ECE 397: Machine Learning Project Report

Leonardo Lopez
ll10@illinois.edu

Abstract

Over the course of the semester, I primarily worked on developing algorithms to aid in the detection of specific features within liver tissue scans. The two primary features I looked into detecting were fatty liver cells and red cells. For the fatty liver cells, I researched image processing methods that could be applied to the project. Using these image processing methods and manual detection, I was able to develop a set of golden images that could be used in the future for more advanced machine learning models that can detect these fatty liver cells in a wider variety of liver tissue samples. For red cells, I applied similar methods, comparing images to a "golden" image and adjusting threshold masking values to accommodate for the differences in shades of colors between each image. Over the course of this paper, we will go over my findings.

1 Fatty Liver Detection

To develop a set of **golden images**, i.e. the standard images to base future detections on, I had to use some combination of image processing techniques to alleviate the need for brute-force manual detection over the large dataset of liver tissue samples. Luckily, I found a method to detect bullet holes on a target which translated well to the detection of white fatty cells on tissue samples.[1] In order to detect the fatty liver cells, I first created a binary image to remove the non-fatty liver cell components of the tissue sample. Because the fatty liver cells were white, this was rather simple to do. **Using the Python image processing library, OpenCV**, I was able to create a binary representation of the image by applying a mask. [Figure 2] The mask assigns a **1** to RGB values that fall within a certain range of colors, and **0** otherwise. Playing with RGB values, I determined the range to fall between (200, 200, 200) and (255, 255, 255). To get better detections, I had to remove noise from the image and create a clearer, more concise binary representation of the sample. OpenCV contained the necessary functions to do this, and by playing with **kernel values**, I successfully cleared a lot of the noise in the image. [Figure 3] After clearing up the image, I could then extract the **contours**, outlines of shapes within the image. Analyzing each contour, I could determine their "likeness" to the features of a fatty liver cell. More specifically, I looked at the area and "circularity", or shape factor, of each contour. The equation is as follows:

Figure 1: Liver Tissue Sample

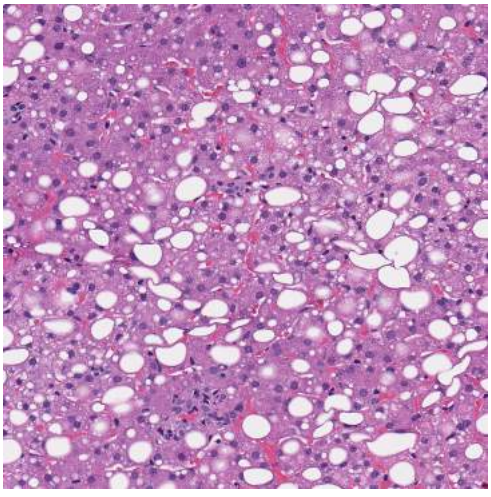
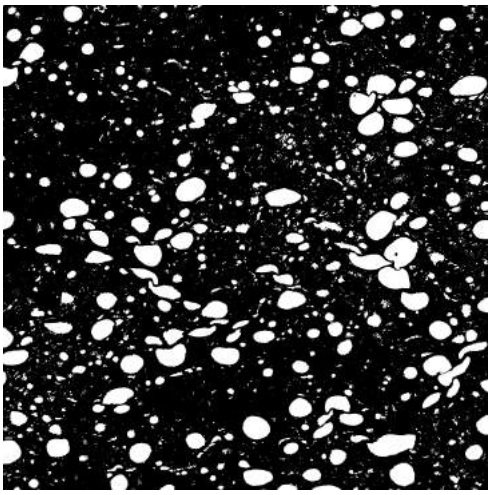


Figure 2: Binary Representation



$$circularity = \frac{4\pi \cdot Area}{Perimeter^2}$$

Where values close to 1 represent a shape similar

to a circle.[2] Using these features, I could detect things that look like fatty liver cells. [Figure 3] However, because we want to identify these cells as best we can, I tightened the conditions of the features, thus decreasing the number of fatty liver cells found, but being much more accurate. The rest of the detections could be done manually by using a simple photo editor, thus giving our end result. [Figure 4]

