10/12)</td><td colspan="2">Project Proposal</td></tr>
<tr><td>Week 5 (Sep 17/19)</td><td rowspan="2">Data collection and initial preprocessing.</td><td>Staff Review</td></tr>
<tr><td>Week 6 (Sep 24/26)</td><td>Staff Review</td></tr>
<tr><td>Week 7 (Oct 1/3)</td><td rowspan="2">Development of basic visualization interfaces.</td><td>Midterm</td></tr>
<tr><td>Week 8 (Oct 8/10)</td><td>Break</td></tr>
<tr><td>Week 9 (Oct 15/17)</td><td rowspan="2">Integration of interactive features.</td><td>Break</td></tr>
<tr><td>Week 10 (Oct 22/24)</td><td>Milestone</td></tr>
<tr><td>Week 11 (Oct 29/31)</td><td>User testing and feedback incorporation.</td><td>Peer Review</td></tr>
<tr><td>Week 12 (Nov 5/7)</td><td rowspan="3">Potential optional features</td><td></td></tr>
<tr><td>Week 13 (Nov 12/14)</td><td></td></tr>
<tr><td>Week 14 (Nov 19/21)</td><td>Peer Review</td></tr>
<tr><td>Week 15 (Nov 26/28)</td><td rowspan="2">Final adjustments and preparation for presentation.</td><td>Break</td></tr>
<tr><td>Week 16 (Dec 3/5)</td><td>Wrap-up</td></tr>
</table>

Image overlay



Prediction

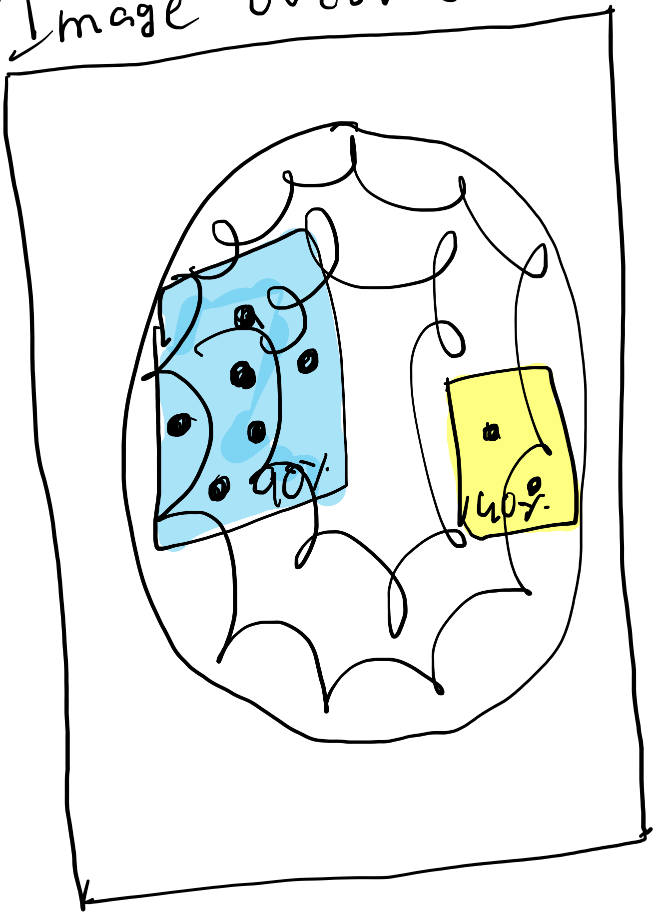
: 90 %

: 40 %

Select image 3

(Can select any image from the dataset to visualize the model performance)

Image overlay



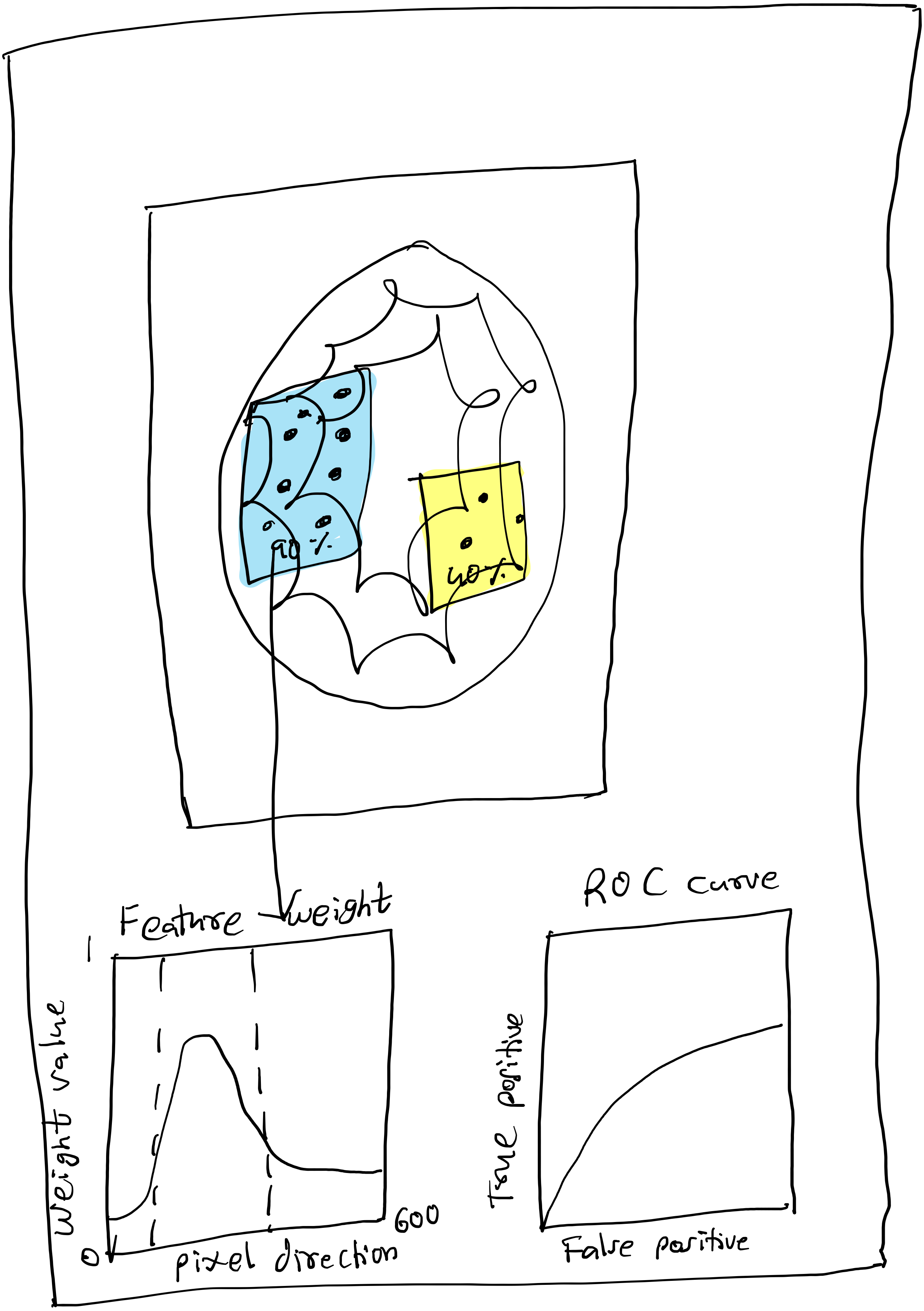
Control pannel

Hue

Light intensity

Grey Scaling

Saturation



90%
40%
Feature weight
1
0
600
Weight value
pixel direction
ROC curve
True positive
False positive