Figure 3: Noise Removal and Feature Analysis

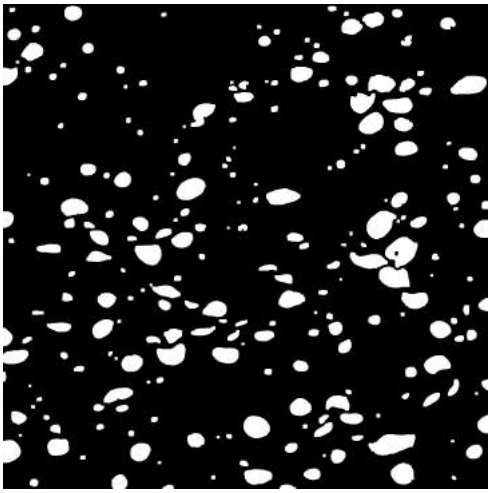
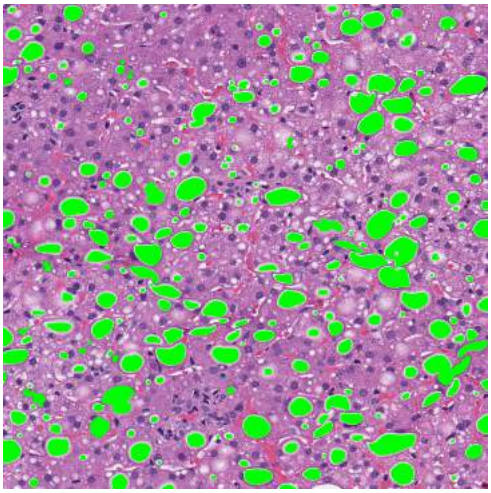


Figure 4: Fatty Detection



2 Red Cell Detection

The second piece of my research project involved segmenting red cells in the tissue samples and analyzing their RGB values. This was done very similarly to the fatty liver cells, with the added complexity of determining the range of red in which to create a masking out of. Finding this range proved to be difficult as the range of colors be-

Figure 5: Golden Tissue Sample

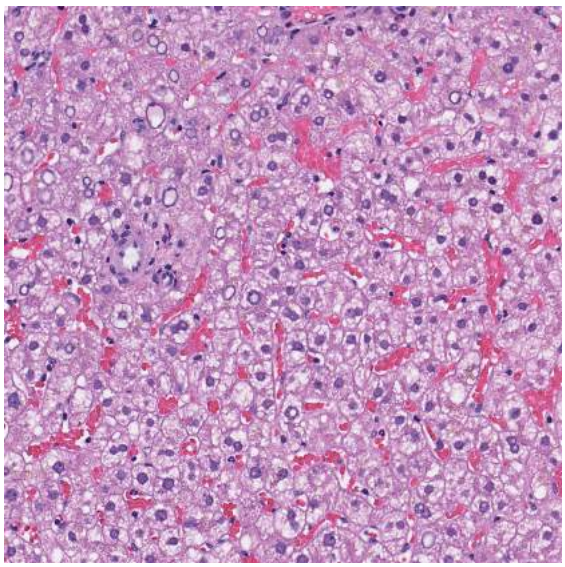


Figure 6: Segmented Red Cells of [Figure 5]

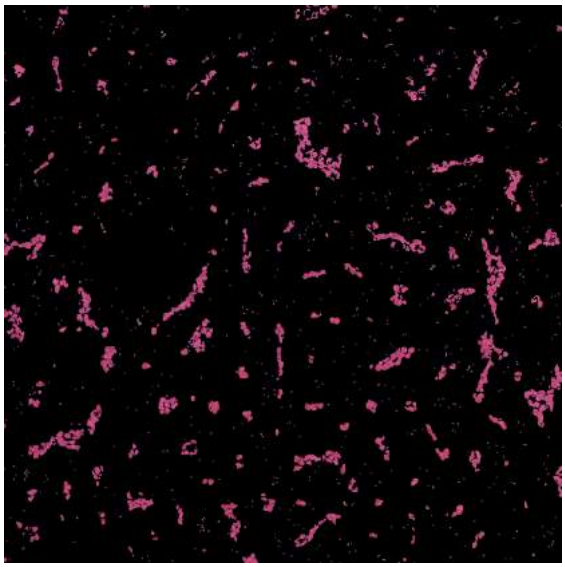
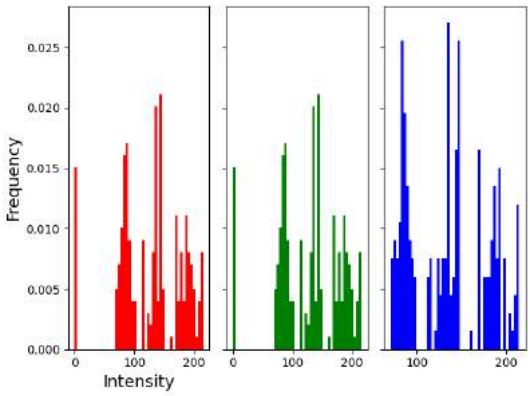


Figure 7: RGB Analysis of [Figure 6]



tween images could vary greatly. That is, some images would be lighter or darker than others. To get a good detection on all images, I first found a range that worked on one image [Figure 5] by playing with values until I was satisfied with the results. [Figure 6] Using this one good detection, I was able to expand the range to other images by comparing their properties to it. I was able to do this by analyzing the average **intensity** of RGB values within each image and comparing them to my one good detection. In order to modify the range appropriately, I simply applied a ratio, scaling the threshold values up or down depending on how intense the colors were compared to my single golden detection. Afterwards, I could analyze and plot the resulting RGB values of the red cells by extracting each color channel from the masked image. The resulting histograms plotted the frequency vs. intensity of each RGB channel seperately. [Figure 7]

2.1 Limitations and Potential Improvements

While this process worked very well for many of the images, the algorithm does struggle on particularly light images. [Figure 8] While I have not examined the causes of poor detections, I can come up with a few reasons as to why these poor detections occur. It is possible that the range needs to be adjusted more strongly depending on the intensity or brightness of the image. Looking at the results [Figure 9] and the image itself, we can see that the red cells are very light and much more close in RGB values to the other components of the tissue sample. A more complex algorithm than applying a ratio may exist to tighten this range better. Another way may be to apply multiple features or increasing the variance between pixels to allow for an easier separation. That being said, these are just ideas. More research into the issue would be required to improve upon the detection.

Figure 8: Lighter Tissue Sample

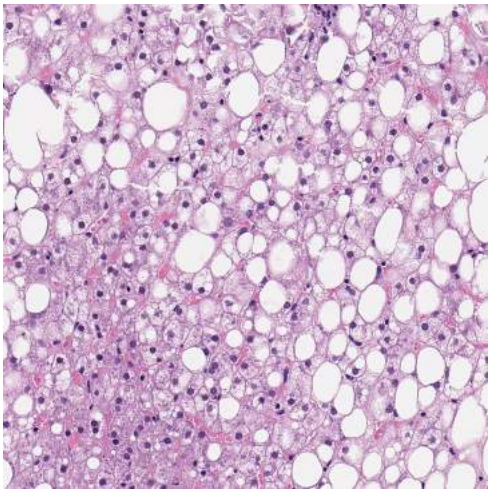
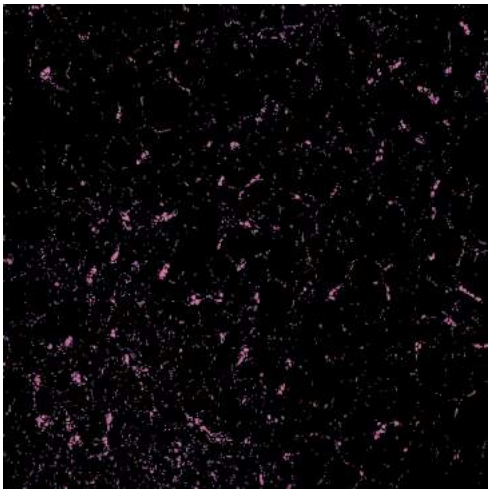


Figure 9: Segmented Red Cells of [Figure 8]



3 Conclusion and Acknowledgments

I had a fun time helping with the detections of specific cells within the tissue samples. I hope the work I've done helps in the future of this project. Through the applications of these techniques, I believe a good machine learning model can be made. Lastly, I'd like to thank Professor Han for allowing me to work with him this Spring semester and Yashuo Wu for her guidance throughout the project.

4 References

[1] "How to detect bullet holes on the target", https://stackoverflow.com/questions/33321303/how-to-detect-bullet-holes-on-the-target

[2] "Shape factor (image analysis and microscopy)", https://en.wikipedia.org/wiki/Shape_factor_(image_analysis_and_microscopy)